Examining the post-Covid inflation surge with a dynamic nonlinear signal weighting algorithm (work in progress)*

Douglas K.G. Araujo[1]

[1]Bank for International Settlements, douglas.araujo@bis.org

Abstract

This paper introduces an nonlinear algorithm that estimates the time-varying importance of the signal of each covariate to minimise an estimation or forecast loss. Using a large and varied macro-financial dataset, I estimate the contribution of each variable to U.S. inflation. Most signal comes from lagged prices, followed by interest rates and foreign exchange rates. Periods of high inflation tend to be better explained by signals from interest rates and exchange rates and especially by an increase in the weights related to labour market. However, even this flexible model does not entirely represent post-Covid high inflation. A quantile version of the model shows that this period is likely explained by tail outcomes. Looking at the weights for each quantile, left and right tails tend to be explained by different variable patterns. Median inflation depends more on lagged prices than other groups, while tail outcomes are also assocaited with variables related to interest rates and exchange rates, and labour markets. The importances of labour market variables and more generally of a broad range of variables increase in the right quantile during high inflation periods.

1 Introduction

Which variables contain the most economic signal to explain inflation outcomes? Do the signal contributions of each variable vary over time? And do different sets of variables explain different quantiles of the distribution of inflation outcomes? This paper explores these questions empirically for U.S. inflation, introducing a flexible time series tool that estimates time-varying importance weights for each covariate. I estimate this model with a large, long-term dataset and use it to examine more closely the recent post-Covid high inflation period. In short, I find that this period is associated with different variables compared to normal inflation periods, behaving instead like the highest quantiles of inflation.

Using the monthly FRED-MD dataset (McCracken and Ng (2016)) as a source of comprehensive macrofinancial data for the U.S., I find that most of the economic signal to explain inflation outcomes is related to lagged inflation metrics and interest rates. Remarkably, the 1970s high inflation episode is associated with a dramatic change in the variables carrying more economic signal, but these weights have largely stabilised since then. The same pattern holds generally across quantiles, although the important variables are different: higher inflation outcomes are usually associated with a richer set of variables compared to the left and centre of the distribution.

This breakdown is possible due to a novel way (to the best of my knowledge) to parameterize econometric time series models. The core methodological contribution is a recurrent neural network (RNN) architecture set up to minimise its loss function by competitively weighing its normalized input variates at each time step. The RNN acts as a flexible, non-linear stable function that maps each time step's inputs and past information

*Preliminary, not for distribution. This work represents the views of the author and not necessarily represent that of the Bank for International Settlements.

to that period's inflation observation. The model is strongly regularized by a assigning non-negative weights to covariates in each period, with weights summing to one. The weights are themselves a function of the covariates and of past weights. A completely novel aspect is that the model also optimises a global stringency parameter that determines how sharp to differentiate salient signals.

In addition to the benefits from regularization, the model architecture introduced in this paper directly addresses one of the biggest known limitations of neural network methods: their "black box" nature. In fact, once the model is estimated, the input series can be used to reveal a panel dataset of weights, where the weights for each variable vary over time, summing to one at each period. Another magnitude of potential interest is the signal competition stringency parameter, referred to as "temperature" in the computer science literature. Understanding this parameter might be easier with an analogy to Bayesian models; in such case, the temperature could be a hyperparameter governing how tight the prior is that only a few variables in a large dataset contribute signal to estimate the outcome. While by itself the parameter might not mean much, its usefulness is revealed when the same data estimates a quantile regression. For example, in the current application I show that inflation outcomes in the left and center of the inflation distribution prefer to extract signal from fewer variabels, while the signal for higher inflation outcomes comes from a somewhat larger set of variables.

This dynamic nonlinear variable selection model is applicable widely in economics and finance. While the current application estimates the contribution of each variable to model a contemporaneous value, but other applications can use the same model in forecasting settings (Aquilina et al. (forthcoming)).

The core idea of this paper is to use neural networks' recognised empirical performance in modelling complex data generating processes to address their own biggest shortcoming, lack of transparency. For all their flexibility to learn complex relationships in data, neural networks are often castigated for being inescrutable "black boxes". But as I show in this paper, it is possible to set them up in ways that reveal which of the myriad observed economic signals in covariates is actually relevant to explain the dynamics of inflation.

This network requires surprisingly few building blocks. Consider first a comprehensive m-dimensional vector of macroeconomic data, $x_t \in \mathbb{R}^m$. Define G_π to be a nonlinear map from x_t and its past realisations to an outcome variable. Throughout this paper, this will be inflation, defined here as the marginal change in the relevant price index, $\pi_t = (\Pi_t - \Pi_{t-1})/\Pi_{t-1}$. Formally, $G_\pi : \mathbb{R}^m \times \cdots \times \mathbb{R}^m \to \mathbb{R}$. Normally this map is estimated with a range of econometric techniques, such as linear regressions, factor models, lasso, random forests, etc. Instead of taking this direct route, I take a detour. Define a second map, G_Ω, that spans the same domain but has a range of an m-dimensional simplex, as in $G_\Omega : \mathbb{R}^m \times \cdots \times \mathbb{R}^m \to \Delta^{m-1}$, with $\Delta^{m-1} = \{\omega \in \mathbb{R}^m | \omega_i \geq 0 \forall i \in \{1, 2, ..., m\}, \sum_{i=1}^m \omega_i = 1\}$. Instead of using computational methods to estimate $\hat{G}_\pi$ directly, I focus instead on a version that maps *competitively weighted* vectors to inflation, $G_{\tilde{\pi}} : \Delta^{m-1} \odot \mathbb{R}^m \times \cdots \times \Delta^{m-1} \odot \mathbb{R}^m \to \mathbb{R}$, with $\tilde{x}_t = G_\omega(x_t, ..., x_0) \odot x_t$. By leveraging the battle-tested abilities of gradient descent to estimate high-dimensional models where the number of parameters greatly surpasses that of the data, I estimate the system $G = \{G_{\tilde{\pi}}, G_\Omega\}$ jointly in a consistent way.

The simplex is obtained by softmax. This function can also be parameterized with the addition of a so-called "temperature" parameter, $\lambda > 0$, like so: $\text{Softmax}(x_i; \lambda) = \frac{e^{x_i/\lambda}}{\sum_{j=1}^{|x|} e^{x_j/\lambda}}$. Higher values of λ flatten the simplex, while lower values make each small difference more salient. Simply using the Softmax formula without parameterizing the temperature is equivalent to setting $\lambda = 1$.

However, there is no a priori economic reason to dogmatically assume that (a) the true λ is 1, or (b) the same amount of simplex smoothness applies to all inflation quantiles, ie $\lambda_{q=5\%}$ might very well be different from $\lambda_{q=50\%}$. In this section I describe two variations of the algorithm, one where the network learns the optimal λ that applies throughout the time series, and another where the network learns λ_t from the data (ie, from a map $G_\lambda : \mathbb{R}^m \times \cdots \times \mathbb{R}^m \to \mathbb{R}^+$).

