

平成 30 年度海外行政実態調査報告書

根拠に基づく政策形成と統計的会計検査：
イギリスおよびフィンランドの教育・統計
データベース調査を基に

Evidence-Based Policy Making and Application of Statistical Methods to Audit:
A Survey of Education Database and Audit in England and Finland

特別研究官 田中 隆一（東京大学社会科学研究所教授）

平成 31 年 3 月

調 査 課

はじめに

公共政策を効果的かつ効率的に実施してゆくことの重要性は昨今の財政状況を反映して高まっており、会計検査における「経済性・効率性・有効性」の観点としても重要視されている。公共政策の有効性についての検証は、質的な検証のみならず、過去の同一事業との比較や、他の類似事業との比較といった視点も重要なため、当該事業の成果に関しての量的な評価が求められることになる。そのため、調査データや統計的資料の利用可能性が低い場合には、正確性や合規性といった観点からの検査に比べて、有効性からの事業評価が困難な場合もある。

公共政策の有効性の量的な評価のためには、その事業のインプットおよびアウトプット・アウトカムに関するデータの収集・蓄積が重要である。また、すでに行われた公共政策の事後評価を行い、その結果を蓄積することは、将来の政策選択にとって重要な基盤となる。政策の事後的な検証を可能とするデータベースを構築し、それを活用した検証結果を蓄積してゆくことは、根拠(エビデンス)に基づく政策形成(Evidence-based Policy Making; EBPM)推進上の最重要課題である。

証拠(エビデンス)に基づく政策立案とは、政策目的を明確化させ、その目的のため本来に効果が上がる行政手段は何かなど「政策の基本的な枠組み」を証拠に基づいて明確にするための取組のことである。政府は「経済財政運営と改革の基本方針2017」(平成29年6月9日閣議決定)等においてEBPM推進体制の構築を図ることとしており、限られた資源を有効に活用し国民により信頼される行政を展開するためには、EBPMを推進する必要があるとしている(平成30年1月12日 内閣官房行政改革推進本部事務局資料)。

本報告書では、EBPMや会計検査における統計的手法の活用が比較的進んでいる事例としてイギリスとフィンランドに着目し、データベース整備の現状、政策評価の実施状況、政策立案への反映状況といった視点から両国における関係省庁等にヒアリングを行い、実態を調査した。また、両国の会計検査院に対するヒアリングを通じて、会計検査とEBPMの関係について調査し、今後の日本における根拠に基づく政策形成を推進する上で会計検査院の果たしうる役割についても考察した。

なお、本報告書における見解は筆者個人のものであり、会計検査院の公式見解ではない点は付記しておく。本調査の実施においては、国内外の様々な関係機関の方々の協力を得た。

この場を借りて関係各者に対してお礼申し上げる。また、今回の調査に際して様々な支援と協力を賜った会計検査院とその職員の方々に対してもここに記して謝意を表したい。

目次

1 イギリス	1
1-1 調査の概要	1
1-2 制度的基盤としての統計データ	2
1-2-1 データの収集（教育省）	2
1-2-2 データベースの構築・管理（国家統計局）	6
1-3 政策評価と結果の蓄積のための組織（WWC、EEF）	8
1-4 大学や研究者との連携（ロンドン大学教育学研究所）	16
1-5 統計的会計検査の例（イギリス会計検査院）	20
2 フィンランド	28
2-1 調査の概要	28
2-2 制度的基盤としての教育統計データベース	28
2-2-1 データの収集・管理（統計局）	28
2-2-2 データのアクセシビリティ（教育庁）	30
2-3 大学や研究者との連携（VATT）	32
2-4 統計的会計検査の例（フィンランド会計検査院）	34
3 おわりに	50
【参考文献】	53
【付録：調査日程】	54

1 イギリス

1 - 1 調査の概要

厳しい財政状況のもと、イギリスにおいては、効率的な公共政策の実施の必要性が強く認識されている。そのような状況を象徴的に反映する動きとして、2010 年、イギリスの第 1 次キャメロン政権が発足した時に、What Works Network というイニシアティブが立ち上げられた。これは政府や関係機関が政策をはじめとする意思決定を行う際に活用することのできる、質の高い根拠（エビデンス）を生み出し¹、共有し、活用することを目的とするネットワークである。このネットワークの理念は、政策をはじめとする意思決定を行う際に、政策担当者は既存のエビデンスを知っているべきであり、もしそのようなエビデンスがない場合には、質の高い統計的な手法を採用して政策立案の判断をすべきであるという考え方に基づいている。このような根拠に基づく政策形成の重要性は、各省庁における政策形成においても意識されており、会計検査においても重要な視点となっている。

今回のイギリス訪問においては、EBPM を推進する上で必要不可欠なインフラであるデータベースの整備状況およびその活用とエビデンスの蓄積、さらには政策形成へのフィードバックの有無に関して、特に教育政策分野に関連する諸機関を対象にヒアリングを行った。具体的には、データの収集に関して教育省および国家統計局を訪問し、教育データベース構築の経緯および現状、さらにはデータの管理や利用方法に関するヒアリングを行った。次に、教育政策のエビデンスの生成および蓄積の先進的な組織である What Works Centre のサットン・トラスト教育基金に対してヒアリングを行い、政府から独立した機関として、エビデンスをどのように生み出し、蓄積してゆくのかについてのヒアリングを行った。また、教育データベースのユーザーである大学研究者としてロンドン大学の教育学研究所にある教育改善科学センターの Rebecca Allen センター長にインタビューを行い、大学や研究者と政策立案当局との関係性や課題についてヒアリングを行った。最後にイギリスの会計検査院において、教育政策関連の検査を行っている部署に対してヒアリングを行い、会計検

¹ 根拠を「生み出す」（あるいは「作り出す」）というのは、既に実施された政策の効果検証を行い、将来類似の政策の実施が検討される際の判断材料を提供するという意味であり、実施予定の政策にとって都合の良い情報を集めるという意味ではない。当然、政策効果の検証結果には、「効果がない（弱い）」というものも含まれるが、そのような根拠も政策立案において重要な判断材料として採用されるべきである。

査においてデータや統計的手法がどのように活用されているのか、またそのような手法を活用することを可能とする体制や人材はどのようなものなのかについてのヒアリングを行った。

1 - 2 制度的基盤としての統計データ

1-2-1 データの収集(教育省)

National Pupil Database (NPD) はイギリスの学校教育に関する包括的なデータベースであり、そこに格納されているデータは教育省 (Department for Education) により収集されている。このデータベース構築の背景や経緯、現状や課題について、また、デジタル政府のための教育データ整備について、それぞれ、教育省の担当者 (Iain Bradley 氏、Jeremy Gould 氏) にヒアリングを行った。ヒアリングの結果は、次のア、イのとおりである。

ア NPD 構築の背景・経緯、現状・課題について (Iain Bradley 氏へのインタビュー)

(1) NPD 構築の経緯

NPD は、学校への予算配分を決定することを目的として、1999 年に運用開始された。NPD に収められている情報は、もともと学校への予算配分を決めるための情報であるため、予算を決める上で必要となる児童・生徒の属性情報 (例えば特別支援教育の必要性など) が必要とされる。運用開始当初は集計データのみが利用可能であったが、その後、児童・生徒の個票データもデータベースの中に含めるようになり、利用可能となった。NPD の中には、年3回行われる学校センサスの情報も含んでおり、教育予算算定の根拠となると同時に、教育政策を評価する上でも重要なデータベースとなっている²。

² 国家全体のデータベースに児童生徒の個人レベルの情報を含めることに対しては、個人情報保護の点から意義が唱えられる可能性もあったが、もともと予算配分決定の根拠として NPD を運用していたため、NPD に児童・生徒レベルの個票データを含めること自体に対してそのような点からの異議は唱えられなかった。

(2) データの内容

NPD は学校基本調査をはじめとして、五つの教育段階に関する学力調査など 25 の調査を束ねた包括的データベースとなっている。NPD の中には、生徒数や生徒の基本属性（例えば性別、年齢、民族的属性）はもちろんのこと、学力成果についての情報も含まれている。各学校において教員が採点している通常のテスト結果は含まれてはいないものの、全国統一テストの結果や Examining board GCSE (general certificate of secondary education) という中等教育を受けるためには受けなければならない達成度テストの結果も、その他の属性とともに NPD の中に含まれている。その他の情報として養子縁組情報や特別支援教育の必要性に関する情報も含むが、これらの情報は Children in Needs Census と呼ばれる調査で収集されたものである。これらのデータの NPD 上での紐付けは、生年月日と住所を用いて行われ、NPD 上で作成された ID が付与されている。NPD に含まれる主な情報は表 1 にまとめてある。

ID	児童生徒 ID および学校 ID
基本属性	性別、年齢、英語が第二外国語か、民族的属性等
その他属性	給食費免除、特別支援教育の必要性、居住地郵便番号、貧困状態、養子縁組情報等
テスト結果	5 歳(Early years foundation stage profile) 、 7 歳(KS1) 、 11 歳(KS2) 、 14 歳(KS3) 、 16 歳 (KS4 または GCSE) 、 18 歳 (KS5)
学校属性	学校規模、生徒数等

表 1：NPD に含まれる主なデータ

NPD に含めるデータの選択は、政治的な論点となりやすいので、最終的な決定権は教育相が持っている。ただし、NPD に含まれるデータは教育省関係者や校長会の代表と協議され、その案を教育相が了承することにより決められる。例えば、国籍や出身国、英語の流暢さなどは移民政策に関連する項目なので、以前は重要なテーマであったが、2017 年の Brexit の決定以降は関心が薄れているとのことであった。また、ウィンドラッシュ世代（1948 年

から 1970 年代初頭にかけて旧植民地であるジャマイカなどカリブ海の国々から受け入れた移民) の子供かどうかに関連するような情報を NPD に収録することは違法である。家計に関する情報は English Indices of Deprivation 2015 から得た家計情報から IDACI (Income Deprivation Affecting Children Index)³ を郵便番号ごとに集計し、それを住所情報を用いて紐付けした上で収録している。

(3) データ収集およびセキュリティ

データの収集はコンピュータシステムを通じて行われるが、様々な API (Application Program Interface) システムが並存して非効率であるので、政府デジタルサービス (Government Digital Service) の協力のもと統一化を図っているとのことであった。データの入力、校長によって権限を与えられた事務職員により入力される。なお、NPD に収録されているデータは予算配分の基礎となる資料なので、NPD への情報の登録は教育予算の配分を受ける上で必須となっている。例えば、学校基本調査への回答は 1996 年の教育法第 537A 条において義務付けられている。

NPD は有益ではあるが、どう管理するのが大切である。集められたデータの管理はクラウドサーバで行われており、システムの安全性を検証するために、定期的に侵入テストを無作為に行う。なお、NPD に収録されているデータそのものは匿名化されていない。

NPD は公共の利益になっているというコンセンサスはあるものの、どのように管理し、活用してゆくのかは常に論点となっている。特に、NPD における個人情報保護は市民の権利に直結する問題であり、その分ロビー活動も活発に行われている。

(4) データ利用方法

NPD を利用したい研究者は、1 日コースを受講後、テストを受け、合格すれば 5 年間有効なライセンスが付与される。現在はこのプロセスの管理に 8 人の職員が当たっている。

(5) 政策形成との関連

³ IDACI は郵便番号によって定義された地域における低所得家計にある 16 歳未満の子供の割合である。

NPD は 教育予算の配分を決める基礎資料なので、そもそも教育政策の企画立案のためにあると言えるが、近年では、アカデミースクール（国費で賄われる中学校）の教育成果の事後的検証にも使われている。また、低所得層の教育格差問題の解消のための補助金支給額を決定する際にも使われており、予算配分を決定する際の基礎的資料としての役割が大きい。

イ デジタル政府のための教育データ整備（Jeremy Gould 氏へのインタビュー）

〔データ収集およびセキュリティに関して、前記のとおり、政府デジタルサービスと協力しながら進めているということであるので、Deputy Director of Digital Services の Jeremy Gould 氏に、教育データのデジタル化についてヒアリングを行った。〕

8 年前に政府のデジタル化のための大きな動きとしてデジタル経済法（Digital Economy Act 2010）ができた。この法の下で、政府サービスのデジタル化を推進するために様々な分野のデジタル化が図られた。その中でも教育分野はデジタル化が最も遅れている分野の一つであった。

まずは、教育省内の業務を見渡すために、教育省の中のどこにどのようなデータがあるのかをリストアップし、重複するデータやテーマに基づいて各サービスのクラスター化を行った。その結果、271 あったサービスラインが 57 のクラスターまでに集約された。

データベースを構築する際に、ユーザーフレンドリー⁴なものを作ることを目指し、ユーザーのニーズに応えるデータベースの構造を明確化した。学校が教育関連のデータを入力するシステムは、それまで五つの異なるシステムが並存しており、入力作業に重複が発生するという非効率なシステムとなっていたので、その統一化を図ることも行われた。また、教員、生徒、保護者のデータの接続を行った。

別の教育上の課題としては、学校教員の志望者の不足があった。どの時点で教員志望の大学生が教員志望から漏れるのかを明らかにするため、既存の UCAS（Universities and

⁴ ここでいうユーザーとは校長先生を含む学校関連の人々を指している。なお、Gould 氏は、「ユーザーフレンドリーというよりもユーザーセンタードと言うほうが適切かもしれない」と述べていた。

