OECD publishing

USING DIGITAL TECHNOLOGIES TO IMPROVE THE DESIGN AND ENFORCEMENT OF PUBLIC POLICIES

OECD DIGITAL ECONOMY PAPERS

February 2019 **No. 274**

going digital

OECD
BETTER POLICIES FOR BETTER LIVES

Foreword

This report analyses how digital technologies can improve the design and enforcement of public policies in domains including competition, education, environment, innovation and taxation. This report reflects feedback from a number of OECD bodies, including the Committee for Digital Economy Policy, the Committee on Industry, Innovation and Entrepreneurship, the Working Party on Integrating Environmental and Economic Policies and the Governing Board of the Centre for Education Research and Innovation.

This paper was approved and declassified by written procedure by the Committee for Digital Economy Policy on 16 November 2018 and prepared for publication by the OECD Secretariat.

This report was authored by Nick Johnstone (OECD Directorate for Science, Technology and Innovation), Shardul Agrawala, Elisabetta Cornago, Tobias Udsholt (OECD Environment Directorate) and Stéphan Vincent-Lancrin (OECD Directorate for Education and Skills).

This publication is a contribution to the OECD Going Digital project, which aims to provide policymakers with the tools they need to help their economies and societies prosper in an increasingly digital and data-driven world. For more information, visit

www.oecd.org/going-digital. #GoingDigital

Note to Delegations:

This document is also available on O.N.E under the reference code:
DSTI/CDEP/GD(2018)7/FINAL

Executive Summary

To date, the profound impact of digital transformation in the private sector has not been mirrored by equally significant changes to how policy is designed, implemented and evaluated or how governments interact with their citizens. However, the application of digital technologies holds the potential to reshape existing policies, enable innovative policy design and rigorous impact evaluation, and expand citizen engagement in local and national policy making. The extent to which this potential is realised will depend on whether governments prove willing and able to scale the use of digital technology, and how successfully privacy concerns and digital security vulnerabilities are addressed. From the first pioneering experiences of policy applications, it is possible to identify four areas in which digital transformation promises to improve policy making.

The first area is the improved efficiency of enforcement and targeting of existing policies. The increased possibility to monitor outcomes directly, for example thanks to advanced and linked sensor technologies, and the availability of data that were previously imperfectly observable, or only observable at significant administrative cost, enables more effective enforcement of existing rules and lowers the cost of policy targeting. In the area of finance, financial flows can now be tracked at a level of granularity and periodicity that was not previously possible and allow for the better enforcement of existing financial market regulations and improved public finance management. In agriculture, remote sensing and digital land parcel identification systems allow countries to grant direct subsidies to farmers and to enforce other regulatory measures related to the sustainability of agriculture. Yet the increased complexity of policy coupled with the need to address privacy concerns remain significant challenges to the more widespread use of policy targeting and differentiation, for example in social and education policy.

The second area in which digital transformation holds the potential to improve public policy is in improving policy design and evaluation. Digital technologies broaden the suite of policy instruments available to governments and can lower the cost of policy experimentation and evaluation. In cities, digital cameras that automatically register the license plate of vehicles entering a congestion zone have made it more feasible to design, implement and enforce congestion pricing schemes. Merging congestion data with public transport use data from smart cards can be used to evaluate the effect of urban policies on travel behaviour. In education, being able to track all students over their study path has allowed some countries to uncover patterns that were at odds with study design and led to the experimentation of new study paths. Governments have also started using online laboratory experiments as low cost approaches to test the impact of alternative labelling schemes, enabling more effective policy outcomes upon roll-out.

The third area of promise is the potential of digital transformation to reshape government-citizen interaction and expand stakeholder engagement. Many OECD countries are making more data freely available to enhance accountability in the public sector and allow for evaluating the effects of current policies. Making pollutant release and transfer registers publicly available online can facilitate civil society oversight on regulated entities, making compliance efforts more transparent and breaches more open to public scrutiny. As an example of improved engagement of regulators with the private sector, the creation of regulatory sandboxes in Fintech has allowed companies to test the introduction of new technologies in a controlled and monitored environment and facilitate regulator-firm engagement.

The fourth area of promise is the role of simple digital tools to achieve a better and more participatory policy design: virtual meetings or digital public consultations can help engage all stakeholders within and outside the government. In other areas (e.g. nature conservation) artificial intelligence has been used to develop low-cost image recognition technologies which can give a broad cross-section of stakeholders a valuable role to play in contributing to the attainment of important public policy objectives.

However, digital transformation poses a number of challenges: the increased granularity of data and increased data-sharing between government agencies and across public-private partnerships can generate digital security vulnerabilities and concerns over individual privacy. While the potential for digitalisation to improve public policy settings is considerable, this needs to strike the right balance between the broader public benefits of enhanced sharing of data, and individuals' and organisations' legitimate concerns about the protection of privacy.

Further, the insufficient public infrastructure to link disparate sources of data is a key bottleneck. This raises the question of interoperability of data systems within and across different areas. Finally, while the availability of more data usually helps to improve policies, it is not a panacea and comes with risks that will need to be tackled over the next decade. This has meant that public sector adoption rates remains low relative to the private sector, with the majority of policy examples identified only implemented at the local level or representing pilot projects. Mainstreaming digital best practices across national institutions might be the largest challenge to be tackled.

Table of contents

Tables

1. Introduction

Digitalisation is having a profound impact on social and economic activity, from the revolution in internet search and e-commerce, to the proliferation of social networks, the emergence of new production technologies like 3D printing, and growing deployment of artificial intelligence. Most of these changes have been driven by the private sector, although built on a very long history of public investment in R&D in the underlying technologies.

Through a range of e-government and open data initiatives, governments are also harnessing digital technologies to improve public access to services and information, and to reduce costs associated with public administration and the delivery of public services.[1] However, on the policy front, the impact of the digital transformation on economic and social activity has not been mirrored by equally significant changes in how public policy is designed and implemented.

Nevertheless, the combined adoption of new digital technologies, increased reliance upon new data sources, and use of advanced analytic methods hold significant potential to improve the effectiveness and enforcement of existing policies, enable innovative policy design and impact evaluation, and expand citizen and stakeholder engagement in policy making.

The extent to which this potential is realised will depend on whether governments prove willing to adopt and able to scale the use of such technologies, obtain reliable access to relevant data which is often in the hands of private actors, and how successfully concerns such as privacy and cybersecurity are addressed.

This paper examines some of the ways in which the digital transformation can benefit policy makers in developing and implementing regulations and other policy instruments. It is argued that by reducing administrative and transaction costs associated with all stages of the policy cycle, digitalisation can result in much more effective and efficient policy settings.

The remit of public policy domains that are examined in this paper is fairly broad and includes, among others, competition, education, environment, innovation, and taxation. Applications of digitalisation in the context of national security and law enforcement, while arguably more sophisticated and widespread than elsewhere in government, are however not examined in this paper given the very limited availability of such information. Note also, that this paper does not seek to address the more general issues associated with "digital government" which is concerned primarily with the use of digital technologies and methods to improve access and reduce costs associated with public administration.[2]

While many policy applications of digital technologies remain in the pilot-phase, it is possible to identify four areas in which digitalisation holds promise for better public policy:

- Improving monitoring, supervision and enforcement of existing policy measures. The increased availability of data on the outcomes arising from different policy interventions that were previously imperfectly observable, or only observable at significant administrative cost, enables improved monitoring and supervision and more effective enforcement of policies.

- Allowing for the implementation of new and more efficient policy instruments, and facilitating policy evaluation and experimentation. Digital technologies facilitate data collection at a more granular level, potentially allowing for the use of "first-best" policy measures which target the desired policy objective directly, rather through a proxy outcome. They can also allow for the implementation of more differentiated policy instruments (at the spatial, temporal, technological or socio-economic level) in which marginal benefits and costs are more closely aligned. They can also allow for information feedback, often in real time, which can facilitate policy evaluation and lower the cost of policy experimentation, helping government iterate towards efficient policy settings.
- Predicting emerging risks and opportunities and behavioural responses to policy interventions. Digital technologies can allow policy makers to be more pro-active and reactive in tracking and responding to fast-changing phenomena, whether they be risks or opportunities. At the same time, advanced analytics can help to "predict" responses from those affected by policy interventions in a more robust manner than was the case previously.
- Enhancing government-citizen interaction and expanding stakeholder engagement. Many OECD countries are making more data freely available to enhance accountability in the public sector. The public availability of pollutant release and transfer registers can facilitate civil society oversight on regulated entities, making compliance efforts more transparent and breaches more open to public scrutiny. Online consultation platforms can also ease the engagement of regulatees throughout the policy process, while policy sandboxes can help regulatees assess consequences of planned innovations. Citizen and stakeholder engagement can also facilitate outsourcing or crowdsourcing of certain data collection, which can facilitate policy formulation, as well as compliance and enforcement.

In some cases, the digital transformation of the private sector activities has itself necessitated a "digital response" from policy makers. For example, just as the digitalisation of certain sectors may pose challenges for the international tax rules, the use of digital technologies, big data and advanced analytics by tax authorities can help to facilitate and promote compliance. In other cases, even if the economic activity has not been radically transformed by digitalisation, the use of digital technologies by public authorities can allow for the implementation of more efficient policies, ensure their enforcement, and evaluate their outcomes. As such, the use of digital technologies by policy makers is potentially relevant for every policy sphere.

The remainder of this report is organised as follows: Section 2 looks at how digital technologies can improve the effectiveness of existing public policies and interventions. Section 3 analyses the potential for improving policy design, implementation and evaluation through use of digital technologies. Section 4 looks at the role of big data and advanced analytics in predicting emerging risks and responses to policy interventions in a targeted manner. Section 5 evaluates the potential for greater government-citizen interaction and expanding stakeholder engagement through new digital platforms. Section 6 provides a concluding assessment of the prospects and the challenges facing the use of digital technologies in public policy.

2. Improving monitoring, supervision and enforcement of existing policy measures

Digital technologies can improve monitoring of imperfectly observable outcomes in a number of policy domains. Where it is possible to attribute actions (or actors) to such outcomes, this may also lower the costs and improve the reliability of compliance and enforcement activities. This section discusses these two aspects in turn.

2.1. Improving the monitoring and supervision of imperfectly observable outcomes

Digital technologies can increase the capacity of governments to monitor policy-relevant outcomes that were previously imperfectly observable, or only observable at significant cost. In the absence of the capacity to precisely monitor the status of economic, environmental, or social outcomes which are the targets of public policy objectives, the government is "flying – at least partly – blind". The application of digital technologies can help to overcome this by facilitating data collection and provision of more comprehensive and timely statistics which can empower evidence-based policy making. The increased capacity to monitor imperfectly observable outcomes is particularly relevant for three broad domains: highly data-intensive economic activities; illicit activities and the shadow economy; and where non-market outcomes dominate.

The financial sector is probably the best example of a highly data-intensive economic activity where big data, and digitalisation more generally, can play an important role in helping monitor financial flows at a level of granularity and periodicity that was not previously possible. This is commonly referred to as RegTech,[3] the regulatory handmaiden to the development of FinTech, which has transformed financial markets. The demand for such capacity on the part of financial regulators is itself a function of both the rise in complexity of financial market regulation and the extent to which financial markets have become digitalised and potentially less readily observable through traditional methods.[4]

While it is important to emphasise that digitalisation of financial markets is not a new phenomenon, the transformation of the sector has been marked in recent years. Table 2.1 indicates which elements of the financial services sector have been affected by digital technologies (OECD, 2018h). As this sector has become digitalised, public authorities have been required to turn to digital technologies in order to "observe" developments in financial markets. As the scale, granularity and velocity of the markets accelerates, traditional methods of supervision become increasingly inadequate, and supervisory authorities have turned to the use of advanced analytics in order to better assess developments in financial markets.