These time-varying weights on each covariate's signal are related to the concept of "attention" in the computer

science literature. Perhaps more widely known in economics due to its association with the transformer neural network architecture (Vaswani (2017)), "attention" is at its essence the attribution of weights to certain data points in a way that is, itself, data-dependent. While the vast majority of attention-based applications are in univariate time series (for example, of words), some applications also attribute attention to each covariate. Arik and Pfister (2021) introduce TabNet. Somepalli et al. (2021) introduce SAINT with attentions both over time and across the covariate space. In this paper, in contrast, the weights (which can also be interpreted as attention) are assigned at the cross-section of covariates, and the temperature (λ) itself responds to the data, helping the model better decide whether to be overall restrictive in how wide the weights (or attention) are spread out.

This variable selection also results in a strong and dynamic regularization, which is another desirable feature of empirical models (Giannone, Lenza, and Primiceri (2021)). Unlike other widely-used regularization methods, such as random forests, the class of models described in this paper adapts the degree of regularization dynamically.

To estimate this model, I use the FRED-MD monthly dataset compiled by McCracken and Ng (2016). This large dataset consists essentially of curated series that together represent a broad overview of economic activity and financial conditions in the United States.

As much as this analysis can bring up interesting features of the data, its limitations should also be recognised explicitly. First, the measure of variable importance requires specific interpretation. For example, is not a statement on the influence of each variable on inflation, nor a decomposition of forecast errors. Rather, these weights correspond to the relative learned priority to explain current inflation, while also optimising the ability to explain all subsequent inflation given subsequent data (due to how the model uses today's data to construct a long memory chain). Another drawback is that these decompositions are not exactly estimated, but must be learned. This is, of course, counteracted with a well-designed cross-validation system to prevent overfit on the latter periods to be studied.

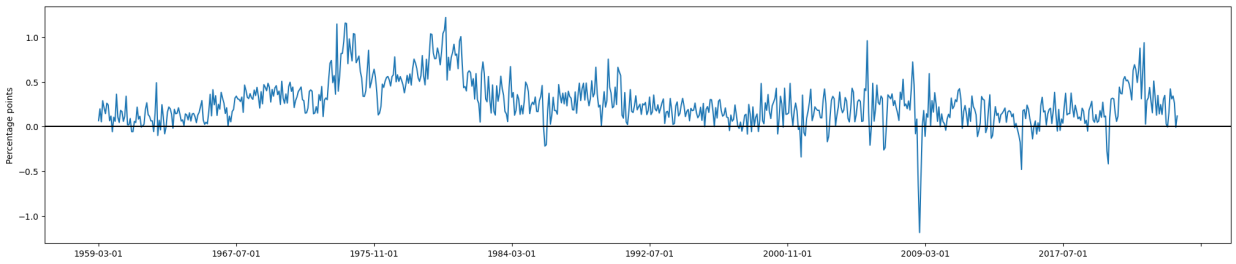
This paper builds on an extensive prior literature that aims to pin down determinants of inflation, in particular, the search for determinants of the 2021-2022 high inflation produced a lively soul-searching in the literature. Bernanke and Blanchard (2023) study a model that compares wages and prices, given labour slack and output, and find that while labour pressures did contribute to increase the overall level of inflation, the bulk of the overshoot can be attributed to disruptions in goods markets leading to disbalances in prices and thus also in supply of goods. (See also Bernanke and Blanchard (2024).) Focusing on the recent episode of high inflation, Reis (2022) credits policymakers' inability to combat the inflation surge on errors of perception or judgement by policymakers (nascent high inflation to factors such as the infrenquent nature (and scale) of the shocks, deterioration in the distribution of inflation expectations, and an increased determination to fight low inflation rather than prepare for high inflation.) The quantitative procedures that I introduce in this paper are meant to complement such structural analyses, offering a new perspective. In particular, the quantile inflation variable explainability is associated with other work, such as inflation at risk (Banerjee, Mehrotra, and Zampolli (2020), Lopez-Salido and Loria (2024)). Gründler (2023) estimate a switching-point VAR and find that the transmission of inflation expectations to price increases is stronger after Covid, but not sufficiently different to constitute a different regime. While my goal is not to forecast inflation but to explain its contemporaneous dynamics, it is useful to compare these findings with other works in the literature. Using other types of nonlinear models, namely random forests, Kohlscheen (2023) finds that measures of inflation expectations are key contributors to forecasting performance.

2 Empirical application

In this exercise, I focus on the personal consumption expenditures price index ("PCEPI" henceforth). The inflation metric π_t is the first difference of log PCEPI (Figure 1). The goal is to estimate G to map a range of macroeconomic data (the FRED-MD dataset, transformed as suggested by McCracken and Ng (2016)) up

to time t to π_t. In other words, I do not seek to *forecast* inflation but rather to uncover the relevance of different economic signals to inflation movements.

Figure 1: Personal consumption expenditure monthly inflation



This dataset includes variables on output and income; labour market; consumption and orders; orders and inventories; money and credit; interest rate and exchange rates; the stock market; and prices. The only change to the dataset that I implement manually is that the price variables (listed in Group 7 of McCracken and Ng (2016)) are lagged, except for crude oil price, which is retained for the same date. This avoids making the model trivially assign a contemporaneous value of inflation to movements in its subcomponents, while maintaining a potentially rich inertia dynamics, where the aggregate inflation responds differently to lagged components.

To find the configuration that best captures economically-relevant signal for inflation from this dataset, I slice the dataset into different disjoint periods. First, I reserve all observations after end-2019 (59 monthly observations) to a testing set that is never used to calibrate the parameters of the estimates. This enables us to check empirically the quality of the model with a large and economically interesting period of pseudo out-of-sample data, a useful prior step to interpreting the variable weights. The remaining periods, from April 1959 to December 2019 (729 monthly observations), are used in the training process with cross-validation as follows. First I define the number of broadly equal-sized disjoint training periods, F. For each $f_j \in \{1, \dots, F-1\}$, I use the periods in $f_1, \dots, f_j$ to train the algorithms and the data in f_{j+1} to calculate the loss as a form of validation data. The model configuration with the lower mean loss across folds $f_2, \dots, f_F$ is the preferred configuration.

I use $F = 4$, obtaining the following periods in Table 1. In the first fold, the model optimises forecasting performance using data from April 1959 to August 1971, when the average monthly PCE inflation was circa 0.21 percentage points (p.p.; the standard deviation was 0.15 p.p.); then, the performance of this version of the model is tested against data ranging from September 1971 to September 1983, a period broadly corresponding to the high inflation of the 1970s when monthly inflation averaged 0.57 p.p. (standard deviation: 0.25 p.p.). Then the same neural network configuration is trained with all data up until September 1983 and validated against October 1983 to October 1995, with a more normalized inflation environment. In the next fold, the model trains using data up to October 1995, a validates its predictions in the low-inflation period that includes both the Dot-Com bubble and the lead-up to the Global Financial Crisis, namely November 1995 to November 2007. Finally, the last fold tests predictions of models that train using all data up to November 2007, validating those predictions during December 2007 to December 2019, broadly corresponding to the post-Global Financial Crisis period that includes unconventional monetary policies. Note that each fold's validation period broadly covers very different periods, each broadly with 12 years.

