NPL
National Physical Laboratory

NPL Report IEA 36

A pilot study into the application of Social Network Analysis on NPL's publication data

Callum Dias

September 2026

National Physical Laboratory (NPL)

A pilot study into the application of Social Network Analysis on NPL's publication data

Callum Dias

We are the UK's National Metrology Institute (NMI), a world-leading centre of excellence that provides cutting-edge measurement science, engineering and technology that underpins prosperity and quality of life in the UK.

ISSN: 2633-4194

https://doi.org/10.47120/npl.IEA36

National Physical Laboratory

Hampton Road, Teddington, Middlesex, TW11 0LW

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

Abstract

This pilot study applies Social Network Analysis (SNA) to the National Physical Laboratory's publication dataset (2000–2024) to better understand internal scientific collaboration, which is an area not captured by traditional citation-based indicators. Using network analysis and its respective metrics, this study maps internal co-authorship patterns across within NPL over five time periods, assessing how connectedness and influence have evolved over time. In addition to this, an assessment can also be made on the distribution of these factors across NPL's nine departments.

Findings show a rise in co-authorship activity over the period, with an increase in interconnectivity across the internal network. What was also seen for NPL across the board was an increase in connectedness, reflecting expanding interdepartmental bridging, while also shifts in the level of influence held by NPL scientists. With regards to the distribution of connectedness and influence, there are significant differences between the authors, particular among less-connected authors. There are notable differences within NPL, with the profile of departments such as AES, MMN, Time, TRM, and CBS indicating distinctive strengths, including high influence or strong brokerage roles.

The study highlights potential implications for organisational resilience, showing how SNA can identify key knowledge brokers, potential vulnerabilities, and links between collaboration structures and innovation outputs. It provides the context and background for deeper analysis in a forthcoming 2026 study.

Contents

With thanks to:

Antoine Vernet, CONA

Thomas Rose, CONA

Carlos Galera Zarco, CONA

Executive Summary

This report presents a pilot study exploring the application of Social Network Analysis (SNA) to the National Physical Laboratory's (NPL) publication dataset (2000–2024). Building on recent bibliometric work by Science-Metrix, this study addresses a gap in NPL's monitoring and evaluation framework: understanding the structure of internal scientific collaboration. While citation-based indicators illuminate research uptake, they give limited insight into how collaboration patterns shape scientific influence, resilience, and departmental performance. SNA offers a complementary, quantitative lens through which to analyse these dynamics.

What is SNA

Social Network Analysis (SNA) is a graph-theory-based methodology that maps and quantifies the relationships between entities to understand how network structures drive outcomes like information flow and organisational efficiency. It utilizes nodes to represent entities and edges to define the varied relationships between them, with these elements visualized through sociograms to map structural patterns, including clusters and influential actors.

With these tools, SNA has a range of applications. It identifies key influencers and strategic "bridges" within a system by using centrality measures to pinpoint individuals who control the flow of information. It maps how resources, knowledge, or even diseases spread through a network, allowing for targeted interventions and improved communication resilience. Furthermore, it can identify structural bottlenecks and isolated "silos".

Purpose and Rationale

NPL's research outputs are central to the impact of the National Measurement System (NMS), requiring robust tools to evaluate performance, with traditional indicators capturing only part of the picture. SNA helps address two longstanding R&D evaluation challenges highlighted in recent DSIT-commissioned work:

- **Intangible benefits** (e.g., knowledge flows, influence structures), and
- **Low observability of collaboration** beyond formal outputs.

By mapping co-authorship networks, this study enables NPL to assess how scientific influence is distributed, how collaboration has evolved over time, and where structural vulnerabilities or strengths lie.

Methodology

The study analyses approximately 24 years of NPL publication data using four core network metrics: Degree, Degree per Paper, Betweenness Centrality, and Eigenvector Centrality. These metrics capture breadth of collaboration, collaboration intensity, bridging roles, and relative scientific influence respectively. To evaluate differences in structure, measures to assess the concentration of connectivity and influence were applied both between and within departments. Results are assessed over five time periods and across nine scientific departments.

Headline Findings

Collaboration Trends: Collaborative activity has grown, as seen in the increase in degree and degree per paper over time, while Betweenness centrality has risen as well. However, between from 2010 to 2024, the growth in these metrics is at a lower rate than NPL's employment growth, which has seen the organisation doubled in size. This has been driven by programmes such as the National Timing Centre and the UK Telecoms Lab, both of which are large-scale engineering projects, unlikely to drive publications and therefore, internal co-authorship. It should be noted that this study doesn't account for external networks, which could be growing during this time.

Concentration level differences: The centrality metrics (Eigenvector and Betweenness) see greater concentration in a few individuals than the degree metrics, suggesting that connectedness and influence is particularly focused in a few people. Within departments, the centrality metrics see particular concentration among authors with less papers, meaning disparities arise mainly among junior, less-connected authors. Differences between departments are notable, with some exhibiting high influence, while others act as key knowledge bridges with high Betweenness.

Link to Performance Metrics: Betweenness centrality shows a positive relationship with Collaborative Innovation output, underscoring the importance of connectedness for research productivity. No comparable relationship was found for Measurement Services, indicating that academic collaboration structures are largely distinct from commercial delivery.

Operational Implications

The findings can help contribute to addressing organisational challenges:

- **Resilience issues:** Departments with high levels of concentration may be vulnerable if key individuals leave, while identifying opinion leaders can support targeted communication and direction-setting.
- **Principal-agent challenges:** SNA increases transparency in collaborative behaviour, helping address information asymmetries typical in public R&D governance.

This pilot demonstrates that SNA offers a valuable, complementary perspective on NPL's research activities. It provides quantifiable insights into collaboration patterns, structural strengths, and potential vulnerabilities that traditional metrics overlook. However, it should be highlighted that a department's score (and its relative strength) shouldn't be seen merely as a good or bad thing. It is merely a representation of its activity and can be explained by a range of different factors.

While limited by reliance on publication data and the absence of a peer baseline, the analysis establishes a strong foundation for future studies. A follow-on report by UCL CONA in 2026 will extend this work with deeper methodological detail and case studies. Together, these analyses will strengthen NPL's evidence base and support more targeted, resilient and effective research management.

Introduction

The report details a pilot study into the use of Social Network analysis (SNA) on NPL's publication data since the turn of the century. This follows on from Yildiran & King (2025), which saw Science Metrix commissioned to run a bibliometric study on NPL's publication output. This study scored NPL's papers on a range of measures, providing useful information but one dimension that was neglected was the structure of NPL's co-authorship network. This paper uses the data from the above paper to analyse this network at an NPL level, while also running the analysis at a more granular level. This section will provide the rationale for why NPL needs to assess its science-based output, why this method goes beyond traditional bibliometric measures and will also provide context for the rest of the report.

What is SNA?

Social Network Analysis (SNA) is a methodological approach that examines the relationships and interactions among individuals, groups, or organizations. As outlined in Wasserman & Faust (1994), SNA is rooted in graph theory and conceptualizes social structures as networks composed of **nodes** (actors) and **ties** (relationships), enabling researchers to visualize and quantify patterns of connectivity. This approach provides insights into how relational structures influence outcomes such as information dissemination, organizational efficiency, and social influence.

Key Concepts of SNA

- **Nodes:** Nodes represent the entities within a network, which may include individuals, teams, organizations, or even technological units such as computers (Scott (2017)). For example, in an organizational network, employees are nodes connected by communication ties.
- **Edges/Ties**: Ties denote the relationships or interactions between nodes, such as friendships, collaborations, or transactions. These ties can be **directed** (e.g., one-way communication) or **undirected** (e.g., mutual friendship), and vary in strength (Granovetter, 1973).

- **Visualization**: SNA often employs **sociograms**[1], graphical representations where nodes are depicted as points and ties as lines. Freeman (2000) saw how visualization tools help identify clusters, bridges, and isolated actors, making complex networks more interpretable.

SNA Metrics

Along with the sociograms, SNA also generates metrics to assess the importance of nodes within a network. For this report, three metrics were used that track different things in different ways:

- **Degree** measures a node's influence by calculating its total number of direct connections, identifying the most "well-connected" actors within a system. In an undirected network, such as co-authorship, this metric quantifies mutual relationships.
- **Eigenvector centrality** determines a node's influence by prioritising the "quality" of its connections, meaning a node gains more prestige when it is linked to other highly connected actors. This recursive approach distinguishes it from connection counts like degree, making it an essential metric for identifying key hubs and high-status influencers within a network.
- **Betweenness centrality** identifies nodes that function as strategic bridges or "gatekeepers" by measuring how often they lie on the shortest paths between other actors. High scores in this metric reveal individuals who control the flow of information and resources, maintaining structural importance even if they have a low number of direct connections.

What SNA Reveals

Identifying Key Influencers

Centrality measures - such as degree, betweenness, and eigenvector centrality - identify actors who hold strategic positions within a network. For instance, as in

[1] A sociogram is a graphic representation of social links that a person has. It is a graph drawing that plots the structure of interpersonal relations in a group situation.

Borgatti et al. (2009), individuals in a corporate setting with high betweenness centrality often act as bridges connecting otherwise disconnected departments.

Tracing Information Flow

SNA details how information, resources, or even contagions spread through networks. Public health researchers, for example, use SNA to model disease transmission pathways and design targeted interventions (Valente, 2010).

Revealing Network Structure and Bottlenecks

Network metrics such as density and clustering reveal structural properties, including formal hierarchies and informal subgroups. Cross et al (2002) identified that finding bottlenecks (nodes that control critical pathways) can inform strategies to improve communication and resilience.

Uncovering Inefficiencies and Silos

SNA highlights isolated clusters or “silos” that impede collaboration. In organizations, addressing these inefficiencies can foster cross-functional teamwork and innovation (Krackhardt & Hanson ,1993).

How to assess NPL’s science outcomes: As with any form of good management practice, assessing the scientific outcomes of research organisations would be vitally important to assure that an organisation is operating well. This could take the form of internal or external analysis, with the University of Birmingham’s Research Portal and the Research Excellence Framework for Higher Education institutions as good examples respectively. This provides accountability for public funding and assures that benefits are being generated by the investment. Although NPL’s main measure is through its economic impact channelled through its userbase, that impact originates in the scientific research needed to support and develop the labs that comprise the National Measurement System (NMS) programme[2]. This programme is designed to provide “the UK with an infrastructure of laboratories that deliver world-

[2] https://www.gov.uk/government/publications/national-measurement-system/uk-national-measurement-system

class measurement science and technology and provide traceable and increasingly accurate standards of measurement." As with any significant government programme, a monitoring and evaluation (M&E) plan is essential to ensure that taxpayers see value for money. The Magenta Book (2020) provides an outline for all programmes, but the NMS has its own evaluation framework, as detailed in King et al (2023).

This report is the Evaluation Framework for the NMS as a whole, of which NPL is the majority recipient of funding. As part of this framework, the logic model for the NMS details four main impact mechanisms:

1. Investing in world-leading **research** into the field of metrology alongside other areas of measurement science.
2. Accelerating the **innovation** activities of other organisations through engaging with them in deep collaborative R&D projects.
3. Underpinning the national quality infrastructure through delivering unique measurement services that provide the technical basis for **trade and standards** within the UK.
4. Helping organisations get closer to the technological frontier by fostering the adoption of best practice through a range of active **knowledge transfer** activities.

Each of these mechanisms are important, and link to the NMS's business case, but the one to focus on for this study is **Research.** The monitoring activity for this theme sees a range of tools listed to best assess how the NMS labs do with regards to this theme, such as indicators, case studies, peer review and interviews. When looking at indicators, there are three primary measures:

- 1.1 No. of Academic Collaborators
- 1.2 No. of peer reviewed papers
- 1.3 Citations

Issues with Citation scores: The first two metrics are useful when assessing Research activities and outputs, but the third involves far more complexity. By their nature, citation scores attempt to reduce the multifaceted concept of a research paper's impact or quality into a single number, but this leads to various biases, contextual limitations, and challenges in their calculation and interpretation. These can include:

- Field Variation: Citation practices vary widely across fields. STEM fields often generate more citations and have shorter "citation half-lives"[3] (the time it takes for citations to accrue) than humanities or social sciences.
    - To allow for fair comparisons across different fields, citation counts must be "normalized" against field averages. These normalized scores (e.g., Field-Weighted Citation Impact or FWCI) are more statistically robust but less intuitive and too opaque for many people to interpret.
- Time Lags: A paper's impact can take years to become fully apparent, and metrics that use a short window (e.g., a two-year impact window) may undervalue "slow burn" research.
- Nature vs. Impact: A high score indicates a work has *attention* or *influence*, not necessarily that it is high-quality, solid, or positively received. Papers can be heavily cited because they are controversial, flawed, or used as an example of a "bad actor".
- "Matthew Effect"[4]: Already famous researchers or journals tend to attract more citations, amplifying existing status regardless of the specific paper's quality.
- Data Sources: Different databases (e.g., Scopus, Web of Science, Google Scholar) have different coverage, especially for non-English language publications or specific disciplines, leading to different scores for the same work.

The Science-Metrix study: Given the above, it was decided that the citation analysis should become a broader study. This was commissioned and written up by NPL, but the analysis was done by Science Metrix. This saw NPL (and similar international National Measurement Institutes[5]) assessed on a range of different measures, which included FWCI but also looked at International Collaboration Rate, Growth Rate in publications and alternative measures of uptake[6]. The findings from this study were very useful in assessing what NPL's outcomes were with regards to

[3] https://clarivate.com/academia-government/blog/a-closer-look-at-cited-and-citing-half-lives/
[4] https://www.sciencedirect.com/topics/psychology/matthew-effect
[5] The five NMIs primarily compared were the National Institute of Standards and Technology (NIST - USA), National Metrology Institute of Germany (PTB), Korea Research Institute for Standard and Science (KRISS), the National Institute of Metrology of China (NIM) and NPL.
[6] These include Normalised Share of Publications Cited by Patents, the Private sector and by policy literature respectively.

