Speech

Supervisor-centred automation – the role of human-centred automation in judgement-centred prudential supervision

Speech given by
James Proudman, PRA Senior Advisor

To have been given at a conference on the 'Impact of AI and Machine Learning on the UK economy', organised by the Bank of England, CEPR and Imperial College

26 March 2020

INTRODUCTION

It's March 2030. The phone rings[1]. It is the PRA CEO's office. "How worried should we be about BigBank and SmallBank's vulnerability to the widget market?" "Give me an hour" you say…

Accessing the shared data repository from your laptop, you extract BigBank's and SmallBank's current credit and derivative exposures to the widget market, and double-check there are no other regulated entities with exposures as large. You then estimate the impacts of a 20% fall in widget prices on last night's current and projected capital ratios for the two banks, and then – with an eye to informing macro-prudential effects - for the system as a whole. You calculate indirect exposures by estimating the effect a 20% fall in widget prices will have on BigBank's and SmallBank's counterparties. To complement your quantitative analysis, you analyse various textual sources - including recent internal analysis, investor assessment and social media commentary on BigBank's and SmallBank's widget-lending business, and compare it with relevant past episodes. Thinking worst case, you then retrieve data on the banks' stock of liquid assets last night: you then simulate what could happen to liquidity positions under different assumptions – derived in part from historic events and in part from AI learning from those same events - about social media reactions, operational resilience and equity prices. You decide to cross check your own thinking against that of BigBank's and Small Bank's boards, and extract from the board minutes the bank's own evaluations and their precautionary steps.

An hour later, you call back, "No need to worry, and here's why…"

If only things today were this easy but - as I know from my previous career as a prudential supervisor - this is not how supervision currently works. Prudential supervision is aimed at promoting the safety and soundness of regulated financial firms by identifying the key risks those firms face, and implementing strategies to mitigate them. Much of what makes the job difficult and slow stems from problems accessing, aggregating and analysing data versus regulatory standards, as well as finding and processing the needle-like evidence from the giant haystack of firms' management information. Supervisory data – and information more broadly - are only available with a lag; they are hard to extract, and to compare; they are difficult to manipulate and harder to interpret. Supervision can at times feel like detective work.

But the hard graft of supervision does not have to be as hard as it is today. Advances in technology are rapidly changing both the world around us, and our ability to analyse those changes – through the application of increased computing power to mixed structured and unstructured data sets and text - freeing up time for the forward-looking, judgement-based supervision, which is where the true value add of supervision lies.

---

[1] Or whatever device we are using by then.

In these remarks, I look to develop some of the themes that I first explored in a speech a couple of years ago about the possibilities offered by artificial intelligence[2], and explore how far it might be reasonable to expect technology to change the way prudential supervision is done over the next 5 to 10 years – towards what one might label 'supervisor-centred automation'. In particular, advances in technology pose 3 practical questions about the way we do prudential regulation and supervision – centred around supervisory access to timely and accurate regulatory data; the introduction of a machine-executable rulebook; and the application of technology within prudential supervision itself.

ACCESS TO DATA

The first question is whether advances in technology can help improve regulatory data and the way it is collected.

In the example I used in the introduction, my imaginary supervisor of the future was able to draw on up-to-date data, specified to answer a precise question, at the click of a button. So, my first question is aimed at exploring what are the constraints on accessing regulatory data in real – or near real – time, and indeed at different levels of granularity, depending on the question at hand. This question is currently under active consideration at the Bank of England and PRA.

At the risk of over-simplification, the current process works something like as follows: firms extract underlying 'input' data for a particular date (such as end quarter) from their core business servers; they then combine and manipulate that input data to transform them into 'regulatory' data that meet regulators' definitions and standards; they then transpose that data into whatever electronic format is needed to submit to the regulator. Subject to meeting accuracy, governance and timeliness requirements, the PRA leaves it broadly up to the firms themselves to work out how best to execute these steps.

Significant advances in technology have transformed our ability to gather, store and analyse data in multiple forms from multiple sources, be they structured or unstructured, more quickly and more cheaply. So now is a good time for regulators to stand back, and ask whether – while holding the line on prudential standards - we are collecting data in the most efficient and cost effective way, from the perspective of both regulators and those we regulate. And if we are not, how we should implement change to advance our statutory objectives most effectively, both for us and those we regulate. As the Governor explained at Mansion House last summer, in response to the Future of Finance report, "this is the new frontier of regulatory efficiency and effectiveness. [We are] exploring how new technologies

---

[2] "Cyborg supervision – the application of advanced analytics in prudential supervision" 19 November 2018

could streamline firms' compliance and regulatory processes while improving our ability to analyse relevant data"[3].

