Panel on the future of high frequency data

Contributor: Björn Fischer, Adviser, DG-S Data Integration

I would like to start first on the question of the future of the high frequency data used in the assessment of the Covid-19 crisis.

Looking structurally at this question, one probably needs to confess that we are at this workshop at the peak of a typical cycle. As labelled by Seung-Pyo Jun et al[1], it is the 'peak of inflated expectations', a period in which positive outlooks dominate. Such a period is followed by the 'trough of disillusionment' dominated by negative challenges. When being able to addressing these challenges positively, successful technologies and theories are developed. We have seen similar cycles after the emergence of Google Trends and I have personally gone through this cycle when being tasked in the 90ies to evaluate, whether we could derive reliable early estimates of industrial production using high frequency energy consumption corrected for productivity and technological progress, going through pains similar to those presented by James Stock this afternoon.

If we are thus at such a peak of inflated expectations, it may make sense to look more structurally in the future, whether those high frequency indicators will remain relevant for central banking. In that respect, I will leave the aspects related to now-casting economic indicators to Domenico, who contributed with his theoretical and practical work considerably to technological progress in that domain.

In order to play a permanent role in decision making at the ECB or central banks in general, high frequency data need to be relevant in the policy process over and beyond this inflated expectations peak and probably over and beyond now-casting.

In this respect I see 3 arguments in favour of why these data will play an increasing role.

First, and most obvious, high frequency activities like high frequency trading, change the behaviour of markets and economic agents and thus need to be modelled and assessed appropriately to appraise for example their risks to financial stability.

Second, high frequency data belong to the standard toolset of central banks not least to overcome endogeneity problems in assessing monetary policy, in particular in times of unconventional monetary policy measures.

As you are certainly aware, I refer here to event studies based on high frequency data that emerged during the past crisis, helping to identify the impact of policy announcements referring to unconventional monetary policy tools like APP, forward guidance and alike on longer-term yields. The event-study methodology provides some value-added as compared to lower-frequency

[1] Seung-Pyo Jun, Hyoung SunYoo and San Choi, 2018, Ten years of research change using Google Trends: From the perspective of big data utilizations and applications, Technological Forecasting and Social Change Volume 130, May 2018, Pages 69-87.

methodologies. As outlined by Thornton[2] they first, by being model-free are not impacted by model uncertainty. Second, the high-frequency element of it avoids running into endogeneity problems that are often observed for models based on lower frequency data. Thus, event studies became a standard tool in the toolset of central bankers already during the previous crisis for assessing the announcement effect of non-standard monetary policy measures and remained so in this crisis. Furthermore, as outlined by James Stock in his presentation, could be applied as well for other purposes, close to the topics of this workshop.

Third, high frequency data, including those used in Covid-19 crisis and discussed during this workshop can support to further overcome the insularity of our macroeconomic profession and allow assessing economic and social phenomena better

As Andrew Haldane put it[3], macroeconomic modelling has been under intense scrutiny since the Great Financial Crisis, when serious shortcomings were exposed in the methodology used to understand the economy as a whole. Criticism has been levelled at the assumptions employed in the dominant models, particularly that economic agents are homogeneous and optimizing and that the economy is equilibrating.

In overcoming those criticisms, we need to overcome the insularity of our macroeconomic profession by learning from other professions. One important topic in this respect is in my opinion raised by Robert Shiller's work on narrative economics[4] that is based on the insights gained in other disciplines that persons react significantly to narratives. Thus Robert Shiller promotes the idea that in macroeconomics, narratives might be considered as exogenous shocks to the economy and thus deserves to be modelled properly. Without doubt, this would require the use of diverse high frequency datasets, including network analysis using social media and news. High frequency data thus does not only support tracking lower frequency information but deserves attention on its own.

Thus, given all these arguments, I am strongly convinced that these high frequency data will stay with us over and beyond the current positive expectation peak.

If one is convinced about this, your final questions to us panellists become definitely relevant. Which data will be used and what is the role that governments and international institutions should play in the provision of these data?