Colleges Admissions Service) データベースを整備した。この作業を行う上で、各個人を識別する ID を使うことが教育省により許可されなかった⁵ため、既存の教員志望者のデータベースや教員養成系大学への志願情報のデータとの結合を行った。また、デジタル化の別の例として、教員の研修方法に関するビデオを作成して、教育上の課題を教育省幹部に説明する際に用いたこともある。

個人情報の収集と個人情報保護のトレードオフは常にあるので、厳格なガイドラインを作成し、法的な要件を明確にした上でデジタル化を進めることは大切である。特に、省内でのデータシェアさえも認められない教育分野においては、これらの重要性は非常に高いものである。

ウ 教育省についてのまとめ

数多くある教育関連調査を NPD という一つのデータベースとして集約し、予算配分をはじめとする教育政策の基礎資料として活用している点は非常に参考となった。NPD には児童生徒の基本的な情報のみならず、教育成果に関する情報も含まれているため、教育政策の評価分析を行うことが可能となっている。実際に、数多くの研究者が NPD を活用して研究を推進しており、教育政策に関するエビデンスの蓄積のためのインフラを提供していると言える。

1-2-2 データベースの構築・管理(国家統計局)

国家統計局 (Office of National Statistics) は、教育分野をはじめとして、様々な政策分野のデータを管理している。そこで、同局において公共サービスの生産性に関する推計業務を行っている Fred Foxtton 氏に対して、どのようなデータを収集管理しているのか、またデータの利用や利用法、さらには利用例としてレポート“Public Service Productivity Estimates: Education 2013”などについてのヒアリングを行った。その内容は次のとおりである。

⁵ Gould 氏のデジタル化チームには、政府機関である政府デジタルサービス出身者が多く、同じ教育省であっても、傘下にあるというより、教育関連の NPO という位置づけだったとのことである。

(1) 国家統計局が管理する教育関連データ

国家統計局が管理するデータは、他の省庁が収集したデータであり、教育関連データは教育省により収集されたものである。特に、NPD は、イングランドにおける教育成果に関するデータを集めていて、国家統計局の中にある virtual microdata lab という安全な IT 環境を通じて研究目的の二次利用提供を行っている。ただし、後述の国家統計局のレポートは、公表された集計データを用いて書かれたものである。

(2) レポート“Public Service Productivity Estimates: Education 2013”について

このレポートは、公教育の生産性を計測したものであり、公共サービスの生産性を測ることを勧告したアトキンソンレビュー（Atkinson Review）に応える形で作成されたものである。アトキンソンレビューとは、イギリス政府の生産高と生産性を測定したものであり、オックスフォード大学 Nuffield College の Tony Atkinson 教授が、国家統計局、財務省、保健省、およびイングランド銀行から派遣されたチームの支援を受けてレビューしたものである。同レビューの主な目的は、イギリス政府の生産性を測定するために、使用できる方法とアプローチを推奨することである。

レポート“Public Service Productivity Estimates: Education 2013”は、一般的な枠組みと原則を勧めることに加えて、健康、教育、公の秩序と安全と社会的保護の主要な機能領域を測定するための実用的な解決策も提示したものである。

本レポート自体は、アトキンソンレビューに応える形で出版されたものではあるが、国家統計局は 1998 年以降、公教育サービスの生産性計測をすでに行っていたため、教育サービスの生産性計測を新たに始めたというわけではない。また、公教育の生産性の計測はこのレポート以降も作成されているが、公教育単体としてのレポートの作成は行っていない。現在は全公共サービスの一部として公表されている。

教育サービスの生産性の計測結果は、説明責任というアトキンソンレビューの視点からは活用されているが、翌年度以降の政策形成や予算計上の際に考慮する材料としてどの程度用いられているかについては、はっきりとしたことは分からない。ただし、政策形成の議論に対して質の高いエビデンスを提供するという使命を持って行っていることは確かであ

る。イギリスの公共サービスの生産性は歴史的に低く現在もそうであるが、要因は一つではなく、さらなる分析が必要である。公共サービスは国民経済の2割を占めており、公共部門の生産性を計測することで、どのように生産性を高めていくのかについての知見を得ているとは言える。

生産性の計測の際には、生産物の量だけではなく質に対しても考慮している点は特徴的である。例えば、公教育の生産物を測る際に、卒業生の数のみならず、出席日数や学力調査の結果、GCSE（General Classification for Secondary Education）の結果も考慮し、加重した上で生産物を計測している。質の計測の改善は国家統計局全体で引き続き行われていて、他の分野に関しても行われている。その中でも公教育サービスの質の計測は最も重要なことのひとつとして議論されている。

(3) 国家統計局と外部部局との関係

国家統計局の職員の中で、経済関連のレポートを作成している職員は主に学部卒業レベルであるが、より専門的な知識を活用するために、国家統計局の出資により ESCoE（Economic Statistics Centre of Excellence）という機関を設立し、国家統計局外の経済学研究者を受け入れている。先述の生産物の質の計測の改善においては ESCoE の研究者と共同で行われている。

国家統計局自体は政府からの独立機関であり、WWC（後述）や会計検査院と特別な連携関係があるというわけでもない。内閣は最も重要な国家統計局のユーザーではあるが、独立性を持っており、社会的なインパクトを考慮した上で統計の発表を行っている。また、デジタル経済法という政府の電子化の下で国家統計局の役割は大きくなっている。

1 - 3 政策評価と結果の蓄積のための組織(WWC、EEF)

イギリスには、EBPM を推進するイニシアティブとして What Works Network があり、このネットワークには、政策分野別に七つの What Works Centre (WWC) がある。その中で、教育分野を担うのが公益財団のサットン・トラスト教育基金 (Sutton Trust / Education

Endowment Foundation : EEF) である。WWC や EEF はイギリスにおける EBPM を象徴する存在のように思われる。それぞれのインタビュー結果を紹介する。

ア WWC について

What Works Network は、イギリスの EBPM を推進するイニシアティブであり、このネットワークは、政策分野別に七つの What Works Centre (WWC) から成り立っている。七つの政策分野とは、健康と衛生、教育、犯罪減少、幼児教育、地域経済成長、高齢者問題、福祉であり、全体で 2000 億ポンド以上の公共支出を伴う政策範囲がカバーされている（表 2 を参照）。各センターは、主に次の五つのことを行う。①既存の政策の有効性に関するエビデンスの収集、②未開の分野に関して、エビデンスに関する質の高い包括的なレポートの作成、③共有されている成果指標に対する政策の有効性の評価、④エビデンスを共有するための環境整備、そして⑤政策当局者や実務家に対するエビデンス活用の勧奨である。これらのセンターは、政府からの資金と、経済社会研究評議会（Economic and Social Research Council）や国営くじ基金（Big Lottery Fund）といった非政府系の基金で賄われているが、基本的には政府から独立した機関として位置付けられている。

センター名	政策分野
National Institute for Health and Care Excellence (NICE)	健康と衛生
Sutton Trust/Educational Endowment Foundation	教育
College of Policing What Works Centre for Crime Reduction	犯罪減少
Early Intervention Foundation	幼児教育
What Works Centre for Local Economic Growth (hosted by LSE, Arup, Centre for Cities)	地域経済成長
Centre for Ageing Better	高齢者問題

センター名	政策分野
What Works Centre for Wellbeing	福祉

表2：What Works Centre に参画している機関と政策分野

イ EEF について

次に EEF の設立の経緯や政府からの独立性、エビデンスの生成および蓄積方法について、EEF の Jonathan Kay 氏および Maisie Monroe 氏にヒアリングを行い、政府の予算編成や政策形成との関係性について調査した結果を述べる。

(1) 設立の経緯および概要

WWC の七つのセンターのうち、教育分野に最も関係しているのは、公益財団のサットン・トラスト教育基金（Sutton Trust／Education Endowment Foundation：EEF）⁶である。この公益財団は、2011 年に教育省から 1250 億ポンドの資金を得て設立された。EEF 自体は 2011 年に設立され、イギリスの貧困層と富裕層の教育格差を埋めることを目的とする組織である。そのために、教育政策のインパクト評価や、それに基づくエビデンスの作成を行っている。特に、質の高いリサーチデザインの基で収集された分析結果の包括的レビューを行い、それらの結果を活用するためのレポートの作成を行っている。

EEF は、教育上の有効な取組を発掘するために、プログラムへの資金拠出やその評価への資金の拠出を行っている。この評価自体は、大学の研究者によって独立に行われる。現時点で 100 以上の学校のプロジェクトに資金を拠出している。

EEF は、教育、心理学、統計学、政策評価等の専門家で構成されている。2018 年にはフルタイム、パートタイム合わせて 35 名のスタッフがいる。スタッフの約半数は以前教員と

⁶ サットン・トラストはイギリスの教育的慈善団体で、社会的流動性の向上と教育上の不利益の解消を目的としている。この慈善団体は 1997 年に教育慈善家の Sir Peter Lampl によって設立され、幼児、中等、高等各教育機関へのアクセスを支援する幅広い実践的プログラムに資金を提供している。

して働いていた。評価チームのスタッフのうち、3 名が博士号（Ph.D.）を保有しており、専攻分野は経済学、教育学、心理学となっている。他のスタッフも少なくとも修士号を取得している。なお、図 1 は 2018 年 9 月時点での EEF の組織図である⁷。



図 1：サットン・トラスト教育基金の組織図（2018 年 9 月）

評価チームは、教育成果を図るためのアウトカム指標を何にするのか、無作為化比較試験（RCT）⁸をどう実装するのか、などについて決める。プログラム評価に当たっては、2018 年時点でプログラムマネージャー 7 名、評価マネージャー 2-3 名、事務職員 1 名、資金調達スタッフ 1 名が働いている。

⁷ 評価活動の内容を決める部署とその評価を行う部署、さらに外部への発信のためのレポートを作成する部署が分かれている点は印象的である。

⁸ 無作為化比較試験（Randomized Controlled Trial: RCT）とは、政策効果を検証するための統計的手法の一つ。ある政策の効果を検証するためには、政策の影響を受けたグループ（処置群）と、政策の影響を受けないグループ（対照群）の比較が必須である。無作為化比較試験では、処置群と対照群の割り振りを（くじ引きなどにより）無作為に行うことで、処置群と対照群の事前の均質性を保ちつつ施策後の成果を比較することで施策の効果を検証する方法である。

(2) EEF の活動

EEF はこれまで、100 の評価活動を行い、570 億ポンドの予算を用いて 4,900 の学校における 62 万人の児童に対するプロジェクトを実施してきた。また、1 万件以上の研究成果に基づいて、教育手法の費用と効果の評価を行っている。

様々なテーマのもとでプロジェクトが実施されているが、各プロジェクトはその進捗に応じて次の表 3 にある三つの段階に分けられる。

第 1 段階	実施が決まっているプロジェクトにおいて、具体的な政策介入方法についてのアイデアを募る。
第 2 段階	プロジェクトへの実装方法が決まり、その効果を計測する。
第 3 段階	プロジェクトの遂行が完了し、その効果は第三者の独立機関により評価されエビデンスとして蓄積される。

表 3：プロジェクトの段階

さらに、蓄積されたエビデンスに基づき、各教育施策に関する効果の確からしさについて、次の表 4 の基準に基づいて評価される。さらに、それぞれのテーマにおける成果指標の改善に効果的な方法がエビデンスの確からしさとともに公開され、その実装方法はツールキットとして配布される。これらのツールキットやエビデンスをまとめたガイドやレポートは、学校現場の教職員により、学習成果の改善のための資料として活用される。

非常に限定的 ：一つの研究から得られた量的な評価結果であり、効果サイズが報告されているか、またはそれを計算できる。
限定的 ：学力や認知能力、またはカリキュラムの成果指標へのインパクトに関する量的なエビデンスに関するメタ分析または総合的レビューが少なくとも一つはある。
中程度 ：学齢期の生徒の認知能力、またはカリキュラムの成果指標へのインパクトに関する実験的な手法を用いた分析結果のメタ分析が複数ある。
頑健 ：実験的手法によって注意深く設定された対照群があり、学校で収集された児童生徒の達成度データを用いた分析に関するメタ分析が三つ以上あり、そのうちのいくつかは効果の異質性に関する分析も行われている。
非常に頑健 ：良好な生態学的妥当性を有し、かつ結果測定において学校の科目分野でのカリキュラム測定または標準化された試験結果を用いた研究が大半を占める、少なくとも五つの頑健で新しいメタ分析があり、それらの結果が一貫している高品質のエビデンスがある。

表4：評価方法の基準

(3) EEF の政府からの独立性

EEF は、サットン・トラストのような非政府系団体からの資金によって運営されているが、設立当初には教育省からも多額の資金提供を受けている。しかしながら、設立後の支援は少なくなっているため、政府から独立に教育政策評価を行うことができる。なお、EEF の外部委員の中には政府から送られてくる委員もいるが、会議で発言することはできても、意思決定プロセスに影響が及ぶことはない。

評価のための資金に関しては、特定の研究対象として政府から研究資金の提供はあるものの、プログラムの参加や評価は各協力校との関係に基づいて行われる。イギリスの教育システムの特徴の一つとして、各学校の自律性というものがあるが、評価プログラムへの参加は各学校が独自に判断できるので、EEF が協力校に直接声かけを行い、プログラムの評価を行うことができる。学校教育の質に関しては Ofsted（後述）による調査があり、それによって質の担保並びに政府および学校の説明責任を果たしている。また、EEF は学校単位