Table 2.1. Application of new technologies to financial services

DIGITAL TECHNOLOGY	FINANCIAL ACTIVITIES AND SERVICES							
	Payment services	*Advisory & agency services Planning*	*Investment & trading*	*Lending & funding*	*Insurance*	*Security*	*Operations*	*Communica-tions*
Distributed ledger technology	x	x	x	x	x	x	x	x
Big Data		x	x	x	x	x	x	x
Internet of things					x			x
Cloud computing				x			x	
Artificial intelligence		x	x		x			x
Biometric technology					x	x		
Augmented / Virtual reality		x	x					x

Source: OECD (2018h).

Imperfect observability of policy outcomes is also a particularly pervasive challenge in environmental and natural resource management. Much before the current digital revolution, there has been a long record of using remote sensing technologies to address this challenge, going back at least to the launch of Landsat-1 (then called Earth Resources Technology Satellite) in 1972, followed more recently by the era of Earth Observing Systems (EOS) with the launch of the Terra satellite in 1999 (Melesse et al., 2007). Remote sensing has been extensively used to monitor, among other things, the status of forest resources (Lefsky et al., 2002), larger mammals and birds in open habitats (Leyequien et al., 2007), capture fisheries (Santos, 2000), and ground water resources (Huang et al., 2016). Another example, at a more local level, is the detection via remote sensing of increased turbidity in the Northern Poyang Lake, the People's Republic of China (hereafter "China"), which resulted in a ban on sand mining (Wu et al., 2007; de Leeuw, 2010). Perhaps the most dramatic example, however, was the detection of the hole in the ozone layer through remote sensing in 1985, which ultimately resulted in the Montreal Protocol and follow-up agreements. More recently, remote sensing and digital land parcel identification have been used by European Commission inspectors in order to verify farmers' eligibility for EU Common Agriculture Policy direct aid: in 2010, these technologies enabled 70% of the required inspections (OECD, 2014c).

While past experience with the application of remote sensing technologies is quite extensive, these are not the primary focus of this paper. Rather, we focus on applications of newer-generation digital technologies, which have radically changed the available policy toolbox to monitor imperfectly observable environmental phenomena.

A number of new-generation technologies have been applied to e.g. monitor pollution and natural species conservation. For example, cheaper identification of pollution emissions can be achieved with infra-red sensors. This approach has already been applied by the US Environmental Protection Agency (EPA) to monitor methane leaks from tanks and wells (Ziegler et al., 2015).[5] Big data from Google Street View Cars has been used to enable high resolution air pollution monitoring in Oakland, California (Apte et al., 2017). In the area of conservation, tracking individual animal species is now possible with small, satellite-connected devices, which can replace bulky collars emitting radio signals (Pimm et al., 2015).

New analytical techniques can also be leveraged for monitoring purposes: one example is machine learning, which lies "at the intersection of computer science and statistics, and at the core of artificial intelligence and data science" and aims at constructing computer systems that "automatically improve through experience" (Jordan and Mitchell, 2015). For instance, Israel has put in place systems for online monitoring of pollutant emissions from industrial facilities. The potential of machine learning techniques is now being explored in conjunction with online pollutant monitoring to predict, and ultimately prevent, future pollution episodes (Laster, 2018).

Perhaps most promising, developments in advanced image processing techniques which allow for reliable identification in "unstructured environments", when used in conjunction with machine learning, have the potential to transform conservation policy, tracking wildlife populations and monitoring biodiversity loss. For example, such methods have been used to conduct the first ever census of the Grevy's zebra in Kenya (Berger-Wolf et al., 2016). Similarly, in the case of agriculture artificial intelligence has been used to diagnose a disease, pest or nutrient deficiency affecting a crop, with very significant implications for agriculture extension programmes (Tibbets, 2018).

However, improved observability is not just facilitated by the use of advanced sensors and image processing. A number of other digital technologies can help ensure the traceability of products and services which have the characteristics of credence goods, insofar as their underlying attributes are not visible. Littlefair and Clare (2016) discuss the use of DNA barcoding and "high-throughput" sequencing techniques to detect market substitutions and increase the transparency and security of the food chain. More generally, the use of technologies which allow for product traceability and authentication can play a significant role in reducing product counterfeiting. In this respect, the European Union has devoted considerable effort to applying digital technologies to combat illegal trade (Poiret, 2016).

Obviously this has implications for tax revenue as well as money laundering,[6] but potentially also for consumer health and safety since medicines are one of the most commonly traded counterfeit goods (OECD, 2017f). The 2011/62/UE Directive made the authentication through bar-coding of medicines obligatory. Consumers can use applications on their cellphones to ensure authenticity (Poiret, 2016). While such methods reduced administrative costs (i.e. for customs officials), according to Poiret (2016) "the use of new technologies would (also) help to engage consumers, who would be able via their smartphone or their computer, to download certified information." In a review of emerging "digital" solutions to combat fake medicines, Mackey and Nayyar (2017) identified five distinct categories of technology: mobile, radio frequency identification, advanced computational methods, online verification, and blockchain technology (see also Hoy, 2017).

In a related vein, blockchain technology can be used to certify the characteristics of assets in markets where potential information asymmetries between market participants are pervasive. The case of financial assets is well-known (OECD, 2018h and Berryhill et al., 2018), but potential applications are much wider and can affect a number of different policy domains. For example, blockchain technology is currently being used to enhance traceability of food products (e.g. provenance of organic food) in extended supply chains in China.[7] However, there are a number of other potential domains in which blockchain technology can be used to help buyers identify the provenance and characteristics of goods and services which are produced over extended supply chains.

While in some cases this can relate to the characteristics of the good itself (e.g. food quality), in other cases blockchain can be used to certify the manner in which it was produced. For example, blockchain can be a tool to help assure due diligence with respect to employment conditions (i.e. occupational health and safety, child labour) in internationally fragmented production processes. This can be complemented by digital technologies used to monitor difficult-to-observe employment conditions. For example the deployment of the electronic tachograph can be used to monitor working time of professional drivers.[8] In the construction sector – where accident rates are particularly high – "real-time hazard management" enables tracking the location of workers, materials and equipment in real-time (Guo et al. 2017).

However, the integration of digital monitoring tools in professional settings is a process that should be managed carefully. Social dialogue between employers and employee representatives is important in establishing the objectives of the use of digital technology and the parameters under which it takes place. Particularly, clear guidelines on privacy, non-discrimination and re-purposing are important in ensuring that the full potential of digital technologies is realised, improving job performance and quality, while preventing potential misuse.

In the education field, blockchain is seen as a possible way to securely record higher education credits and diplomas, and facilitate their cross-border recognition across higher education institutions. The internationalisation of higher education has indeed led to increased possibilities of fraud, not to mention the need to exchange quickly information when students want to transfer from one institutions to another (OECD, 2005; Vincent-Lancrin et al., 2015).[9] In 2017, the Massachusetts Institute of Technology (MIT) started an experiment supplementing its traditional diploma with digital diplomas using the blockchain technology. This new trend may ease the administration of student mobility, but also change the way employers receive and use the degrees and credentials of their (potential) employees.

Another possible application of blockchain would be the use of digital technologies for the protection of intellectual property rights, an area that appears relatively underdeveloped compared to other applications, and could yield significant benefits.[10] The case of land title registries is another promising area: in Ghana, BenBen provides a blockchain-based digital land registry, which has reduced the average time for confirmation of land entitlement from one year to 3 months (see Berryhill et al. 2018).

Even without the use of modern technologies, digitalisation allows (or will allow) for a more accurate targeting of some types of policies. For example, in the Netherlands, the support to schools with a large share of disadvantaged students is currently based on school self-reported data. The on-going digitalisation of students' administrative data will allow for a better targeting of these policies, and hopefully contribute to their enhanced effectiveness.

Digital technologies have also been affecting science and innovation policy, an area in which monitoring impacts has been fraught with challenges. However, research and innovation activities increasingly leave digital "footprints" that can be analysed using big data, natural language processing, machine learning and several other types of digital technologies. These footprints are emerging as building blocks for the statistical infrastructure on science, technology and innovation, a topic featured prominently at the OECD Blue Sky Forum on Science and Innovation Indicators in 2016 (OECD, 2017g). Policy makers need new tools to monitor the outcomes from policy initiatives in the field of science and innovation. Such tools are emerging in a number of countries. The Norwegian Ministry of Education and Research is investing in developing the Current Research Information System in Norway (CRIStin) that collects and collates data on

Norwegian research. The system strengthens monitoring and assessment of research funding and provides insights on research activities in Norway.

The digital transformation can positively improve monitoring and supervision of imperfectly observable outcomes, enabling the collection and analysis of increasingly granular datasets. However, the creation of such datasets, together with data-sharing between government agencies and across public-private partnerships can generate digital security vulnerabilities and concerns over individual privacy. The benefits of enhanced monitoring capabilities for public policy need to be balanced with the broader public benefits of enhanced data-sharing, and with individuals' and organisations' legitimate concerns about the protection of privacy.[11]

2.2. Facilitating compliance and enforcement

In conjunction with improving the monitoring of imperfectly observable outcomes, digital technologies can, under certain conditions, allow for improved compliance and regulatory enforcement. For example, if a link can be established between observed conditions and specific actions (and actors) through the data generated by digital technologies, then they can replace (or more usually complement) traditional enforcement methods. This depends upon the extent to which a "trace" can be attributed, and thus upon the density of the monitoring system in relation to potential sources of observed outcomes. As noted by Glicksman et al. (2016), "the question is whether it is possible to draw a causal link, with sufficient statistical support, between an observation […] at a particular location and time and a violation attributable to a particular action".

Even if it is not possible to attribute causality in a legally-enforceable manner, digitally-powered monitoring can be used to identify risk factors, thus helping regulators better target their enforcement activities, focussing on cases in which the probability of non-compliant behaviour is relatively greater. For example, on the specific case of tax fraud and tax evasion, OECD (2018g) notes that "This data [transaction and income data, behavioural data generated from taxpayers interactions with the tax administration, operational data on ownership, identity and location, and open source data such as social media and advertising] can be used as individual sources or in combination to enable partial or full reporting of taxable income and to uncover under-reporting, evasion or fraud. It can also be used to understand better taxpayer behaviour, to measure the impact of activities and to identify the most effective interventions, both proactive and reactive".

Technology solutions can detect tax evasion by way of under-reported sales income, such as installing a tamper-proof "black box" in point of sales machines, which can also provide real-time transaction reporting to the tax authority. The tax authority has assurance of the integrity of the data, as well as the access to all of the transaction records for audit purposes. This has resulted in substantial increases in tax revenue in a number of countries using these tools as described in OECD (2017b). For example the Danish Business Authority is using machine learning to identify fraudulent firms based on VAT and corporate income tax returns.