Table 1: Periods for each training fold

Table 1

Fold	Training dates	Training, avg.	Training, s.d.	Validation dates	Validation, avg.	Validation, s.d.
1	1959-04 to 1971-08	0.209088	0.149314	1971-09 to 1983-09	0.566871	0.253429
2	1959-04 to 1983-09	0.385546	0.273737	1983-10 to 1995-10	0.253754	0.152107
3	1959-04 to 1995-10	0.342015	0.248159	1995-11 to 2007-11	0.174363	0.174598
4	1959-04 to 2007-11	0.300389	0.243004	2007-12 to 2019-12	0.118961	0.206772

The best model configuration in this cross-validation exercise is then used to train the model using all data up to December 2019, yielding the best estimate $\hat{G}$ within the parameter search space. Intuitively, the best configuration amongst those tested is the one that is better able to generalize the relationships between the input data and their histories and inflation. Assuming this has occured, the estimate $\hat{G}$ will provide the best (pseudo) out-of-sample breakdown of the high inflation period. Retaining the period of interest out of the training samle is crucial to avoid misinterpreting weights that are assigned due to overfitting rather than to (what might be) a true generalization of the relationships between data.

The true π_t and the $\hat{\pi}_t$ predicted using RNN-DNVS trained up until end-2019 are plotted in Figure 2.

Figure 2: Predicted inflation

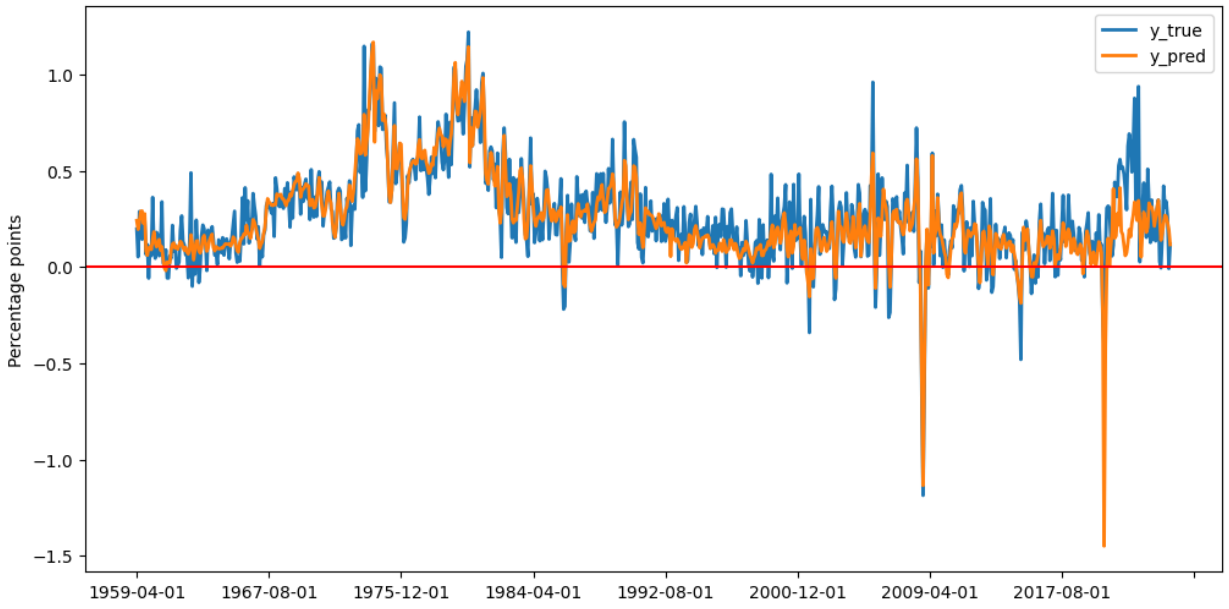
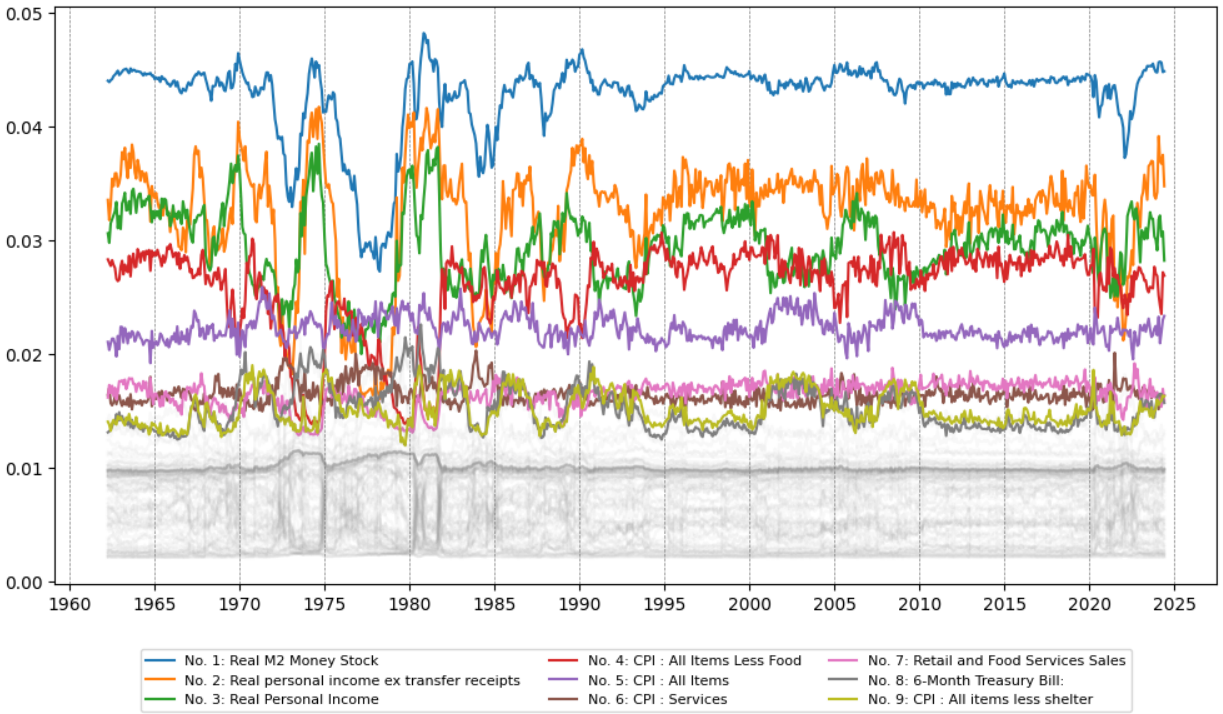


Figure 3: Top variable weights for inflation



The variables with the highest weight is the real growth in M2 money, followed by real personal income growth variables. The next highest weighted variables are lagged CPI inflation measures, then retail sales, 6-month Treasury yields and another component of CPI. All other variables in Figure 3 are plotted in light gray to convey their general movement while not focusing on lower-weight variables (note the thicker gray line corresponding to a number of variables that are concentrating around those weights).