での結果の公表は行わない。プログラム評価のための間接費用は全体の 5%までであり、共同事業の場合には 7%までと決まっている。

(4) 公共政策の有効性検証という側面の有無

EEF は様々な独自プロジェクトを実施しているが、EEF が行っているのはあくまでも教室での取組に関する効果測定であり、EEF のエビデンスのユーザーは基本的に学校の先生が主となっている。教育省の政策はもっと上位レベルのものであるため、EEF のエビデンスが教育省の政策決定に直接影響を与える形にはなっていない。EEF のエビデンスやガイダンスが、政府の予算編成に影響を与えることは基本的にはないと認識している。

しかしながら、時には重要な政策的含意のあるテーマについてのエビデンスを生み出すこともある。例えば、習熟度別クラス導入の有無が学力形成に対して与える影響について分析したことがあるが、それによると習熟度別クラスを導入した学校では、そうでない学校に比べて学力成果が少しだけ低くなるという結果も見られた。また、地域の資金を用いて「マジック朝食」⁹の効果を RCT で調べたところ、プラスの効果が検出された。これらの結果は、さらに上位の教育政策形成に対しても少なからず影響を与えたのではないかと考えられる。

(5) システマティックレビューの方法

エビデンスの提供を行う際には、同一教育政策の効果分析に関する系統的レビューが必要である。この系統的レビューは教員へのフィードバックを行う際のツールキットの基礎となる。系統的レビューの実施に当たっては、その研究の経過年数（発表されてから何年経っているのか）、実験室で行われた研究か教室で行われたものかなどに応じてそれぞれの研究結果のウェイトを計算し、既存の結果を全て用いる方法を採用している。

⁹ マジック朝食（Magic Breakfast）は、慈善団体による貧困家庭の子供たちに朝食を提供するプログラム活動のことである。

(6) 学校の協力要請方法

上述のとおり、イギリスの学校は高い自律性を有するため、評価プログラムへの参加は各学校が独自に判断できる。実際には、大学の研究チームとつながりのある協力校もプログラムへ参加することが多いが、そうでない学校の参加も受け入れる。現在では、イギリスにある3万校のうちの約3分の1が何らかのプロジェクトに参加している。協力校に対しては、協力に対する謝金が学校に対して支払われ、各学校はその資金を学校運営に使うことができるため、プログラム参加のインセンティブが生まれる。この謝金支払はプログラム終了後に行われるため、協力校の途中での脱落は少ない。

(7) フィードバック：ツールキットの活用

WWCによって蓄積されたエビデンスは、ツールキットという形で教育関係者にフィードバックされる。会計検査院の“Evaluation in Government”というレポートによると、現場の先生の約3分の2はEEFのツールキットを参照しているとのことである。ただし、これは校長先生（head teacher）に占める割合であり、クラス担任の先生だと約30%が参照しているということである。

(8) アメリカ WWC との類似性と異質性

アメリカのWWC (What Works for Clearinghouse)は、既存の研究結果のレビューをRCTに基づいた研究に絞り込んでストイックに行っているという印象がある。そのため、イギリスのWWCに比べ長い歴史を持つにもかかわらず、多くのエビデンスが蓄積されているという印象は薄い。他方、イギリスのWWCは独自に様々な教育プロジェクトを実施してその評価を第三者機関が行い、そうした経験をエビデンスとして蓄積するという全く異なったアプローチを取っている。EEFでは、エビデンスの厳密性ととも、エビデンスの利便性も重視しており、エビデンスを使うという視点が重要である。

また、世界規模でエビデンスを収集するため、国際的な協力関係を結んでおり、オーストラリア、スコットランド、南米諸国、スペインのカイシャバンク（La Caixa Banking

Foundation) ともエビデンスの共有を行っている。日本については、日本財団および三菱UFJ リサーチ&コンサルティング株式会社と協力している。

ウ WWC、EEF についてのまとめ

EEF は教育方法に関するエビデンスを幅広く収集し、それを系統的レビューによりまとめ、ツールキットとして現場の教員にフィードバックするという体制を構築している。このようにして集められたエビデンスは、政策形成にとって重要なものもあるが、教育省における政策は予算配分等のより上位なものであるため、EEF のエビデンスが教育予算配分上の教育政策形成に直接用いられているというわけではないとのことであった。しかしながら、学校教育現場における「ベストプラクティス」を探求し、フィードバックを行うという点においては、教育予算形成と合わせて重層的な評価体制がとられているというふうに考えることもできる。

1 - 4 大学や研究者との連携(ロンドン大学教育学研究所)

教育データベースのユーザーとして、大学をはじめとする研究機関がある。ロンドン大学ユニバーシティカレッジ (University College London) の教育学研究所 (Institute of Education: IOE) の Rebecca Allen 教授を訪問し、教育データベースのユーザーの視点から、NPD の構築および維持上の問題点、また、研究に用いる際に留意すべき点などについてヒアリングを行った。

Allen 教授は、NPD を用いた実証分析を数多く発表しているが、NPD の研究利用体制を確立するためのイニシアティブにも深く関わっていた研究者の一人である。また、同教授は、NPD の研究利用に関する講義も数多く担当し、NPD の利用申請の方法や利用可能なデータの種類、さらには NPD の限界や課題について広く講義を行っている。例えば、サウザンプトン大学の the National Centre for Research Methods (NCRM) で行われた講義の内容はインターネットで確認できる¹⁰。そこでは、①NPD で利用可能な中心変数、②NPD

¹⁰ <https://www.ncrm.ac.uk/training/show.php?article=6225>

以外の行政データや社会調査データと NPD の接続方法、③分析に用いる変数の作成方法、④政策効果の定量的分析に NPD を用いる可能性と限界について講義を行っている。これらの講義は大学の研究者を中心とするユーザーグループの拡大のみならず、大学における研究教育を通じて NPD のより広いユーザー層の獲得に貢献していると考えられる。

以下にヒアリングの内容を紹介する。

(1) NPD の利用について

NPD に格納されている児童・生徒の個票を研究に利用するというイニシアティブは 2002 年に始まったが、それ以前は学校レベルの集計データのみが利用可能であった。ロンドン大学教育学研究所、ロンドン大学経済政治学院 (London School of Economics and Political Science: LSE) の教育経済学センター (Centre for economics of education) およびブリストル大学を中心とする 10 名程度の研究者グループ (ユーザーグループ) で、児童・生徒の個票データの研究目的の二次利用へ向けてのインフラづくりを始めたが、それ以前はそのようなインフラはなかった。現在は 100 名を超える研究者がユーザーグループとして関わっている。

研究のための NPD 利用を申請する際には、まず国家統計局において 2 日間にわたって開催される情報利用および統計分析の知識に関する研修を受講し、テストに合格して証書を授与される必要がある。証書を授与された研究者は、研究の目的および公益性、分析の手法並びに分析に必要な変数 (データ) を申請する必要がある。これは、申請書に基づき、必要とされるデータの (個人情報保護の点からの) 機微性の確認が必要とされるためである。また、分析に必要な変数リストに基づいて利用可能となるデータを作成するためにもこの情報が用いられる。申請書は審議会において審査され、貸与が決定されるとウェブを用いたセキュリティ上安全なシステムから利用可能となる。

(2) FFT Datalab について

NPD の他にも、FFT Datalab というデータ利用のためのリサーチラボが収集している学校や自治体の教育データベースがある。このリサーチラボは、慈善団体 (Fischer Family Trust) からの寄付により 20 年以上前に設立されたものである。このリサーチラボでは、技

術系企業の協力のもと、NPO や地方自治体において収集された教育関連のデータの収集を行い、研究利用のためのインフラを提供している。

FFT Datalab の特徴の一つとして、地方自治体から提供を受けた教育データの分析結果をデータ提供元の各自治体にフィードバックする点が挙げられる。現在、各地方自治体の教育データは、共通のデータ転送システムを通じて FFT Datalab に送られてくる。そのデータの分析結果は、他自治体を含めた全体の傾向および当該自治体の特徴という形で各自治体にフィードバックされ、各自治体はその分析結果を参考にしつつ教育政策の立案を行う。自治体からのデータ提供は任意であるため、全ての地方自治体が FFT Datalab にデータを提供しているわけではないものの、分析結果のフィードバックを受けることができることにメリットを感じている自治体も多く、新規自治体の獲得の上でも分析結果のフィードバックが重要な役割を果たしていると考えられる。

NPD の研究利用のためのインフラづくりは、当初 10 名程度の研究者グループにより始められた。このグループの中には、当時の教育省の統計担当者と個人的なつながりの強い研究者が含まれており、データ利用に向けての動きの初期においては、学術研究者と行政担当者とのネットワークが重要であった。また、同様のことは FFT Datalab においても当てはまり、特に地方自治体のデータ提供の初期においては、研究者との個人的な関係の強い行政担当者がいた。

(3) 教育省や EEF との関係

IOE は、教育省をはじめ EEF とも強いつながりを持っており、例えば、IOE において EEF に関するワークショップを開催したりすることもある。

IOE と教育省は研究プロジェクトへの IOE からの募集と、採択されたプロジェクトへの助成という形でのつながりはある。また、このようなプロジェクトの結果は少なからず EBPM の推進に役立っているという認識を持っている。ただし、FFT Datalab と教育省との関係はそれほど強くはないという印象を持っている。これは FFT Datalab において行っている活動の一部が教育省で行われている事業（例えば、地方自治体の教育行政データの収集）と重なる部分があるためであろう。

大学や研究者との連携についてのまとめ

研究利用可能な教育行政データベースの最も先進的な事例の一つであるイギリスのNPDも、その生成過程の初期においては研究者と行政担当者との個人的なつながりが大きな役割を果たしていたという点は、日本における状況とも近いと言える。日本においても、自治体における機微性の高い情報を用いた実証研究を行う際には、まずは研究者とデータ管理者（例えば行政側の担当者）との間の信頼関係に基づいた共同作業として開始されることが多く、後述のフィンランドにおける軍データの利用などとも共通の特徴であると言える¹¹。

イギリスのケースの特徴としては、イギリスは、個人的なつながりを持っていた研究者グループがそのデータベースの整備を行い、研究のための二次利用の窓口を他の研究者に向けて大きく開く努力を行ってきた点であろう。それにより、質の高い教育行政データベースを活用した研究が蓄積され、教育政策の立案上有益な情報を提供することを可能にしているのではないと思われる。日本では行政データをデータベースとして構築し、それを用いた研究を行うことはあっても、その行政データそのものを研究目的のための二次利用に広く開いた例はまだ多くはなく、今後推進されるべき方向性の一つであると思われる。

地方自治体のデータを統一のフォーマットでセキュアシステムを用いて収集し、その分析結果をフィードバックするという FFT Datalab の取組も非常に参考となる。学校の自律性の高いイギリスにおいては、学校単位での参加の決定を行うこともできる。日本においてはこのような取組は教育行政の窓口が地方自治体の教育委員会となっているため、データシェアシステムへの参加のためには、各学校間のコーディネーションが必要となり、イギリスよりも困難となることも考えられる。ただし、一旦、基礎自治体レベルの教育委員会で各学校長の合意形成がなされ、教育関連のデータベース構築が行われれば、各自治体の教育委員会に対してデータの提供を要請するだけで済むため、包括的なデータ収集という意味ではよりうまく機能する可能性もある。日本においては、基礎自治体レベルでの教育行政データベースの構築の取組がいくつかみられるが、複数の自治体が参画するコンソーシアムを形成することによって、データ分析の精度が上がり、各自治体にフィードバックされる分析結果の有用性も上がる。また、フィードバックされる情報の有益性が高まると、その他の自

¹¹ 例えば、著者は東京都足立区との連携のもとで教育委員会のデータベース作成に従事し、足立区への教育政策上のフィードバックを行うと同時に、学術研究も遂行している。このような連携が可能になった背景の一つに、著者が足立区の区民（行政）評価委員会活動に長年従事していたという背景がある。

治体の参画インセンティブも高まるであろう。このような取組の推進において、データ収集および管理のノウハウを蓄積している大学のデータアーカイブ等を活用することも一つの方法であろう。

1 - 5 統計的会計検査の例(イギリス会計検査院)

イギリス会計検査院（National Audit Office : NAO）において、統計的手法を活用した有効性分析についてヒアリングを行った。特に、教育分野における検査事例として、“Ofsted’s inspection of schools (2018)”と“Retaining and developing the teaching workforce (2017)”に注目し、これらの担当者である Laura Brackwell 氏（Director, Education value for money audit）、Mark Parrett 氏（Audit Manager）および Charlie Acton 氏（Analyst）にヒアリングを行った。さらに、有効性検査の位置付けと役割についても質疑を行った。両レポートについて、ヒアリングに基づいたレポートの概要を紹介する。

（１）学校格付け機関による学校調査 “Ofsted’s inspection of schools (2018)”

ヒアリングに基づいたレポートの概要

Ofsted (the Office for Standards in Education, Children’s Services and Skills)とは、子供や若者に対する教育の水準や生活の質の向上を目的とし、学校教育の実体調査、格付けおよびその公表を行うことを目的とする組織である。本レポートの検査の着眼点は、Ofsted によって実施されている学校調査の方法はVFMを生み出しているかというものである。検査の際の着眼点は、Ofsted の配置および調査方法は効率的かどうか、また Ofsted による調査は学校教育に対してプラスの効果を生み出しているのかというものであった。

検査は主に表5にある次の五つの方法によって行われた。

① インタビュー	検査方法やデータ、トレーニングやその質に対して責任のある Ofsted のスタッフおよび教育省の職員を対象
② 文献調査	Ofsted の運用に関する法律、2016-17 年および 2017-20 年の Ofsted の戦略的優先事項を定めた文書、Ofsted の公開されている共通の調査フレームワークとハンドブック、Ofsted がその調査プログラムをどのように計画しているかに関する文書、Ofsted の内部管理情報および年次報告書
③ Ofsted と教育省のデータ分析	Ofsted が調査期間中に完了した調査、Ofsted の品質保証プロセスからのデータ、2010 年から 2011 年までの Ofsted の支出データ、および Ofsted のコストモデリングツールからのデータ分析
④ 六つの学校の Ofsted による調査の視察	Ofsted の調査方法を理解
⑤ 校長先生へのアンケート調査	公立学校の校長先生に対するオンライン調査

