Self-driving neural networks for term structure modeling

Sicco Kooiker Janneke van Brummelen Julia Schaumburg Marcin Zamojski

discussion by Sarah Mouabbi

The views expressed are those of the authors and should under no circumstances be interpreted as reflecting those of the Banque the France or the Eurosystem.

Brief overview

- Model:
    - Factor model of the yield curve: L, S, C
    - Replace fixed loadings with shallow neural networks that vary over time
    - Preserve interpretability using identification restrictions on loadings

- Findings:
    - Outperforms standard benchmarks (RW, NS, Diebold-Li) on average and uniformly across maturities
    - Strongest gains at short forecast horizons and medium maturities

Key strands in term structure literature I

- Parametric curve fitting (e.g. Nelson–Siegel, Svensson)
  - functional form for yields across maturities
  - small number of interpretable factors
- Non-parametric/flexible fitting
  - avoids fixed functional forms (splines, kernels, or machine learning)
  - fit the cross-section of yields as flexibly as possible
- Forecasting: predict future yields
  - RW, Nelson-Siegel
  - Time-varying parameters, machine learning, neural networks

Key strands in term structure literature II

- No-arbitrage asset pricing models: satisfy absence of arbitrage across maturities
  - link yields to risk factors
  - decompose yields into expectations and risk premia
- ZLB: address constraint
  - shadow-rate models
  - ZARG
- Stochastic volatility models
  - time-varying uncertainty and volatility clustering
  - risk management
- Spanned vs. unspanned factors: what information is captured by the yield curve?

Better understand the role of key ingredients

Factors

- Important identification restrictions added to preserve L, S, C
- What is the exact added value in terms of predictive power?
- These factors summarise and isolate segments of the yield curve but bear little economic interpretation

Factor loadings

- How important is the time-variation in loadings?
- Are shallow neural networks the optimal way to introduce time variation?
- Large variation can also be harmful (Figure 7)

Your hybrid setup might be even richer than presented

Beyond forecasting

- Model outperforms on average but also uniformly across maturities
- This might suggest little tension to incorporate no-arbitrage restrictions
- No-arbitrage implies coherence across maturities
- Possible extension: asset pricing
- Offering more fundamental economic interpretation (expectations and risk premiums)

Exploit full potential of model

- Data: starting date (1961 vs 1985) and frequency (daily vs monthly)
- Fit of ZLB period and performance of implied time to lift-off
- Capture stylized facts over time: ST rates more volatile than LT rates (pre/post ZLB)
- Inclusion of survey data (known to help predictive ability)

Overall, great contribution!

- Results speak for themselves
- Better understand the source of your success (NS; TVP)