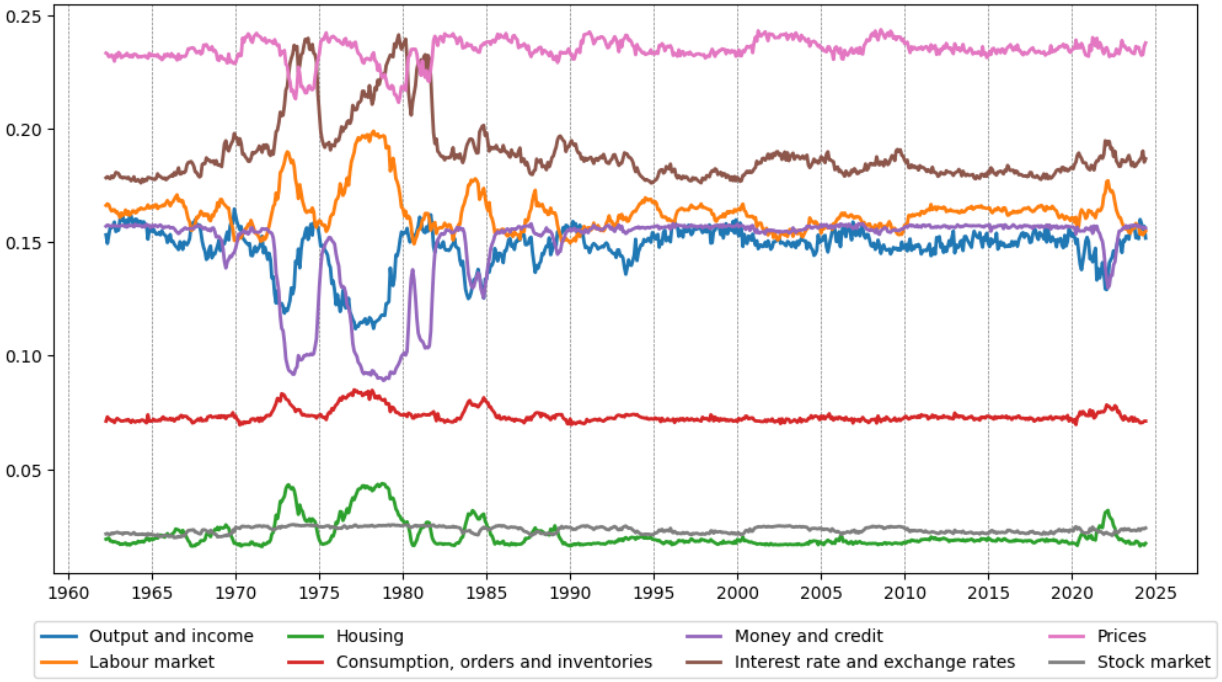
Instead of analysing the variation of weights over time for each of the top-weighted variables, it is perhaps more instructive to note general patterns. First, there is considerable volatility in the weights in the 1970s inflation episode, followed by a decades-long stability in the weights. The recent inflation also appears to be associated with a historically significant reshift in the weights, although not at the same magnitude as during the 1970s. What is significant here is that this change in the weights for the recent period is identified in a pseudo out-of-sample manner, that is, the model was only estimated with data prior to the Covid period, and still it was able to model a redistribution of the weights between variables.

Taking an even broader view, Figure 4 presents the weights of similar variables summed together, using the grouping of McCracken and Ng (2016). It is now easier to see that the most relevant signals are present in past values of inflation (pink line).[1] During the 1970s inflation, this variable group was somewhat less relevant. The second most important group of variables are the measures of interest rates and exchange rates (brown line); during the 1970s, and now at a smaller scale, these variables seem to contain more signal to explain contemporaneous values in inflation. Next, variables related to labour markets (orange line) follow a similar pattern as the previous group, with remarkable stability except for marked increases in importance during the 1970s high inflation and a more attenuated increas in the recent bout of inflation. On the other end of the relevant spectrum, variables related to housing and stock prices contribute much less signal to

[1]Note that the model does not know in advance which variables are related to one another; the stationary time series of each one are the only covariates used in the model.

explain contemporaneous inflation.

Figure 4: Sum of weights for variables by group



An interesting feature of this model is that while it is estimated on a multi-decade period and was cross-validated across different inflation periods, it is not really able to completely reconstruct inflation during the post-pandemic period. This suggests that beyond flexible selecting the signal that matters most for the mean inflation, that period is associated with extreme tails of the distribution of inflation. The next model addresses this observation, estimating quantile regression versions of the model seen above.

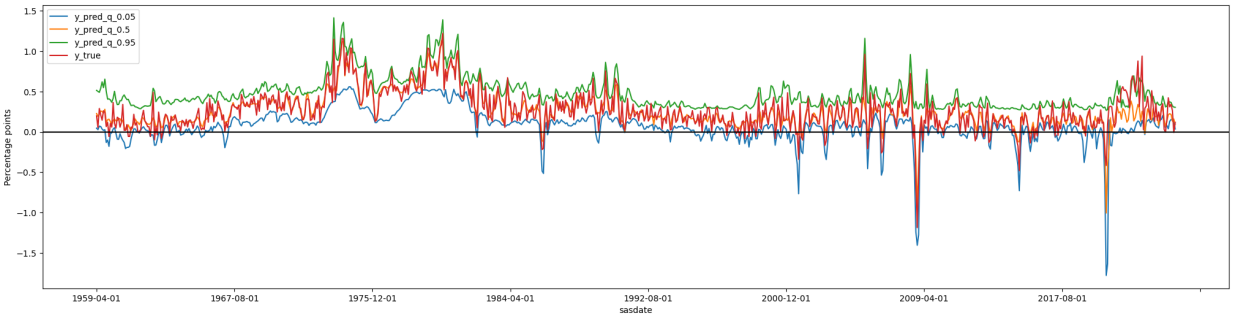
3 Quantile models

The DNVS model is easily adaptable to quantile regressions: the output associated with each quantile is weighted by a separate set of weights.[2] This enables estimating answers to questions such as: Are the variable weights the same across quantiles? Are they similarly stable? Given the remarkable forecast errors in the model above, this section explores a quantile regression. To keep the model simple, the only quantiles studied are $q = \{0.05, 0.50, 0.95\}$, however, any relevant combination of quantiles is valid (and more quantiles would probably help the model's performance).