表 5：Ofsted 調査の段階

上記の「⑤校長先生へのアンケート調査」の活用は、会計検査における統計的手法の応用である。2018 年 1 月、DJS リサーチという外部団体に依頼して、公立学校の校長先生に対してオンライン調査を実施している。2015 年 9 月から 2017 年 8 月の間に Ofsted によって調査された約 7,800 の公立学校を対象として、Ofsted の調査の全体的な有効性に対する回答（優れている、良い、不十分で改善が必要な 3 段階）を集め、785 校から有効回答を得ている。

質問では、学校の最新の Ofsted 調査に関して、①調査チームの規模と専門知識、調査の期間、および調査の一環として調査担当者が検討した調査証拠、②調査中および調査後に学校に提供されたフィードバック、③Ofsted による指摘の方法、④調査の公正さと負担、について回答してもらった。会計検査院は、回答の偏りを考慮するために、小学校と中学校の数および Ofsted による学校の格付け結果を用いて加重した上で結果を計算している。また、

平均的な回答に合わせて信頼区間を報告することによって、どの程度それらの回答が信頼できるものかについても言及している。

Figure 13

Headteachers' views on the fairness of inspection outcomes

Headteachers of outstanding and good schools were more likely to think that the outcome of their most recent inspection was fair

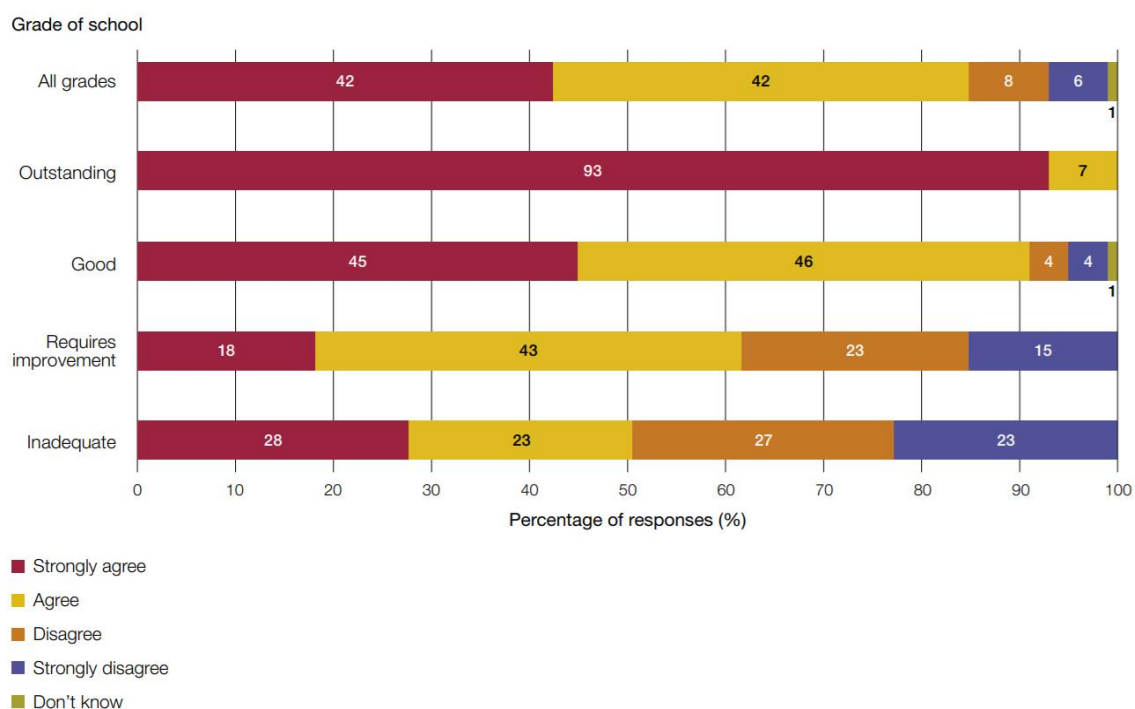


図2：調査結果の公平性に関する校長先生の見解（NAO(2018)より転載）

校長先生へのアンケート調査結果のうち、調査結果の公平性に関する校長先生の見解を Ofsted による格付け結果ごとにまとめたものが図2である。84%の回答者が、学校での最近の調査の結果は公平であったと回答している。また、格付けが高いほど、最新の調査の結果が公正であると考えた回答者の割合が高くなっている。

そのほかにも、校長アンケートでは、44%が調査で学校の改善につながったと答え、28%がそうではないと答えたことから、半数弱の学校が Ofsted による調査が役に立ったと答えている。さらに、回答者の 71%が、調査訪問中および調査終了時に調査担当者が有用なフ

ードバックを提供したとも答えている。これらの結果は、Ofsted の調査の有効性を裏付ける数量的な根拠となっている。

(2) 教員の確保と養成 “Retaining and developing the teaching workforce (2017)”

ヒアリングに基づいたレポートの概要

教育省は、学校教育の質を担保する上で質の良い教員の採用、トレーニング、確保が最優先であると考えている。そのために、教育省は学校が教員の養成を行うことを期待しており、それをサポートするため様々な金銭的補助プログラムを行っていて、2016-17 年では総計約 3570 万ポンドの助成を行っている。会計検査院は、この目的が達成されているのかを検査している。教育省はどのようにしてこの目的を達成し、そのシステムが効率的に機能するようにしているのか、学校による教員確保や採用をどのようにサポートしているのか、教員の質を向上させるために何を行っているのかに着目して検査を行っている。

検査は主に表 6 にある次の五つの方法によって行われた。

① インタビュー	教員確保養成プログラムの各担当者（教育省内）を対象
② 文献調査	自己評価書、内部文書、他機関により出版されたレポート
③ 教育省のデータ分析	学校教員全数調査、児童生徒数予測、学校による教職員への支出、教育省による教員供給モデル、Ofsted による学校調査結果
④ 二つの小学校と三つの中学校の訪問調査	小中学校を実際に訪問し、実態を調査
⑤ 校長先生へのアンケート調査	公立学校の校長先生に対するオンライン調査

表 6：教員養成プログラムの検査の段階

前記（1）の Ofsted への会計検査と同様に、本件の「⑤校長先生へのアンケート調査」の活用は、会計検査における統計的手法の応用であると言える。会計検査院は、2017 年 1

月-2月、DJS リサーチという外部団体に依頼して、公立学校の校長先生に対してオンライン調査を実施している。10,000 人以上の校長先生を対象として、小学校と中学校の比率を考慮し、また、標本の代表性を考慮した上で調査を行い、258 人の小学校長と 201 人の中学校長から有効回答を得ている。

質問では、①教員の質を向上させ、教員を採用し維持することに対する障壁とインセンティブは何か、②教師のための公式および非公式のトレーニングの費用、③質の高い教師を募集した経験およびその方法、④教育省および教員養成大学からの支援やプログラムへの参画の有無について回答してもらった。会計検査院は、回答の偏りを考慮するために、小学校と中学校の数を用いて加重した上で結果を計算している。また、平均的な回答に合わせて信頼区間を報告することによって、どの程度それらの回答が信頼できるものかについても言及している。

Figure 7

School leaders' views of the factors that affect teacher retention

School leaders consider that workload is the most significant barrier to teacher retention

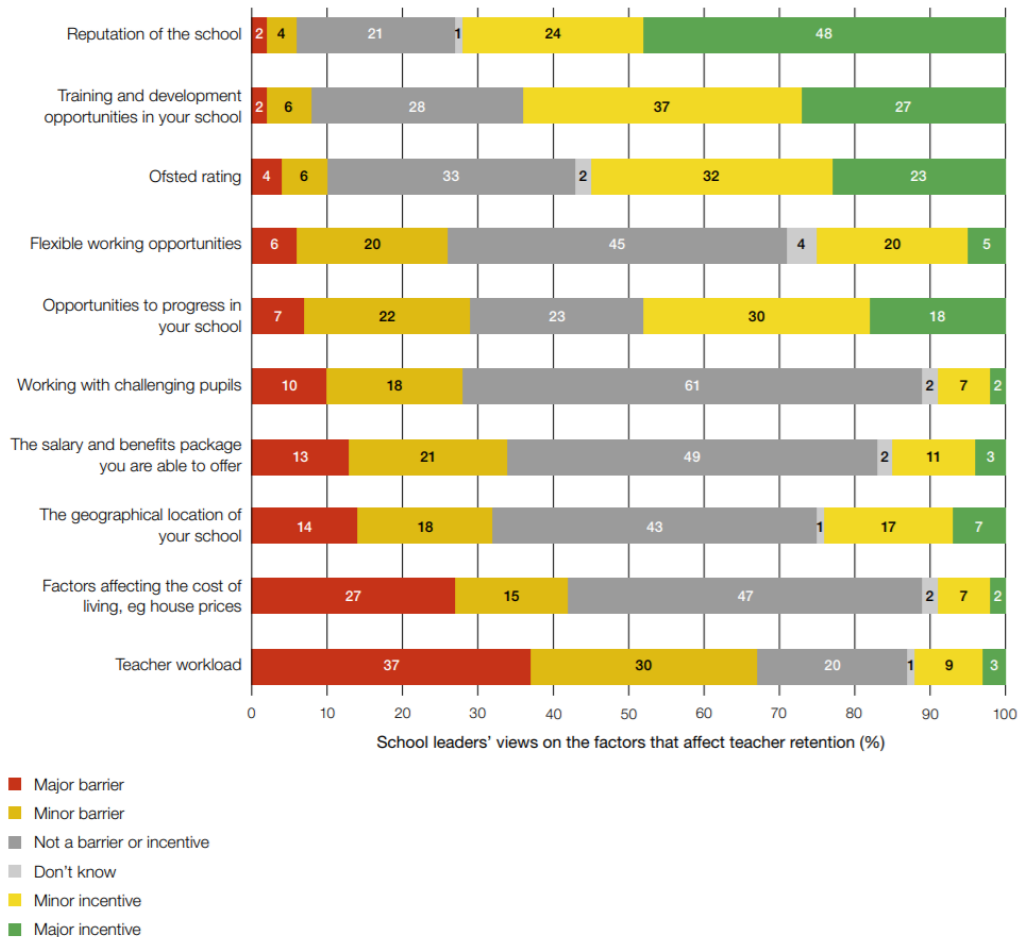


図3：教員確保の要因に関する校長の見解（NAO（2017）より転載）

図3に教員確保のための要因に関する校長先生の見解がまとめてあるが、このアンケート調査から分かったことは次のとおりである。

まず、教員が自らの訓練と能力開発をあまり行っておらず、訓練や能力開発のための時間と費用が教員の質を向上させるための障壁となっていることである。アンケート調査の結果では、少なくとも94%の校長先生が、時間と費用が教育労働力の質を向上させる上での障壁となっていると回答している。また、学校における教育活動の大幅な効率化を図る必要性は、教員の能力開発を支援することをより困難にする可能性があることを示唆している。このことは、作業負荷が教師の定着にとって大きな障害であるというアンケート結果とも

整合的であり、調査では教務をはじめとした事務作業を含む作業量が、教員を維持・確保するための最も重要な要素であることがわかった。回答者の 67%が、作業負荷の重さは定着の障壁であると報告している。

さらに、学校が必要とする教員の質を確保することはますます困難になっていることも明らかとなった。2015/16 年の間に、教員ポストを学校で必要な経験と専門知識を備えた有資格の教員で埋めることができたのは半分にも満たなかったとの回答であった。学校は一般的に残りのスロットを経験の浅い教員、あるいは異なる専門分野の教員で埋めたが、約 10 分の 1 のケースでは、教員のスロットを埋めることができていなかった。

会計検査院は、これらの結果に基づいて、教育省はこの問題に対して一貫性のある対策がとれているわけではなく、学校と教育の専門家と一緒に、教育労働力を改善するためにどのように協力できるかを共有していないこと、および既存の教育労働力を支援するための教育省の介入は比較的小規模であり、教員を維持・確保するのに十分な支援を学校に提供することができていないという結論に達している。

(3) 会計検査院の検査レポートについてのまとめ

Ofsted の学校調査に関する検査レポートにおいて、校長先生へのアンケート調査を有効性検査に活用している点は、Ofsted による調査の直接的な受益者の一人である校長先生に対する直接的かつ幅広い検査を可能にしている。さらに、インターネット調査で 10%程度の回収率となっており、サンプルセレクションの影響が懸念されるが、加重平均を計算することで母集団を意識した統計的処理が行われている。特に検査の結果としての Ofsted による格付けも考慮している点も統計的有効性検査の結果の信頼性を高めていると言える。

上記の検査レポートと同様に、教員確保・養成に関する検査レポートでも、校長先生へのアンケート調査を有効性検査に活用している点、加重平均を計算することで母集団を意識した統計的処理が行われている点は、統計的有効性検査の結果の信頼性を高めていると言える。これらの検査レポートで示されたように、独自のアンケート調査を用いた統計的推論は、その政策の受益者に直接かつ幅広くアクセスして有効性に関する情報を収集することのできる手法であり、会計検査における有効性検査にも活用できるものであると考えられる。