Research, but what isn't considered is the relationships between the co-authors in the publications.

Importance of collaboration: As detailed in a range of prior papers (with Veugelers (1998), Annique Un et al. (2010) and Annique Un et al. (2014) to name a few), there are significant benefits to collaboration in the R&D space. This would suggest that monitoring NPL's collaborations would be incredibly useful as a tool to better understand patterns in its research output. To do this, following discussions with Department of Science, Innovation and Technology (DSIT) analysts[7], the use of Social Network analysis (SNA) was suggested as a tool to enable greater understanding of the nature of our collaborations.

Working without a baseline: One flaw this study does have is the lack of a baseline. Without equivalent data from similar organisations, the analysis cannot produce a real benchmark to compare to. To account for this, this paper is primarily focused on a comparison of the departments over time, rather than NPL as a whole. This is done by both comparing the headline metrics and by using inequality measures to assess the concentration of the metrics' scores in a few individuals. These include measures like Gini to compare across the different metrics, and Theil T and L to assess within- and between-groups across the top and bottom of the distribution of the metrics. A future report, developed by the analysts who ran the study, will endeavour to develop their own baselines.

Concentration level: Assessing concentration will allow for a range of aspects to be considered. For example, higher concentration in NPL's co-authorship networks suggests dependence on a few key individuals, while lower concentration indicates more distributed influence and connectivity. This matters because opinion leaders - central, trusted figures who shape others' views - play a major role in directing scientific activity. Departments with greater concentration therefore face resilience risks if these individuals leave, yet they may also be easier to redirect strategically and have a stronger unique identity.

[7] The NMS programme is delivered through DSIT, which provides funding and oversight to the NMS labs. DSIT is also the sole shareholder of NPL, which is listed as a public corporation.

Assessing NPL lifecycle: Fundamentally, this study is positive in nature; it looks to report and detail the structure of NPL's internal co-authorship network as it has developed over the past 20 years, going deeper than standard organisational structures. As part of this, the departments can be seen as existing along a spectrum of diffusiveness, with some departments acting as key centres of knowledge, while others are far more dispersed. This ties into the age of fields at NPL as well, with older, more-established areas able to coalesce around themselves, while younger areas must work more broadly to start with. In turn, the impact of age and diffusiveness leads to the following trade-off:

- Older, less diffuse departments tend to have a strong identity, while lacking broader links to the rest of NPL.
- Younger, more diffuse departments tend to have strong links to the rest of NPL, while lacking an identity unto themselves.

Broader operational uses: Beyond assessing the trade-off detailed above, the information generated by SNA has further benefits as well. Given the fact that the outputs of SNA produce maps of NPL's internal collaborative network, it allows for management to see how NPL's researchers engage with one another, both now and in the past. The network metrics can also be analysed with NPL's output components, developed in Nayak et al (2014), assessing the relationship between connectedness and NPL's output types. Bringing this all together allows for two main potential issues to be considered, the principal-agent problem in R&D activity and the need to improve resilience in the science areas. To help solve these issues, this study could provide useful information to support strategic decision-making on problems such as succession risk, collaboration planning and research strategy.

Limitations: It is important to highlight what this study doesn't do as well. SNA is a valuable but limited tool, which can highlight patterns and trends in organisational collaboration, yet it cannot uncover the underlying causes. This requires deeper statistical work and discussions with research leaders. This is particularly important when accounting for lower and higher scores. These shouldn't be regarded as a weakness or strength, rather they represent the work done by a department which may require different levels of collaboration over time. Specific to this study, the analysis relies solely on NPL publication data, which captures only a fraction of real collaboration. Most R&D never becomes published work. Furthermore, other forms of collaboration are also absent, as are external co-authors. These omissions mean the network offers only a partial picture, though it still provides a useful, unbiased check on anecdotal understanding.

This paper will continue as follows: Section Two will provide important context, both on SNA generally and why it's useful as part of R&D evaluation. Section Three will detail the data used in this report and the approach used as part of the analysis. Section Four will detail the results from the headline stats and results, while Section Five will look at the departmental breakdowns, both the main results and the concentration analysis.

Why use it for R&D evaluation

The use of this methodology was initially detailed in Rosemberg et al. (2024), with further details found in Technopolis (2014). The report *"What Methods Work for Evaluating the Impact of Public Investments in RD&I"* was commissioned by DSIT to address a critical need for robust evaluation methodologies in R&D policy. Traditional evaluation approaches often struggle with challenges that come with R&D activity such as long lags between investment and impact, attribution complexity, and intangible knowledge flows, so this study aimed to synthesize international best practices and provide guidance for designing credible, context-sensitive evaluations of R&D programmes.

Key Findings

The study reviewed 107 evaluation reports from advanced economies and conducted detailed case studies on over 20 examples. Its findings highlight several persistent challenges in RD&I evaluation:

- **Lagged effects of research investments**
    - The benefits from the intervention can often take several years to appear

- **Intangible benefits**
    - Interventions may generate assets that may be difficult to both measure or quantify (such IP, improved reputation or increased networks)
- **Low observability of research impacts and knowledge flows**
    - As a specific extension of the prior way, the impact from the research-based outputs of the intervention is often difficult to estimate.
- **Skewness of impact**
    - Estimated outcomes may be heavily influenced by a small number of very successful projects, not accounting for the large number of failed projects, which can be expected in R&D activity.
- **Additionality/attribution and contribution**
    - Determining an intervention's additional impact beyond what would have occurred anyway, attributing those outcomes to specific interventions amid many influencing factors and understanding the specific role of an intervention within a broader context of contributing elements are very difficult when there is no natural counterfactual.

The report concludes that there is no single "silver bullet" for RD&I evaluation. Instead, high-quality evaluations require a pluralistic approach that blends multiple data sources and methods, tailored to programme context. Essential components include a clear Theory of Change, mixed-methods design, counterfactual analysis, and transparent reporting for diverse stakeholders.

Bringing this back to NPL, the prior challenges outlined should be considered in more detail, which that report does. Technopolis (2014) goes this in detail but the issues that are most pertinent for NPL are *Intangible Benefits* and *Low observability of research impacts and knowledge flows.*

As currently constructed, the NMS (and NPLs) current evidence portfolio is made up of primarily economic evidence in a range of different forms. This includes econometric analysis of supported firms, macroeconomic analysis on the impact of the UK's measurement infrastructure and a range of other quantitative studies. There are also a range of surveys, quantitative case studies and other piece of analysis to complement the primarily economic evidence portfolio. This is in-line with the NMS's evaluation framework, which is detailed in King et al (2023), and details the ideal shape of the evidence portfolio.

Currently the portfolio is well placed to account three main factors:

Additionality, Attribution & Contribution

- King & Olakojo (2023) accounts for these by using quasi-experimental methods, which Rosemberg et al (2024) outlines as an approach to account for additionality, attribution & contribution.
- King & Nayak (2025) utilises a macroeconomic approach but approaches the cost-benefit analysis by accounting for a realistic scenario in two main ways. Firstly, the lower bound of benefits is utilised as the expectation is that the UK would take traceability from another NMI, rather than deprive itself of access to an NMI. Secondly, the return generated by the NMS is split into two forms:
    - Average rate of return – which is calculated if the NMS was scrapped in its entirety
    - Marginal rate of return – 66% of the average rate of return, which is calculated if the NMS sees a proportion of its funding reduced.
        - Given the longevity of the NMS, we do not assume that it would be scrapped in its entirety and even if it was, we assume that the UK would take traceability from another country.
        - By doing so, we can assess the benefits that the NMS's measurement infrastructure in a manner that makes them attributable to itself and are additional to the programme.

Skewness of Impact

- King & Olakojo (2023) accounts for this by assessing average number of new jobs generated by NMS supported firms and the average wage premium paid to incentivise job-switchers.
- King & Nayak (2025) accounts for this by looking at an economy-wide model, accounting for company-level differences.

Lagged effects of research investments

- Following evidence from our most recent customer survey (Katanguru et al, 2025), we have found that products created by firms supported by the NMS generate benefits for seven years.

However, the remaining two factors are not accounted for in the same way. The evidence portfolio only has a few examples of the methodological approaches outlined in Rosemberg et al (2024) tagged to *Intangible Benefits* and *Low Observability*, both of which were published recently. These two reports (Jiang et al (2025) and Yildiran & King (2025)) are examples of bibliometric analysis that address these factors, but this can only be viewed as a start. Furthermore, these factors require greater variety in approach to better account for them, which is what provoked an investigation into the potential for a social network analysis (SNA) project; particularly given the fact that SNA covers both above factors. On top of this, the development of those bibliometric papers saw the construction of a joined-up dataset which could be used for SNA. This will be detailed later, but its existence made an SNA project more feasible.

One example of the use of SNA in public R&D evaluation was seen in the assessment of the European Cooperation in Science and Technology programme (COST). This was an intergovernmental initiative established in 1971, which aims to foster research networks across Europe and beyond, with a strong emphasis on inclusivity, interdisciplinarity, and support for young researchers.

The evaluation of the COST (European Cooperation in Science and Technology) found difficulty in observing the intangible and often informal outcomes of research collaboration, such as the formation of new professional relationships and the evolution of research networks. To address this, the evaluation team used SNA to map and analyse the structure and dynamics of the networks formed through COST-funded activities, including meetings, training schools, short-term scientific missions (STSMs), and conference grants.

The use of social network analysis was found to be incredibly useful when trying to co-ordinate a range of networking events (meetings, training schools, STSMs, etc) into a single dataset, while applying the metrics to the econometric analysis. It should be noted that this evaluation was incredibly resource-intensive and relied heavily on secondary monitoring data for it to be feasible, which is expected when evaluating a programme with the complexity and scale of COST, though not feasible for the NMS in the same way at this point in time.

The study also finds that this model can be applied in the UK, where multiple R&D programmes may also shape existing research networks. Focusing on how a specific programme alters network characteristics compared with normal patterns can reveal its effectiveness, especially when complemented by counterfactual analysis or in-depth case studies. It also provides examples of potential data sources[8] which could support detailed examination of network structure and composition.

Additional benefits to SNA

Resilience benefits

One major area of focus for NPL's management concerns the need to plan for the future, particularly concerning the senior scientists. Generally, organisations need to plan for senior staff leaving because these roles often hold critical knowledge, strategic insight, and key relationships that are not easily replaced. Without succession planning, departures can disrupt operations, delay decision making, and weaken organisational stability. Proactive planning helps maintain continuity, preserves institutional knowledge, and ensures that capable people are ready to step into leadership roles. This strengthens resilience, allowing the organisation to adapt smoothly to change and continue delivering its goals without interruption

This is even more important for R&D institutes and for metrology. Due to the nature of metrology research, the knowledge to operate instrumentation, the development of new methods and the realisation of units that have been built up in senior scientists who have been working in their respective fields for long periods of time. As NPL is the UK's NMI, there are very few other places in the UK that offer space for metrology research, which leaves NPL particularly vulnerable when compared to universities for example. Broader academia can find replacements for older faculty to fill gaps in their ranks, but as the market for metrology scientists is so small[9], the

[8] Researchfish is an online platform used by research funders and organisations to track the outputs, outcomes, and impact of their research investments.
Gateway to Research (GtR) is a public web portal and database developed by UK Research and Innovation (UKRI) to provide transparency and open access to information regarding publicly funded research and innovation in the UK.
[9] Most countries only have one NMI, limiting the size of the metrology field, though there are some notable exceptions. For example, France has one NMI and 9 designated laboratories. However, even the largest organisations are small when it comes to universities. For example, the five NMIs assessed in Yildiran et al (2024) employ between 500-4000 people (of which scientists are a

need to account for potential departures is far more important. While these individuals are the backbone of NPL's operations, the organisation must proactively plan for succession and potential turnover, whether due to retirement or career progression.

Currently there is work being done to account for and mitigate against scientists leaving but this study would provide a range of different tools to support this work, assessing which individuals are key nexus points, while also assessing which areas are more vulnerable to specific people leaving the organisation than the other. Following up on this last point, the ability to break down the data by department will allow for comparisons, measure across a range of different measures of connectivity, which is possible thanks to the range of metrics detailed below.

Principal-Agent benefits

The other potential area which SNA can benefit concerns solutions to the principal–agent (P-A) problem in the public sector. This arises when government bodies, as principals, delegate responsibilities to agents whose objectives and information differ from their own. This dynamic is especially pronounced in public R&D systems, where uncertainty, long time horizons, and opaque research processes exacerbate information asymmetries. Huffman & Just (2000) demonstrate that public funders cannot observe researchers' true effort, quality, or innovative capacity. Furthermore, research outcomes depend on scientists' best outputs rather than average performance, principals face significant monitoring challenges. They also show that scientists tend to be more risk averse than administrators, and that different funding mechanisms (such as formula funding or competitive grants) create distinct incentive distortions. These factors make it difficult to design contracts that reliably elicit optimal research behaviour.

Further evidence of P-A issues arise when looking at research governance. Braun & Guston (2003) highlight that public R&D operates through "nested principal–agent relationships" spanning several different management layers, such as ministries, funding agencies, universities, and laboratories. At each layer, information is lost and

proportion of that). In comparison, University College London employs 1200 people in its *Faculty of Maths & Physical Sciences* alone.

incentives are reshaped. Because policymakers cannot directly observe creativity or effort, they rely heavily on ex post indicators such as publications or patents, a dependence that risks steering researchers toward easily measured outputs at the expense of broader scientific impact.