As a first step in the Bank of England's strategy, we have published "Transforming data collection from the UK financial sector"[4]. This is a 'green' discussion paper, in which are set out a range of options for the future of regulatory data collection, and invite feedback from industry and others.

At this stage, the Bank of England and the PRA are agnostic about the end outcome – in the sense that we do not have a strong view at this stage about which technology will best promote our objectives, recognising that these differing objectives may impose certain trade-offs. For example, some technology solutions may be more efficient, but come at greater fixed cost – potentially creating a trade-off between our safety and soundness objective, and our secondary competition objective. There may also be trade-offs between the different benefits on offer. Some technology solutions may offer shorter production times for regulatory data, but come at the expense of accuracy. There may also be broader benefits that flow from greater standardisation at the operational level, reflecting data's role as public good.

How far should we move away from the existing architecture of standardised data requirements and reporting towards more flexible solutions? At the extreme, there is the "pull" model of data collection, implicitly described in the example I used in the introduction to these remarks - in which regulators would be able to pull data, at any level of granularity, directly from firms' systems in real time, with no intervention on the part of firms.

This revolutionary approach would eliminate altogether the need for regulatory data reporting by firms. But it would come with other costs attached, both financial and otherwise, and the technological obstacles to overcome remain significant. It is not yet clear how it could be implemented in such a way as to ensure responsibility for data accuracy and congruence remained where it currently lies – that is, with regulated firms. It is not yet apparent, for example, whether the public benefits of real-time, or near to real-time, data access for prudential supervisors would exceed the costs of delivering it.

To map out more clearly the costs and benefits of this and other approaches, the Bank/PRA wants active industry-wide collaboration on the way ahead, with the intention of agreeing and communicating on a collective approach to the reform of data collection over a 5-10 year horizon.

THE RULE BOOK

---

[3] "New economy, new finance, new Bank", speech by Mark Carney, 21 June 2018
[4] Bank of England Discussion Paper, 7 January 2020

The second of my two questions is whether removing manual intervention in regulatory compliance and reporting – through the application of machine executable rules – would speed up processes underpinning prudential regulation, compliance and supervision. At the PRA, we are currently working on a proof of concept.

There is nothing in principle to stop a set of rules being transposed into a machine-executable code in a way that could be read and processed by a robot. But, there are two key challenges to be overcome: one substantive and one practical.

The substantive challenge is whether the nuances of natural language can ever be satisfactorily replicated in numeric code. Principles-based regulation relies on the application of judgement; and sometimes it is appropriate to write regulations in a way that requires the regulated to meet both the spirit, as well as the letter, of the rules. This is particularly the case in the financial services industry, in which history shows that incentives can arise to restructure contracts in such a way as to circumvent the letter of rules.

To date, our experimental approach at the PRA to the question of whether rules can be code-able has been what might best be described as iterative. That is to say, the approach is to begin with the most black-and-white rule, or blocks of rules, and see how easy it is to code up in a way that satisfies policy makers and supervisors; if it is, then proceed to the next most black-and-white rule, and see on, until you find a point of nuance that cannot be coded in an acceptable way. That way, we should, in theory at least, be able to work out how much – or how many blocks, if any - of the rule book is code-able.

The second – practical – challenge is how to introduce machine-executable rules in such a way as not to affect adversely the PRA's secondary competition objective – that is to say, not to provide an advantage to larger firms that might have greater access to technology resources. Introducing machine-executable rules amongst regulated firms in bite-sized chunks, for example, might make it more manageable for smaller firms to digest. For some smaller firms, it could even prove to be a competitive advantage.

To the extent that some or all of the rule book remains written in natural language, complying with it could be made simpler and quicker for regulated firms by making what Sam Woods described at the Mansion House last year as the "mighty forest of UK financial regulation" easier for a machine to read"[5]. For example, by more consistently applying meta-data and tabs to not just the rule book, but also the related library of supervisory expectations, it would become easier and quicker for a more or less intelligent search engine to find and collect together all the relevant and related pieces of regulatory and supervisory text.