会計検査における有効性検査の結果は、政策評価の結果でもある。このような検査結果を蓄積することにより、今後同様の教育政策を行う際の潜在的な有効性についての情報を提供することもできるはずである。特に、統計的有効性検査によって収集・分析された検査結果を蓄積してゆくことは、EBPM の基礎である根拠の積み増しにおいても重要な役割を持つ。場合によっては独自にアンケート調査を実施することによって、検査に活用することは、有効性検査の遂行と同時にエビデンスの蓄積という点からも正当化されるものである。

最後に、このような統計的有効性検査が可能となるためには、会計検査院内に統計的手法や調査方法に明るい人材のいることが大切である。教育分野関連の VFM チームには、会計士や、統計学、経済学や社会科学のバックグラウンドを持つ調査官がおり、上述のような社会調査を設計する際には、院内の別のチームから経済学や統計学を専攻した人材を招聘することもあるとのことであった。EBPM 推進という点からも、このような検査を実施する技能を持つ検査職員の育成および招聘は重要であり、イギリス会計検査院はこのような取組を意識した上で行っているといえる。

2 フィンランド

2-1 調査の概要

北欧諸国の中でも最も早く少子高齢化が進んでいるフィンランドでは、近年、財政状況が悪化しており、福祉の高水準を維持するのが困難になってきている。フィンランドでは2012年に、会計検査院の管轄下に独立財政機関を設置し、①独立した財政政策の監視および評価、②財政政策規則の監視、③財政政策の情報ベース、有効性および統制の検査並びに評価を行っている。EBPMを明示的に制度化した上での政策立案は必ずしも行われているわけではないが、財政状況の悪化から、効率的かつ有効性の高い政策を優先的に実施してゆく必要性は高く、データベースの整備状況や会計検査における統計的手法の積極的活用など、EBPM推進のための前提条件は比較的整っている。

今回のフィンランド訪問においては、EBPMを推進する上で必要不可欠なインフラであるデータベースの整備状況およびその活用とエビデンスの蓄積、さらには政策形成へのフィードバックの有無に関して、特に教育政策分野に関連する所管を対象にヒアリングを行った。具体的には、データの収集に関して国家統計局および教育庁を訪問し、教育データベース構築の経緯および現状、さらにはデータの管理や利用方法に関するヒアリングを行った。

次に、教育データベースのユーザーである研究者としてフィンランド政府経済研究所において、教育データの研究利用と、分析結果の政策へのフィードバックの現状について、学術研究者としての見解のヒアリングを行った。最後に会計検査院に対してヒアリングを行い、会計検査においてデータや統計的手法がどのように活用されているのか、またそのような手法を活用することを可能とする体制や人材はどのようなものなのかについて調べた。

2-2 制度的基盤としての教育統計データベース

2-2-1 データの収集・管理(統計局)

統計局（Statistics Finland）における教育データの収集および管理について、Mr. Mika Tuononen（Chief Advisor, Population and Social Statistics/ Education, Culture and Time Use）並びに彼の同僚である Mr. Marko Ylitalo および Mr. Mika Witting に対してヒアリングを行った。その内容は次のとおりである。

(1) 教育データの収集について

フィンランドの教育データは、主に教育文化省（Ministry of Education and Culture）によって収集され、統計局はその管理を行っている。まず、生徒に関するデータは業務データ（registration data）として収集されているもので、これらのデータは各学校で入力されたものをファイル転送システムを使って集めており、約 300 の自治体にある 770 の教育機関からデータを収集している。このシステム自体は 1990 年代末に導入され、統計局におけるウェブでのデータ収集の最初の例である。これらのデータの登録は様々な補助金を受け取るために必要な手続となっており、教育関係機関にはデータを登録するインセンティブがある。

(2) 統計局において管理しているデータについて

統計局では、様々な教育データの管理を行っている。各教育段階の業務データに加えて、高等教育に関するデータとしては最終学歴の登録データがある。これはフィンランドの個人ごとの最終学歴に関するデータであり、様々な問題（例えば移民問題）について議論をする際の重要な基礎的情報となっている。そのほかにも、大学生の登録データからは大学生の専攻選択の実態についての情報を得ることができる。

(3) データセキュリティについて

収集／管理している教育関連のデータは、個人 ID で管理されている。ただし、統計局内でも個人 ID へアクセスできる人数は限られている¹²。個人 ID で管理されているデータは匿名化されてはいるものの、個人特定のリスクに対する明確な基準は設定されていない。

データの収集・管理についてのまとめ

教育関連のデータは、生徒レベルの個票データを含めて、統計局で一元的に管理する体制に移行しつつあると言える。それによって、教育以外のデータとの接続も可能となり、様々な問題解決のための定量的分析を行うことができるようになってきた。ただし、データセキュリティの問題に万全を期するべく、個人 ID へのアクセスなどは限られた人数の職員にのみ認められるため、二次利用による分析の広がりを実現するまでには、まだ課題が残されている。

2-2-2 データのアクセシビリティ(教育庁)

教育に関連した統計データは教育文化省が直接収集しているが、国内外の様々なステークホルダーに対して教育データベースのサービスを提供するために、教育庁（Finnish National Agency for Education）が教育データポータルサイト Vipunen を整備している。Vipunen の成り立ちについて、教育庁の Ms. Riikka Koivusalo (Special Advisor) に対してヒアリングを行った。その内容は次のとおりである。

Vipunen の概要について

Vipunen は 2013 年に EDUFI（現在の教育文化省）によって立ち上げられた API（Application Program Interface）である。使用言語はフィンランド語、スウェーデン語、および英語からなるが、一部の情報はフィンランド語のみによって提供されているものも

¹² これはフィンランド統計局内のルールというよりは、EU からの要請でもあるようである。

ある（今後英語化する予定）。教育庁はこれらのデータベースの利便性を高めるために、教育文化省および統計局との連携のもと、Vipunen の管理運営を行っている。なお、図 4 は Vipunen のトップページ（英語版）である。



図 4 : Vipunen ホームページ (<https://vipunen.fi/en-gb/>)

Vipunen には国民の全ての教育段階のデータが集約されている。例えば、特別教育の必要性、高等教育の受験者および合格者、卒業生の保有学位、卒業後の進路などである。また、学校教育に関与している教員情報も含まれる。これらの教育関連データ以外に人口・職業分布などの情報が含まれており、教育関連データと関連づけて分析することができる。本ペー

ジのユーザーは、自らの興味関心に基づいたデータの集計方法をウェブサイト上で入力することによって、集計結果表を作ることができるようになっている¹³。

教育に関するデータの提供元は、主に教育文化省の KOSKI（児童生徒に関する業務データ）¹⁴および Studyinfo（申請者データベース）¹⁵からなり、どちらも統計局が管理している。

Vipunen は、様々なデータの利用可能性の向上を図ると同時に、テーマに焦点を絞った分析結果を広く発信してゆくこともその重要な役割の一つである。その役割を果たす活動の一環として、Vipunen にはスタンダードレポートとアナリシスレポートという2種類のレポートが掲載され、公開されている。アナリシスレポートは、スタンダードレポートよりも詳細な分析がなされているものと位置付けられている。こういったレポートは、教育文化省における政策形成のための基礎資料として使われることもある。例えば、幼児教育と初等教育に関するレポートや、2018年2月の人口構成と教育レベルに関するレポートは、議会や研究者等のリクエストに応える形で出版されたものであった。

データのアクセシビリティについてのまとめ

教育関連のデータポータルサイトである Vipunen の取組は、より多くの利害関係者がデータにアクセスすることを可能とし、データに基づく議論を行うことを可能とするインフラとしても参考になるものであった。日本の政府統計に関しては、統計ポータルサイト (<https://www.e-stat.go.jp>) がその役割を担っているが、より詳細かつ分野横断的なデータを提供するという点では参考になる点があった。

2-3 大学や研究者との連携(VATT)

フィンランド政府経済研究所(VATT Institute for Economic Research: VATT)の Tuomas Pekkarinen 博士と Mika Kortelainen 博士に対して、教育データの研究利用と、分析結果の

¹³ 日本における「政府統計の総合窓口（e-stat）」の教育分野に特化したものに相当すると考えることができる。

¹⁴ <https://confluence.csc.fi/display/OPHPALV/Koski>

¹⁵ <https://studyinfo.fi/wp2/en/>

政策へのフィードバックの現状について、学術研究者としての見解のヒアリングを行った。その内容は次のとおりである。

(1) VATT の成り立ちについて

VATT は、財務省の計画局（Department for Planning）によって 1970 年に設立された。公共政策や財政学を専門とする約 30 名の専任の経済学者が所属しているが、様々な大学のコンソーシアムとしての役割も果たしている。政策評価を行うことは重要なミッションのうちの一つであるが、政策形成に役立つ学術研究を行うことも重要なミッションとなっている。専任研究者の平均年齢は 40 代前半である。

(2) 行政データの利用可能性について

フィンランド統計局の管理しているデータを学術目的で利用するためには、利用申請の手続（申請料は 1 回当たり 150 ユーロ）を行う必要がある。利用申請においては、研究の目的とその社会的有用性について明記する必要がある。利用申請の対象には、個人レベルの行政データも含まれている。1960 年代からは国民年金番号に紐付けて管理されているデータもあり、それらも利用申請の対象となる。なお、個人情報保護の点から、個人を直接識別できる情報の開示は違法であるため、匿名化されたデータの利用のみが可能である。直接識別が可能な ID がなければ、たとえ付随情報から間接的に個人が特定できたとしてもそれは違法ではない。なお、初等教育に関する児童個票データはない¹⁶。

(3) 詳細な個票データの研究利用が可能になった経緯について

フィンランド統計局の管理している業務統計や様々なデータは、利用申請によって研究利用が可能となるが、各自治体の保有しているデータや、その他の省庁が個別に管理している業務統計に関しては、個人的なネットワークを使ってデータ利用の可能性を探る必要がある。

¹⁶ このことは、フィンランド統計局の情報とも整合する。

〔Kortelainen 博士〕 ヘルシンキの小学校の学力データを用いて分析を行った論文を執筆した際、そのデータの利用は個人的なネットワークからヘルシンキの地方自治体に交渉をした上で可能となった。

〔Pekkarinen 博士〕 フィンランドの徴兵時の適性テストである the Finnish Army Basic Skills test の全数個票データを用いて、学校制度の変更が生徒の学力や非認知能力の形成に対して与えた影響を評価 ("School Tracking and Development of Cognitive Skills") した際、その論文で用いたデータが利用可能になったのも、知人が関係者にいて、そこから話を積み上げた結果である。

大学や研究者との連携についてのまとめ

一般的に、北欧諸国では個票レベルの行政データの利用可能性が高い印象があるが、フィンランドではそのような印象が特に強かった。フィンランドにおいて個票レベルの行政データを用いた分析が盛んな理由の一つとしては、政府の保有するデータの機微性に関する国民の考え方を反映している部分もあるようである。特に、納税データは個人情報ではないという認識が、北欧諸国では共有されている印象があった。また、統計局では様々なデータを個人 ID によって紐付けし、分析することを可能とするデータベースの管理を進めているが、その研究目的での利用のためには、厳格な利用申請の手続を経る必要があるという点もイギリスをはじめとする諸外国と共通である。さらに、利用申請の対象から漏れているデータに関しては、個人的なネットワーク等を通じて利用可能性を開いてゆく必要があるという点も共通していた。

2 - 4 統計的会計検査の例(フィンランド会計検査院)

フィンランド会計検査院 (National Audit Office of Finland : NAOF) の Deputy Auditor General である Marko Mannikko 氏は、2018 年 3 月 26 日、パリで開催された OECD の Auditors Alliance の会合において、“Innovative approaches and case studies in performance audit” (有効性検査における革新的アプローチと事例研究) と題する講演を行っている。その中で、フィンランドは、財政政策検査 (fiscal policy audit) を推進しており、①財政政策

の意思決定の効率化、②情報の信頼性、③財政政策機関の機能と目的の達成（有効性検証）を目的とした検査を行っていることが述べられている。財政政策検査は 2012 年に始められた独自の検査であり、よりマクロ的な視点からの検査を行うのが特徴である。検査の手法としてはすでに報告・発表されている結果の再計算や既存制度の評価、財務諸表の分析、および統計的分析であり、財政政策の信頼性や有効性、透明性をマクロ的視点から柔軟に検査する方法となっている。これらの検査に基づく所見は主に行政的手続や意思決定のための情報基盤の改善に関するものであり、EBPM の基盤となるものである¹⁷。

上述の Mannikko 氏の報告では、財政政策検査について三つの事例を紹介している。一つ目の検査事例は "Innovative public procurement procedures" と題し、公共調達により良い手法を見つけるための検査を行ったものである。そこでの検査の視点は、従来の会計検査で最重要視されていたコンプライアンスよりも良い手法の探索に重点をおいた検査となっている。二つ目の検査事例は "Effectiveness of structural fund programs" と題し、EU による地域インフラ整備補助金の有効性について、補助金対象地域の変更という“自然実験”を用いて検証したものである。三つ目の検査事例は "Using unconventional data utilized in performance audit" と題し、移民世帯の児童生徒の学力形成について、国際比較可能な学力調査である PISA（The program for international student assessment）のデータを活用して分析し、公教育政策の有効性を論じた例である。今回の現地調査では、上記の二つ目と三つ目のレポートの責任者である調査官に対してインタビューを行い、どのような背景でこのような検査レポートを作成したのかを調べた。

