RBA Conference Remarks

Dominic Jones[1] – Discussant for "Central Bank Communication of Uncertainty" (Klodiana Istrafei)

Thank you for the opportunity to discuss this paper. The following discussion is built upon my reading of the paper – any and all mistakes of interpretation are my own.

Central bankers have some internal view of the world and its likely future states – you might call this a narrative. This narrative underlies policy. The core argument of the paper is that we can *measure* the confidence that central bankers (and presumably other policymakers) have in their narrative. It turns out that this is informative for monetary policy.

The measurement is of something I've understood to be a "surprise-confirmation" axis. Some passage, for example, might express surprise, indicating that a policymaker would be less confident in their narrative. Or, it might express confirmation, and so they would be more confident in their narrative.

In some sense this is a surprising dimension to think about policy with: if we wanted to relate communication with policy, why not use a measure of hawkishness or dovishness? Or, in another direction, why not measures of policy uncertainty? The authors are certainly not ignorant of these: the key contribution that the authors make is demonstrating that this "surprise-confirmation" axis is both measurable and enlightening, even given other measures.

The authors use a (mostly) traditional dictionary approach, with a slight relaxation of the bag-of-words assumption. This assumption treats documents as unordered sets of words (think scrabble bag, but words). The dictionary then defines the words you're interested in – pull them out of the bag, add, multiply, divide, whatever, and so measurement. The paper relaxes this a bit by considering *patterns* of words (e.g. "than … expected") rather than individual ones.

Most importantly, the dictionary is purpose-built. I think this is probably the most important thing to do when you're using a dictionary. Off-the-shelf dictionaries are often built for a *completely different task* than whatever they're applied to [1]. The dictionary is also surprisingly small! Though the patterns may cover many more concrete examples – the dictionary itself fits in a few tables in one page the appendix.

At this point, I have a few comments on the approach. First: smaller dictionaries will be more vulnerable to (even small) semantic shift. This occurs when the way we use language changes over time. The way we spoke in the 1800s, generally, is quite different to how we speak today. Should we worry about this for central bank speak? Probably not: we tend to be quite conservative, at least with our communication.

Second: is the dictionary complete? Are there passages of text that have a semantically similar meaning but a different form? Third: small dictionaries are going to be subject to more variance driven by sampling fluctuations. If I was writing the latest Statement on Monetary Policy, for example, I might tire of using the same word ("consistent").

Finally, I'd be curious to see what an out-of-sample validation exercise looks like when we think of the dictionary as a "trained" classifier. As opposed to "downstream-only" validation (i.e.: does the measurement line up with our intuition? If we put it in a regression, do the parameters make

[1] The views expressed here are the author's and do not reflect those of the RBA.

sense?), it would be good to see how the dictionary performs on actual text – validated by a human [2]. Error analysis would also be useful to understand the limitations of the dictionary.

Now I have some questions for discussion about the results. I'll split this more-or-less how the paper does: first, some discussion about how central bankers talk, and then some about the connection with monetary policy.

First on how we talk.

I found the differences between the measures of economic policy uncertainty and the measures of surprise-confirmation interesting. Uncertainty and surprise are certainly (apologies) related. What's driving the difference? If the dictionaries are small enough, term-level differences could provide a mechanical explanation – but why might the relationships differ across country?

I'm also curious on how the dictionary would fare on sentences meant to convey forward guidance. We'd use a lot of the same language – do the control words in the dictionary effectively filter these sentences? Probably. Should we?

Finally, one of the empirical findings of the paper is that central bankers express more confirmation than surprise. Does this line up with our historical forecast errors?

Now on to the connection with monetary policy.

Before I mentioned that if we're interested in policy, the hawkish-dovish axis is a cleaner measurement. Of course, I understand it's not as clean if you want to get at the *narrative* central bankers hold. But I'd be curious about the shared and distinct information between hawkish-dovish and surprise-confirmation.

Finally, if it takes some time for confidence to build (especially considering costly policy), or policymakers have high smoothing parameters, what do the lead-lag relationships look like?

That's all I have. Thank you again for the opportunity to discuss this stimulating paper.

References

[1] Chan, C.- hong, Bajjalieh, J., Auvil, L., Wessler, H., Althaus, S., Welbers, K., van Atteveldt, W. and Jungblut, M. 2021. Four best practices for measuring news sentiment using 'off-the-shelf' dictionaries: a large-scale p-hacking experiment. Computational Communication Research. 3, 1 (Apr. 2021), 1–27.

[2] Justin Grimmer and Brandon M. Stewart. 2013. Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. Political Analysis 21, 3 (2013), 267–297. DOI:https://doi.org/10.1093/pan/mps028