Across these examples, asymmetric information remains the central challenge: researchers invariably know more about their work than administrators can observe. While social network analysis (SNA) cannot eliminate this asymmetry, it can mitigate its effects by making collaborative networks more transparent. By mapping collaboration structures, identifying bottlenecks, and revealing influential actors, SNA can strengthen monitoring, improve information flow, and support better-aligned incentives, thereby helping public R&D function more effectively despite inherent uncertainties.

Data & Analytical Approach

Analysts

As NPL is inexperienced in SNA projects, the academics at UCL's Centre of Organisational Network Analysis (CONA) were brought in to run the analysis and provide the initial results. They will also be writing up their own analysis soon, with their report far more on the approach, tools and findings, while also doing some specific case studies (such as departmental deep-dives and dependency analysis).

Key Definitions

In order to assess the interconnectedness of the network, a series of metrics will be calculated for each of the scientists within the network, but before this, two important definitions are needed

- Node – an entity within the network being analysed
    - For this analysis, a node is an NPL author.
- Ties/Edges – the line that connects two nodes, representing their relationship
    - For this analysis, an edge is the paper that respective co-authors share.

Having defined the two key concepts above, there are three main metrics that have been generated as part of the analysis.

Degree

Degree[10] refers to the number of direct connections (edges) a node has to other nodes. It is the simplest and most intuitive measure of how "central" or "well-connected" a node is in a network. Nodes with higher degree are typically more influential because they have more direct relational ties.

[10] https://cambridge-intelligence.com/keylines-faqs-social-network-analysis/

Degree is calculated differently depending on whether a network is directed (where edges are single direction) or undirected (where edges are multi directional). Given the nature of co-authorship on a paper, this network is undirected.

The formal mathematical definition of degree centrality[11] for a node i uses the adjacency matrix $\mathbf{A}$, where:

- $A_{ij} = 1$ if there is an edge between nodes i and j
- $A_{ij} = 0$ otherwise

For an undirected network, degree centrality is:

$$C_D(i) = \sum_{j=1}^{n} A_{ij}$$

This formula simply sums all edges connected to node i. For this analysis, the degree is the number of co-authors each author has on their papers in the period.

Eigenvector centrality

Eigenvector centrality is a network measure that identifies the most influential nodes by considering both the number of connections a node has and the importance of the nodes it is connected to. Unlike degree centrality, which simply counts connections, eigenvector centrality is recursive; a node is important if it is connected to other important nodes.

This makes it especially useful for identifying influencers, prestige nodes, or hubs in social, economic, or information networks. Effectively, this focuses on the “quality” of the connections that an author has; does a respective author work directly with people who also have high numbers of co-authors. A variant of eigenvector centrality, PageRank, is used to identify the importance of webpages in Google searches[12].

[11] https://bookdown.org/omarlizardo/_main/4-2-degree-centrality.html
[12] https://cambridge-intelligence.com/eigencentrality-pagerank/

As outlined here[13] the mathematical definition of eigenvector centrality c for node i is:

$$\lambda c_i = \sum_j g_{ji}\, c_j$$

Where:

- g_{ji}— Entry of the **adjacency matrix** G.
- c_i— **Eigenvector centrality score** of node i.
  Represents how influential or important node $i$is in the network.
- λ— The **largest eigenvalue** (principal eigenvalue) associated with the adjacency matrix.
  This scaling constant ensures the set of $c_i$values is consistent and non-trivial.
- j— Index representing all nodes in the network whose edges point **toward** node i.

The centrality of node i equals a constant (λ) times a weighted sum of the centralities of all nodes that point to it. That is, the centrality of each node i is proportional to the sum of the centrality of its neighbours.

Betweenness centrality

Betweenness centrality is a measure of how often a node lies on the shortest paths between other pairs of nodes in a network. It captures the extent to which a node acts as a bridge, broker, or gatekeeper in the flow of information, resources, or influence.

A node with high betweenness centrality:

- Controls pathways between others
- Can influence communication or information flow
- Is structurally important even if it has few direct connections

The concept was formally defined by Freeman (1977) and is widely used in sociology, network science, biology, infrastructure networks, and communication studies.

[13] https://ocw.mit.edu/courses/14-15-networks-spring-2022/resources/mit14_15s22_lec3/

As outlined in Barthelemy (2022), the mathematical definition of Betweenness Centrality $g(i)$ is as follows:

$$g(i) = \frac{1}{\mathcal{N}} \sum_{s \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}}$$

Where:

- $\sigma_{st}(i)$ is the number of shortest paths going from node s to node t through node i.
- σ_{st} is the number of shortest paths going from node s to node t.
- $\mathcal{N}$ is a constant which counts the total number of pairs in the graph (=(N-1)(N-2), where N is the total number of nodes in a network).

Betweenness Centrality vs Eigenvector Centrality

When considering both Eigenvector and Betweenness Centrality, the following comparisons can be made:

	Betweenness Centrality (BC)	Eigenvector Centrality (EC)
What it measures	A node's role as a bridge or intermediary on the shortest paths between other nodes	A node's overall influence based on the influence of its neighbours.
Focus	Quantity of shortest paths a node lies on	Quality of a node's connections
How it works	A node with high BC acts as a bottle neck for information flow between different parts of the network	A node connected to other highly-connected and influential nodes will have a higher score.
Example analogy	The "gatekeeper" or "connector" who is essential for communication between specific groups.	The "influencer" who is connected to other important influencers, making their opinion valuable.

Key Characteristic	High scores indicate a node is crucial for connecting disparate parts of the network.	High scores indicate a node is connected to other important, high scoring nodes
Underlying concept	Measures how often a node lies on the shortest path between two other nodes.	Measures the relative importance of a node within the network based on its connections to other important nodes.

Table 1- Comparison of Betweenness and Eigenvector Centrality

Data

The data is the same dataset used in Yildiran (2024), Yildiran and King (2025) and Jiang et al (2025). These saw an assessment of NPL's bibliometric output and a comparison across the world's national measurement institutes (NMIs).

The data that sat behind those analyses drew on three major bibliometric databases[14] which together provide extensive coverage of global research outputs. A complete list of NPL publications from 2000–2023 was assembled by merging records from these sources and removing duplicates using Digital Object Identifiers (DOIs). These DOIs were then used to evaluate coverage and metadata consistency across the databases, after which Dimensions was queried to retrieve key publication information such as citation counts and journal titles.

The data also had to be indexed by NPL author, which was doable with what was provided, but required rigorization to prevent double counting within the data. This allowed for the classification of papers by NPL's departments, which are defined in Nayak et al (2024).

Accounting for random connections

In terms of the approach, a combination of descriptive statistics (including measures of centrality) coupled with comparing observed networks to appropriate random

[14] Web of Science, Scopus, and Dimensions.

networks were used to estimate whether observed values were due to properties of the observed network or arising naturally from the structure of (similar) networks.

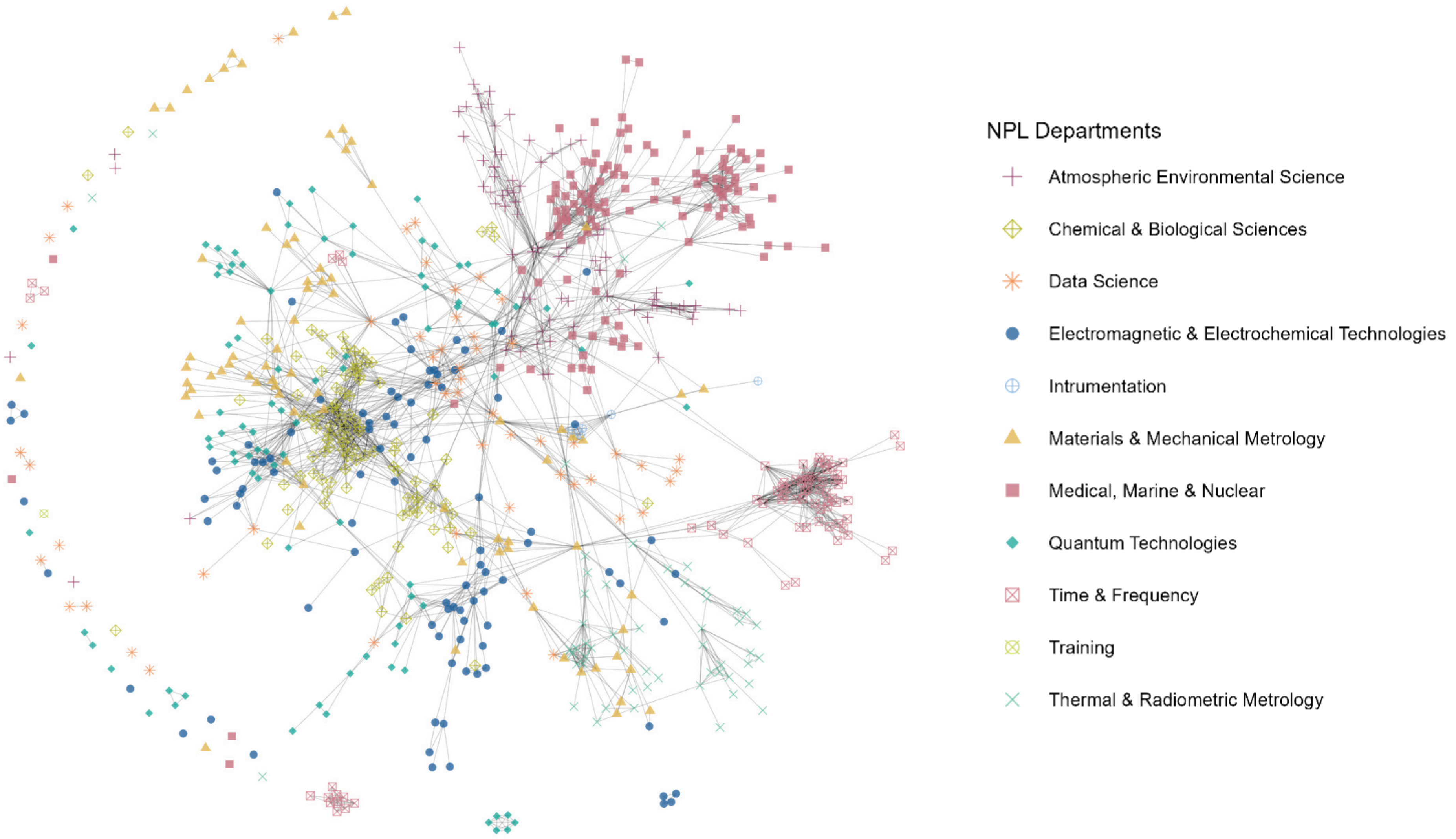
For example, by using random networks with the same degree distribution as the observed network, an estimate can be made concerning whether properties of the network are likely due to the behaviour of the individuals involved or if the properties of this network are a result of the structure of the network (in this case – size, density, and degree distribution).

Adjusting the Degree metric

Beyond the three-headline metrics that have been outlined above, an adjusted version of degree has also been developed, which has been done to account for publication quantity. For degree, there is potential for authors who publish many papers would see their degree value boosted through output of papers, not by number of collaborations. To account for this, **degree per paper** was developed, this sees the degree metric divided by number of papers by the same author in the same period. This accounts for authors who have higher authorship numbers, scaling the number of connections to the number of papers.

The map

Figure 1 - Full co-authorship network for NPL between 2020 and 2024



The aforementioned future study led by CONA will dive into the graphs and maps, but the focus of this study will be on the metrics as they can often be under-utilised.

Lack of a baseline

The biggest difference between our study and the COST study detailed before is the lack of a baseline, as NPL has never done this before and there is no equivalent data from a comparable organisation currently available. Despite this, there are ways to assess the data without a traditional base line. This can be done in two main ways

- By Period (2000-2004, 2005-2009, 2010-2014, 2015-2019, 2020-2024)
- By Department (as defined in Nayak et al (2024):
    - Atmospheric Environmental Science (aes)
    - Chemical & Biological Sciences (cbs)
    - Data Science (ds)
    - Electromagnetic & Electrochemical Technologies (em)
    - Materials & Mechanical Metrology (mm)
    - Medical, Marine & Nuclear (mmn)
    - Quantum Technologies (qt)
    - Time & Frequency (time)
    - Thermal & Radiometric Metrology (trm)

Beyond just comparing the metrics across the above dimensions, economic inequality measures have also been used to assess the differences in concentration between and within the departments as well.

Resilience vs Mobility in the Co-Authorship network

For NPL, concentration in its co-authorship network has significant implications, particularly concerning resilience of the departments:

- If there are higher levels of concentration in a department, it can be assumed that that department's network is more dependent on specific individuals.
- If there are lower levels of concentration, it can be assumed that the department is less dependent on specific individuals.

The importance of concentration in this context concerns the ability of individual scientists to influence their respective department. This falls in line with Opinion Leadership Theory, which originates from the landmark research of Lazarsfeld, Berelson, and Gaudet (2021)**.** Their study showed that ideas and influence often spread not directly from mass media to the public, but through *key individuals* ("opinion leaders") who interpret, filter, and relay information to others. This approach was applied to academia in Doehne & Herfeld (2023), who show how early adoption of the *Theory of Games* at the Cowles Commission was driven not only by scientific merit but also by the influence of academic opinion leaders embedded in organisational hierarchies. Their analysis of 488 papers' acknowledgements reveals that figures exerted disproportionate sway because they were both scientifically respected and organizationally central, enabling them to steer research agendas and shape intellectual trajectories.