The case of financial markets is particularly interesting. Just as digitalisation of financial markets (FinTech) has made the supervisory role of financial regulators more complex, it has made it more difficult for both the regulator and regulatee to be aware of non-compliance. The costs of demonstrating compliance with new demands has increased significantly since the Great Financial Crisis, and according to "Let's Talk Payments" the annual spending by financial institutions on compliance in the United States is in excess of

USD 70 billion (Arner et al. 2016.) Such demands have also increased as a consequence of the OECD's Financial Action Task Force, and in particular the provisions which relate to anti-money laundering and the requirement of "knowing one's customer".[12] Against this backdrop the implementation of RegTech can be seen as both a means to reduce compliance costs (for both the regulator and the regulatee) and increase the reliability of enforcement actions in a complex market environment (Arner et al. 2016). Digital technologies not only provide opportunities for the regulator to reduce costs of enforcement, they also can reduce the administrative costs for demonstrating compliance for the regulatee.

In the area of competition policy enforcement, the situation is similar. Indeed, the amount of evidence through which competition authorities must wade is vast. Moreover, with the digitalisation of markets and office practices the material is increasing rapidly. Identifying and interpreting such a massive amount of evidence is increasingly reliant upon digital technologies and practices.[13] As in the case of financial markets, digital technologies can provide opportunities for the regulator to reduce costs of oversight and increase quality, they also can reduce the administrative costs for demonstrating compliance for the regulatee. It has been proposed that the wider application of blockchain technology by firms is likely to be helpful for competition authorities in cases of merger control, cartel investigations and monitoring of commitments in abuse of dominance (Tulpule, 2017).

The analysis of big data and advanced analytics has already been used in a number of settings to identify possible cases of collusion. More specifically, "behavioural screens" relying upon the analysis of big data have been used as a complement to "structural" screens to identify possible cartel-like behaviour. Screens use commonly available data such as prices, market shares, bids, transaction quotes, spreads, volumes, and other data to identify patterns that are anomalous or highly improbable. The best-known example of the use of such a screen was that related to Libor, in which Abrantes-Metz et al. (2012) applied econometric techniques, revealing such anomalous behaviour. Such digital-based behavioural screens are usually not sufficient in and of themselves for compliance and enforcement. Rather they are used to "flag" cases for which the regulatory authorities should devote resources to investigate further through the identification of cases in which there is a relatively higher probability of there being cartel-like behaviour. For example, the Korean Fair Trade Commission's Bid Rigging Indicator Analysis System (BRIAS) screens every public tender, flagging on average more than 80 cases per month, warranting further investigation by the Korean Fair Trade Commission (OECD, 2014b). More specifically, behavioural screens helped identify suspicious patterns in bids for subway/train contracts, and this information was used to seize incriminating documents and prosecute using more traditional methods (OECD, 2014b). Interestingly, the BRIAS system has also allowed the Korean Fair Trade Commission to cast a wider net, subjecting relatively smaller contracts to regulatory oversight. Thus, while the workhorse of compliance and enforcement in the competition space remains paper-based, "digital evidence" has become an increasingly important complement.

In terms of benefit fraud, Berryhill et al. (2018) have noted that potential applications of blockchain in areas such as social security and pension entitlements could be considerable. At the firm level, the State Secretariat for the Information Society in Spain is using graph visualisation and natural language processing in aid programs for the promotion of R&D in the ICT sector: the objective is to track grant applications and avoid double funding of projects or companies by different national funding agencies.

In the area of fisheries regulation, the "landing obligation" under the Common Fisheries Policy (CFP) in the EU increased demand for the use of digital technologies. Since fishers were now required to land all regulated species subject to a Total Allowable Catch (TAC), rather than discarding them (which has become common practice in recent years), an affordable means of documenting catches by species was required. A number of tests have been undertaken to assess the reliability of remote electronic monitoring (REM) linked with closed circuit TV (see e.g. Catchpole et al. 2017, Plet-Hansen et al. 2017). While such methods have not yet been adopted on a wide scale, they do raise the prospect of significantly increasing the quality of fisheries management and reducing waste and over-exploitation.

In the case of Australia, The Australian Fisheries Management Authority (AFMA) uses digital technologies to monitor compliance with Western Central Pacific Fisheries Commission (WCPFC) conservation measures. Traditionally, logbooks on-board observers, information on landings, and port inspections had been used to monitor compliance. However, starting in 2009 the AFMA began trialling electronic monitoring in a range of commercial fisheries under its management, and from mid-2015 it was introduced to all vessels in Australia's Pacific tuna fishery. The system uses hardware (cameras, gear sensors and GPS) and software to collect and transmit fisheries information in an automated manner that avoids any potential for tampering (Larcombe et al. 2016).

Monitoring non-point sources of pollution from agriculture for enforcement purposes is another area where digital technologies hold promise (OECD, 2018c). For example, the US EPA's "Next Generation Compliance Strategy" is assessing how digital technologies can be used to better enforce regulations related to point-source agricultural enterprises (OECD, 2018c). The EPA's New England Regional Laboratory uses a network of solar-powered water quality sensors to identify cases which warrant field-level enforcement action due to relatively higher probabilities of non-compliance (Glicksman et al. 2016).

In the area of natural resource management, the Real-Time System for Detection of Deforestation (DETER) programme in Brazil is an illustrative example. Under this programme the enforcement agency (IBAMA) receives information at a high level of periodicity from satellite images. This allows IBAMA to distinguish between naturally-occurring cases of reduction in forest cover and those arising from human intervention. Enforcement agents are dispatched to regions where there is some evidence that the latter may be occurring, significantly saving on enforcement resources through targeted action. The cost savings are considerable, and the effects on deforestation have been marked. According to a study by Assunção et al. (2013), the programme "helped avoid over 63 800 square kilometres of Amazon forest clearings."

Deploying digital technologies has also eased the detection of illegal waste dumping in Lima: through a joint project between the Peruvian environment ministry and USAID, ten vultures were kitted with GoPro video cameras and GPS trackers.[14] While this enables vultures to identify illegal waste dumps and provides GPS coordinates to environmental authorities, it ultimately is the municipalities' responsibilities to tackle dumping and fly-tipping.

In terms of working conditions, a number of cases cited in Section 2.1 indicate how digital technologies can be used lower the costs of monitoring. In principle this can be linked to regulatory obligations such as those related to occupational health and safety risks: increased easy of monitoring is, in this sense, particularly beneficial in those parts of the economy where compliance and enforcement regimes are particularly costly to implement (e.g. SMEs, fisheries).

In the area of education, the digitalisation of some educational processes, the legal standardisation of data collections and the introduction of individual student identifiers have opened new enforcement possibilities. For example, in Hungary, the admission process to higher education is entirely determined by computers, allowing for a fair and transparent application of the system's central admission rules by all actors in the country.

While digital technologies can facilitate enforcement and compliance in myriad ways, the case of waste also illustrates how digitalisation of economic activities has complicated the regulator's task. For example, the advent and fast expansion of online sales has created new free-riding opportunities: consumers are able to buy more easily from sellers in other countries which are not registered with national Extended Producer Responsibility (EPR) schemes, hence avoiding producer and retailer/distributor obligations and costs. This imposes policy challenges in the form of the reduced competitiveness of domestic producers, obligated to comply with domestic waste regulations, and the increased cost imposed of reprocessing waste to the required standards (OECD, 2018f). Thus, cross-border e-commerce allows for the purchase of products deemed to be unsafe in one country being purchased by residents of that country.[15] Digitalisation can also lead to adverse consequences: for example, according to some observers the digitalisation of the admission process to higher education is one factor explaining the slight decrease in the actual entry rates to higher education in Hungary, as it also makes the system more rigid. This is at odds with the overall policy goal of the country to increase the populations' participation in higher education.

A particular advantage of digitised data is that they can be examined not only for compliance with regulations as discussed but also for anomalous data that is misreported by accident or fraudulently. Misreported data can be detected through a variety of techniques, which can rely on conditions for data to add up or correlate within or across units; or by purely statistical tools. An obvious example is tax reporting, which generates large amounts of data; several countries have used data analytic techniques to identify potential fraudulent tax reporting. These methods can also be applied to environmental data (Rivers et al. 2015): for example, Chen et al. (2012) identify discontinuous jumps in reported air quality index values that correspond to thresholds for awards in China. As another example, Foremny et al. (2017) provide strong evidence that municipalities in Spain manipulated their population figures in order to qualify for grants, and find that improved audits corrected the tendency to manipulate. Improved data-collection through the application of digital technologies can hence support detection of malpractice and facilitate compliance and enforcement.

3. Allowing for the implementation of new and more efficient policy instruments, and facilitate policy evaluation and experimentation

Digital technologies facilitate data collection at a more granular level, potentially allowing for the use of "first-best" policy measures which target the desired policy objective more directly. They can also make differentiated and more targeted policy instruments more viable to implement, allowing for greater efficiency by equating marginal benefits and costs of policy interventions. Finally, they can provide information feedback, often in real time, which can facilitate policy evaluation and lower the cost of policy experimentation, helping government iterate towards efficient policy settings.

3.1. Allowing for improved policy design and differentiation

In some cases, digital technologies can enable the use of the "first-best" instrument which was previously too costly (or technologically infeasible) to implement. By allowing the regulator to target the policy objective directly – rather than some proxy – digitalisation can allow for the implementation of more efficient policy instruments.

For example, agri-environmental policies have traditionally relied upon so-called "pay-for-practices" regulation since the outcome or performance could not be observed directly. With digital technologies, it is possible to move closer to observed outcomes, allowing for the introduction of more results-oriented policies (OECD, 2018c). However, the implications can be much more far-reaching. Across the OECD, tradable permit systems are replacing much less efficient technology-based standards: this has been facilitated by digital technologies which can help attribute integrity to the asset (i.e. the permit), yielding significant economic benefits.[16]

The use of digital technologies can also be used to tailor policy design to specific geographic features. The Norwegian government uses drones to construct orthoimagery of swamps in their efforts to reverse the drainage of wetlands. An elevation model based on drone imagery is used by the Norwegian Environment Agency to plan the restoration of swamps and can be later used for the long term monitoring of the wetlands (Næss, 2018).

The benefits of digitalisation for policy design arise not only from the capacity to have more direct incidence on the policy objective, but also from the possibility to differentiate policy settings across vectors for which the costs and benefits differ. As such, access to such granular information is empowering policy makers to implement more efficient policy instruments, which may not have been technically possible or economically viable if built upon analogue technologies. Relevant vectors include temporal, spatial and technological differentiation. These may intersect in a number of different ways.

Temporal differentiation can be advantageous where external costs and benefits vary across time. This variation may arise at different levels of periodicity – minute-to-minute or over much longer periods of time. Predicting and adjusting policy to changes in supply and demand may be beneficial in regulating a sector such as transport where usage varies widely depending on time.

Improved data-sharing between agencies also enables the potential for targeted service delivery based on socio-economic characteristics. In the UK, the water and energy regulators, OFGEM and OFWAT, have launched a data sharing scheme to collaborate on the identification of vulnerable customers. A recent report found that where a customer needs additional help from their energy company, they will likely need the same from their water company and vice versa (UKRN, 2017).

Further, spatial differentiation of policies may be appropriate where geography is relevant to policy implementation. Latvia uses hydrological and meteorological data for targeted policy responses to flood risk management. Using the Flood Risk Information System, policy makers identify which areas are vulnerable to floods and for the spatial target flood mitigation measures (OECD, 2018d). Spatial targeting can also be combined with socio-economic differentiation: by overlaying spatial and socio-economic and spatial datasets, particularly vulnerable groups or individuals can be offered special services. The US EPA has developed EJSCREEN, a digital screening tool which maps environmental and demographic indicators. This allows the identification of overlap between areas of high pollution and areas of vulnerable populations. EJSCREEN has been made publicly available, allowing state and local governments to better integrate such information into the policy process (OECD, 2018d).[17]

In other cases the level of differentiation may be a function of the technological characteristics of the good or service responsible for the external cost or benefit. In Stockholm and Milan, city authorities exempted hybrid and electric vehicles which fulfil a number of technical specifications from payment of the congestion charge (Comune Milano, 2018; Transport Styrelsen, 2018).