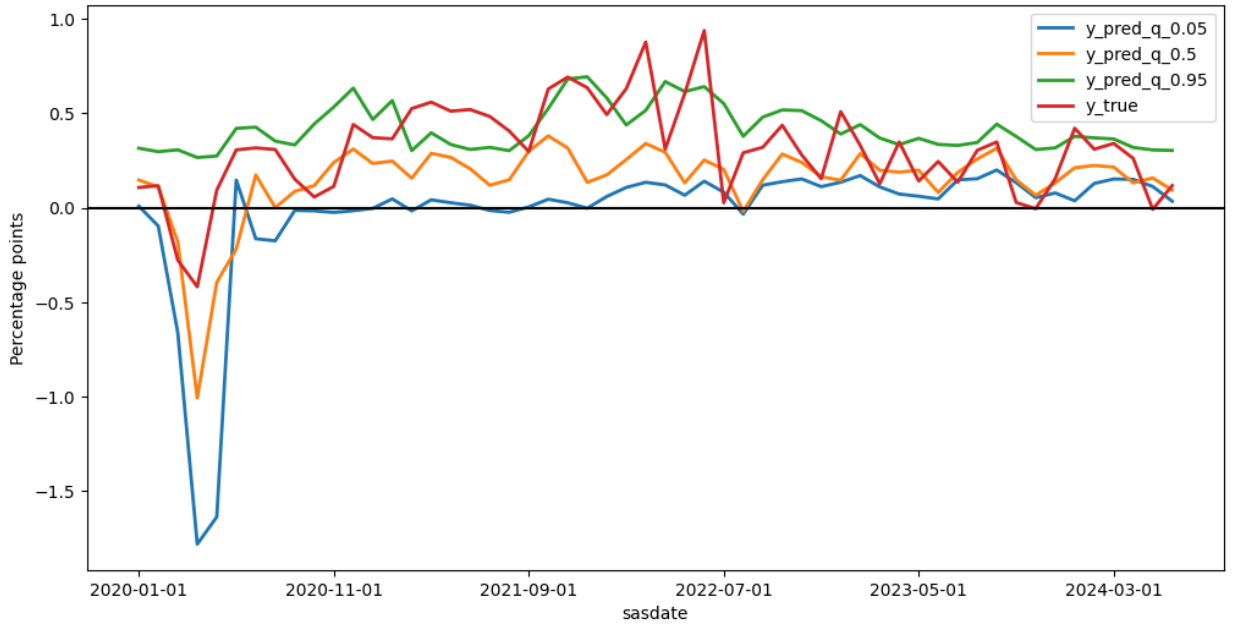
[2]The weights can be estimated separately directly from the data or from a common transformation applied to the data. In the current implementation, I use the latter option: the weights for each quantile are all estimated from features extracted jointly from the covariates, which helps the model further "learn" the best rules to extract relevant information from the data.

Figure 5: Quantile prediction and true inflation

(a) All periods



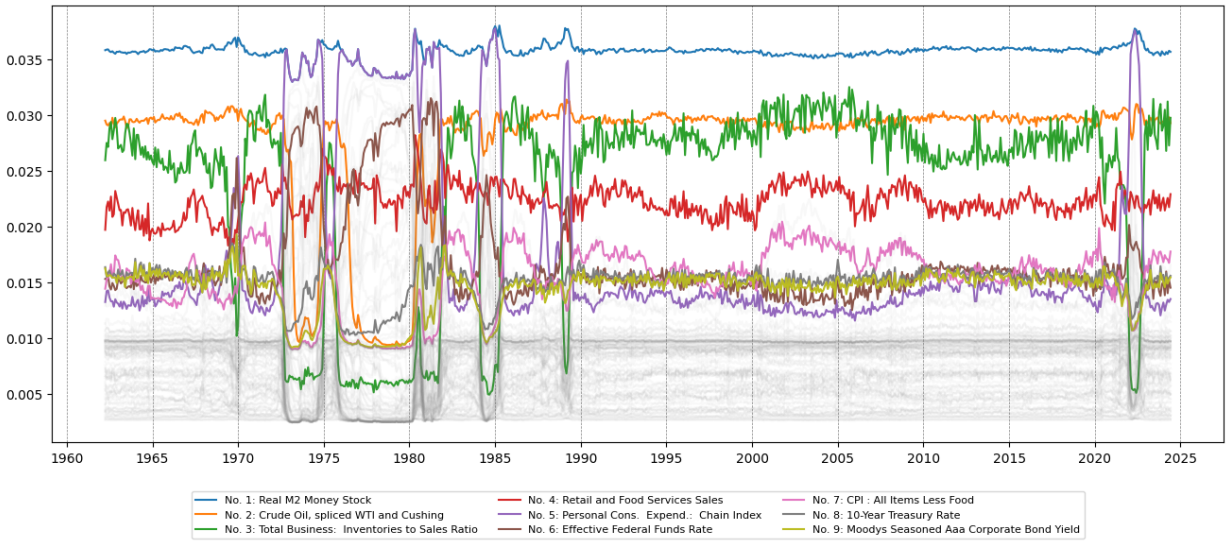
(b) Out-of-sample period only



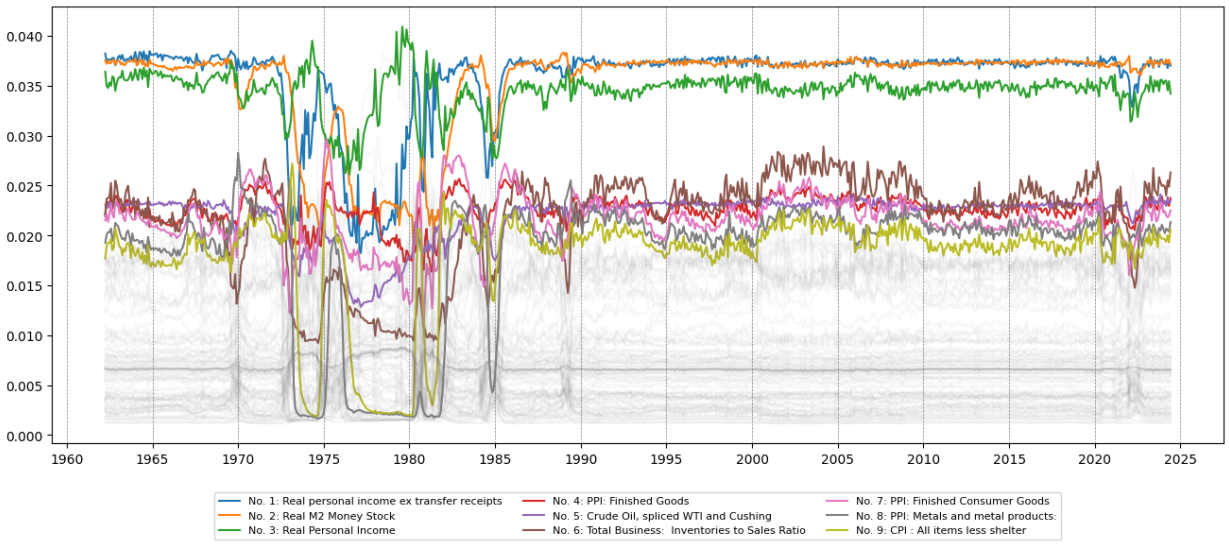
As posited above, the post-Covid high inflation period is closer to the tails than to the median prediction. Note that these quantile metrics are exclusively out-of-sample, based on parameters estimated up until end-2019. The parameters to set up the RNN were the same as defined for the mean inflation value using cross-validation, to avoid overly optimizing for the quantiles. Still, the quantiles over time appear indeed to represent the dynamics of inflation quite well.