Given this, the importance of identifying key individuals is important to help shape and direct a scientific programme but should be noted as one of many different factors. For example, the NMS shaped by a range of different stakeholders, such as the customers of the NMS programme's laboratories, DSIT as NPL's owner and the key individuals detailed above as a small selection.

Identifying key individuals also helps to resolve issues with regards to the resilience of a department. If department x saw greater levels of concentration with regards to co-authorship, it can be assumed to have a greater number of opinion leaders. This makes department x at greater risk than a department y (with less concentration) if those individuals were to leave. But there is also a trade-off. Having more opinion leaders does make a department far easier to change direction. Rather than trying to convince a greater number of less connected scientists to adjust direction, the focus could be on the opinion leaders which would help shift the rest of the area. Naturally, a balance between both is ideal but prior to this study, there is no objective, numerical information on where any of the departments are. For this analysis, we are using two main measures of concentration.

Concentration measures

For this analysis, inequality measures have been used to assess the concentration levels of the metrics for NPL as a whole and its departments.

Gini

The Gini Index (also called the Gini coefficient or Gini ratio) is a measure of statistical dispersion used to quantify the degree of inequality in a distribution—most commonly income or wealth within a population.

It ranges from 0 to 1, where:

- 0 represents perfect equality (everyone has identical income/wealth), and
- 1 represents perfect inequality (one person has all the income/wealth and everyone else has none).

The measure was developed by the Italian statistician Corrado Gini in 1912 and is mathematically related to the Lorenz curve[15], representing twice the area between the Lorenz curve and the 45-degree line of perfect equality. The Gini coefficient can also be interpreted as the expected relative difference between the incomes of two randomly selected individuals, scaled by mean income.

This index is widely used by economists, governments, and international organisations to compare inequality both within and across countries, and it is also applied in fields such as sociology, ecology, health sciences, and engineering.

As outlined in Guandalini (2022), the formula for Gini is as follows:

$$G = \frac{Cov(X, F(X)}{\mu}$$

where:

- X is the score on each of the respective metrics,

[15] A Lorenz curve is a graphical representation used to measure the distribution of wealth or income within a population by plotting the cumulative percentage of total income against the cumulative percentage of the population. It illustrates inequality by showing the deviation of actual distribution from a theoretical "line of perfect equality," which is represented by a 45-degree diagonal.

- $F(X)$is the cumulative distribution (rank) of X,
- μ is the mean of the distribution, and
- $\mathrm{Cov}(X, F(X))$is the covariance between score and its rank.

Theil Index

The Theil index, as outlined here, developed by Dutch econometrician Henri Theil (1924–2000), who applied information theory to economics. Theil interpreted inequality as a form of redundancy: the difference between the maximum possible entropy (which corresponds to perfect equality) and the observed entropy in a real income distribution. This framing connects the Theil Index directly to Shannon's entropy and the broader foundations of information theory. is a statistic used to measure economic inequality and other distributions (such as research funding or water usage) within a population or across regions.

Theil formalised the idea that inequality can be seen as the "expected information content" needed to transform population shares (priors) into income shares (posteriors). This makes the Theil Index not only a statistical measure but also a conceptually rich tool grounded in probability and information.

The Theil Index subsequently became recognised as a special case within the Generalized Entropy (GE) family of inequality measures, which are notable for their decomposability. This positioned the index as one of the most theoretically robust tools in modern inequality research.

The index ranges from 0 (perfect equality, where everyone has the same share of income/resources) to an upper bound that depends on the population size (ln N, where N is the number of individuals).

Its primary advantage over other measures, such as the more intuitive Gini coefficient, is its additive decomposability. This means that the total inequality can be broken down exactly into inequality between different subgroups and inequality within those subgroups (e.g., inequality between states vs. inequality within each state)[16]. This is important to flag as there are significant differences between the two types of inequality, using the gender pay gap as an analogy:

- Within-group inequality looks at the differences between those in the same group, within the analogy, this would be looking at men and women as collective groups. Naturally, income levels will be distributed among them, with significant differences flagged and investigated, but one could expect that within-group inequality would be similar between groups.
- Between-group inequality looks at the differences between the groups, independent of the income distributions each group sees. If a large difference can be seen between men and women, it can be viewed as a systematic, structural difference between the two.

In NPL's case, it can be expected that inequality exists within all the departments, given the range of experience level and tenure of scientists. But what is interesting is if there are any differences between the departments with regards to the network metrics.

It should also be noted that the Theil index has two versions, T & L, with their difference lying in their sensitivity to different parts of the income distribution and their respective formulas as special cases of the Generalized Entropy (GE) family of measures.

- **Theil T (GE(1))** reflects redundancy relative to maximum entropy and is more sensitive to **top-end inequality**, where large incomes weigh heavily in the log-ratio term.
- **Theil L (GE(0))** reflects mean log deviation and has a geometric-mean interpretation and is more sensitive to **low-end inequality**, because its formula emphasises deviations below the mean.

The formulae for both versions of the Theil index are as follows:

Theil T:

$$T_T = \frac{1}{N}\sum_{i=1}^{N}\frac{x_i}{\mu}ln\left(\frac{x_i}{\mu}\right)$$

[16] Details on the decomposition will be in the annex

Theil L:

$$T_L = \frac{1}{N}\sum_{i=1}^{N} ln\left(\frac{\mu}{x_i}\right)$$

Where:

x_i = characteristic (e.g. income, metric score) of individual i

$N =$ population size

$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$= mean income

Having outlined the methods, the analysis can be presented, with the first section detailing the headline stats over time and the inequality/concentration measures.

Headline Stats

Headline Stats over time

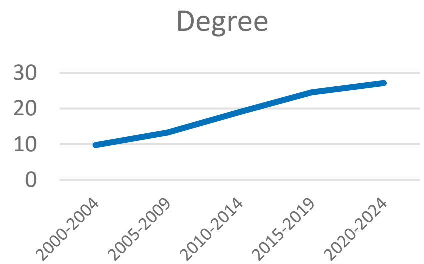


Figure 2 - Degree over time

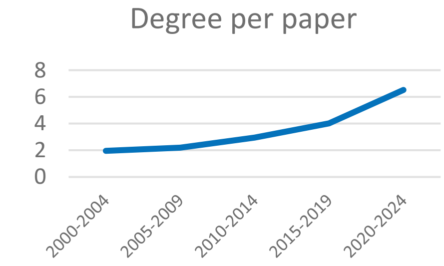


Figure 3 - Degree per Paper over time

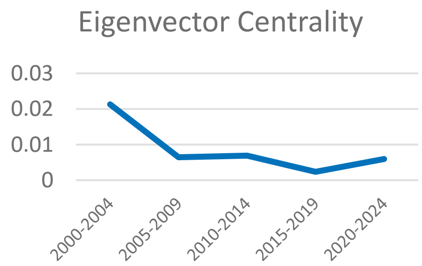


Figure 4 - Eigenvector Centrality over time

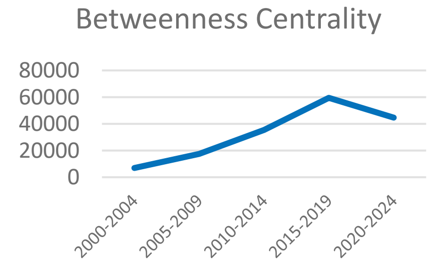


Figure 5 - Betweenness Centrality over time

Headline Inequality stats

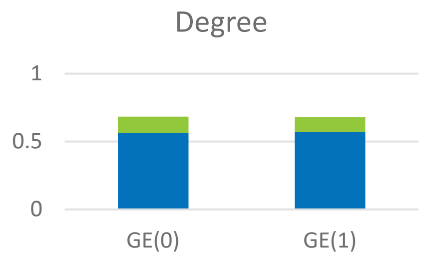


Figure 6 - Degree by Theil T and L

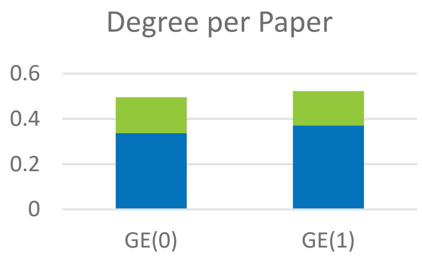


Figure 7 - Degree per Paper by Theil T and L

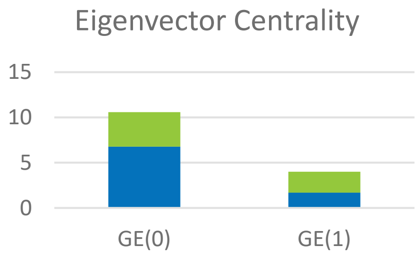


Figure 8 - Eigenvector Centrality by Theil T and L

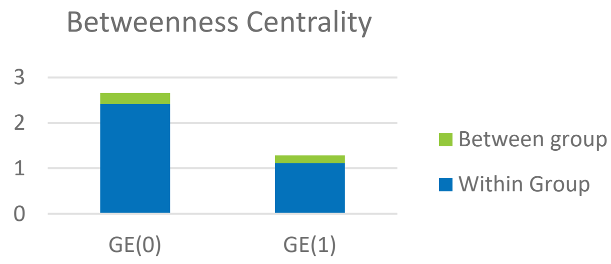


Figure 9 - Betweenness Centrality by Theil T and L

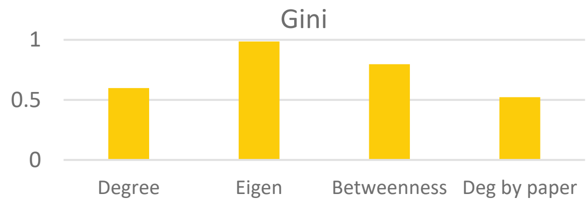


Figure 10 - Gini scores by the four-network metrics

Findings

Headline stats

Starting with the degree metric, what can be seen in an increasing trend across all the time periods. It should be noted that this is coupled with the fact that NPL's employment has grown over the same period, more than doubling from 2006 to 2024[17]. This suggests that the increase in degree could be driven by an increase in authors, rather than authors becoming more collaborative. However, when the count of papers is accounted for, the increase in degree per paper seems to suggest the latter point is more likely to be true than the former.

On eigenvector centrality, the metric saw a sudden decline after the first period then then a broad flatlining for the rest of the period, suggesting that potentially influential authors stopped publishing after the first period. On the other hand, betweenness has increased over time, with the first three periods seeing considerable growth over the time, though this has been curtailed in the most recent window.

Headline Inequality Stats

Focusing on Gini to start with, what can be seen is that all the metrics are unequal, though more so than the other. Given the fact that the two-centrality metrics represent "influencer" and "connector" types within an organisation, it isn't surprising that these are very unequal. The two-degree metrics are less (though still quite) unequal.

Looking at the Theil indices, the following can be seen

- Starting with the degree metrics, both see similar levels of inequality across its distribution, with similar breakdowns of between-group and within-group inequality. However, degree per paper does see proportionally more between-group inequality than degree.
- Moving onto the centrality metrics, a very different distribution can be seen. Most of the inequality resides in the lower part of the distribution, with only a small proportion of total inequality residing in the upper part. When decomposing, similar patterns can be seen with regards to the split of

[17] As can be found in NPL's financial statements.

between- and within-group, though eigenvector centrality does see a larger proportion of between-group inequality

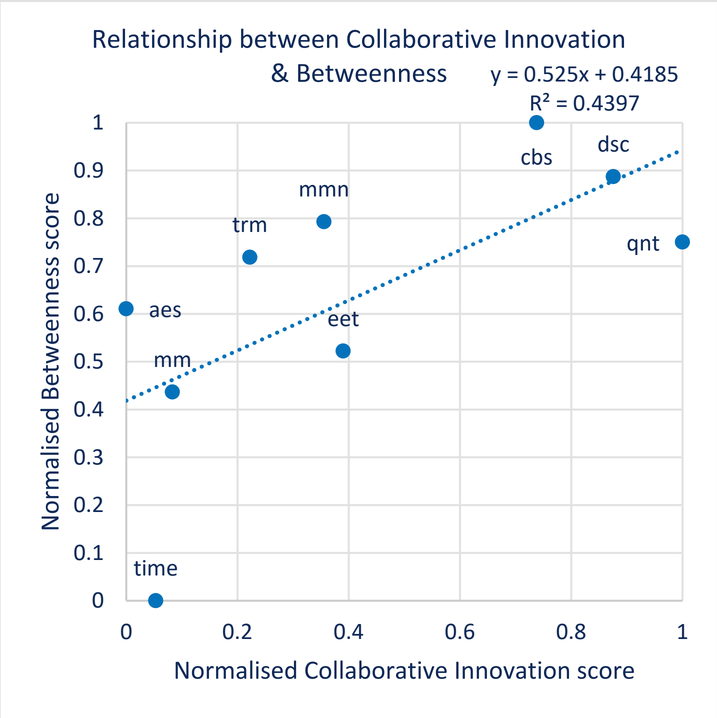
Network and Activity-Output metrics

In order to assess the real-world effects of the network metrics, an analysis was ran using data from Nayak et al (2024). This paper saw the construction of statistical components using the NMS metrics data that is submitted to DSIT each year, of which two main ones were developed:

- Collaborative Innovation – comprised of the "number of academic collaborators, number of business collaborators, and number of peer reviewed papers, [they] represent activities and outputs that pertain to R&D and innovation".
- Measurement Services – comprised of "income from measurement services and reference materials, and publication of new or amended standards with an NMS contribution. Both these variables reflect activities and outputs that pertain to pertain to core measurement capabilities".