The case of road congestion charges is illustrative of the implications of differentiation more broadly. The use of advanced image processing technologies in conjunction with automatic payment systems can lower transaction costs for the implementation of road charging schemes, obviating the need for less efficient means for addressing congestion, e.g. imposing restrictions on vehicle use by license plate numbers. Since the external costs of congestion vary across the day, such differentiation allows for policy settings which are more efficient. In Singapore and Stockholm, for example, time-differentiated road pricing is used to reduce peak time traffic; in New York, dynamic parking fees are used to reduce the number of cars coming into congested areas.

However, variation across space and time in external costs and benefits can arise in other policy fields as well. In the case of agricultural policies, the introduction of digitally-enabled and spatially-differentiated policies allow the policy maker to strike a better balance between marginal benefits and costs at a disaggregated level (OECD, 2018c).

Policy differentiation may have some limitations: it may conflict with the need of individuals and markets for salience and long-term, stable policy signals. Digital technology enables a tailored policy design approach that differentiates across time and space, and between people or groups. However, this level of granularity may lead to opaque policy and interfere with the ability of market actors to respond to incentives. Policy differentiation is also likely to bear higher implementation costs; hence its rollout should be based on cost-benefit analysis reflections. The availability of real-time social and economic data entails the potential for frequent policy adjustment to address changing market conditions, yet the continuous reformulation of policy may undermine stable, long-term policy signals.

3.2. Facilitating policy experimentation and evaluation

Digital technologies can be leveraged to simplify policy experimentation and ease evidence-based policy making. Ex ante policy experiments can, in theory, provide a cost-efficient way to test different variations of policy measures in small samples and assess their impacts prior to the general roll-out of the most cost-effective variation.

Randomised control trials (RCTs) remain the holy-grail for evaluating the impact of policies, building on the revolutionary impact they have had in fields such as medicine to provide evidence on the efficacy and safety of treatments. However, even in medical research, RCTs face several challenges, including high cost, generalisability of findings, confounding effects of selection and long-time delays. Big data drawing on medical records is being increasingly used in medicine to address some of these limitations of RCTs, and in some cases the two approaches are being used in conjunction (Angus, 2015).

Harnessing the possibilities offered by digitalisation, private companies like Google have dramatically increased the use of RCTs in testing alternate designs of particular web products. Digitalisation can make RCTs, or A/B testing, both inexpensive and fast. In 2011 Google conducted seven thousand A/B tests, while Facebook reportedly runs thousand A/B tests per day, more than the number of such tests run by the entire pharmaceutical industry in a year (Stephens-Davidowitz, 2017). A/B testing has applications beyond tech companies as well. It was, for example, used to test alternate versions of the campaign website of then US Presidential candidate Barack Obama (Stephens-Davidowitz, 2017). An obvious question, therefore, is whether such RCTs or A/B testing has applications in public policy as well. While actual examples of policy applications are hard to identify, there is nevertheless considerable potential in instances where the impact of particular policies or instruments is critically dependent upon the *framing* of the information provided. For example, the design of product labels, or information about energy conservation, and so on. A/B tests could offer a low cost way to effectively test the impact of alternate ways of framing the information, thereby potentially increasing the intended impact (Haynes et al. 2012). It is, however, premature to speculate whether big data approaches could be similarly applied to evaluate policy design, as a substitute or in conjunction with RCTs or quasi-experimental approaches.

On a different track, digitalisation can also facilitate ex-ante policy experimentation in simulated/lab settings. The European Commission's online lab experiments have enabled the testing of variations on energy efficiency labels across multiple EU member states to identify the most effective label format. While the limitations of lab experiments should be acknowledged, this format has allowed the testing of the same policy initiatives across different geographical contexts by exploiting the same online platform. The results of such experiments have been reflected in the updated regulation (London Economics and IPSOS, 2014; Leenheer et al., 2014; Ecorys, Tilburg University and GfK, 2014).

Moreover, greater access to more granular data can facilitate the use of other standard policy evaluation methodologies used outside of the context of controlled experiments, such as regression discontinuity, difference-in-differences and propensity-score matching methods. A challenge in the application of such methods is the identification of a "control" group, since unlike RCTs they are not defined by construction. Big data can facilitate the process by giving the evaluator access to a much richer set of relevant information, to identify relevant control groups and with respect to the underlying data on their characteristics.

Not surprisingly, there has been a mushrooming of institutions applying the principles of scientific labs (experiment, testing and measurement) to public policy. The underlying principle of these institutions is to apply experimental methods used in the hard sciences while leveraging digital technologies to identify optimal solutions to public policy objectives. While much of the work of "policy innovation labs" such as NESTA in the UK and MindLab in Denmark has focussed on social policy, education and health, applications in all areas of public policy are emerging. Collaborations with data-oriented firms also create new opportunities to conduct and evaluate randomised experiments (see Einav and Levin (2014) for a discussion). For example, the London-based Behavioural Insights Team recently launched Predictiv, an online tool enabling public institutions to easily develop and roll out online randomised controlled trials with the objective of testing pilot policy interventions.[18] This allows to easily test a policy on a smaller scale prior to e.g. country-wide roll-out.

In education, as countries develop new generations of information systems, the large amount of administrative data routinely collected is increasingly, albeit still insufficiently, used beyond its statistical purpose to inform policy evaluation (OECD, 2015c). One of the major innovations in this area is the development of longitudinal data systems that follow students over time, sometimes from kindergarten to the labour market. The mere descriptive power of these data can change the policy discussion. For example, in the United States, the comparison of the strength and growth of inequity patterns based on the micro-data of 120 million students across US districts has shifted the policy discussion from inequity levels to the variability of outcomes across and within states (Reardon, 2017). The recent ability to link educational performance data to household income data has revitalised the discussions about inequality and resources that go to schools serving a majority of low income students (e.g. Duncan and Murnane, 2015).

Another example lies in the study of low completion rate of community college attendance. Using state longitudinal data over a 6-year period, Crosta (2014) showed that the common idea about enrolment patterns in U.S. community colleges was inaccurate: only 1% of the students take this in principle two-year programme in two years, and students switch so often between part-time and full-time status during their studies that the very notion of "full-time" and "part-time" enrolment has become questionable. The possibility offered to also link these patterns to graduation (or lack of graduation) has opened new avenues for designing programmes and interventions to improve the outcomes of community college students. These two examples show how administrative data is being used for programme and policy evaluation to address inequalities and achievement gaps in education.

Combining both the "information" and the "technology" elements of digitalisation, perhaps the most promising area in the policy evaluation space is the use of big data generated by embedded digital technologies (e.g. smart cards required to access bike-sharing or public transport) to assess the effects of changes in policy-induced settings (OECD, 2016c). While such information can be harvested at low- or no-cost, access to datasets of this magnitude and granularity can enable empirical assessments of policy effects. One example is the ongoing OECD study which, building on big data, aims to assess the impact of congestion charges on the use of bike-sharing in Milan.[19] Other applications relate are the evaluation of alternative measures related to consumer product safety.

3.3. Lowering the costs of policy implementation

Integrating automated data collection and compliance mechanisms into policy design can substantially lower costs of enforcement for government and costs of compliance for individuals and businesses. The underlying principle of this "compliance by design" approach is to ensure that reported information is correct when filed, reducing the need for, and cost of, ex-post auditing. Compliance by design regulation is well-developed in the domain of tax policy in the area of employment income where in many countries withholding is done by employers on the basis of rules and information provided by the tax authority. As a result compliance rates are high, the burden on the individual taxpayers concerned very low and evasion extremely difficult. However, technology is now being used to expand compliance by design mechanisms to the taxation of businesses by exploiting the increasing rate of digital business transactions.

For example, Brazil is among a number of countries that have mandated e-invoicing: electronically sending, receiving and storing invoices between suppliers and buyers (either business to business or business to government). This has helped establish a national digital bookkeeping system on the basis of which the Brazilian tax authority can now review, assess and act on information almost instantly which in turn strongly influences compliance behaviour (OECD, 2017h). Jacobs (2017) provides a discussion of how the digitalisation of tax enforcement can result in lower rates of tax evasion, through the linking of data on earnings, capital incomes, consumption expenditures, and bequests.

The use of digital protocols and payment systems can lower administrative costs for governments and transaction costs of compliance for citizens. For example, the design of automatic payment systems for road charging schemes can lower both the costs of policy implementation and the transaction costs of compliance by users. This obviates the need for less efficient means for addressing congestion, such as restrictions on vehicle use by license plate numbers. In London, cameras automatically detect and record the number plates of vehicles entering the congestion charge zone. Vehicle-owners can choose to register in a central directory and be charged via direct debit automatically each month (London Transport, 2017). In another example, the UK has established online systems to change, transfer or cancel environmental permits.[20] Thanks to the digitalisation of its admission procedures to higher education, and to the legal possibility to link students' and graduates' individual data to their subsequent income tax data, in Hungary the Education Authority can implement the country's conditional subsidisation policy of higher (e.g. the duty to exercise medicine for a certain number of years in Hungary for the publicly subsidised students) can now easily and cheaply be enforced by the government, while keeping some flexibility for citizens.

Digitisation of citizen and government data offers one pathway to substantial cost reductions. In New Zealand, the Marlborough District Council has digitalised its resource consent system in order to allocate and track the use of natural resources in a simpler and more efficient way (Hearnshaw, 2018). The new system includes record of water use, forestry, stream crossing and marine farming, however, it is often difficult to estimate the reduction in administrative burden derived from the introduction of digital technology. For instance, a number of cost-benefit studies of e-government in Europe estimate cost reductions from user time savings but do not, as a default, calculate benefits from cost savings to government (European Commission, 2012). While difficult to estimate, substantial direct cost savings are likely to come from administrative time saving, greater revenues and "efficiency gains due to the reduction of the number of transactions and improved data/information quality" (European Commission, 2012, p. 34). The Indian government operates the world's largest biometric ID system, Aadhaar, which digitises the transfer of benefits and schemes to residents and aims to cover 1.3 billion people. Given the scale of the program, cost-savings are difficult to quantify but substantial benefits are expected from fraud reduction, reduced leakage and efficiency gains (World Bank, 2016).

While the introduction of digital technologies may substantially lower the administrative costs for government, it may also reduce the salience of price signals for users. This involves two distinct components: first, as automation of payment systems involves a shift from manual payments and associated transaction costs to automatic payments, behavioural responses to price changes may be weakened. The general trend towards digital technology-enabled automatised payment systems can be found in many domains: automatic highway payment enabled by transponders or by drive-through gates with cameras; energy bills sent by e-mail and related payments automatically withdrawn from bank accounts. Second, as administrative costs of managing information are cut through digitising records, a larger volume of data is available to users. However, the sheer volume of information available on all metrics of consumption patterns poses a challenge to a boundedly rational individual. Cognitive limits to the ability to interpret and respond to a vast collection of data points imply diminishing marginal returns to the provision of additional information.