Figure 6: Variable weights by inflation quantile

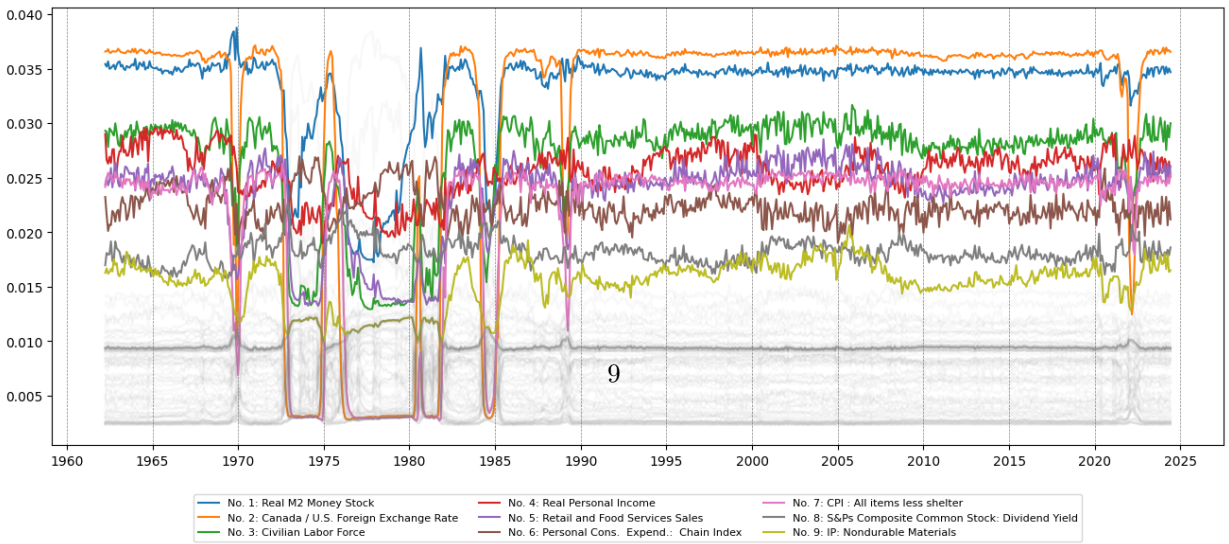
(a) Quantile: 0.05



(b) Quantile: 0.50



(c) Quantile: 0.95



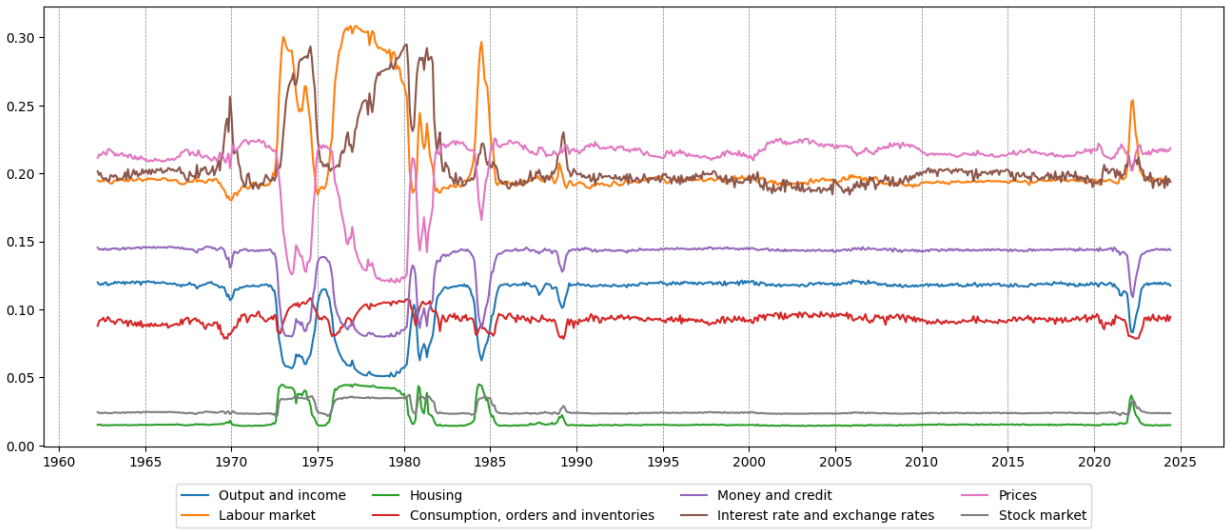
These different variable weight distributions also reflect how stringent the competition is between variables to present signals for estimating inflation quantiles. Namely, inspecting the learned parameter λ by quantile reveals two interesting facts. To estimate the 5% quantile, the model uses $\lambda_{q=0.05} = 0.765$, and the corresponding values for the other quantiles are $\lambda_{q=0.05} = 0.563$ and $\lambda_{q=0.05} = 0.718$.[3] First, all three values are below one, suggesting that for the dimensionality of the covariates in this model, inflation appears to be prefer signal from fewer rather than more variables. Second, the estimation of the median relies on a more concentrated signal pattern due to the lower λ value than the tails, which appear to be relatively more "agnostic", with a wider aperture to search signals. One potential reason for this could be that inflation tail outcomes might depend on a broader set of conjunctural variables that might not necessarily be the same every time, so the nonlinear map learned by the neural network performs better if it retains signal from a broader array of variables. Indeed, the actual weights for both quantile extremes suggest support to this interpretation, since weights of the top variables effectively go down sharply during the high inflation period. Offsetting this, the group of lowly-weighted variables rises in conjunction (note the thick gray line close to 1% weight), indicating the model captures signal from a much broader array of variables.

Having said that, grouping the variables according to topic washes out some of this remarkable variation in high inflation periods seen at the tails (Figure 7). One explanation - to be explored - could be that these periods see substantial *intra-group* variation but not so much between groups. In any case, this grouping enables a comparison between the different quantiles. Both 5% and 95% quantiles see a prominent share of the weights assigned to "Phillips Curve" variables: past prices, labour market and interst rates and exchange rates. The model estimating median inflation, on the other hand, receives signal primarily from past inflation values. Another observation is the role of labour market variables as a tell-tale of more unhinged inflation processes: signals from labour market variables become noticeably more important both in the 1970s and the post-Covid high inflation periods, although at different scales.

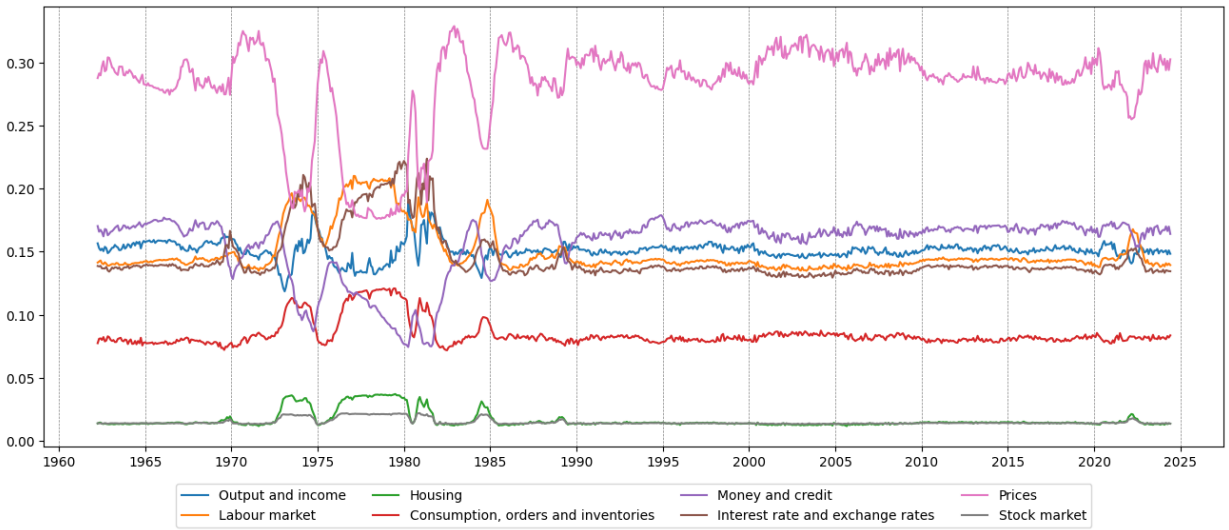
[3]The estimated temperature for the mean inflation model shown above is 0.637.