These components were developed at a department level for NPL, from 2015 to 2023, which makes them comparable to the network stats for their last 2 periods (2015-2019, 2020-2024). After taking averages for the network metrics, all four were assessed in comparison to the components detailed above, with eigenvector centrality, degree and degree per paper seeing no relationship with the component. However, when assessing betweenness centrality with regards to the two components, a relationship can start to be seen between it and collaborative innovation. For simplicity, both metrics were Min-Max scaled[18] to ease analysis and aid presentation. To assess the relationship, the variables are first scattered, then regressed, with Betweenness acting as the outcome variable and the factors as independent variables:

[18] Min-max scaling or min-max normalization is the simplest feature scaling method and consists in rescaling the range of features to scale the range in [0, 1] or [−1, 1]. Selecting the target range depends on the nature of the data. The general formula for a min-max of [0, 1] is given as: $x' = \frac{x - \min(x)}{\max(x) - \min(x)}$, where x is an original value and x' is the normalized value



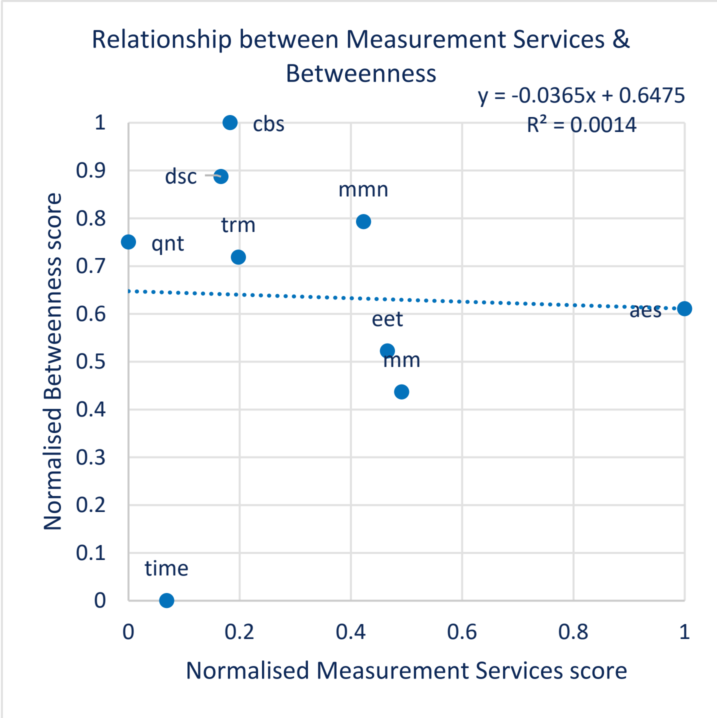


Figure 11

Figure 12

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.794066
R Square	0.63054
Adjusted R Square	0.507387
Standard Error	0.207304
Observations	9

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	0.440058477	0.220029239	5.119962155	0.050431484
Residual	6	0.25784867	0.042974778		
Total	8	0.697907147			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	0.140083	0.190708871	0.7345389	0.49032	-0.326564719	0.606730875	-0.326564719	0.606730875
Collaborative Innovation	0.777215	0.243155114	3.1963735	0.018685	0.182235441	1.372193701	0.182235441	1.372193701
Measurement Services	0.523309	0.297271067	1.760376	0.128828	-0.204087256	1.250704936	-0.204087256	1.250704936

Table 2 - Regression of Betweenness on Collaborative Innovation and Measurement Services

As is self-evident from the R^2 values in the scatter plot and coefficient in the regression, there is effectively no relationship between Measurement Services and betweenness. This suggests that the ability to deliver work to the private sector doesn’t influence a department’s connectedness on it academic papers. On the other hand, there appears to be a positive relationship between collaborative innovation and betweenness. This should be expected and confirms the assumption that greater connectedness leads to greater output.

While the findings of a regressions based on so few observations shouldn’t be considered too deeply, what can be seen is the relationship detailed above, but strengthened while accounting for measurement services as well. This shows how greater betweenness could lead to greater output.

Departmental Stats

The following section will go through the departmental breakdowns of the metrics detailed before, in the following structure:

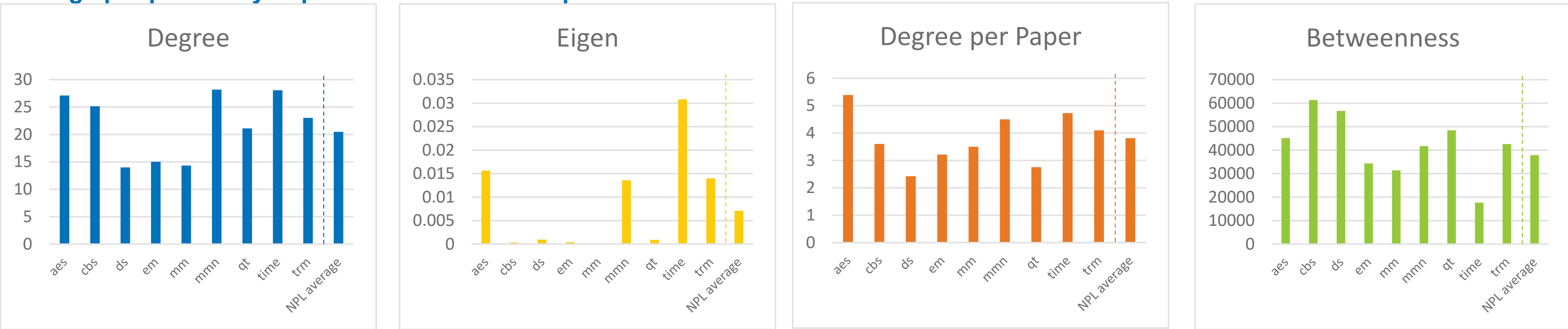
- Headline Departmental stats
    - Average by person by department across the 5 periods
    - Average change per person by department across the 5 period
- Inequality analysis by metric and department
    - Degree
    - Degree per paper
    - Betweenness Centrality (Betweenness)
    - Eigenvector Centrality (Eigen)

Unlike above, the results will be presented on a department-by-department basis, detailed using all the stats and analyses detailed below. The focus will be on the points of difference for each respective department when compared to their peers and an NPL average[19].

[19] The annex will have a metric-by-metric analysis.

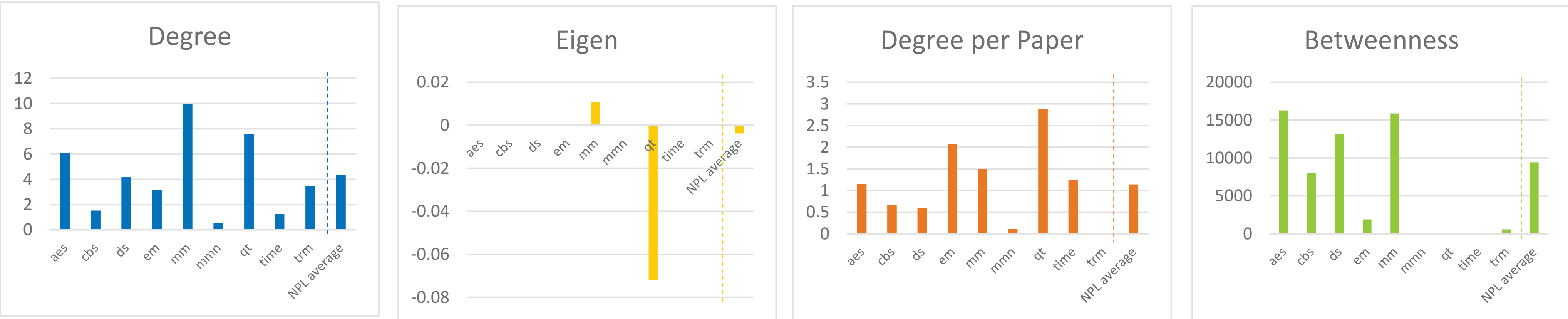
Headline Departmental Stats

Average per person by department across the 5 periods



Figures 13 to 16 – Average scores of the Network Metrics by Department across the five periods (2000-2004, 2005-2009, 2010-2014, 2015-2019, 2020-2024

Average change per person by department across the 5 periods



Figures 17 to 20 – Average Change in Network Metrics by Department across the five periods (2000-2004, 2005-2009, 2010-2014, 2015-2019, 2020-2024

Degree Inequality

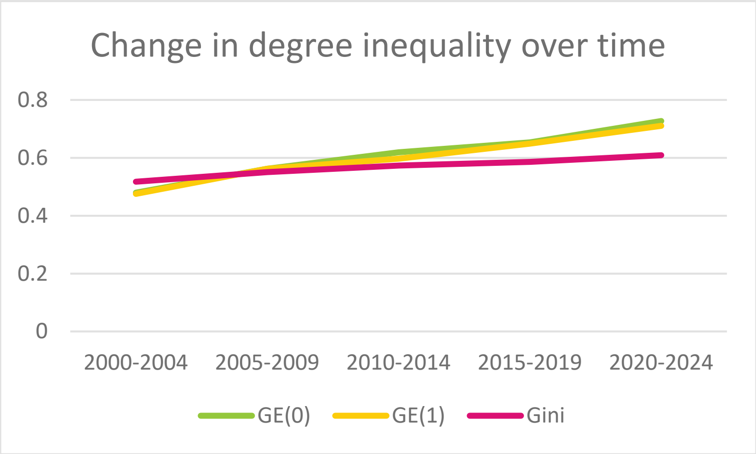


Figure 21

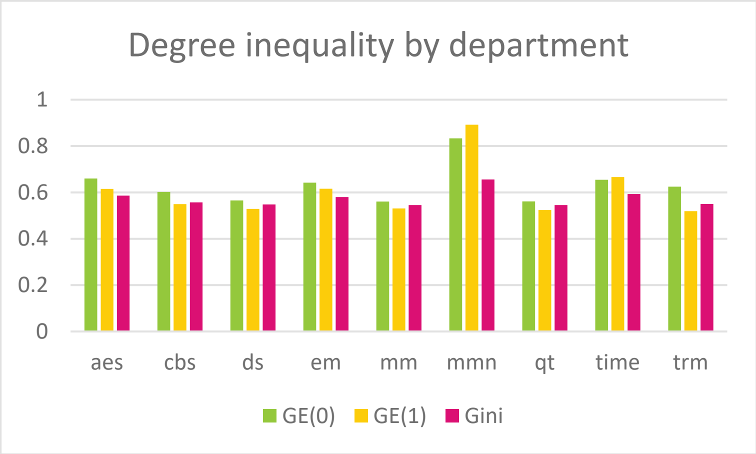
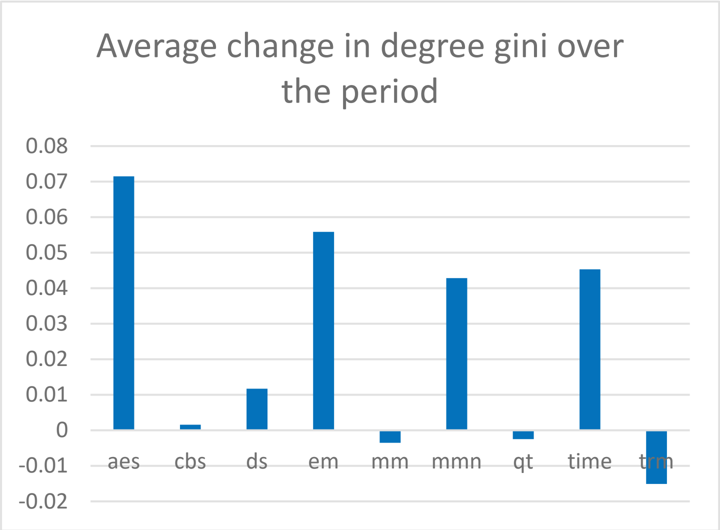
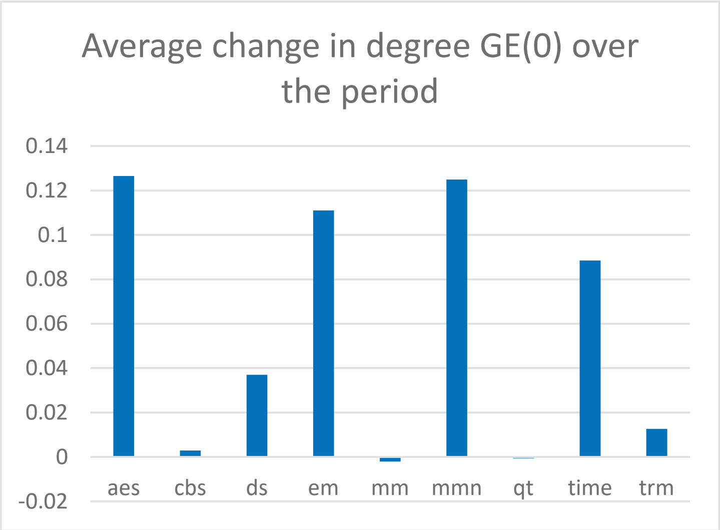
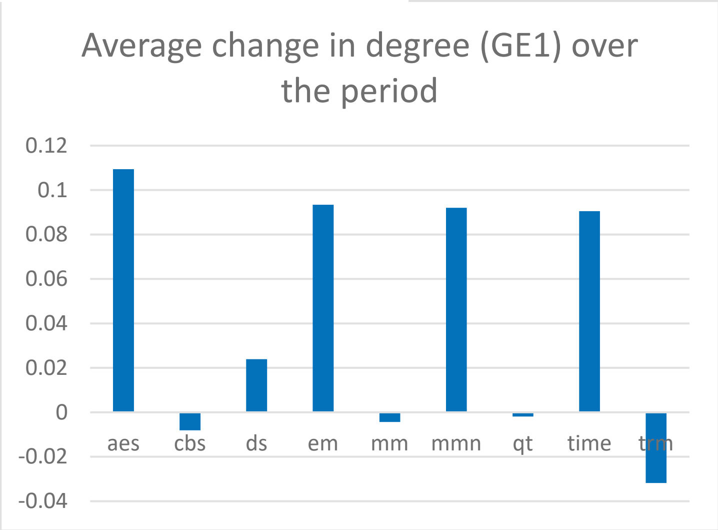