4. Monitoring and predicting emerging risks, opportunities and behavioural responses

Digital technologies can allow policy makers to be more pro-active and reactive in tracking and responding to fast-changing phenomena, whether they be risks or opportunities. At the same time advanced analytics can help to "predict" responses to policy interventions in a more robust manner than was the case previously.

4.1. Monitoring fast-changing phenomena and emerging risks

Digital technologies can be an important tool in tracking fast-changing phenomena such as developments in financial markets, as well as emerging risks (e.g. disease outbreaks) and opportunities (e.g. emerging technologies).

In providing information to public authorities on faster time-scales than conventional tools, digital technologies therefore have the potential to reduce the response time and improve targeting. As long ago as 2004, Singapore created the Risk Assessment and Horizon Scanning program to collect and analyse large-scale datasets in order to manage risks associated with terrorist attacks, financial crises and infectious diseases (Kim et al. 2014).

In the area of public health, Hay et al. (2013a and 2013b) propose spatially continuous maps of infectious disease, which are updated continuously as new occurrence data become available using a machine learning element. Such methods have been used to investigate the potential of using Google Trends or Google queries data to measure the spread of influenza (Preis and Moat, 2014). Similar methods have been applied to model the spread of the 2014 Ebola epidemics in West Africa.[21] Using mobile phone data gathered thanks to an agreement with local mobile phone carriers, researchers have mapped regional population movements. This enables public health officials to predict which places might be at increased risk of new outbreaks, based on their connections to the outbreak location. Ultimately, this type of data can support decision-making on where to focus preventive measures and health care.

A complementary and more targeted approach was used by the Centers for Disease Control and Prevention (CDC) and Liberia's Health Ministry. To respond to the crisis, Liberian authorities set up a special hotline for residents to ask Ebola-related questions and report cases: the geo-data generated by these calls can be layered over census information to identify outbreak areas.[22] These are examples of how big data from sources which would normally not inform health policy (e.g. geolocated metadata from cell phone use) can play a substantial role in emergency response, by providing accurate and granular information in a timely way.

However, there are many documented cases of the hazards associated with such exercises, with the most well-known being the case of Google Flu Trend's (GFT) failure in predicting the 2013 flu season, when GFT predicted "more than double the proportion of doctor visits for influenza-like illness than the Centers for Disease Control and Prevention, which bases its estimates on surveillance reports from laboratories across the United States" (Lazer et al., 2014). Citing this case, in which traditional public health surveillance methods far outperformed methods drawing upon advanced analytics, Khoury and Ioannidis (2014) note that "big error" can plague "big data".

In order to fully benefit from the potential of these technological developments, efforts need to be made to improve international and inter-jurisdictional co-operation. In the domain of health, the increasing movement of people and microbes across jurisdictions entails a need to exchange health information. The Centre for Disease Control has established protocols for intra- and inter-jurisdictional health information sharing which rely heavily on digital record keeping.[23]

Policy makers are also making increasing use of big data to get early warnings of emerging market conditions which may generate significant risks to economic stability. Ultimately, this can enable faster, more accurate policy responses. The clearest examples relate to financial markets, where the scale and speed of transactions in the market can lead to sudden unforeseen shocks. The availability and granularity of data should potentially allow for improved capacity to identify emerging risks ("distress situations") at an early stage through methods such as machine learning. However, the capacity for such methods to yield useful insights is constrained by their relative infrequency. Wall (2016) notes that "machine learning applied to more granular data cannot be of much help when there are few, if any, large shocks during the period in which the data are obtained." This lack of variation is compounded by the fact that those few large shocks which do arise will likely affect most assets in a given portfolio.

Big data analytics are not only useful with respect to the identification of risks, but can also be used to detect patterns of emerging research areas, technologies, industries or policy issues. This can support short-term forecasting of issues of policy concern (OECD, 2017g). For example, in the United States a number of agencies, including the White House Office of Science and Technology Policy, the National Science Foundation and the National Institute of Health established the Big Data Research and Development Initiative. One of the main objectives of this initiative was to advance state-of-the-art big data technologies in order to accelerate discovery in science and engineering (Kim et al. 2014).

4.2. Improving policy prediction and nowcasting

Another form of differentiation which can improve policy settings relates to policy "prediction" according to the characteristics of the affected population, whether at the level of the individual or at a more aggregated level. Many of the technologies used in this domain are similar to those discussed above with respect to the identification of emerging risks. Access to big data, combined with new statistical techniques such as machine learning (ML), can open the road to a radically different way of targeting policy interventions.[24] ML algorithms can be used to predict the sub-sample of the potential target population that is most likely to be "policy-compliant", i.e. responsive to the intervention (Mullainathan and Spiess, 2017).[25] Moreover, these algorithms can make greater use of data structured in non-traditional formats which standard techniques are less well-suited to address.

Private companies already use big data and predictive modelling to engage in such kind of targeting – however there are clearly significant challenges to governments applying predictive modelling in a similar vein. Predictive models based on machine learning, for example, might be particularly suited for the design of tax rebates aimed at providing economic stimulus if targeted to households with the highest marginal propensity to consume (Einav and Levin, 2014).

Other examples of policy domains in which ML could be or has been used to predict outcomes include teacher quality (Rockoff et al., 2011), regulatory inspections (Kang et al., 2013), social policy (Chandler et al., 2011) and tax rebates (Andini et al. 2017). The information derived can be used to potentially mitigate issues of moral hazard and adverse selection, and thus increase the returns associated with policy interventions in these different domains. Moreover, as Kleinberg (2016) notes "as we expand our notion of what is predictable, new applications will arise". The potential of ML can be harnessed on the emergence of longitudinal administrative data. For instance, longitudinal data following students over time could allow to identify students at risk of adverse educational outcomes: not completing high school on time (Lakkaraju et al., 2015), dropping out of one's higher education programme (Zeng et al., 2017), having a lower academic growth than expected based on past performance, etc. While this identification still needs to be supplemented by human judgment, it allows educational authorities to target their remedial interventions at an earlier stage than what could be done before.

While digital technologies have allowed for policy prediction, inaccurate predictions or misidentification can pose a challenge. While this also happens with more traditional forms of screening, it is often more politically sensitive when done on a systematic basis by an algorithm. In education, it is increasingly common to assign students to schools based on an algorithm: this is the case in France and England for example. As part of Race to the Top, a US federal education policy, states were asked to evaluate the quality of teachers in part based on their value added on student test scores – and some districts based decisions such as promotion or dismissal on value added models. Beyond political opposition, this has created some unease in the research community, sometimes as an over-simplified response to the complex problem of educational improvement (Raudenbush, 2015), but also for the mere technical limitations in the quality of the data or in the involved estimation techniques given small samples (Ballou and Springer, 2015). For example, some estimates could show that the value added of the same teacher could be rated very differently depending on the size of their classes (and thus on the size of the sample used for estimation).

In addition to the role it can play in terms of policy prediction, digital tools and methods can be used to generate more timely estimates of important macroeconomic variables. Using digital tools to monitor and predict the present and very near future is known as "nowcasting" and can enable better real-time data collection and a more rapid policy response. For example, online-generated data is widely used to generate estimates of the rate of inflation. One prominent example is the Billion Price Project (BPP) which collects price information over the Internet to compute a daily online price index and estimate annual and monthly inflation (Cavallo and Rigobon 2016). Initially developed for Argentina in 2006 in the face of concerns about the accuracy of official data, the project now generates measures for 22 countries (see also Carrière-Swallow and Labbé, 2013).

Recent academic studies have used online activity of Internet users to estimate unemployment rates. Askitas and Zimmermann (2009), in a study of the German labour market, analysed the prediction power of keywords relating to job searches and found that changes in the unemployment rate could be detected much earlier than in official statistics. D'Amuri and Marcucci (2012) undertook similar exercises using internet job search data in the United States. Pavlicek and Kristoufek (2015) find that job-related Google searches in the Czech Republic and Hungary can be used nowcast unemployment rates.[26] This type of approach is adopted, for instance, by the European Central Bank (Rodrigues and Speciale, 2017).

Online big data can be used to quantify the number of job offers versus the number of job seekers. Both China and Italy nowcast job vacancy rates by scraping data on vacancies from job search websites (Hammer et al., 2017). The Australian Department of Education, Employment and Workplace Relations (DEEWR) has created the Internet Vacancy Index (IVI) based on new vacancies from four online recruitment websites. This complements the newspaper-based Skilled Vacancy Index (SVI) also created by the DEEWR (Reimsbach-Kounatze, 2015).

Big data is also being leveraged to provide a more granular picture of job market matching: to this end, the Minnesota Department of Employment and Economic Development has developed the Graduate Employment Outcomes data tool which links postsecondary graduate outcomes to wage records.[27] This allows the observation of matching patterns by region, institution type and field of study. Similarly, the Bank of England have used data from online job adverts and techniques from machine learning to analyse how mismatches in demand and supply of labour affect occupational and regional productivity and output (Turrell et al., 2018). This type of data analysis can help to identify the types of skills required in segmented labour markets in the face of automation.

Such initiatives can help governments monitor economic activity in real-time, helping to address emerging issues. For example, information on short-term changes to the estimated unemployment rate can enable more responsive policy and be used to shift resources to required areas. Moreover, when combined with the digitalisation of tax administration, public expenditures, and allocation of benefits (so-called "financial management information systems") governments have a much more accurate and timely view on their fiscal stance. This can allow them to manage the business cycle in a more prescient and nimble manner (Misch et al. 2017).

Such initiatives are likely to be of particular value at inflection points, when the cost of significant lags in data collection can be particularly great. However, public institutions have to use caution: data sources informing the policy process should be both reliable and stable. Benoît Coeure, member of the ECB Executive Board, summarised this as follows: "Just as there are concerns about 'fake news' dominating social media, there is a risk of 'fake', or at least poor quality, statistics driving out better quality ones in public discourse" (Rodrigues and Speciale, 2017).

5. Building more participatory relationships between government and stakeholders

The digital transformation offers the potential for more participatory relationships between government, citizens, civil society, trade unions and the private sector. For governments, digital platforms provide a low-cost means of eliciting perceptions and engaging stakeholders in policy design, monitoring and implementation. However, the potential for more direct engagement with government, also increases demand for a more user-responsive and accountable government. The use of digital tools for stakeholder engagement (e.g. web-based surveys) has been uneven across countries. More widespread adoption of these technologies holds significant scope for improving how governments interact with their citizens.

There are three broad areas in which digital technologies are reshaping government interaction with citizens and facilitating greater stakeholder engagement: open data initiatives; crowdsourced data collection and monitoring; and stakeholder engagement.

5.1. Developing open data initiatives

Open government data (OGD) is the practice of making government data publicly available on digital platforms. OGD can increase accountability and transparency of government while fostering public participation in policy making. OECD governments are at different stages of promoting open government data to generate socio-economic impact: the OGD frameworks developed by Korea, France and Great Britain are particularly advanced (OECD 2017e). For OGD to be effective, it is necessary to complement data availability with an emphasis on data reusability. For instance, to facilitate the re-use of OGD by other users, the Norwegian public agencies make data available in standardised, machine-readable formats through the *data.norge* portal (European Commission, 2017). In Denmark, a data catalogue (*Virk Data*) has been created with open government data from various public institutions. The catalogue is "living", and will expand gradually as more data is made available.[28]

Such cases can empower firms and individuals to make more informed decisions. For example, Virk Data is promoting the use of "public" data in Danish business, arranging hackathons with partners from the business community and universities. [29] Virk data also promotes IT-solutions based on open data, demonstrating how companies are using open data to develop new business models.