さらに、統計的手法を活用した会計検査の例として、政府による景気予測に偏りがないかを統計的仮説検定により分析した "Statistical analysis in studying forecast reliability and the methodology of fiscal policy audit" の執筆者の一人である Senior Economist の Arto Kokkinen 氏に対してもインタビューを行い、同様の調査を行った。

以下では、統計的会計検査の例として、（１）政府景気予測指標の妥当性検証、（２）制度変更を自然実験として用いた EU の地域振興補助金政策の有効性検査、（３）外部調査を活用した教育政策の有効性検査を紹介する。

¹⁷ ここでの財政政策評価の説明は Ms. Pälvi Polojärvi (Senior Auditor) へのヒアリングに基づくものである。

(1) 政府景気予測指標の妥当性検証

“Statistical analysis in studying forecast reliability and the methodology of fiscal policy audit (2018)”の執筆者の一人である Senior Economist の Arto Kokkinen 氏に対してインタビューを行った。その内容は、次のとおりである。

(ア) インタビューに基づくレポートの概要

本レポートのトピックは、会計検査院内に設置された独立財政機関である Fiscal Policy Evaluation Unit により選定された。そのトピックは「フィンランド財務省が 1976 年から 2016 年の間に発表した GDP 成長率、失業率およびインフレ率の予測値は信頼できるか」というものであった。このことを調べるために、フィンランド財務省が発表した GDP 成長率、失業率およびインフレ率の予測値を調べ、①予測誤差に何らかの固定的な傾向がないか、②その他の 4 機関（フィンランド経済研究所：ETLA、経済協力開発機構：OECD、労働経済研究所：PT、ペレルボ経済研究所：PTT¹⁸）が発表した予測値および実質値と比較し、フィンランド財務省の予測値がそれらのものと乖離していないか、を検査するというものであった。

これらの検査を実施するために、以下の統計的手法が用いられた。まず、予測誤差に何らかの固定的な傾向がないかを調べるために、フィンランド財務省の予測誤差の時系列データを調べた。経済指標の予測には誤差が必ず発生するが、フィンランド財務省の予測誤差に偏りがなく、かつ利用可能な情報を十分に活用した上で一定の精度を持った予測がなされているのであれば、予測誤差は平均的にゼロとなるはずである。しかしながら、仮にフィンランド財務省が GDP 成長率を過大に予測する傾向を持っているとすると、予測誤差は平均的にプラスの値となる。このことを統計的に検証するために、予測誤差の時系列平均を計算し、それがゼロと統計的に有意に異なるかを検定した。

¹⁸ Pellervo Economic Research はフィンランドにある独立系の応用経済研究所である。

次に、その他の4機関が発表した予測値および実質値と比較し、フィンランド財務省の予測値がそれらのものと乖離していないかを調べるためにも同様の手法を用いて統計的検定を行った。本検査の目的、概要、検定方法は表7にまとめてある。

目的	概要	検定方法
予測誤差の固定的な傾向の有無をチェック	フィンランド財務省の予測誤差の時系列データを調べ、予測誤差は平均的にゼロとなっていたかを検定	予測誤差の平均値がゼロであるという帰無仮説 ¹⁹ をt検定 ²⁰
その他の五つの機関が発表した予測値および実質値との比較	フィンランド財務省の予測値とその他機関の予測値との乖離を検定	予測値の乖離の平均値がゼロであるという帰無仮説をt検定

表7：政府景気予測指標の妥当性の検証方法

さらに、経済指標予測に関する時系列モデルを推定し、予測誤差の時系列相関についても検定している。もし予測誤差に固定的な傾向（偏り）がないのであれば、予測誤差の時系列相関はゼロに近いものになるはずである。逆にフィンランド財務省の予測値に固定的な偏りがあるのであれば、予測誤差の時系列相関はゼロとは異なるものとなる。さらに、これらの結果を他の機関による予測誤差と比較することで、財務省の予測のみが偏っていたのか否かを検証している。

¹⁹ 統計的に否定することによって無に帰する仮説のこと。ここでは、予測誤差の平均値がゼロであるかどうかを統計的な基準をもって客観的に評価した結果、ゼロでなかった場合に帰無仮説が否定される。すなわち、帰無仮説が否定されれば、予測誤差の平均値がゼロとは言えず、予測誤差に固定的な乖離があることになる。

²⁰ 帰無仮説が正しいかどうかを判断するために、母集団が正規分布に従うと仮定した上で、標本の平均値が特定の値に等しいかどうかを調べるための手法。

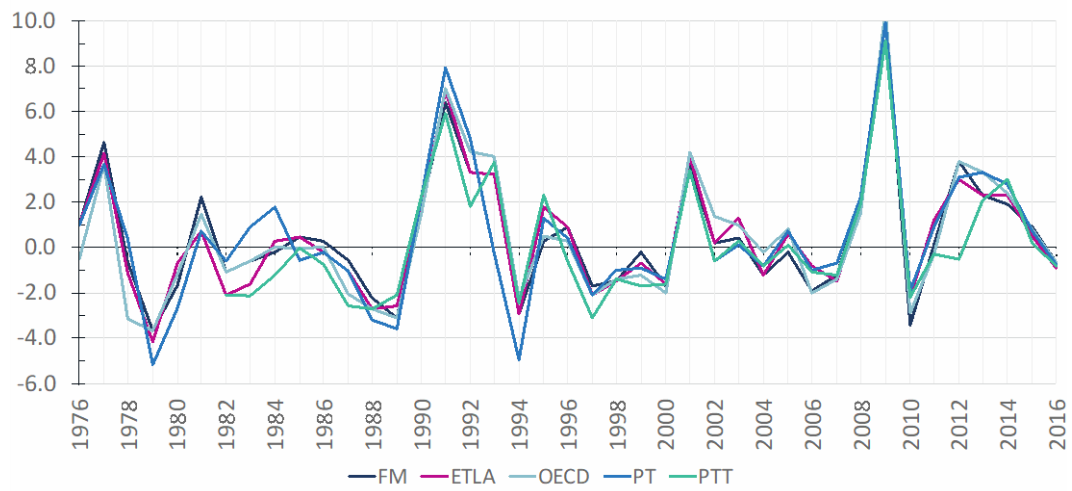


Figure 1: Forecast errors in GDP growth ($t+1$)

図 5 : GDP 成長率の予測誤差 (NAOF (2018)より転載)

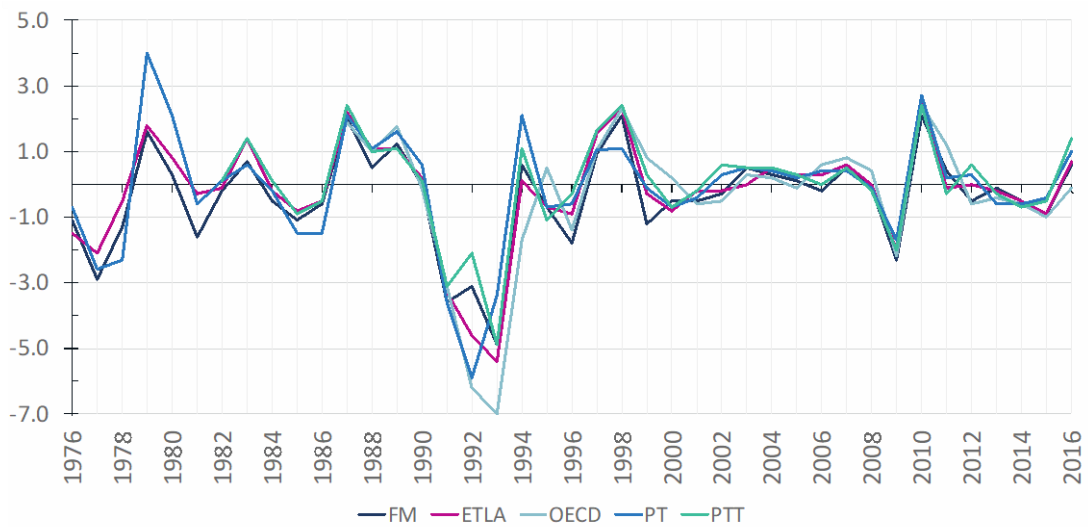


Figure 2: Forecast errors in unemployment rate ($t+1$)

図 6 : 失業率の予測誤差 (NAOF (2018)より転載)

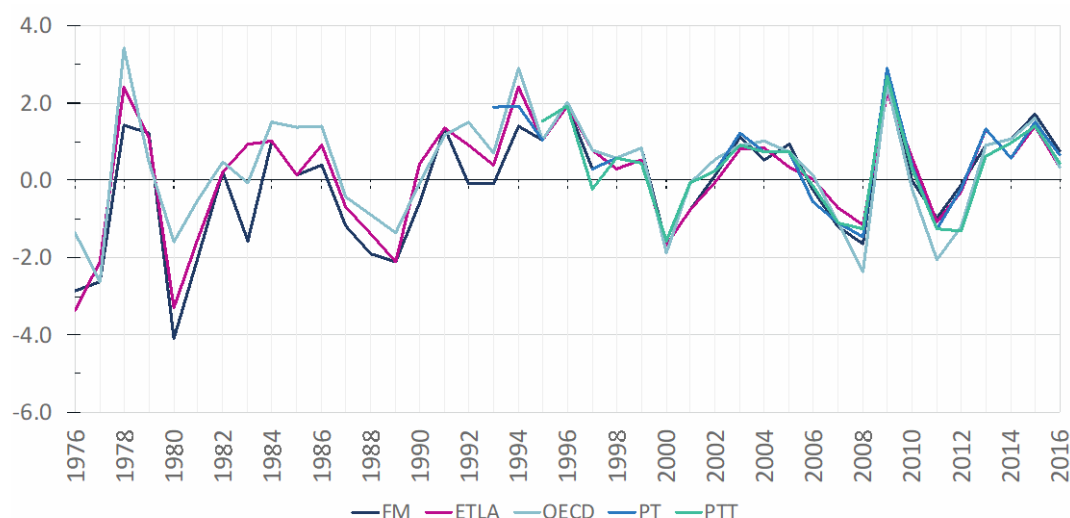


Figure 3: Forecast errors in consumer inflation ($t+1$)

図7：インフレ率の予測誤差（NAOF (2018)より転載）

図5から図7はそれぞれフィンランド財務省のGDP成長率、失業率、インフレ率の予測値とその他の四つの機関の予測値の時系列データを描いたものである。これらのデータに対して統計的検定を実施した結果、財務省の予測誤差の時系列平均は統計的にゼロと有意に異なっておらず、また他の機関の予測誤差とも有意に乖離していたとは言えないこと、さらに予測の精度は他の機関の予測に比べても劣るわけではなく、フィンランド財務省の予測は信頼できるとの結論に達した。

(イ) 検査の背景について

本レポートは前述のように Fiscal policy evaluation unit によってトピックの選定が行われたが、このトピックが選択された大きな理由は EU に対する説明責任という側面も強い。EU のルールでは債務残高が GDP 比の 60%を超えてはならないため、GDP 値が正確に計測されているかを EU の他国に対して説明する責任を負っている²¹。また、財政政策を正し

²¹ 自国民に対する説明責任を超えて、EU 内の他国に対して予測の信頼性を説明するという目的のためには、できるだけ客観的かつ定量的な検証方法である統計的手法を活用することのメリットが大きいという印象であった。

く実施するためには、その予測値に偏りがなく、財務省の利害とも独立である必要があるため、その予測の信頼性を検証することは国内的にも重要な論点となっている。

EU 内のその他の国においても、財務当局の発表している予測値データの信頼性に関する検証は検討されており、欧州委員会（European Commission）やイギリスの予算責任局（Office for Budget Responsibility）でも同様の分析が行われている。

上述のように、本レポートの結果は、財務省の予測には偏りがなく、信頼できるものであるというものであった。もし偏りや信頼性について疑義の残る結果であったとしてもレポートを公表していたか、と問われるのであれば、我々はどのような結果でも（関係各所との調整の上）発表したであろう。

本検査のように、統計的手法を用いて定量的な検査を行うためには、それらの手法に明るい人材が内部にいる、または外部の協力を得る必要がある。会計検査院には、3 名のエコノミストに加えてもう 3 名のスタッフがおり、そのうち 1 名が博士号 (Ph.D.) を持っている。また、外部の学識経験者 10-15 名の協力を常時得ていて、委員会（Scientific Council）を形成している。それらの教授の下で博士課程学生の養成も行われており、将来会計検査に役立つ統計的手法を学んでいる。現在、同委員会には、4 名の社会科学研究者と 7 名の法学研究者がいる。

(ウ) 政府景気予測指標の妥当性検証についてのまとめ

フィンランド会計検査院の検査のための重要な資源は、他の機関によって集められたデータおよびその使用权である。特にフィンランド会計検査院は詳細なデータの利用について強い権限を持っており、関係省庁に対して機微情報を要求することができる。他の省庁によって計算され公表された指標の妥当性を確認するためにフィンランド会計検査院が“再計算”することも多く、その際には元データの提供を要求することができる。このように、フィンランド会計検査院は情報利用に関して比較的強い権限を他省庁に対して持っているように思われるが、Kokkinen 氏によると、その源泉は政府および会計検査院に対する国民からの信頼が背景にあるとのことであった。また、EU という自国民を超えたステークホルダーに対する説明責任があるという点も、統計分析という客観的な手法を用いた会計検査を行う動機の一つとして強く作用しているとの印象を持った。

(2) 制度変更を自然実験として用いた EU の地域振興補助金政策の有効性検査

Principal Performance Auditor である Ville Vehkasalo 氏に対して検査レポート“Effectiveness of structural fund programmes (2016)”の内容についてのインタビューを行った。その内容は次のとおりである。