Figure 22







Figures 23 to 25 – Average changes in Degree's inequality stats over time

Degree per paper inequality

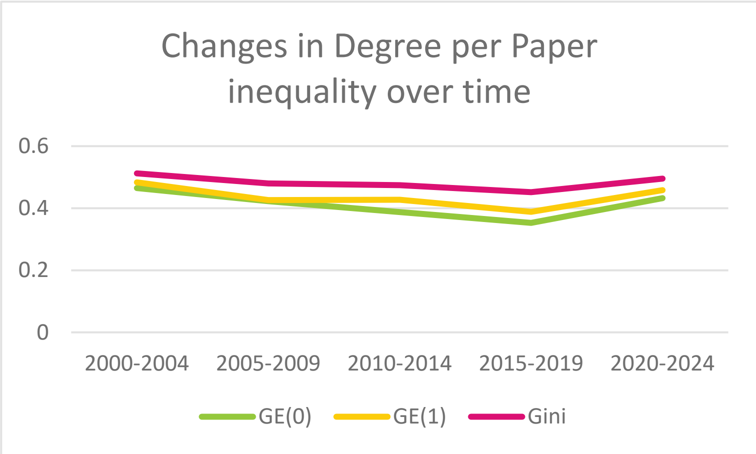


Figure 26

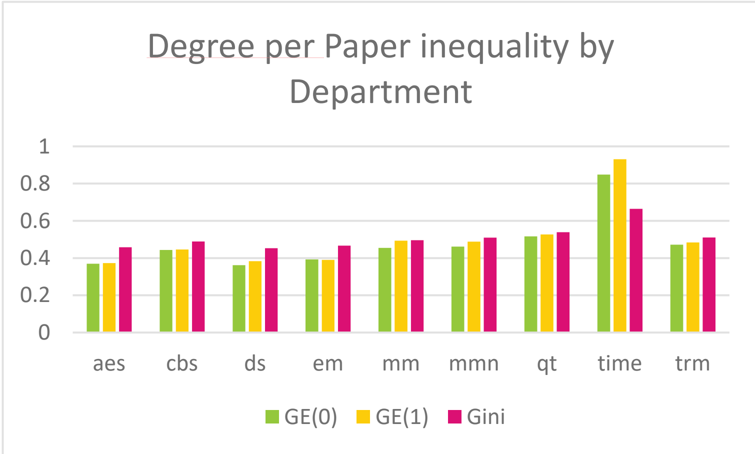
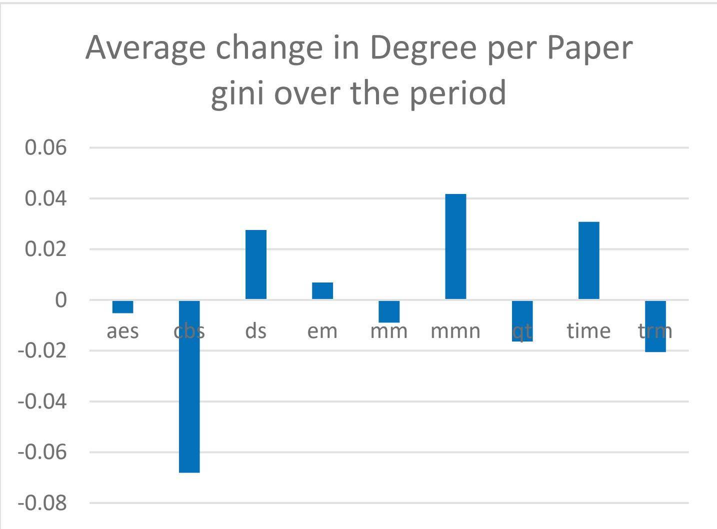
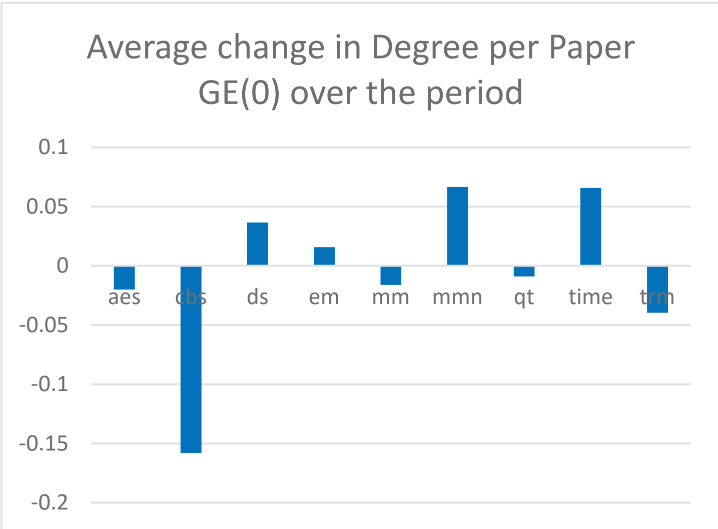
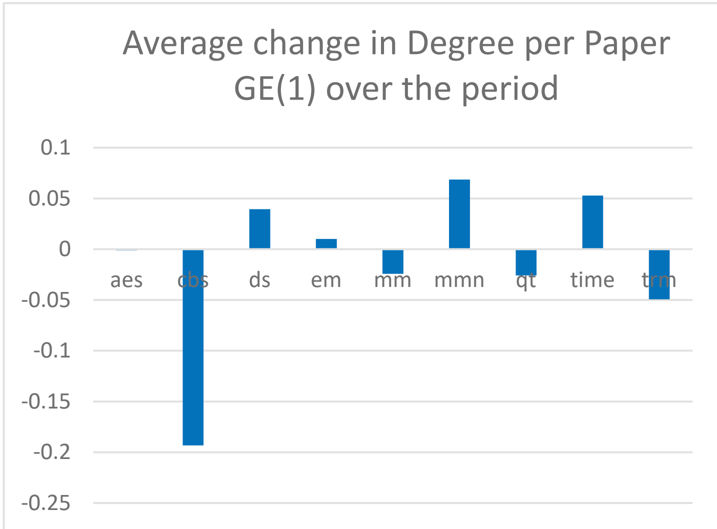


Figure 27







Figures 28 to 30 – Average changes in Degree per Paper's inequality stats over time

Betweenness inequality

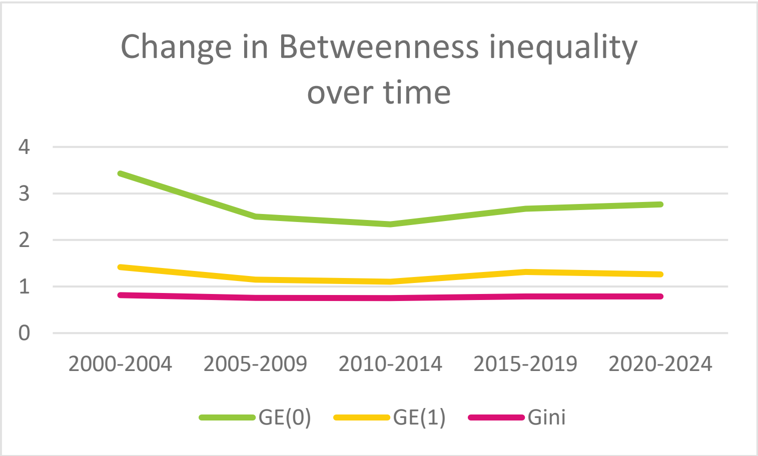


Figure 31

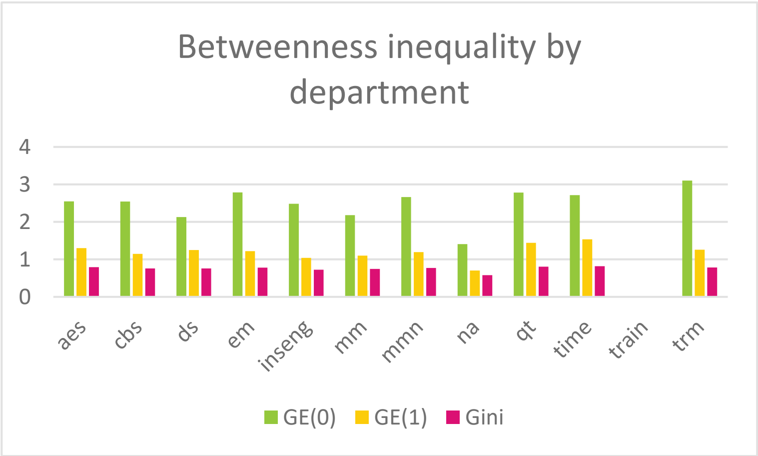
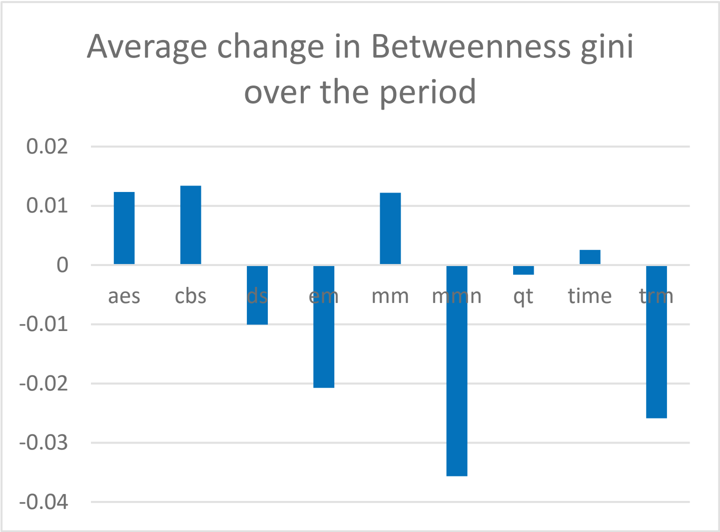
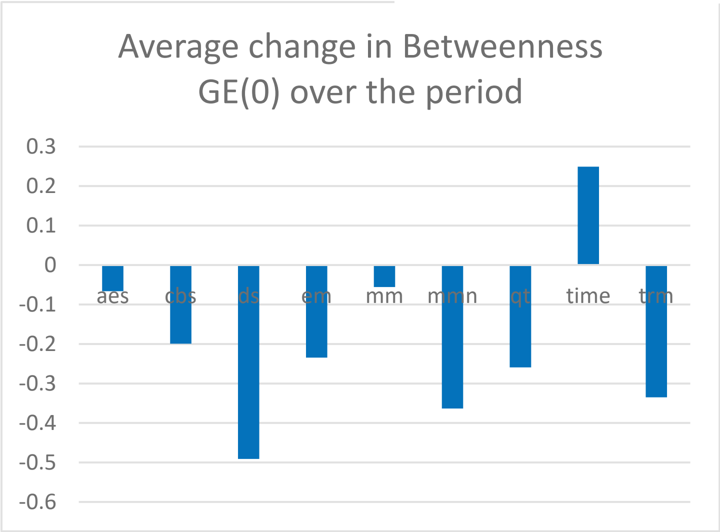
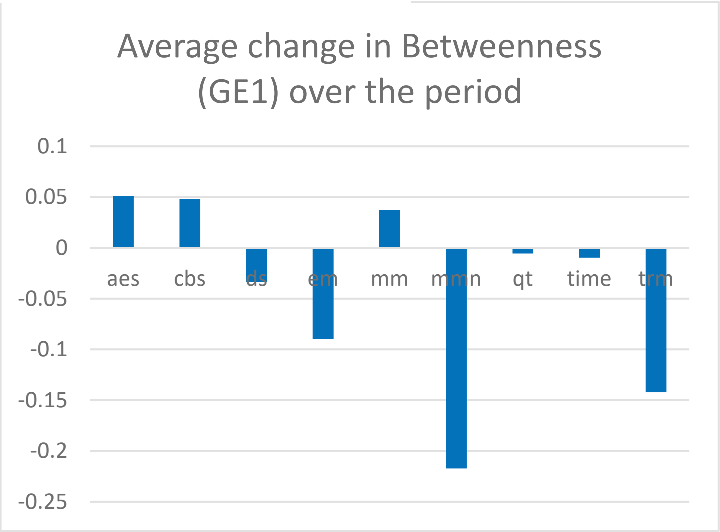