Figure 7: Weights by group for each quantile

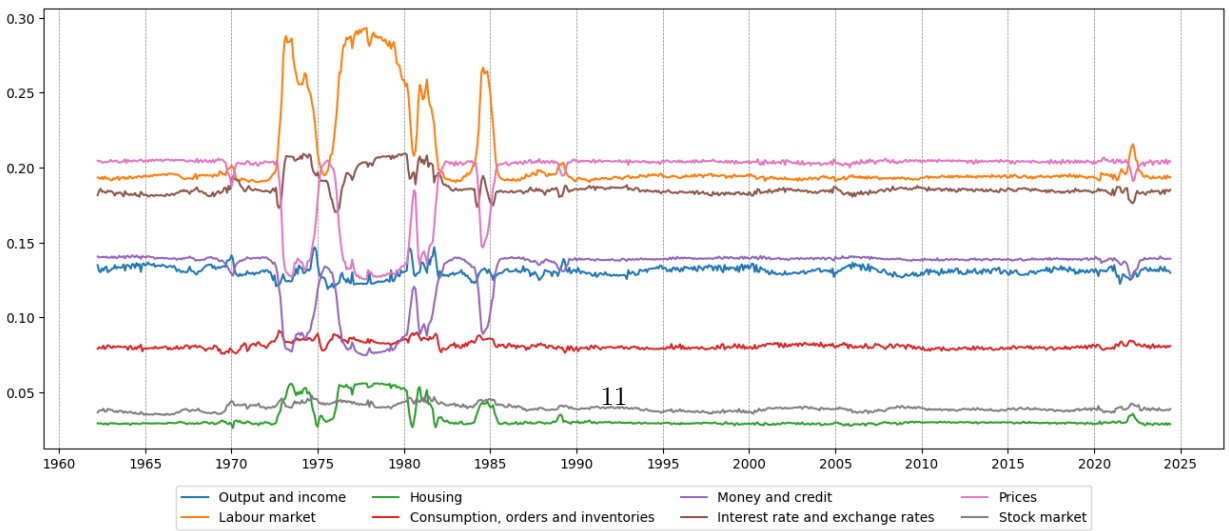
(a) Quantile: 0.05



(b) Quantile: 0.50



(c) Quantile: 0.95

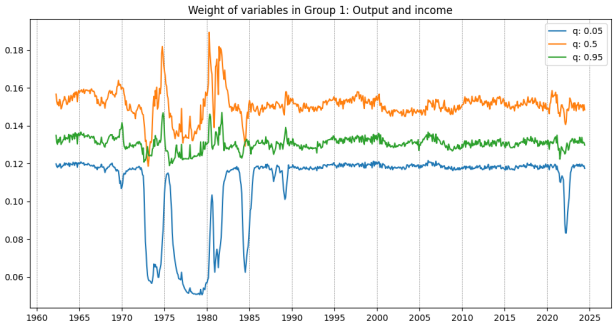


Comparing the different groups is easier with the plots in Figure 8. The first pannel shows the summed weights for variables related to output and income. In the recent post-Covid inflation, the most marked movement was their lower relevant to explain the left tail of inflation. The second panel shows the dynamic weights for variables related to labour market, showing more clearly the findings described above (note also the scale, reaching almost a third of the signal estimating inflation). The third panel, showing the weights related to housing, shows an interesting feature whereby the signals from these variables are generally "dormant" but rise up in importance, especially at the quantiles, during high inflation periods. The fourth panel, related to consumption, order and inventories, is interesting also to inspect because of the evidence of supply chain disruptions in the post-Covid inflation process. While its weight falls for the lower quantile, it retains its importance in the median and higher tail, meaning that when these variables reflected the strains from global supply chains, their signal was receiving approximately the same weight as before for these quantiles.

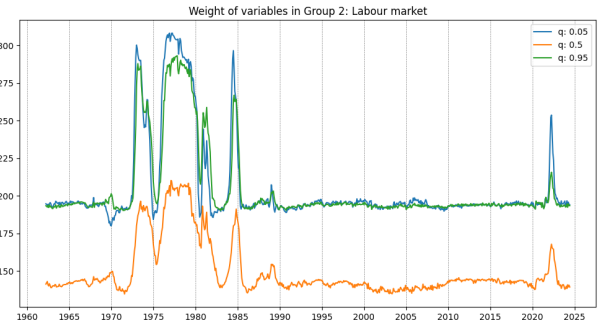
Continuing in Figure 8, the fifth panel shows the weights for money and credit, which are highest and more varying for the median than for the tails. In high inflation periods, the relevance of money and credit drops sharply. The next panel shows the weights for the variables related to interest rates and exchange rates. Their importance are generally higher for the tails compared to the median. But in high inflation periods, these variables are even more important to the lowest quantile. The seventh panel are the lagged inflation components and the contemporaneous oil prices. These variables are more important as a whole to inflation median than to the tails, and this importance falls further during the acute high-inflation period. Finally, the last panel shows the weights related to stock market. Those are generally low, although curiously are higher overall for the 95% quantile.