To answer this question concretely, I would like to quote from the planned activities as announced by the European Commission, which, in February this year, without doubt very late, published its Data Strategy[5]. Domenico with his experience on both sides of the Atlantic might provide here further insights in the comparison of the US with Europe in this respect. I would like to outline this strategy here, to allow as well other panellists and participants of the discussion to raise missing items in what I call here the benchmark strategy for Europe.

---

[2] Daniel L. Thornton, 2014, An Evaluation of Event-Study Evidence on the Effectiveness of the FOMC's LSAP Program: Are the Announcement Effects Identified?, St Louis Fed Working Paper 2013-033B

[3] Andrew G Haldane and Arthur E Turell, 2017, An interdisciplinary model for macroeconomics, Bank of England Staff Working Paper No. 696.

[4] See Robert J. Shiller, 2017, Narrative Economics, NBER Working Paper No. 23075

[5] See https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy

The motivation of the European Commission is recognition that digital technologies have transformed the society with data being at the centre of this transformation. In that respect, the Commissions European data strategy has the vision to create a single market for data.

The actions are based on four pillars, three horizontal and one vertical pillar, ranging from providing a proper legal framework and level playing field and statistical standards, over investment in required infrastructure and skills to prioritisation of strategically important data spaces.

Pillar 1 is labelled a cross-sectoral governance framework for data access and data use

Key actions are

- Proposing a legislative framework for the governance of common European data spaces, Q4 2020

- Adopting an implementing act on high-value data-sets, Q1 2021

- Proposing, a Data Act enabling among other goals business to business data sharing and government to business data sharing, 2021

Pillar two is called enabling pillar: Enabling through Investments in data and strengthening Europe's capabilities and infrastructures for hosting, processing and using data

Key actions foreseen here are in particular Investment in European data spaces, encompassing data sharing architectures and governance mechanisms, as well as the European federation of cloud infrastructures and related services. The amount of combined investments of EUR 4-6 billion is targeted. The first implementation phase is foreseen for 2022;

The third pillar is labelled competences pillar: Empowering individuals, investing in skills and SMEs

The fourth pillar defines common European data spaces in strategic sectors and domains of public interest

Those are of interest for our discussion, as they define data areas of particular interest in Europe that are prioritised over other topics. Concretely, 9 areas are proposed:

1. A Common European industrial data space;
2. A Common European Green Deal data space;
3. A Common European mobility data space;
4. A Common European health data space;
5. A Common European financial data space;
6. A Common European energy data space;
7. A Common European agriculture data space;
8. Common European data spaces for public administration;

9. A Common European skills data space, to reduce the skills mismatches between the education and training system on the one hand and the labour market needs on the other.

These are the rough outline of the European Data strategy of the European Commission that I would like to present here as the current benchmark plan for Europe and as an answer to your question, what governments are currently planning concerning their role in providing data, often high frequency data. A lot of papers in the past 2 days demonstrated that such data can be used effectively not only for the tasks that were described in the European Commission strategy, but in particular show their strength in the creation of interconnected multipurpose datasets for a number of purposes, including now-casting standard economic indicators.

Let me conclude with two statements

A. Yes, I strongly believe that high frequency data and its analysis via AI will play a structural role in the central bank toolbox, over and beyond the current peak of inflated expectations. An appropriate legal and technical infrastructure, interconnected data allowing the creation of multipurpose datasets and machine learning tools applied in a scalable and reproducible way will need to support this process as well as openness to other disciplines from our macroeconomic profession. Without doubt, this process will need a fair collaboration among many actors to be successful. The question, what type of data will be used in the future is open to creativity. What we have seen so far from own experience confirms that datasets are multipurpose datasets and that data will not only cover classical numerical measures or indicators, but as well will include more and more text data not least for sentiment analysis, pictures and videos and other unstructured information.
B. Europe is lacking behind in data sharing and in providing powerful data pools for machine learning activities as compared to other economic areas. At the same time, the European Commission's Data Strategy announced this year provides a concrete starting point to a better future and a benchmark for our discussion.

Thank you very much for your attention.