Figure 32







Figures 33 to 35 – Average changes in Betweenness's inequality stats over time

Eigen inequality

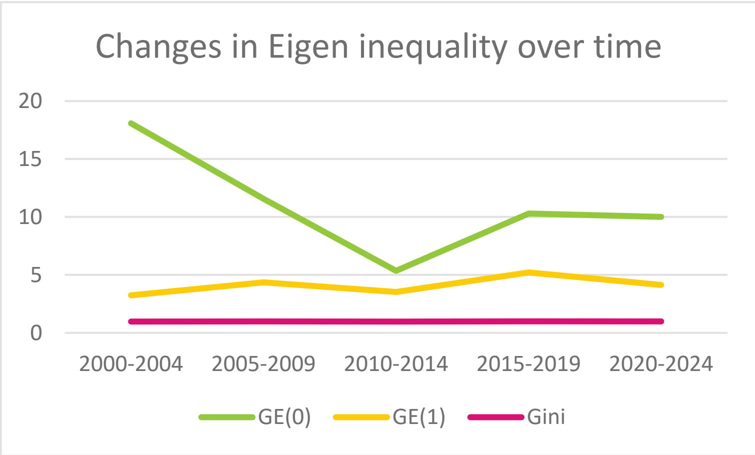


Figure 36

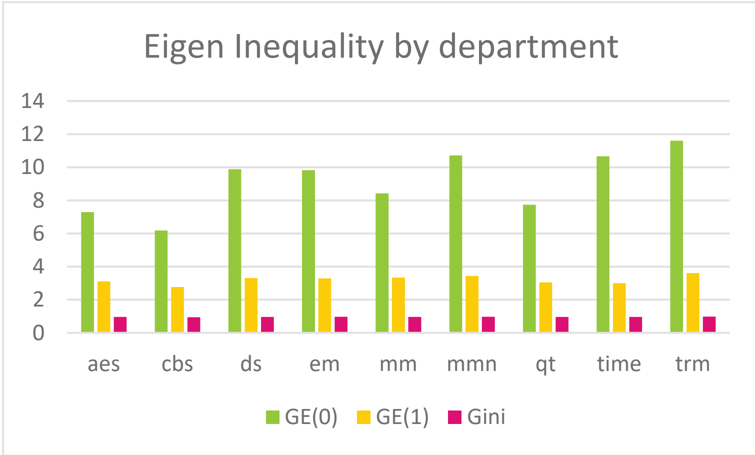
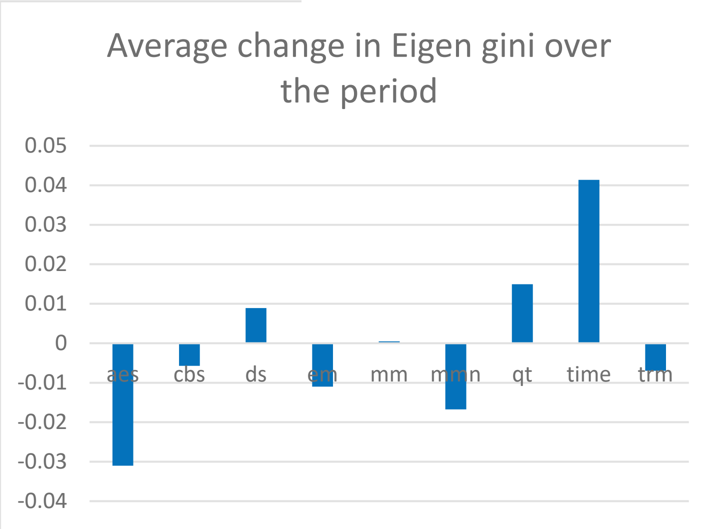
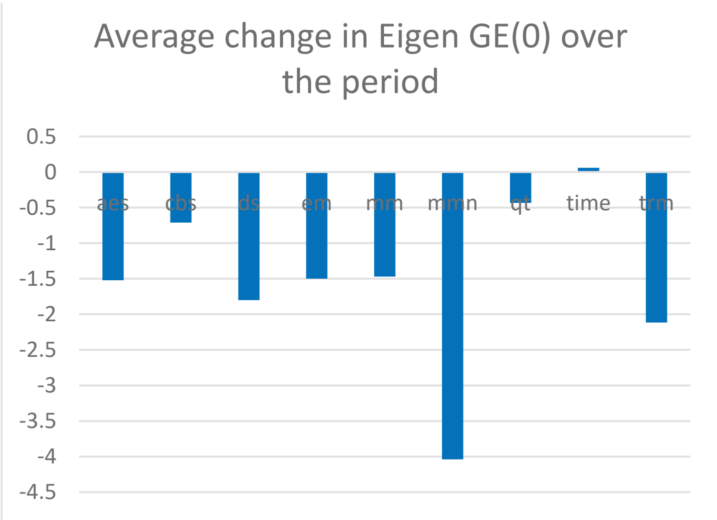
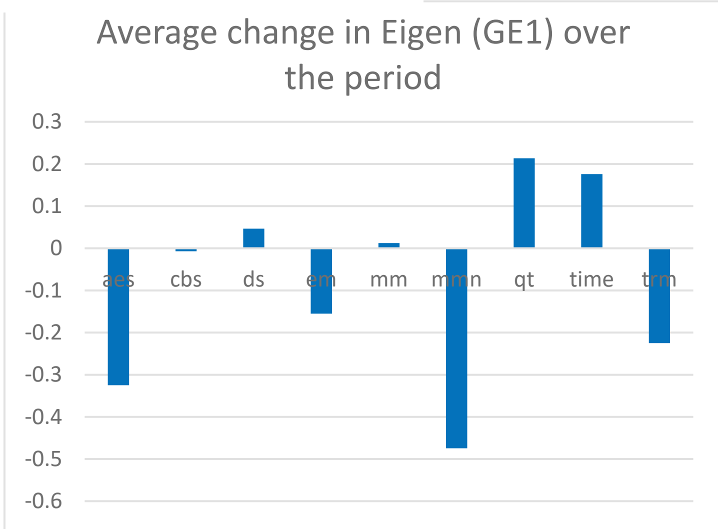


Figure 37







Figures 38 to 40 – Average changes in Eigen's inequality stats over time

Departmental Assessment

AES

This department has scored particularly high on degree and has one of the highest four Eigen scores. Moreover, it is seeing growth in its Degree and Betweenness over the years. Alongside this growth, this department has also seen an increase in its degree and betweenness inequality, though it has also seen a decline in its eigen inequality as well.

Overall, this department appears to have a considerable amount (and growing levels) of co-authorship and connections, while also being one of the four departments with notable “influencers” within their ranks. However, this growth is coupled with increasing inequality among their authors, though the difference between authors with regards to influence appears to be declining. This could be driven by growth in staff, but further investigation would be needed to confirm this.

CBS

They have the highest betweenness score of all the departments, while also scoring high on degree, though this does reduce when accounting for the number of papers. Concerning its inequality, there have been little change in its degree inequality. This does change when accounting for papers, where a notable decline in inequality has been seen. This is also one of three departments that have seen an increase in its betweenness inequality, driven by increases in inequality among the upper distribution of their authors.

Overall, this department appears to have the highest number of connections and high-level number of co-authors on their papers, though the latter fact may be driven by a high output of papers. It is doing well on inequality, seeing a decline in paper-adjusted co-author inequality, though it has also seen the inequality of its connections increase – particularly among their more frequent publishers.

DS

They have the 2nd highest betweenness score, though also having comparatively low degree and degree per paper scores. This suggests that this department tends to work with a broad group of co-authors but not in an intensive manner. On its inequality, its metrics are going in different directions. They are becoming more equal with regards to their betweenness, but they are also facing increasing inequality when looking at the degree per paper and the eigenvector centrality, particularly at the top end of the distribution.

Overall, this department appears to be working in a “breadth, not depth” manner, looking at working with a wide variety of co-authors across a range of seniority levels. With regards to its inequality, it appears to be increase across most of the metrics,

EM

Across the board, this department is on the weaker side of the headline stats, though their degree per paper is better across the board. It should also be noted that despite its current position, it is seeing some increases over time, particularly in its degree-per-paper and Betweenness (though its betweenness is only increasing by a small amount). With regards to inequality, there is a mix of increasing and decreasing levels across the four metrics. With regards to degree and degree per paper inequality, there is an increase over time, particularly with regards to degree. With regards to the centrality metrics, a decline can be seen over the period as well.

Overall, this department doesn’t appear to be particularly strong with regards to the network metrics, but it has seen improvement over time. When diving into the period-level stats, EM has gone from second bottom to middle of the pack with regards to Betweenness, showing serious improvement over time. Its inequality has seen mixed outcomes over time as well.

MM

Much like EM, this department is on the weaker side of the headline stats, while also seeing improvement over time. They have seen the most improvement over time with regards to all the stats bar degree per paper, in which it has seen moderate growth.

They have lowest betweenness inequality of all the departments, with average levels of inequality for the other metrics. On the changes in inequality, there has been none on both kinds of degree inequality, an increase in inequality of betweenness and a decline in eigenvector centrality lower-distribution inequality.

Overall, this department is one that is not particularly well-connected department but is seeing significant improvement over time. Coupled with this, it isn’t seeing an increase in inequality and is on the lower side of inequality more generally.

MMN

Across several metrics, this department is particularly strong, with high scores on degree, degree per paper, and eigen, while also being average on betweenness Despite its high scores, it has seen effectively no change over the period across all the statistics, maintaining its position over the last 24 years. Concerning its inequality, it has high levels of degree and eigen inequality, while also seeing an increase in its degree inequality. The other metrics either see no increase or a decline in its inequality.

This department has maintained a high level of co-authors and influence, while being average with regards to connectedness. However, there is considerable inequality with regards to co-authors, which appears to be increasing.

QT

This department has a mixed return with regards to the metrics. It is one of the lowest scores with regards to degree per paper, is average with regards to degree and above average with regards to Betweenness. Concerning changes over time, it’s a very mixed bag. It has seen an increase in the two-degree metrics over time, while also seeing a massive collapse of its eigen metric when compared to the other departments. With regards to its inequality, it is largely middle of the pack though it is seeing a decline. Degree inequality has been at the same level over the periods, while degree per paper inequality has declined slightly. Betweenness and Eigen inequality have headed in differing directions, with the former declining (particularly in the lower distribution) and the latter increasing (particularly in the upper distribution).

Overall, this department appears to have very mixed results across the board, concerning its network metrics, the changes in these metrics and their inequality.

Time

This department is particularly strong with regards to the degree metrics and eigen, while also seeing has seen very little change over the period in these metrics. In contrast, this department scores particularly poorly on betweenness. Concerning its inequality, it appears quite inequal with regards to degree per paper, while seeing inequality increase across all metrics. When decomposed, an increase in betweenness inequality at the lower end of the distribution can be seen, along with an increase in eigen inequality at the top end of the distribution can be seen.

Overall, this department appears to have high co-authorship and influence, suggesting that it is centralised around a few specific individuals. However, there appears to be few bridges connecting them to the rest of NPL, along with an increase of inequality across the period.

TRM

This department is one of the few that sees a high eigen score, while being slightly above average on all other metrics. It has also seen no changes in its score while the increase in degree is accounted for when adjusting for papers. Concerning its inequality, it is the highest on average on betweenness and eigen inequality, though it has seen a reduction in the inequality for both degree metrics and eigen. Concerning betweenness, there has been a mixed story. At the lower end of the distribution, there has been an increase in their inequality - while also having the highest value. In the upper end, the inequality has decreased. The net effect can be seen when looking the Gini score, which has gone down.

Overall, this department see a mixed story, not unlike several of the departments detailed above. It has some of the best scores while also being quite unequal with regards to connectedness and influence.

Conclusion

To conclude, Social Network analysis is a useful tool in a range of different ways. The metrics it can generate are useful in assessing NPL's science outcomes, providing a useful contribution to the monitoring and evaluation of the NMS programme. They complement citation scores incredibly well, both in their construction and in what they represent. Given the relationship between Collaborative Innovation and Betweenness, it also highlights the importance of collaboration. Beyond this, this study has also been able to assess and compare NPL's departments to each other, accounting for the lack of a baseline by using temporal assessments and inequality measures. This study has also produced outputs that have operational benefits, given the ability to assess potential resilience issues and principal-agent problems within an organisation.

This report's main purpose is to prove important context to couple with a more detailed analytical report, which is currently in the pipeline and should be published later in 2026. This paper will provide in-depth detail on the methods, outcomes and findings of the analysis. However, in the meantime, the above analysis has highlighted a few key aspects of NPL's departments. What can be seen is a broad divergence in the departments' co-authorship structures, with some departments structured as "spiderwebs" (highly interconnected, scattered across NPL's network) with others being more heliocentric ("orbiting" a few influential authors).

It should be highlighted that the scores that are seen for all the metrics, whether they are higher or lower, should not be seen as a strength or weakness. Rather, they are representative of where a given department is now and how that has changed over time. Furthermore, having high or low scores on the network metrics merely represent the present activity of a given department. For example, some departments have the responsibility to deliver significant national infrastructure, such as Time which delivers the National Timing Centre. Time's focus on NTC delivery may have come at the detriment of their NPL-wide collaborative activity, but may have boosted its external collaboration work, which this study does not pick up. Furthermore, departments that are far more concentrated are also able to establish an identity around their academic offerings, when compared to a diffuse department which works in many different things but may lack a *Modus Operandi* that the more concentrated departments have. These differences and changes could be regarded as a form of "speciation", where departments have evolved to best suit their research interests and outputs. This can be reinforced by findings in Nayak et al (2024), with clear trade-off between commercial- vs research-focused work seen across the departments. This could be interpreted as an internal shift from a more interdisciplinary organisation to one that is more multidisciplinary.

It must be noted that this tool does have its limitations and cannot be viewed as a "silver bullet" to solve all potential issues an organisation may have. Fundamentally, SNA captures trends and patterns, but the analysis so far does not and cannot diagnose the causes of said trends and patterns. That requires further statistical analysis, while also conducting discussions with those leading research across an organisation. Nevertheless, SNA provides a useful window to confirm and challenge anecdotal knowledge in a quantitative and unbiased manner.

Furthermore, the data used – NPL's publications – is only a small window into collaboration patterns. Publications are often only the "tip of the iceberg" of the R&D activity, as most R&D does not yield successful results, let along results worthy of writing up. Also, as outlined in Nayak et al (2024) NPL has a range of different avenues of activity that count as "collaboration", including grant-funded collaboration, key comparison activity and shared staff. These are not included in the network at this point in time (and would pose difficulties as these datasets are organisational-level rather than individual-level) but must be considered. Finally, as mentioned above, this network does not include external co-authors, which may obscure potential connections around NPL. This was out-of-scope of this study but may be added to the network in due course.