---

[5] "Credit union meets robot", speech by Sam Woods, 24 October 2019

SUPERVISOR-CENTRED AUTOMATION

The PRA's prudential regime is based around the principle that judgement by supervisors is required to assess key risks to the safety and soundness of the regulated firms, and to design, implement and oversee strategies to mitigate those risks – as illustrated again by my introductory example. And human-centred automation is sometimes used to refer to the situation in which humans can do tasks or make judgements better than machines, and automation is designed around these strengths. The third question I would like to raise, therefore, is how best to introduce – or at least move closer to - human-centred automation into a judgement-centred prudential regime.

There are plenty of opportunities for the application of AI, machine learning or other advanced analytic techniques into supervision. Across the PRA, there are increasing examples of the application of new advanced analytics and machine learning: be it to improve the quality of firm-by-firm peer analysis; monitoring of social media for developments in authorised firms; standardising our assessment of certain credit risk books; to automation of the verification of large exposures rules. Gradually over time, advances in technology and modelling techniques should – I believe – make more possible the type of flexible desk-top simulations of banks' balance sheets imagined in my example - just as an earlier generation of technology enabled, some 25 years ago, quick-fire desk-top simulations of the effect of shocks on the macro economy: but there remain significant technical and practical challenges to overcome.

One area where new technology offers, in my opinion, the opportunity for real productivity gains quickly is in the field of natural language processing. In the hypothetical example, I allowed my imaginary supervisor of the future to cross-check his/her analysis against that of the supervised firm. In practice, chasing down relevant information, facts and arguments within the voluminous management and board information provided by firms is one of the lengthiest challenges faced by supervisors: both finding the correct information to assess a particular problem, as well as making sure that relevant information or warnings are not accidentally overlooked within the mass of otherwise irrelevant information. It can be like searching for the proverbial needle in the haystack.

More important, however, than outlining personal wish-lists of ideas for the application of new technology, is to develop processes that can harness the natural creativity and aptitude of people for change. In the case of applying technology within prudential supervision, processes are needed to overcome the classic management problem where one set of people know that there is a business process problem that needs fixing, but do not necessarily possess the requisite skills to know how to do so (e.g. supervisors); and another group of people have the requisite skills, but do not necessarily understand the business problem that needs fixing (e.g. the technology teams). At Bank of England, we are exploring flexible mechanisms for encouraging and explicitly matching supervisors together with technology experts to brain-storm and jointly design new solutions to quantitative problems.

Such mechanisms are well suited to the task of finding quick fixes to local problems. A more strategic approach, however, is likely to prove necessary to make a reality of a longer-term goal of embedding technology at the heart of how prudential risks are supervised – that is, not simply identifying applications in supervision that would benefit from technology, but fundamentally re-engineering the way we work. To do so is likely to require not just the application of technology, but cultural change more broadly – facilitating training and re-skilling, and creating appropriate incentives.

As well as there being plenty of opportunities for embracing change, there are also, inevitably, challenges too – including, for example, the degree of available bandwidth for already busy supervisory teams – so our approach will be proportionate and delivered over time, in line with appropriate training.

How might this transformation impact the 'day job' and skills of supervisors? The aim is 'human centred automation': using technology to free up supervisors' time to focus on where they can add the highest value – making supervisors' jobs more productive, more insightful and more rewarding. To do so is likely to require some change in skills – incorporating more quantitative and analytical skills – without altering the fundamental ability to apply judgement to assess key risks to firms' safety and soundness and evaluate practical mitigants.

And how might changes like those I have described impact firms' experience of being supervised? For the better, I believe. Smarter, quicker supervision should generate fewer costly ad hoc data requests from firms; and generate better-informed, more timely and more insightful supervision.

CONCLUSION

Technology is rapidly changing the world around us. As prudential regulators, we need to understand the impact of that technology. First and foremost, we need to understand its impact on the firms we supervise – and the financial system as a whole – if we are to understand properly risks to financial stability, and safety and soundness – just as we have always done. There is nothing new in this.

But what is new is the need – and opportunity - to take advantage of the new technology in our own business processes to change the way supervision is done. Providing answers to the questions I have outlined in the preceding remarks will help us to know how far we might, in time, go in introducing technology into supervision, and provide a road map for the future of how prudential supervision could be done. Who knows, we might even one day make a reality of the imaginary anecdote of the introduction.