Figure 8: Weights by quantile for each group

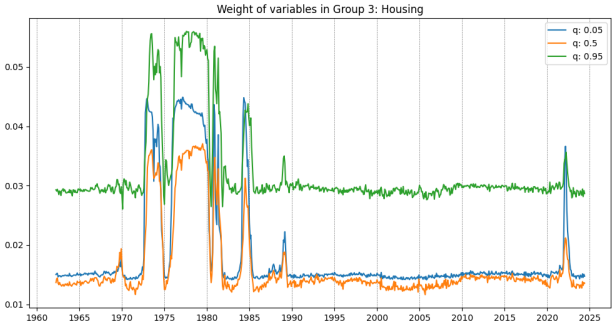
(a) Output and income



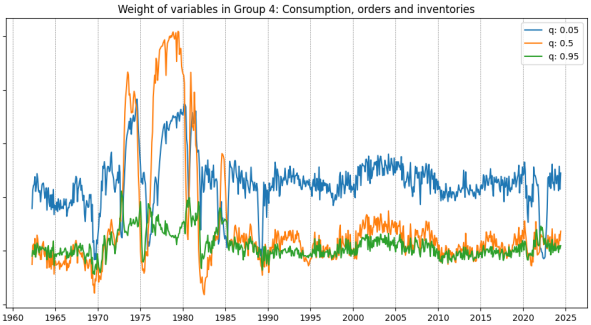
(b) Labour market



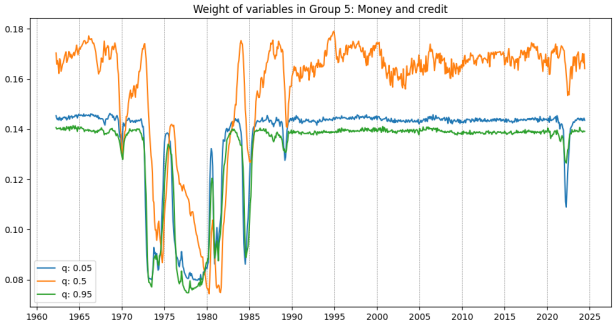
(c) Housing



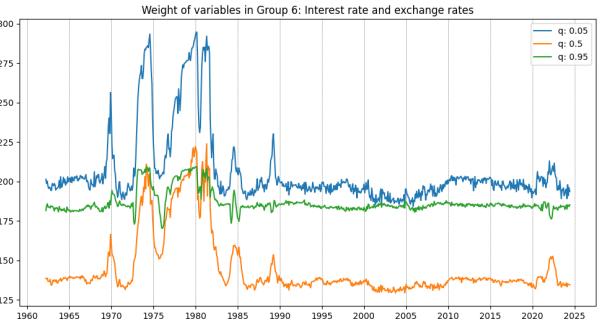
(d) Consumption, orders and inventories



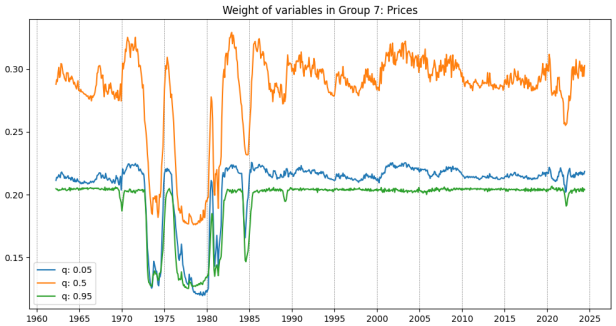
(e) Money and credit



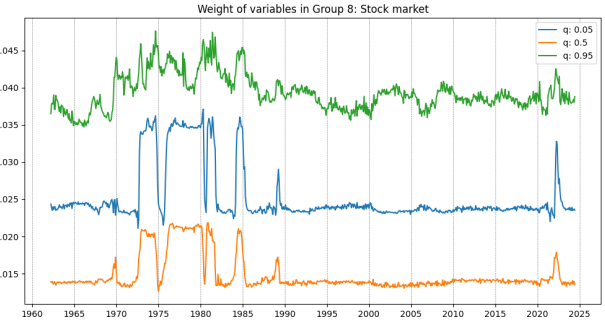
(f) Interest rate and exchange rates



(g) Prices



(h) Stock market



4 Preliminary conclusions

This paper contributes to the growing literature that seeks to understand the post-Covid high inflation process, taking a complemetary stance to other investigations that rely more on structure and stronger functional form assumptions. Rather, these results can serve as an almost "model-free" sanity check: structural models of high inflation bursts as the one experienced after Covid (or, for that matter, in the 1970s) should also preferably explain why inflation dynamics shift gears and starts "picking up more signal" from (or more broadly, correalting more with) other variables.

Beyond a potential role in the important post mortem work to understand the collective failure to forecast and prevent high inflation, such a breakdown of inflation signals can help economic analysis in real time. Economists at central banks, forecasting agencies and others might be interested in knowing the variables that are more associated with inflation readings, almost as soon as they become available.

Finally, just as this model successfully uncovers - by itself, without imposition of any prior structure by the econometrician - a prominent role for variables commonly associated with inflation à la Phillips curve, this type of models can be useful in other areas of economics and finance. For example, it can help detect the sources of forecasting signal; identifying which asset pricing theory tends to be more relevant at each point in time;[4] or identifying which macroeconomic variables tend to better explain monetary policy shocks, bank failures or other interesting economic and financial events that still eludes the profession.

Appendix

4.1 Details of the training process

All input data is sliced in a way that 36 consecutive months of covariates are used to predict a 36-month period of inflation. The models are set to train for up to 1,000 epochs, with early stopping if the validation performance, as judged by the mean squared error, does not improve for 15 epochs. At every 5-epoch window, whenever the validation performance does not improve, the optimizer learning rate is multiplied by a factor of 0.9.

References

Aquilina, Matteo, Douglas Araujo, Gaston Gelos, Taejin Park, and Fernando Pérez-Cruz. forthcoming. "Harnessing Artificial Intelligence to Monitor Financial Markets." BIS Working Papers. Bank for International Settlements (BIS).

Arik, Sercan Ö, and Tomas Pfister. 2021. "Tabnet: Attentive Interpretable Tabular Learning." In *Proceedings of the AAAI Conference on Artificial Intelligence*, 35:6679–87. 8.

Banerjee, Ryan Niladri, Aaron Mehrotra, and Fabrizio Zampolli. 2020. "Inflation at Risk from Covid-19." Bank for International Settlements.

Bernanke, Ben, and Olivier Blanchard. 2023. "What Caused the US Pandemic-Era Inflation?" *Peterson Institute for International Economics Working Paper*, no. 23-4.

———. 2024. "An Analysis of Pandemic-Era Inflation in 11 Economies." *Peterson Institute for International Economics Working Paper*, 24–11.

Giannone, Domenico, Michele Lenza, and Giorgio E Primiceri. 2021. "Economic Predictions with Big Data: The Illusion of Sparsity." *Econometrica* 89 (5): 2409–37.

Gründler, Daniel. 2023. "Expectations, Structural Breaks and the Recent Surge in Inflation." *Economics Letters* 233: 111394. https://doi.org/https://doi.org/10.1016/j.econlet.2023.111394.

Kohlscheen, Emanuel. 2023. "What Does Machine Learning Say about the Drivers of Inflation?" arXiv. org.

[4] I thank Fred Liu for thinking of this example.

Lopez-Salido, David, and Francesca Loria. 2024. "Inflation at Risk." *Journal of Monetary Economics*, 103570.

McCracken, Michael W, and Serena Ng. 2016. "FRED-MD: A Monthly Database for Macroeconomic Research." *Journal of Business & Economic Statistics* 34 (4): 574–89.

Reis, Ricardo. 2022. *The Burst of High Inflation in 2021-22: How and Why Did We Get Here?* Bank for International Settlements, Monetary; Economic Department.

Somepalli, Gowthami, Micah Goldblum, Avi Schwarzschild, C Bayan Bruss, and Tom Goldstein. 2021. "Saint: Improved Neural Networks for Tabular Data via Row Attention and Contrastive Pre-Training." *arXiv Preprint arXiv:2106.01342*.

Vaswani, A. 2017. "Attention Is All You Need." *Advances in Neural Information Processing Systems*.