Reference list

Alessio Guandalini (2022). Things you should know about the Gini index. *RIEDS - Rivista Italiana di Economia, Demografia e Statistica - The Italian Journal of Economic, Demographic and Statistical Studies*, [online] 76(4), pp.4–12. Available at: https://ideas.repec.org/a/ite/iteeco/220412.html [Accessed 3 Feb. 2026].

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network analysis in the social sciences. *Science*, 323(5916), 892–895.

Braun, D. and Guston, D.H. (2003). Principal-agent theory and research policy: an introduction. *Science and Public Policy*, 30(5), pp.302–308.

Cross, R., Borgatti, S. P., & Parker, A. (2002). Making invisible work visible: Using social network analysis to support strategic collaboration. *California Management Review*, 44(2), 25–46.

Doehne, M. and Herfeld, C. (2023). How academic opinion leaders shape scientific ideas: an acknowledgment analysis. *Scientometrics*, [online] 128(4), pp.2507–2533. doi:https://doi.org/10.1007/s11192-022-04623-z.

Freeman, L.C. (1977). A Set of Measures of Centrality Based on Betweenness. *Sociometry*, [online] 40(1), p.35. doi:https://doi.org/10.2307/3033543.

Freeman, L. C. (2000). Visualizing social networks. *Journal of Social Structure*, 1(1).

Gini, C. (1912). *Variabilità e mutabilità*. [online] Reprinted in Memorie di metodologica statistica (Ed. Pizetti E. Available at: https://ui.adsabs.harvard.edu/abs/1912vamu.book.....G/abstract [Accessed 3 Feb. 2026].

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6), 1360–1380.

Haveman, H.A. (2025). Organizations and Economic Inequality. *Journal of Organizational Sociology*, [online] 3(1), pp.1–25. doi:https://doi.org/10.1515/joso-2024-0020.

HM Treasury (2020). *Magenta Book Central Government Guidance on Evaluation*. [online] Available at: https://assets.publishing.service.gov.uk/media/5e96cab9d3bf7f412b2264b1/HMT_Magenta_Book.pdf.

Huffman, W.E. and Just, R.E. (2000). Setting Efficient Incentives for Agricultural Research: Lessons from Principal-Agent Theory. *American Journal of Agricultural Economics*, 82(4), pp.828–841. doi:https://doi.org/10.1111/0002-9092.00084.

Jiang, E., Yildiran, C. and King, M. (2025). A Summary of Science Metrix's Bibliometric Assessment of the National Physical Laboratory's Outputs on Advanced Manufacturing, Data & Digital, Life Sciences & Health, and Energy & Environment (2013-2023) | NPL Publications. *Npl.co.uk*. [online] doi:https://eprintspublications.npl.co.uk/10224/1/IEA30.pdf.

Katanguru, D.R., Qureshi, Z. and King, M. (2025). A survey of UK-based businesses using laboratories funded through the National Measurement System (2023) | NPL Publications. *Npl.co.uk*. [online] doi:https://eprintspublications.npl.co.uk/10151/1/IEA28.pdf.

King, M. and Nayak, S. (2025). A macroeconomic model with imperfect production processes and conformance testing | NPL Publications. *Npl.co.uk*. [online] doi:https://eprintspublications.npl.co.uk/10127/1/IEA27.pdf.

King, M. and Olakojo, S. (2023). NMS business case model: an explanatory note | NPL Publications. *Npl.co.uk*. [online] doi:https://eprintspublications.npl.co.uk/9681/1/IEA15.pdf.

King, M., Olakojo, S. and Gorringe, G. (2023). *Evaluation Framework for the National Measurement System: A framework for assessing the impact of NMS laboratories*. [online] Available at: https://eprintspublications.npl.co.uk/9853/1/IEA20.pdf.

Krackhardt, D., & Hanson, J. R. (1993). Informal networks: The company behind the chart. *Harvard Business Review*, 71(4), 104–111.

Lazarfield, P., Berelson, B. and Gaudet, H. (2021). *The People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign, Legacy Edition*. [online] Columbia University Press. Available at: https://cup.columbia.edu/book/the-peoples-choice/9780231197953/ [Accessed 3 Feb. 2026].

Nayak, S., Yildiran, C., Dias, C. and King, M. (2024). *NPL REPORT IEA 22 VISUALISING THE LOCATIONS OF NPL'S SED DEPARTMENTS AND DESIGNATED NMS LABORATORIES WITHIN AN IMPACT SPACE*. [online] Available at: https://eprintspublications.npl.co.uk/9962/1/IEA22.pdf [Accessed 29 Jan. 2026].

Olakojo, S. (2025). Impact of Innovation Support Delivered through the National Measurement Programme on Business Outcomes: Summary. | NPL Publications. *Npl.co.uk*. [online] doi:https://eprintspublications.npl.co.uk/10152/1/IEA29.pdf.

Rosemberg, C., Obando, C., El-Rayyes, T., Mihaylova, N., Babb, R., Centeno, J., Vingre, A. and Brown, N. (2024). *WHAT METHODS WORK FOR EVALUATING THE IMPACT OF PUBLIC INVESTMENTS IN RD&I*. [online] Available at: https://assets.publishing.service.gov.uk/media/66f56baa30536cb9274827fe/what_works_rdi_evaluations_report.pdf.

Scott, J. (2017). Social Network Analysis. [online] doi:https://doi.org/10.4135/9781529716597.

Technopolis (2024). *WHAT METHODS WORK FOR EVALUATING THE IMPACT OF PUBLIC INVESTMENTS IN RD&I*. [online] Available at: https://assets.publishing.service.gov.uk/media/66f56bb430536cb9274827ff/what_works_rdi_evaluations_annex.pdf.

Un, C.A. and Asakawa, K. (2014). Types of R&D Collaborations and Process Innovation: The Benefit of Collaborating Upstream in the Knowledge Chain. *Journal of Product Innovation Management*, [online] 32(1), pp.138–153. doi:https://doi.org/10.1111/jpim.12229.

Un, C.A., Cuervo-Cazurra, A. and Asakawa, K. (2010). R&D Collaborations and Product Innovation*. *Journal of Product Innovation Management*, [online] 27(5), pp.673–689. doi:https://doi.org/10.1111/j.1540-5885.2010.00744.x.

Valente, T. W. (2010). *Social networks and health: Models, methods, and applications*. Oxford University Press.

Veugelers, R. (1998). Collaboration in R&D: An Assessment of Theoretical and Empirical Findings. *De Economist*, [online] 146(3), pp.419–443. doi:https://doi.org/10.1023/a:1003243727470.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge University Press.

Wei, Y. and Gang, J. (2017). The Research of Incentive Model Based on Principal-Agent for R&D Personnel in System-Transformed Institutes. [online] Proceedings of the Eleventh International CoConference on Management Science and Engineering Management. pp.1645–1652. doi:https://doi.org/10.1007/978-3-319-59280-0_138.

Yildiran, C. (2024). *Bibliometrics: a Pilot Study for NPL's Publications over 2000-2023*. [online] Available at: https://eprintspublications.npl.co.uk/9968/1/IEA23%20.pdf [Accessed 29 Jan. 2026].

Yildiran, C. and King, M. (2025). A summary of Science Metrix's bibliometric assessment of the National Physical Laboratory's metrology research outputs (2013 - 2023) | NPL Publications. *Npl.co.uk*. [online] doi:https://eprintspublications.npl.co.uk/10128/1/IEA26.pdf.

Annex

Metric-by-Metric findings

Average per person by department across the 5 periods

What can be seen is substantial differences across departments in their network characteristics. When looking at Degree and Degree per Paper, AES, CBS, MMN and Time all stand out with notably high Degree scores. AES, MMN and Time also show strong performance on Degree per Paper, indicating that their connectivity remains high even after accounting for publication volume. In contrast, CBS records a comparatively low Degree per Paper, largely because its high paper output dilutes the ratio. Meanwhile, DS, EM and MM exhibit lower Degree scores overall, though EM and MM recover somewhat when Degree is normalised by paper count.

Turning to Eigenvector and Betweenness centrality, a diversity of score distributions can be seen. Eigenvector centrality appears to be driven predominantly by four key departments, AES, MMN, Time and TRM. This suggests the presence of a more “heliocentric” structural pattern rather than a diffuse, web-like network. Betweenness centrality sees CBS, DS and QT rank highly, indicating their roles as bridging units within the network, whereas EM, MM and Time sit at the lower end of this measure.

Average change per person by department across the 5 periods

When looking at the changes in the metrics, interesting patterns can be seen across all of them. On the Degree and Degree-per-Paper changes, a varied outcome across departments can be seen. For overall Degree, MM, QT and AES all sit above the average, while Time, MMN and CBS fall below it. When Degree is normalised by paper output, a slightly different pattern emerges, with QT, EM and MM increasing at an above average. TRM, MMN, DS and CBS saw a below average change over the periods.

These shifts highlight several notable differences. Among departments with above-average Degree scores, AES drops to a lower-than-average position once Degree is measured per paper, suggesting that its strong connectivity is partly volume-driven. For those at the average level on Degree, EM moves upward when normalised, indicating stronger connectivity relative to its publication count, while TRM and DS move lower. Among departments below average on Degree, Time shows a meaningful improvement on the per-paper metric.

The Eigenvector and Betweenness centrality measure changes reveal even greater disparities. For Eigenvector centrality, MM saw a modest increase, whereas QT experiences a substantial decline, pointing to a shift in its relative influence within the network. In contrast, Betweenness centrality shows large increases for AES, MM and DS, suggesting that these departments have become more significant as bridging nodes. Meanwhile, EM, MMN, QT, Time and TRM exhibit minimal movement, implying stability in their roles within the network's pathways.

Following on from the headline stats, a deeper look is taken onto each of the metrics inequality, with a respective focus on:

- Their changes over time
- The headline departmental inequality scores, focusing on:
    - Theil T
    - Theil L
    - Gini
- Assessing the changes in these metrics over the timeframe.

What was seen was the following:

Degree inequality

This metric is the second least unequal overall, with most of its variation stemming from inequality within groups rather than between them. This pattern holds across both the upper and lower ends of the distribution. One department, however, stands out: MMN exhibits noticeably higher inequality than its peers.

Over time, several departments, notably AES, EM, MMN and Time, appear to be becoming more unequal relative to others, which have largely maintained stable levels of inequality. In contrast, TRM shows a modest reduction, indicating a slight shift toward a more even distribution.

Degree per paper inequality

This metric emerges as the most equal among those assessed, suggesting that paper counts themselves are driving a notable share of the inequality observed in Degree. Even so, there is a slightly higher level of between-group inequality compared with other metrics.

Overall, most departments cluster closely together, with Time standing out as the only clear outlier. Examining how the metric evolves over time reveals three distinct patterns: CBS and TRM show signs of becoming more equal; DS, MMN and Time appear to be growing more unequal; and AES, MM, QT and EM remain largely unchanged across the observation period.

For the next two metrics, it should be noted that the inequality is driven primarily at the lower levels of the distribution (the authors with lower publications and lower numbers of co-authors).

Betweenness inequality

There is very little variation across departments on this metric overall, with most of the inequality originating in the lower end of the distribution rather than the top. When looking at changes over time, however, a few departments stand out. AES, CBS and MM show the largest increases in inequality, reflected in rising Gini values driven primarily by widening gaps at the upper end of the distribution. Time and QT also experience increases in inequality, though to a smaller extent; for these departments, the rise is mostly due to shifts in the higher part of the distribution, with Time additionally showing some growth in inequality at the lower end.

Eigen inequality

This metric sees a substantially higher level of within-group inequality, driven largely by disparities at the lower end of the distribution. Although between-group inequality represents only a small share of the total, it is still notable and contributes meaningfully to the overall pattern.

Clear differences emerge between departments. Time, TRM and MMN appear as the most unequal, forming the upper tier of inequality. DS and EM follow as the next group, exhibiting moderate levels of inequality. In contrast, AES, CBS, MM and QT show the lowest inequality, positioning them as the most equal departments in the comparison.

Looking at changes over time, the lower end of the GE distribution has declined across all departments, indicating a broad reduction in lower-tail inequality. However, three departments - DS, QT and Time - have become more unequal overall, a shift driven primarily by widening gaps at the upper end of the distribution.

Theil decomposition calculation

To decompose Theil T and L into the between- and within-group inequality, the following calculation is done, coming from Liao (2016):

When the population is partitioned into Kgroups, the **between-group component** of the Theil T index is:

$$T_{\text{T between}} = \sum_{k=1}^{K} y_k \ln\left(\frac{\bar{x}_k}{\bar{x}}\right)$$

Where:

- y_k= income share of group $k$relative to total income
- $\bar{x}_k$= mean income of group k
- $\bar{x}$= mean income of entire population

The **within-group component** is:

$$T_{\text{T within}} = \sum_{k=1}^{K} y_k \left[\sum_{i=1}^{n_k} y_{ik} \ln\left(\frac{x_{ik}}{\bar{x}_k}\right)\right]$$

Where:

- y_{ik}= income share of individual $i$within group k
- x_{ik}= income of individual $i$in group k
- n_k= group size

Additivity

Theil's decomposition satisfies:

$$T_T = T_{\mathrm{T\ between}} + T_{\mathrm{T\ within}}$$

Meaning total inequality splits cleanly into between- and within-group contributions.

A similar formulation for the Theil L is performed by switching the role of the income share and population share variables, but it is not given here to save space.

A pilot study into the application of Social Network Analysis on NPL's publication data

Providing the measurement capability
that underpins the UK's prosperity
and quality of life

To find out more about NPL:

npl.co.uk

To get in touch:

npl.co.uk/contact

020 8977 3222

Please consider the environment before printing this document