(ア) インタビューに基づくレポートの概要

本レポートにおける検査の着眼点は、EU が拠出している地域インフラ整備補助金政策 (European Regional Development Fund programmes) は地域経済振興に対して有効かというものである。この補助金政策の効果検証を行うため、フィンランド内で補助金を受けている地域と受けていない地域における雇用、賃金、失業率といった様々な経済指標の比較を行った。しかしながら、補助金を受けている地域は、本来経済状況の芳しくない地域が選定されているため、単純な地域間比較では補助金を受けている地域の経済指標の方が補助金を受けていない地域に比べて悪いという結果となりがちである。

このような問題を回避するために、本レポートでは、補助金を受けた地域 (処置群) と、補助金を受けていない地域 (対照群) の比較を注意深く行うためのリサーチデザインを明示的に議論している。具体的には、2007 年以前は補助金を受けていなかったが EU の政策変更によって 2007 年以降に補助金の受給対象となった地域と、2007 年以前も以後も補助金を受けていない隣接地域とを比較することで、補助金効果の識別を試みている。このような効果の識別方法は、自然実験を用いた「差の差の検定法」と呼ばれるものである。

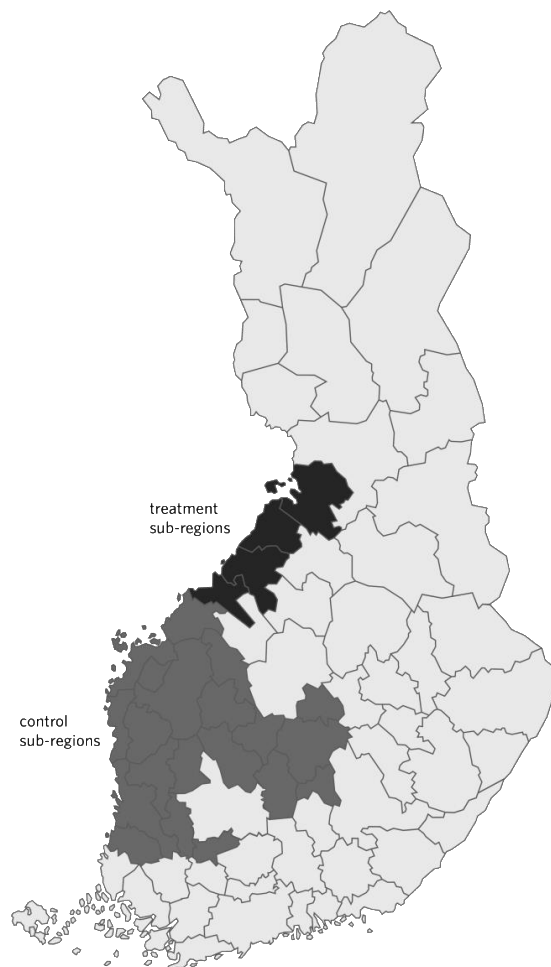


図 8：補助金受給政策の影響を受けた地域（処置群）と比較対照群
(Vehkasalo (2018)より転載)

図 8 はフィンランドの地図であるが、地図中の黒い地域は 2006 年までは補助金の対象ではなかったが、2007 年以降に補助金の対象となった地域（処置群）である。一方、灰色の地域は 2007 年以降も補助金の対象とはなっていないが、2006 年までは黒い地域と似た経済的状況にあった地域（対照群）である。

	2000-2006 (事前)	2007-2013 (事後)
黒 (処置群)	A	B
灰色 (対照群)	C	D

(注) 図の網掛け部分は、補助金の対象となった地域および期間を表す。

表 8：差の差の検定法

差の差の検定法では、黒い地域と灰色の地域の 2007 年前後の変化を比較する。表 8 にあるように黒い地域の 2006 年までの平均的な所得を A、2007 年以降の所得を B とする。同様に灰色の地域の 2006 年までの平均的な所得を C、2007 年以降の所得を D とする。もし地域振興補助金政策が、その地域の所得の引き上げに有効であるのであれば、

$$(B-A)-(D-C)>0$$

となる。このように、政策変更の前後の差を地域間で比較することによって、政策の影響として経済指標の改善があったかを調べることができる。

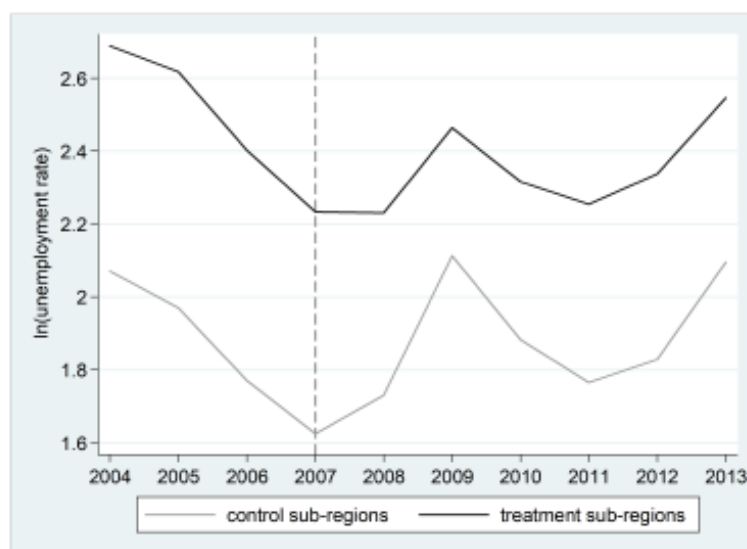


図 9：補助金受給政策の影響を受けた地域（処置群）と比較対照群の失業率の推移
(Vehkasalo (2018)より転載)

上記の方法によって比較した結果、EU による補助金は受給地域の経済指標、特に失業率の低下に効果を持つという結論に至っている。図 9 は、黒い地域の失業率の推移（黒線）と灰色の地域の失業率の推移（灰色線）を描いたものである。2007 年から 2009 年まで、灰色の地域（対照群）の失業率は急速に上昇しているが、黒い地域（処置群）の失業率はそこまでは上昇していない。これは地域補助金政策の効果として失業率の上昇が抑えられたためと解釈することができる。

(イ) 検査の背景について

北欧諸国の会計検査院の調査官で構成されるネットワークがある。その 2015 年の会合で、EU の地域振興政策の効果検証が話題となった。かつて EU の政策変更によって地域補助金の対象地域が変更されたことがあり、その結果、自然実験的状况が生まれていたことから、フィンランドで検証をするということが決まった。なお、本レポートは 2018 年に学術誌 (Region) から出版されている。

自然実験を用いた差の差の検定法を用いた同様の方法で政策の有効性を検査したレポートは、少なくとも EU 域内ではまだないのではないかと。(Vehkasalo 氏は) Licentiate という博士号と修士号の間となる最終学歴 (学位) を持っているが、同様の検査を行うことができる人材は検査院内で 4 名程度はいる²²。

(ウ) 制度変更を自然実験として用いた EU の地域振興補助金政策の有効性検査について のまとめ

本レポートは統計データに厳密な識別戦略を用いて公共政策の有効性を検証した貴重なレポートである。今後このような質の高い有効性検証をさらに進めてゆくことは全世界的に推進されるべきであろう。現時点では EU の政策形成に対して本レポートが与えた影響はあまりないということであったが、会計検査に限らず、このようなエビデンスに基づいて政策を形成することこそが最重要であるといえよう。

²² 会計検査院のホームページによると、職員の総数は約 150 名とのことである。
<https://www.vtv.fi/en/naof/questions-and-answers/>

(3) 外部調査を活用した教育政策の有効性検査

Principal Performance Auditor である Tanja Kirjavainen 氏に対してインタビューを行った。主に彼女の担当した検査レポート "Immigrant students and the effectiveness of basic education (2015)" について、また、教育政策の有効性検査の実施状況一般についても併せてヒアリングを実施した。その内容は次のとおりである。

(ア) ヒアリングに基づくレポートの概要

フィンランドでは近年外国人比率が高まってきている。人口に占める外国人比率の高まりとともに、義務教育段階の学校においても、移民の子供たちの比率が高まってきている。図 10 は 2005 年から 2014 年までのフィンランドの義務教育学校における外国籍の児童生徒数およびその比率の推移を示したものである。この図によると、2009 年以降、外国籍の児童生徒数およびその比率は急速に伸びていることが見て取れる。

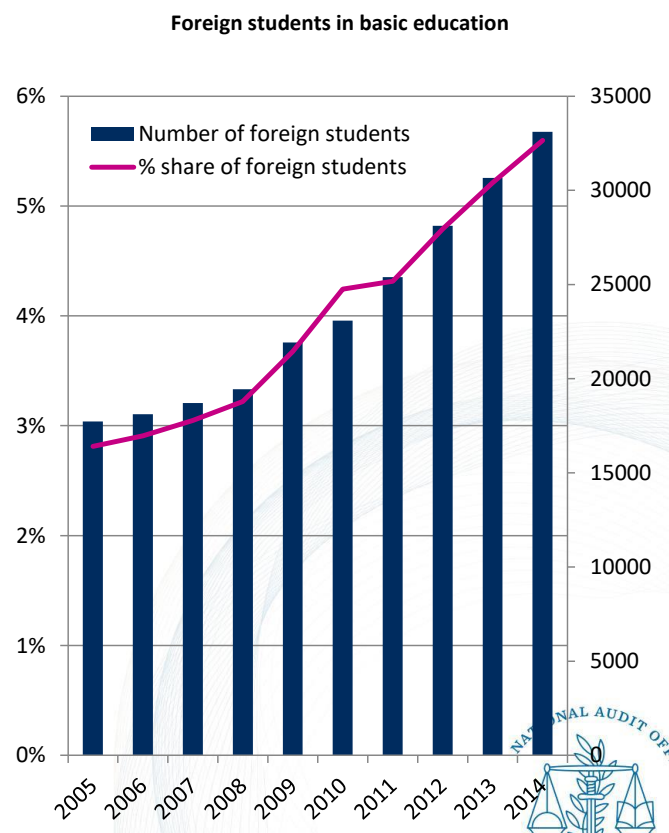


図 10：フィンランドの外国籍生徒比率（NAOF (2015)より転載）

本レポートにおける検査の着眼点は、学業成果や進学準備といった視点から、義務教育段階の学校教育は移民の子供たちに対してネイティブの子供達と同等の教育サービスを提供することができるのか、というものであった。特に、親とともに移住してきた移民の第1世代の子供達のみではなく、親のフィンランドへの移住後にフィンランド国内で生まれた移民の第2世代の子供の学業成績にも着目した。

フィンランドのネイティブと移民の子の学力を測る指標として、フィンランドで2012年に実施された国際学力調査（Programme for International Student Assessment: PISA）のデータ（311 学校、8,828 人が対象）を用い、学業成果その他の指標を移民かネイティブかで比較した。数学、読解力、科学それぞれの教科における PISA テストの平均点を、フィンランドのネイティブ、第1世代の移民の子、第2世代の移民の子ごとに示したものが図11である。

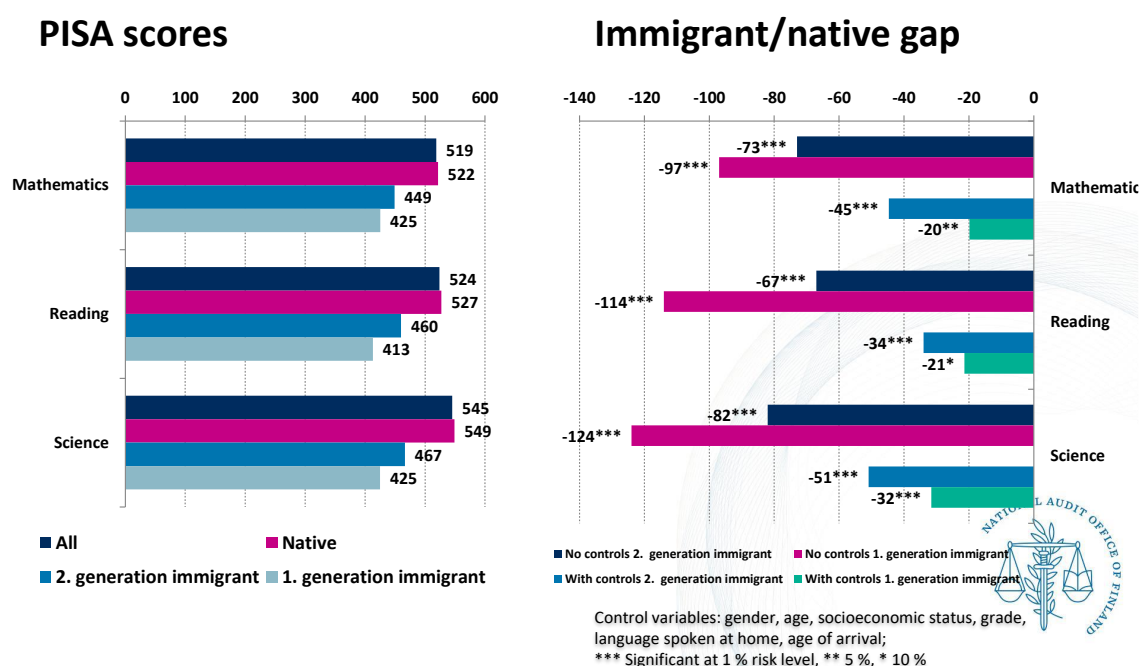


図 11：PISA テストの平均点および移民/ネイティブギャップ（NAOF (2015) より転載）

図 11 の左図は、数学、読解力、科学の平均点を全生徒、ネイティブ、移民の第2世代、および移民の第1世代ごとに示したものである。この図からは、すべての教科においてネイ