The availability of environmental data from industrial facilities such as the European Pollutant Release and Transfer Register (E-PRTR) give the public insight into processes that cause environment pressures (OECD 2018e). Indirectly, they may drive improvements in the environmental performance of industrial facilities as the private sector civil society. Release of data on local levels of air pollution enables individuals to change transport routes to minimise exposure to air pollutants. By providing data, government can act as a platform which facilitates collaboration between institutional and non-institutional actors (OECD 2017e, p. 2). A useful example of this idea is the growing number of transport and weather mobile applications, tailored to individual needs and designed by private companies, but which rely on data collected by public agencies.

Increased availability of information can contribute to greater awareness of individual impacts and broaden the scope of policy interventions. Critically, the extent to which OGD is used and consulted by stakeholders depends on how successful governments have been in standardising the formats and access requirements across different platforms.

5.2. Crowdsourcing data collection and monitoring

Engaging the public in the collection of data has been facilitated by the proliferation of smart devices which reduce administrative costs and enable remote sharing of data. Examples of participative data collection efforts are particularly common in the environmental sphere. Commonly referred to as “citizen science”, recent initiatives include eBird, HotSpotter and Air Sensor Toolboxes (see Leibovici et al. 2017, Clark 2014, and Glicksman et al. 2016, Ziegler et al. 2015). The technologies required in order for such programmes to be successful vary. For instance, eBird, an online database collecting bird observations reported by 100 000 smartphone-wielding citizens, is dependent upon highly advanced image-recognition algorithms. Analogous technologies have been used to prevent double counting when conducting a census of an animal species. In Kenya, researchers have developed Hotspotter, an image recognition software which can identify individual zebra by their barcode-like stripe pattern and body shape. In 2015, hundreds of volunteer scientists and members of the public participated in photographing zebra using GPS-enabled cameras. Images were then run through the image recognition software which led to the identification of 2 350 unique animals (Berger-Wolf et al., 2016). Data on population size and demographic distribution are valuable in assessing the long-term viability of a species.

Engaging citizens in data collection can be complementary to traditional remote sensing approaches. The United States Environmental Protection Agency’s “Air Sensor Toolbox for Citizen Scientists, Researchers and Developers” provides guidance on the use and maintenance of air pollution monitors for so-called citizen scientists. However, “EPA has specific guidelines it must use in establishing regulatory-grade air monitors. No lower cost sensors currently meet these strict requirements or have been formally submitted to EPA for such a determination” (US EPA, 2014). The issue of quality assurance is clearly relevant if third parties are to be used to collect data for regulatory purposes. While governments can now turn to citizens for crowdsourced data, and as social media has lowered barriers for citizens to engage with government on regulatory violations, the importance of assessing potential bias has also grown. As Glicksman et al. (2016) point out “data generated individuals and community groups may be “self-selected with unsure representativeness”. In areas where internet access is not complete, data collected from digital devices may not be representative of the broader population. Limited access or lack of affordability in obtaining and using digital devices may exclude certain stakeholders from the process. Crowdsourced data must therefore be externally validated data with due considerations to the representativeness of the digitally connected population as a share of the larger whole. The issue of errors and potential bias remains a barrier to the direct use of user-generated data in policy making.

5.3. Engaging stakeholders for better policy design

Digital technologies can also be used to heighten engagement between regulator and regulated entities. For instance, ‘regulatory sandboxes’ which allow for the managed introduction of new technology have now been introduced in several countries. The UK Financial Conduct Authority (FCA) has since 2016 operated a regulatory sandbox to allow firms to test the introduction of new financial technologies (FinTech) in a controlled environment (FCA, 2018). Regulatory sandboxes allow financial companies to test innovative products within a pre-defined time period and at a small scale. During the period of testing, regulatory authorities can closely monitor the effect of new technologies. For policy makers, this provides the data necessary to define appropriate consumer protection safeguards. For regulated entities, the risk, and cost of introducing innovative technologies can be reduced. This regulatory approach has predominantly been adopted to enable innovation in finance but it is spreading to other domains. Both the UK and Singaporean energy regulators have launched or are preparing to launch a regulatory sandbox for smart energy innovations.

The different stakeholders may also be within governments. In Canada, the development of an Education and Labour Market Longitudinal Linkage platform will provide Canadian citizens the possibility to make more informed study and career decisions. As the data will be open, they will also contribute to the evaluation of government programmes by a variety of stakeholders, and thus to more possible engagement and contributions to policy design or improvement. Last but not least, it is also providing opportunities for various levels of government, for example provincial and federal, to work together: for example, in order to improve education access, the governments of Canada and of Ontario worked together to integrate the federal government's registered savings plans into the province's online birth registry system, so that the take-up of low income children's access to the Canada learning bond can be enhanced.

6. Conclusion

Digitalisation has importantly reshaped the delivery of public services and, more broadly, certain areas of government operations: this trend has been labelled "e-government". At the same time, the application of digital technologies to policy development and implementation has evolved at a slower pace.

On one hand, we can observe a growing number of – mostly prototype – applications aimed at improving efficiency of current policies and broaden the portfolio of policy instruments. On the other hand, public policy applications have generally been much more limited than in the private domain. Many applications still at the "boutique" stage, although there is some scale-up.

While some applications of mature technologies are now well established (e.g. remote sensing), the potential of newer technologies (e.g. blockchain) and analytical approaches (e.g. machine learning) is still vastly underexploited in policymaking. Public sector adoption rate remains low relative to private sector, with the majority of policy examples identified only implemented at the local level or representing pilot projects. Furthermore, most of these applications will not displace conventional practices and technologies, but rather be used in conjunction with them.

This is a partial view: digitalisation also creates a number of new challenges for policy design, implementation, and enforcement. Many of the challenges are common for public and private applications. These include data access and interoperability of data systems within and across different areas (i.e. allowing for multiple datasets to be linked). While some digital solutions to enforce some policies (resource allocation, admission to schools or other public services) can lead to better acceptability by citizens, in part because of their transparency, their lack of flexibility or unintended consequences may raise new policy issues. In order to fully realise the potential benefits from new data sources, a framework for information management needs to be put in place.[30] Further, the insufficient public infrastructure to link disparate sources of data is a key bottleneck.

Some other issues, meanwhile, are more sensitive in a public policy context: the increased granularity of data and increased data-sharing between government agencies and across public-private partnerships can generate cyber-security vulnerabilities and legitimate concerns over individual privacy. The ownership of digital data is a critical issue, as is the transparency of data analytics. The extent to which digitally-generated data used for policy evaluation is based on a representative sample of the regulated population is another concern when access to digital technologies is less than universal.

Finally, while the availability of more data usually enables to improve policies, it is not a panacea and comes with risks that will need to be tackled over the next decade: in some instances, less data is better than more. Mainstreaming digital best practices across national institutions might be the largest challenge to be tackled. For governments in particular, a greater challenge is to ensure their workforce acquires the skills necessary to best make use of the digital revolution (e.g. data science).

Despite progress in the domain of information sharing, the velocity and volume of the flow of goods and services across borders empowered by advances in digital technology poses an inherent challenge to policymaking which remains rooted to national jurisdictions. Overcoming this challenge necessarily requires international co-operation.

The extent to which the potential of digitalisation in policymaking is realised will depend on whether governments prove willing to adopt and able to scale the use of such technologies, obtain reliable access to relevant data which is often in the hands of private actors, and how successfully concerns such as privacy and cybersecurity are addressed.

Notes

[1] See, for example, http://www.oecd.org/gov/digital-government/Digital-Government-Strategies-Welfare-Service.pdf

[2] The OECD has worked on these issues for some time; recent work is presented at http://www.oecd.org/gov/digital-government/

[3] According to Arner et al. (2017), "RegTech describes the use of technology, particularly information technology (IT), in the context of regulatory monitoring, reporting, and compliance."

[4] In the words of Andy Haldane (Chief Economist, Bank of England): "I have a dream. It is futuristic but realistic. […] It would involve tracking the global flow of funds in close to real time […], in much the same way as happens with global weather systems and global internet traffic. Its centrepiece would be a global map of financial flows, charting spill-overs and correlations." (Quoted in Arner et al. 2016).

[5] See Ziegler et al. (2015) for a discussion of EPA use of digital technologies in different domains.

[6] According to OECD (2017c), counterfeit products represent 2.5% of international trade.

[7] https://www.economist.com/news/world-if/21724906-trust-business-little-noticed-huge-startups-deploying-blockchain-technology-threaten

[8] See Berryhill et al. (2018) and Zambrano (2017) for a discussion of actual and potential applications of BlockChain by public authorities in their regulatory and administrative efforts.

[9] The Groningen Declaration network has started to bring together higher education institutions and accreditation agencies around the world supporting the provision of authentified higher education diplomas (see www.groningendeclaration.org).

[10] Tulpule (2017) cites the use of blockchain in the application of monitoring commitments related to IP in cases of abuses of dominance.

[11] See the OECD Privacy Guidelines and the 2016 Cancun Ministerial Declaration on the Digital Economy. A considerable body of work on such issues can be found at http://www.oecd.org/going-digital/topics/digital-security-and-privacy/

[12] It is interesting to note that if the application of the technology by the regulatee is not mandated, there can be disincentives for investment in such technologies in jurisdiction. This is analogous to the perverse incentives which characterise some "self audit" programmes. See Pfaff and Sanchirico (2000).

[13] The recent case by the EC Competition authorities against Google is one such case. In the words of Commissioner Vestager "before reaching our conclusions we have analysed huge quantities of data. This includes 5.2 Terabytes of actual search results from Google. That's the equivalent of 1.7 billion search queries, or about 460 million copies of my statement here today. It would take me more than 17 000 years to read them out." (see http://europa.eu/rapid/press-release_STATEMENT-17-1806_en.htm)

[14] http://www.minam.gob.pe/notas-de-prensa/gallinazo-avisa-una-iniciativa-que-rastrea-zonas-contaminadas-de-lima-junto-a-las-aves-mas-emblematicas-de-la-capital/

[15] The Committee on Consumer Policy has been doing work in this area.

[16] For an early example see OECD (2004).

[17] See https://www.epa.gov/ejscreen

[18] http://www.behaviouralinsights.co.uk/predictiv/

[19] This work is overseen by the Working Party on Integrating Environmental and Economic Policies.

[20] https://www.gov.uk/guidance/change-transfer-or-cancel-your-environmental-permit

[21] https://www.technologyreview.com/s/530296/cell-phone-data-might-help-predict-ebolas-spread/

[22] https://www.msh.org/sites/msh.org/files/technology_and_ebola_response_in_west_africa_technical_brief_final.pdf

[23] https://www.cdc.gov/phpr/readiness/00_docs/capability6.pdf

[24] The World Bank has recently launched a competition with Driven Data to see how well a household's poverty status can be predicted using ML algorithms: see https://blogs.worldbank.org/opendata/data-science-competition-predicting-poverty-hard-can-you-do-it-better.

[25] By forsaking the need to generate unbiased estimates of the effect of specific predictors – as is the case with traditional econometric techniques - they are able to better predict which household, firm or other agent will respond to a given policy incentive.

[26] Although it should be noted their model performed much less reliably for Poland and Slovakia.

[27] https://mn.gov/deed/data/data-tools/graduate-employment-outcomes/

[28] https://data.virk.dk/what-is-virk-data

[29] https://data.virk.dk/what-is-virk-data

[30] For example, in the United States, the Office of Management and Budget issued a government-wide policy on information management, encouraging interoperability and openness (OMB, 2013). Other countries which have made significant progress in this regard include Denmark and the United Kingdom (OECD, 2017d). The OECD Recommendation of the Council from Enhanced Access and More Effective Use of Public Sector Information (2008) provides key elements for improved information management systems. The Recommendation of the Council on Digital Government Strategies (2013) is more closely targeted on the strategic use and reuse of public sector data, evidence and statistics.

References

Abrantes-Metz, R. D. D. Sokol (2012), "The Lessons from Libor for Detection and Deterrence of Cartel Wrongdoing", *Harvard Business Law Review Online*, Vol. 3:10-16.

Andini, M. et al. (2017), "Targeting policy-compliers with machine learning: an application to a tax rebate programme in Italy" Bank of Italy Discussion Paper Number 1158 - December 2017.

Angus, D. C. (2015), "Fusing randomized trials with big data: the key to self-learning health care systems?" *Jama*, 314(8), pp. 767-768.

Apte, J.S., et al. (2017), "High-resolution air pollution mapping with google street view cars: exploiting big data", *Environmental Science & Technology*, *51*(12), pp. 6999-7008.

Arner, D.W. et al. (2016), "FinTech, RegTech, and the Reconceptualization of Financial Regulation", *Northwestern Journal of International Law and Business*, *37*, p. 371.

Askitas, N. & Klaus F. Zimmermann (2009), "Google Econometrics and Unemployment Forecasting," Applied Economics Quarterly (formerly: Konjunkturpolitik), Duncker & Humblot, Berlin, vol. 55(2), pp. 107-120.

Assunção, J., C. Gandour and R. Rocha (2013), "DETERring Deforestation in the Brazilian Amazon: Environmental Monitoring and Law Enforcement" Climate Policy Initiative CPI Report May 2013.

Ballou, D. and M.G. Springer (2015), "Using Student Test Scores to Measure Teacher Performance: Some Problems in the Design and Implementation of Evaluation Systems", *Educational Researcher*, Vol 44, Issue 2, pp. 77 – 86.

Berger-Wolf et al. (2016), "The Great Grevy's Rally: The Need, Methods, Findings, Implications and Next Steps" Technical Report, Grevy's Zebra Trust, Nairobi, Kenya, Aug. 2016.

Berryhill, J., T. Bourgery and A. Hanson (2018), "Blockchains Unchained: Blockchain Technology and its Use in the Public Sector", OECD Working Papers on Public Governance, No. 28, OECD Publishing, Paris.

Capobianco, A. (2017), "Algorithms and Collusion: Background Note by Secretariat" (internal document).

Carrière-Swallow, Y. and F. Labbé (2013), "Nowcasting with Google Trends in an Emerging Market," Journal of Forecasting, John Wiley & Sons, Ltd., vol. 32(4), pp. 289-298, 07.

Catchpole, T.L. et al. (2017), "The challenges of the landing obligation in EU fisheries", *Marine Policy*, 82, pp.76-86.

Cavallo, Alberto, and Roberto Rigobon (2016), "The Billion Prices Project: Using Online Prices for Measurement and Research." *Journal of Economic Perspectives*, 30 (2): 151-78.

Chassot, E. et al. (2011), “Satellite remote sensing for an ecosystem approach to fisheries management”, *ICES Journal of Marine Science*, Volume 68, Issue 4, 1 March 2011, pp. 651-666, https://doi.org/10.1093/icesjms/fsq195.

Chen, Y. et al. (2012), “Gaming in Air Pollution Data? Lessons from China,” The B.E. Journal of Economic Analysis & Policy: Vol. 13: Iss. 3 (Advances), Article 2. DOI: 10.1515/1935-1682.3227

Choi, H. and H. Varian (2012), Predicting the Present with Google Trends. Economic Record, Vol. 88, pp. 2-9. doi:10.1111/j.1475-4932.2012.00809.x.

Clark A. “Where 2.0 Australia’s Environment? Crowdsourcing, Volunteered Geographic Information, and Citizens Acting as Sensors for Environmental Sustainability”. *ISPRS International Journal of Geo-Information*. 2014; 3(3):1058-1076.

Coglianese, C. and D. Lehr (2017), "Regulating by Robot: Administrative Decision Making in the Machine-Learning Era" *Faculty Scholarship*. 1734.

Comune Milano (2018), Congestion Charge - Area C, https://www.comune.milano.it/wps/portal/ist/en/area_c (accessed 2 April 2018).

Crosta, P. (2014), “Intensity and Attachment: How the Chaotic Enrollment Patterns of Community College Students Relate to Educational Outcomes.” Community College Review 42 (2): 118–42.

D'Amuri, Francesco and Marcucci, Juri, The Predictive Power of Google Searches in Forecasting Unemployment (November 29, 2012), Bank of Italy Temi di Discussione (Working Paper) No. 891. Available at SSRN: https://ssrn.com/abstract=2207915 or http://dx.doi.org/10.2139/ssrn.2207915.

De Leeuw, J. et al. (2010), “The function of remote sensing in support of environmental policy”, *Remote sensing*, *2*(7), pp.1731-1750.

Döll, P. et al. (2014), “Global-scale assessment of groundwater depletion and related groundwater abstractions: Combining hydrological modeling with information from well observations and GRACE satellites”, *Water Resources Research*, 50, 5698–5720, doi:10.1002/2014WR015595.

Duncan, G. and R. Murnane (2016), *Restoring Opportunity. The Crisis of Inequality and the Challenge for American Education*, Harvard Education Press and Russel Sage Foundation.

ECORYS, Tilburg University, and GfK (2014), Study on the effects on consumer behaviour of online sustainability information displays, Brussels, https://doi.org/10.2759/52063.

Einav, L. and J. Levin (2014), “The data revolution and economic analysis”, *Innovation Policy and the Economy*, 14(1), pp.1-24.

European Commission (2017), “Norway ready for full open data services”, https://joinup.ec.europa.eu/document/norway-ready-full-open-data-services (accessed 12 February 2018).

European Commission (2012), “Study on eGovernment and the Reduction of Administrative Burden” European Commission DG Communications Networks, Content & Technology, Brussels.

FCA (2018), “Regulatory sandbox” Available at: https://www.fca.org.uk/firms/regulatory-sandbox (Accessed 5 December 2017).

Foremny, D., J. Jofre-Monseny and A. Solé-Ollé (2017). “‘Ghost citizens': Using notches to identify manipulation of population-based grants”. *Journal of Public Economics*, Vol. *154*, pp. 49-66.

Glicksman, R. L. et al. (2016), “Technological Innovation, Data Analytics, and Environmental Enforcement” *Ecology Law Quarterly*, Vol. 44, No. 1

Gluckman, P. (2017), “Enhancing Evidence-Informed Policy Making” available at http://www.pmcsa.org.nz/wp-content/uploads/17-07-07-Enhancing-evidence-informed-policy-making.pdf.

Guo, B. H. W. et al. (2017), Overview and analysis of digital technologies for construction safety management” University of Canterbury Research Repository” http://hdl.handle.net/10092/14583.

Hay, S.I., et al. (2013a), “Global mapping of infectious disease”, *Phil. Trans. R. Soc. B*, *368*(1614), p.20120250.

Hay, S.I. et al. (2013b), “Big Data Opportunities for Global Infectious Disease Surveillance” in PLOS Medicine, April, Vol. 10, Issue 4.

Haynes, L. et al. (2012), “Test, Learn, Adapt: Developing Public Policy with Randomised Controlled Trials”, Cabinet Office Behavioural Insights Team.

Hearnshaw, E. (2018), *Digitisation in New Zealand*, presentation at the meeting of the OECD Working Party on Integrating Environmental and Economic Policies, March.

Höchtl, et al. (2016), “Big Data in the Policy Cycle” in *Journal of Organizational Computing and Electronic Commerce* Vol. 26, Nos. 1-2, pp. 147-69.

Hoy, Matthew B. (2017), “An Introduction to the Blockchain and Its Implications for Libraries and Medicine” Medical Reference Services Quarterly, 36, No. 3, pp. 273-279.

Huang, J. et al. (2016), “Mapping groundwater storage variations with GRACE: a case study in Alberta, Canada", in *Hydrogeology Journal*, Vol. 24, No. 7, pp. 1663-1680.

International Institute for Finance (2016), “RegTech in Financial Services” IIF, 2016. www.iif.com/topics/.

Jacobs, B. (2017), "Digitalization and Taxation", in: Sanjeev Gupta, Michael Keen, Alpa Shah, and Genevieve Verdier (eds), *Digital Revolutions in Public Finance*, Washington-DC: International Monetary Fund, Ch. 2, 25-55.

Jordan, M. I., and Mitchell, T. M. (2015), Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255-260.

Kaufman, A. et al. (2014), “The Air Sensor Citizen Science Toolbox” Papers from the 2014 HCOMP Workshop (www.aaai.org).

Khoury, M. J. and Ioannidis, J. P. (2014), Big data meets public health. *Science*, 346(6213), 1054-1055.

Kim, Gang-Hoon, S. Trimi and J-H Chung (2014), “Big Data Applications in the Government Sector” Communications of the ACM, March 2014, Vol. 57, No. 3.

Kleinberg, J., Ludwig, J., Mullainathan, S., and Obermeyer, Z. (2015), Prediction Policy Problems. *American Economic Review: Papers & Proceedings*, 105(5):491–495.

Lakkaraju, H., E. Aguiar, C. Shan, D. Miller, N. Bhanpuri and R. Ghani (2015), “A Machine Learning Framework to Identify Students at Risk of Adverse Academic Outcomes”, Conference on Knowledge Discovery and Data Mining (KDD), 2015.

J. Larcombe, R. Noriega and T. Timmiss (2016), “Catch reporting under E-Monitoring in the Australian Pacific longline fishery”. Paper presented at the 2nd Meeting of the Electronic Reporting and Electronic Monitoring Intersessional Working Group, Bali Indonesia.

Laster, Y. (2018), *Developing a Strategy for Digital Transformation: Challenges & Insights*, presentation at the meeting of the OECD Working Party on Integrating Environmental and Economic Policies, March.

Lazer, D. et al. (2014), "The parable of Google Flu: traps in big data analysis". *Science*, *343*(6176), 1203-1205.

Leenheer, J. et al. (2014), Study on the strategic analysis for optimising the role of ICT in EU policy delivery. https://ec.europa.eu/digital-single-market/en/news/study-effects-consumer-behaviour-online-sustainability-information-displays-final-report-and.

Lefsky, M.A. et al. (2002), "Lidar remote sensing of above‐ground biomass in three biomes", *Global ecology and biogeography*, *11*(5), pp.393-399.

Leibovici D.G. et al. (2017), "On Data Quality Assurance and Conflation Entanglement in Crowdsourcing for Environmental Studies", *ISPRS International Journal of Geo-Information*; 6(3):78.

Leyequien, E. et al. (2007), "Capturing the fugitive: applying remote sensing to terrestrial animal distribution and diversity", *International Journal of Applied Earth Observation and Geoinformation*, *9*(1), pp.1-20.

Littlefair, J. E and E. L. Clare (2016), "Barcoding the Food Chain", *Genome*, Vol. 59, pp. 946-958.