ティブの得点が一番高く、次いで移民第2世代、最後に移民第1世代の得点となっていることがわかる。この差の要因をさらに検査した結果、移民の生徒とネイティブの生徒の学力差が大きい理由の一つはネイティブの生徒の得点が世界的にみても高得点であること、さらに移民の（特に第1世代の）生徒は学校教育の経験年数が少ないことが明らかとなった。また、この学力差の原因は移民の子とネイティブの子との間の学校教育に対する積極性の違いによるものではないことも明らかにされている。

さらに驚く発見としては、移民の生徒とネイティブの生徒の学力差は第2世代の方が第1世代に比べて改善しているように見えるが、家計所得や家庭での使用言語など、家庭環境の影響を考慮すると、第2世代は第1世代よりも平均点が下がっていることがわかった。移民の第2世代の平均点が第1世代よりも高くなっているのは家庭環境が改善しているからであって、公教育の有効性という観点からは、第2世代の移民の子弟がフィンランド社会において十分な基礎的教育を受けることができていない可能性を示唆する結果となっている。図11の右側のグラフには、移民の生徒とネイティブの生徒のテストスコアの差を示しているが、性別や年齢、社会経済的な背景や家庭での言語などの影響を取り除くと、すべての教科において第2世代の移民の生徒の方が第1世代の移民の生徒よりもネイティブに比べてテストスコアが低くなっていることがわかる²³。これらの結果に基づき、フィンランド会計検査院は、①教育文化省は言語トレーニングをもっと充実させるべき、②パフォーマンスの良くない移民の子が学力を伸ばすために追加的な施策の必要性の検証が必要との所見を記している。

²³ 図11の右側のグラフにおける各教科（数学、読解、科学）の4本の棒グラフのうち、下二つは家庭環境や性別といった属性の影響を取り除いたうえでのネイティブとのテストスコアの差を表している。この2本の棒グラフのうち、上は移民の第2世代の生徒、下は移民の第1世代の生徒であり、ネイティブとの比較において、第2世代の生徒の方が第1世代の生徒に比べて点数差が大きいことがわかる。

(イ) 教育政策の有効性検査の実施状況について

1) 教育関連の公共政策への検査実施

全体の検査における教育関連の検査（教育文化省が対象となっているもの）の件数は年間に 2-3 件である。検査対象およびトピックは会計検査院がリスク分析という視点から選んでいる。例えば、パフォーマンスに関する検査、移民の子供たちに関する検査、特別支援教育に関する検査などがそのような例である。現在進行中の検査としては、教育の電子化、高校レベルにおける職業訓練などがある。

2) 評価体制および人材育成

教育文化省を担当している調査官のうち、4 名が博士号を取得している。なお、教育部門の有効性検査は教育文化省の出資している Education Evaluation Center という独立機関で盛んに行われており、20-30 人のフルタイムの研究者が初等教育から高等教育までの様々な教育政策の有効性を評価している。また、大学の授業を用いた人材育成プログラムも行われている。

3) データおよび評価における統計的方法の活用

教育政策の有効性分析に用いられるデータは、PISA などの第三者機関によって収集されたものや、フィンランド統計局によって収集されたものなどであり、独自にデータを収集することはない。

有効性検査においては、前述の（２）の EU の地域振興補助金政策の有効性検査のような自然実験的な状況を用いた効果の検証を行うことはあるが、有効性検査のために RCT を行ったことはない。その他の統計的手法の活用例として、移民の子供（いわゆる移民の第 2 世代）は第 1 世代よりも教育成果が悪いという上述の分析結果を受けて、それは同化の失敗なのかを調べるために、言語トレーニングの効果を付加価値モデル、変量効果モデル、マルチレベルモデルを用いて分析しているものがある。

4) EBPM との関連

教育分野において、EBPM (Evidence-based policy making)の推進という点から、KPI の設定、事後評価の実施、評価結果の共有のためのシステム構築の重要性は認識されている。フィンランドにおいて、EBPM の重要性は認識されているが、まだ制度として確立されているというわけではない。しかしながら、フィンランド会計検査院の有効性検査の結果を政策に反映させることは、政治家や役人にとって義務として認識されるべきものである。

(ウ) 外部調査を活用した教育政策の有効性検査についてのまとめ

本レポートは外部データを用いた検査を行っており、その際に、詳細な個票データを活用している点では、前述の（１）および（２）のレポートと同様に貴重なレポートである。イギリスの会計検査院が独自の調査を行ったのとは対比的に、既存のデータを活用した統計的会計検査も有効性の検査と同時に公共政策の有効性に関するエビデンスの蓄積という点からも大いに参考となる事例である。

3 おわりに

本報告書では、EBPM を推進する上で必要なインフラとしてのデータベースの構築およびそれを分析し、エビデンスを蓄積することのできる人材および制度、さらにはそれらのエビデンスを将来の政策立案に反映させるための仕組みはどのようなものかという視点から海外の実態調査を行った。今回の調査では、EBPM や会計検査における統計的手法の活用が比較的進んでいる事例としてイギリスとフィンランドに着目し、データベース整備の現状、政策評価の実施状況、政策立案への反映状況といった視点から各国における関係省庁等にヒアリングを行った結果を紹介した。さらに、各国の会計検査院に対するヒアリングの結果から、EBPM 推進における会計検査院の関与状況を把握することで、今後日本において根拠に基づく政策形成を推進する中で会計検査院の果たしうる役割についても知見を得た。

イギリスでは、教育分野の行政データベースとして NPD を紹介し、NPD が教育政策の形成および研究においてどのように活用されているのかについて、教育省、国家統計局、および大学研究者へのヒアリングから調べた。まず、教育省に対する NPD に関するヒアリングからは、NPD が教育予算配分の基礎資料としての行政データから出発していることが分かったが、それを実際に分析し、活用するためには国家統計局を窓口とするデータの二次利用制度の整備や、大学関連の研究者との連携が大きな役割を果たしていることが分かった。

EBPM の最先端事例として紹介されることの多い What Works Centre の Education Endowment Foundation についても実情をヒアリングし、政府の政策形成とエビデンス形成の関係性についての知見を得ることができた。What Works Centre は EBPM のロールモデルのように日本においても紹介されることがあるが、What Works Centre の活動はより細かなエビデンスの収集および蓄積が主な仕事であり、国家レベルの政策形成や予算配分に対して与える影響が現状では決して大きなものではないということがヒアリングを通じて分かった。ただし、予算編成の効果および公共政策の有効性のマクロレベルでの検証と合わせて、学校教育現場におけるエビデンスの蓄積を行うことで重層的な評価体制を構築することができおり、今後はミクロレベルとマクロレベルの双方のエビデンスが政策形成に活用されることが期待される。

イギリス会計検査院におけるヒアリングでは、会計検査の特に有効性検査における統計的手法の応用として、会計検査院が独自に実施した社会調査を活用した検査事例を紹介した。そこでは、各公共政策の受益者に対して直接オンラインアンケートを実施し、統計的

析を施した上で検査結果に反映させている点は、特筆に価するものであった。このような有効性検査の結果は、政策評価の結果でもある。このような検査結果を蓄積することにより、今後同様の教育政策を行う際の潜在的な有効性についての情報を提供することもできるはずであり、会計検査における統計的有効性検査はEBPMの基礎である根拠の積み増しに大いに貢献し得ると思われ、EBPMを推進してゆく上で会計検査院が重要な役割を持つ可能性を秘めていることが分かった。イギリス会計検査院でこのような統計的有効性検査が可能となった背景の一つには、やはり統計的手法や調査方法に明るい専門的人材の存在が大きいということも明らかになった。

フィンランドでは、教育データの収集および管理、アクセシビリティという点から、教育庁および統計局でヒアリングを行い、データベース構築の現状と課題についての知見を得た。フィンランドをはじめ、北欧諸国は個人レベルの行政データを活用・分析するという点に関しては先進的な国が多い。また、これらの国はEU内にあるという理由により、自国民に対してのみならず、EU内の諸外国に対しても様々な場面で説明責任を果たす必要に迫られているという事情もある。こういった事情を背景にもつフィンランドでは、EBPMとの親和性は高く、データベースの整備・活用、研究者との連携のもとでの分析およびエビデンスの蓄積、政策形成におけるエビデンス尊重の文化が醸成されているように思われる。

詳細な個票データを用いた研究が数多く先行しているフィンランドの経済学者へのヒアリングからは、個人情報保護に関する考え方の違いや、研究と行政との良好な関係を保ちつつ、より良い学術研究と政策形成を同時に生み出してゆくシステムの構築の必要性を強く感じた。国家統計局は個人情報保護や情報セキュリティという点から、安全なシステムの構築を進めていると同時に、研究者はそのルールを遵守しつつ、分析を遂行していることは大いに参考となった。

最後にフィンランド会計検査院では、高度な統計的手法を用いた革新的な会計検査事例の紹介を受け、我が国における将来の会計検査の可能性について考察することができた。特に、公共政策の有効性検査のみならず、フィンランド財務省の発表している統計指標の正確性を統計的に検査するという事例は、統計の信頼性についての会計検査を定量的に行うことの可能性を示してくれており、今後の会計検査における新しい領域を切り開くものであるようにも思われた。

イギリスとフィンランドの調査を通じて受けた印象としては、EBPM的な考え方や取り組みは今日に始まったわけではなく、すでに様々な行政活動の中にビルトインされている

というものであった。しかしながら、このような考え方を実装することを可能とするためには、そのためのインフラとしてのデータの収集および管理、また、エビデンスの生成・蓄積を行うことのできる人材の存在や研究者との連携が意識的に行われる必要があり、両国においてはそれがある程度なされているということが分かった。さらに、両国における会計検査においては統計的有効性検証や正確性検証がなされており、EBPM におけるエビデンスの蓄積という点からも、会計検査院の果たす役割が存在していることが分かった。

イギリスやフィンランドと同様に、日本も少子高齢化、外国人労働力の受け入れ、財政状況の悪化といった様々な問題に直面している。そして、効果的かつ効率的な公共政策を実現してゆく必要性は、これらの国々と同様に、ますます高まってゆくことが予想される。

従来の経済性・効率性・有効性の会計検査では、経済的、効率的、有効であると判断するための基準を会計検査院が自ら設定することが難しいという課題があったかもしれない。一方、本報告書で紹介したフィンランド会計検査院における政府経済指標の予測値の信頼性についての検証は、統計的な手法が会計検査にも応用ができることを示しており、ひいては、会計検査院が公共政策の経済性・効率性・有効性の検査を実施するに当たり、受検庁から得たデータから統計的手法を用いて事業費や事業効果の推定値を計算し、それを客観的な評価基準とする可能性も提示している。科学的手法に基づいたより客観的な検査方法の活用は、専門性を活かした未来志向の会計検査の可能性を開いているように思われる。

統計的検査を実施するためには、統計的手法に精通した調査官の確保・育成や、大学や研究所といった外部機関との連携が必要であるが、その必要性はイギリスやフィンランド以上に財政状況の悪化が進む日本においては特に高い。そのためにも、統計的手法の技術を身につけた調査官の育成は喫緊の課題であると思われる。

統計データの信頼性や正確性検査、使いやすいデータベースを構築する上での効率性検査、未来の政策立案に必要な根拠の生成過程としての有効性検査など、今後日本において根拠に基づく政策形成を推進する中で会計検査院の果たしうる役割は決して小さくないということを、今回の海外実態調査を通じて実感することができた。

【参考文献】

- Harjunen, Oskari, Mika Kortelainen and Tuukka Saarimaa (2018) “Best Education Money Can Buy? Capitalization of School Quality in Finland,” *CESifo Economic Studies*, Volume 64, Issue 2, 1 June 2018, Pages 150–175.
- Manikko, Marko (2018) “Innovative approaches and case studies in performance audit” presented at Launch Meeting of Auditors Alliance, Paris, OECD, March 26th, 2018.
- NAO (2017) “Retaining and developing the teaching workforce”
- NAO (2018) “Ofsted’s inspection of schools”
- NAOF (2015) “Immigrant students and the effectiveness of basic education”
- NAOF (2016) “Effectiveness of Structural Fund Programmes”
- NAOF (2018) “Statistical analysis in studying forecast reliability and the methodology of fiscal policy audit”
- National Office of Statistics (2013) “Public Service Productivity Estimates: Education 2013”
- Sari Pekkala Kerr & Tuomas Pekkarinen & Roope Uusitalo, 2013. “School Tracking and Development of Cognitive Skills,” *Journal of Labor Economics*, University of Chicago Press, vol. 31(3), pages 577-602.
- Vehkasalo, Ville (2018) “Effectiveness of EU Regional Policy: Evidence from a Natural Experiment in Finland,” *Region*, Volume 5, Number 3, 2018, 1-19.

【付録：調査日程】

2018 年

月日	曜日	日 程
11/5	月	AM ロンドン大学教育学研究所 (UCL Institute of Education) PM イギリス会計検査院 (National Audit Office)
11/6	火	AM サットン・トラスト教育基金 (Sutton Trust/Education Endowment Foundation) PM 教育省 (Department for Education)
11/7	水	AM 国家統計局 (Office of National Statistics) ロンドン事務所
11/8	木	AM フィンランド会計検査院 (National Audit Office of Finland) PM 統計局 (Statistics Finland)
11/9	金	AM 教育庁 (Finnish National Agency for Education) フィンランド政府経済研究所 (VATT Institute for Economic Research)