London Economics and IPSOS (2017), "Study on the impact of the energy label – and potential changes to it – on consumer understanding and on purchase decisions", report for the European Commission, https://ec.europa.eu/energy/sites/ener/files/documents/Impact%20of%20energy%20labels%20on%20consumer%20behaviour.pdf.

London Transport (2017), "Congestion Charging Auto Pay User Guide" accessed 18 August 2017 at http://content.tfl.gov.uk/congestion-charging-auto-pay-user-guide.pdf.

Lowman, D. M., et al. (2013), "Fishery Monitoring Roadmap." *Environmental Defense Fund Report*.

Mackey, Tim. K. and G. Nayyar (2017), "A review of existing and emerging digital technologies to combat the global trade in fake medicines", *Expert Opinion on Drug Safety*, Volume 16, 2017 - Issue 5.

Melesse, A.M. et al. (2007), "Remote sensing sensors and applications in environmental resources mapping and modelling", *Sensors*, *7*(12), pp.3209-3241.

Misch, F. et al. (2016) "Nowcashing: Using Daily Fiscal Data for Real-Time Macroeconomic Analysis" in Sanjeev Gupta, Michael Keen, Alpa Shah, and Genevieve Verdier (eds), *Digital Revolutions in Public Finance*, Washington-DC: International Monetary Fund, Ch. 2, 25-55.

Mitchell, A. L. et al. (2017), "Current remote sensing approaches to monitoring forest degradation in support of countries measurement, reporting and verification (MRV) systems for REDD+", *Carbon Balance and Management*, Vol 12, No. 1.

Mullainathan, S. and Spiess, J. (2017), "Machine learning: An applied econometric approach", *Journal of Economic Perspectives*, 31(2):87–106.

Næss, E. (2018), *Digitisation in Norway*, presentation at the meeting of the OECD Working Party on Integrating Environmental and Economic Policies, March.

OECD (2018a), "Smart Data and Digital Technology in Education", OECD, Paris (internal document).

OECD (2018b), "Embracing Innovation in Government", OECD, Paris (internal document).

OECD (2018c), "Digital Opportunities in Agriculture: Some Policy Implications", OECD, Paris (internal document).

OECD (2018d), "Harnessing the Potential of Digitalisation for Better Environmental Policies" Unpublished Survey Results, OECD, Paris

OECD (2018e), "Revised Recommendation of the Council on Establishing and Implementing Pollutant Release and Transfer Registers (PRTRs)", OECD, Paris https://legalinstruments.oecd.org/public/doc/631/631.en.pdf

OECD (2018f), "Extended producer responsibility (EPR) and the impact of online sales", OECD, Paris, https://doi.org/10.1787/cde28569-en

OECD (2018g), *Tax Challenges Arising from Digitalisation – Interim Report 2018: Inclusive Framework on BEPS*, OECD/G20 Base Erosion and Profit Shifting Project, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264293083-en

OECD (2018h), *Financial markets, insurance and pensions: Digitalisation and Finance*, OECD Publishing, Paris, http://www.oecd.org/finance/private-pensions/Financial-markets-insurance-pensions-digitalisation-and-finance.pdf

OECD (2017a), "Making policy evaluation work: The case of regional development policy", OECD Science, Technology and Industry Policy Papers, No. 38, OECD Publishing, Paris, https://doi.org/10.1787/c9bb055f-en

OECD (2017b), "Technology Solutions – Tools to Tackle Tax Evasion and Tax Fraud", OECD, Paris, http://www.oecd.org/tax/crime/technology-tools-to-tackle-tax-evasion-and-tax-fraud.pdf.

OECD (2017c), "Shining light on the Shadow Economy: Opportunities and Threats", OECD Publishing, Paris, http://www.oecd.org/tax/crime/shining-light-on-the-shadow-economy-opportunities-and-threats.pdf.

OECD (2017d), Fostering Innovation in the Public Sector, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264270879-en.

OECD (2017e), Government at a Glance 2017, OECD Publishing, Paris, https://doi.org/10.1787/gov_glance-2017-en.

OECD (2017f), "Fast-track document on more responsible use of medicines for chronic conditions: drivers, impact and policies. The case of non-adherence to medicines for diabetes, hypertension and high blood cholesterol", OECD, Paris (internal document).

OECD (2017g), "Initial scoping report for the CSTP project on *Digital science and innovation policy and governance*", OECD, Paris (internal document).

OECD (2016a), "Big Data: Bringing Competition Policy to the Digital ERA", OECD, Paris (internal document).

OECD (2016b), "Groundwater Allocation", https://doi.org/10.1787/9789264281554-en.

OECD (2016c), "Using Big Data for Environmental Policy Analysis" OECD, Paris (internal document).

OECD (2016d), *Technologies for Better Tax Administration. A Practical Guide for Revenue Bodies*, OECD Publishing, Paris, https://doi.org/10.1787/9789264256439-en.

OECD (2015a), *Data-Driven Innovation: Big Data for Growth and Well-Being*, OECD Publishing, Paris, https://doi.org/10.1787/9789264229358-en.

OECD (2015b), *OECD Regulatory Policy Outlook 2015*, OECD Publishing, Paris https://doi.org/10.1787/9789264238770-en

OECD (2015c), "Innovation strategy for education and training: Progress report no. 10", OECD, Paris (internal document).

OECD (2014a), *Tax Compliance by Design: Achieving Improved SME Tax Compliance by Adopting a System Perspective*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264223219-en.

OECD (2014b), "Ex officio cartel investigations and the use of screens to detect cartels", *Competition Policy Roundtables*, OECD, Paris (internal document).

OECD (2014c), *The Space Economy at a Glance 2014*, OECD Publishing, Paris, https://doi.org/10.1787/9789264217294-en.

OECD (2005), Guidelines on quality provision in cross-border higher education, OECD Publishing, Paris, https://doi.org/10.1787/9789264055155-en-fr.

OECD (2004), "The US SO2 Cap-and-Trade Programme" in *Tradeable Permits: Policy Evaluation, Design and Reform*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264015036-en

Office of Management and Budget (2013), "Open Data Policy – Managing Information as an Asset, Memorandum", Executive Office of the President of the United States.

Pavlicek J, Kristoufek L (2015), "Nowcasting Unemployment Rates with Google Searches: Evidence from the Visegrad Group Countries". PLoS ONE 10(5): e0127084. doi:10.1371/journal. pone.0127084.

Pfaff, A.S. and C.W. Sanchirico (2000), "Environmental self-auditing: Setting the proper incentives for discovery and correction of environmental harm", *Journal of Law, Economics, and Organization*, 16(1), pp.189-208.

Pimm, S.L. et al. (2015), "Emerging technologies to conserve biodiversity", *Trends in ecology & evolution*, 30(11), pp.685-696.

Plet-Hansen, K.S. et al. (2017), "Remote electronic monitoring and the landing obligation–some insights into fishers' and fishery inspectors' opinions", *Marine Policy*, 76, pp.98-106.

Poiret, C. "Digital Revolution and Illegal Trade" European Issues, No. 413, 29 November 2016.

Preis, T., & Moat, H. S. (2014), Adaptive nowcasting of influenza outbreaks using Google searches. *Royal Society open science*, *1*(2), 140095.

Raudenbush, S.W. (2015), "Value Added: A Case Study in the Mismatch between Education Research and Policy", Educational Researcher, 44: 138-141.

Reardon, Sean F. (2017), "Educational opportunity in early and middle childhood: Variation by place and age", CEPA Working Paper, Stanford University.

Reimsbach-Kounatze, C. (2015), "The Proliferation of "Big Data" and Implications for Official Statistics and Statistical Agencies: A Preliminary Analysis", *OECD Digital Economy Papers*, No. 245, OECD Publishing, Paris, http://dx.doi.org/10.1787/5js7t9wqzvg8-en.

Rodrigues J. and A. Speciale (2017), "How central banks are using big data to help shape policy", December 18, Bloomberg Technology, https://www.bloomberg.com/news/articles/2017-12-18/central-banks-are-turning-to-big-data-to-help-them-craft-policy.

Santos, A.M.P. (2000), "Fisheries oceanography using satellite and airborne remote sensing methods: a review", *Fisheries Research*, *49*(1), pp.1-20.

Smith, R. A. and C. D. Kummerow (2013), "A Comparison of In Situ, Reanalysis, and Satellite Water Budgets over the Upper Colorado River Basin" in *Journal of Hydrometeorology*, Vol. 14, June, pp. 888-905.

Stephens-Davidowitz, S. (2017), *Everybody lies: Big data, new data, and what the internet can tell us about who we really are*. Dey Street Books.

Tibbets, J.H. (2018), "From Identifying Plant Pests to Picking Fruit, AI is Reinventing how Farmers Produce your Food" see http://www.downtoearth.org.in/news/from-identifying-plant-pests-to-picking-fruit-ai-is-reinventing-how-farmers-produce-your-food-60079.

Transport Styrelsen (2018), Congestion taxes in Stockholm and Gothenburg, https://transportstyrelsen.se/en/road/Congestion-taxes-in-Stockholm-and-Goteborg/ (accessed 2 April 2018).

Tulpule, A., (2017), "Enforcement and Compliance in a Blockchain(ed) World", *CPI Antitrust Chronicle*, Volume 1, pp. 45, Winter 2017, https://ssrn.com/abstract=2906465.

Turrell, A. et al. (2018), "Using job vacancies to understand the effects of labour market mismatch on UK output and productivity" Bank of England Staff Working Paper No. 737 – July 2018.

UK Regulators Network. (2017), *Making better use of data: identifying customers in vulnerable situations*.

US EPA (2014), "Air Sensor Guidebook" US Environment Protection Agency.

Vincent-Lancrin, S., D. Fisher and S. Pfotenhauer (2015), *Ensuring Quality in Cross-Border Higher Education: Implementing the UNESCO/OECD Guidelines, OECD Publishing*, Paris, http://dx.doi.org/10.1787/9789264243538-en.

Volume 154 (2017), pp. 49-66, ISSN 0047-2727, https://doi.org/10.1016/j.jpubeco.2017.08.011. (http://www.sciencedirect.com/science/article/pii/S0047272717301433.

Wall, L. D. (2016), "Prudential Regulation, Big Data and Machine Learning: Notes from the Vault" Federal Reserve Bank of Atlanta, Nov. 2016.

West, Jarrod and M. Bhattacharya (2016), "Intelligent Fraud Detection: A Comprehensive Review" in *Computers and Security*, Vol. 57, pp. 47-66.

World Bank (2016), World development report 2016: Digital Dividends. World Development Report. Washington, D.C.: World Bank Group. http://documents.worldbank.org/curated/en/961621467994698644/World-development-report-2016-digital-dividends-overview

Wu, G., et al. (2007), "Concurrent monitoring of vessels and water turbidity enhances the strength of evidence in remotely sensed dredging impact assessment", *Water Research*, *41*(15), pp.3271-3280.

Zambrano, R. (2017), "Unpacking the disruptive potential of blockchain technology for human development" (Ottawa, IDRC) https://idl-bnc-idrc.dspacedirect.org/bitstream/handle/10625/56662/IDL-56662.pdf?sequence=2&isAllowed=y.

Zeng, Z., S.C. Chin, B. Zeimet, R. Kuang and C.L. Chi (2017), "Dropout Prediction in Home Care Training", Proceedings of the 10th International Conference on Educational Data Mining.

Ziegler, C. et al. (2015),"US EPA Digital Science: An Evolution" American Geophysical Union, Fall Meeting 2015, abstract #IN31B-1768.