



HAL
open science

Inflation dynamics : policies and determinants

Emanuele Franceschi

► **To cite this version:**

Emanuele Franceschi. Inflation dynamics : policies and determinants. Economics and Finance. Université Panthéon-Sorbonne - Paris I, 2021. English. NNT : 2021PA01E039 . tel-03507761

HAL Id: tel-03507761

<https://theses.hal.science/tel-03507761>

Submitted on 13 Jan 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Emanuele Franceschi

**Inflation Dynamics:
Policies and Determinants**

In partial fulfilment of the requirements for the degree of
Doctor of Philosophy in Economics

at

Paris School of Economics
Paris 1 Panthéon-Sorbonne University
Paris-Jourdan Sciences Économiques

Supervised by

- Fabrizio Coricelli - University of Siena and Paris School of Economics

Thesis Committee:

- Jean-Bernard Chatelain, Paris 1 Panthéon-Sorbonne University and Paris School of Economics
- Anne-Laure Delatte, CNRS and Dauphine University
- Roman Horváth, Institute of Economic Studies, Charles University – Referee
- Raoul Minetti, Michigan State University – Referee
- Chiara Osbat, European Central Bank



Université Paris-1 Panthéon-Sorbonne, UFR d'Économie
Laboratoire de rattachement : Paris-Jourdan Sciences Économiques
Thèse pour l'obtention du titre de Docteur en Sciences Économiques
Présentée et soutenue publiquement le 28 juin 2021 par

Emanuele Franceschi

Dynamique de l'Inflation : Déterminants et Politiques

Directeur de thèse :

- M. Fabrizio Coricelli - Université de Sienne et École d'Économie de Paris

Composition du jury :

- M. Jean-Bernard Chatelain, Professeur, Université Paris 1 Panthéon-Sorbonne et École d'Économie de Paris
- Mme Anne-Laure Delatte, CNRS et Université Dauphine
- M. Roman Horváth, Professeur, Institute of Economic Studies, Charles University – Rapporteur
- M. Raoul Minetti, Professeur, Michigan State University – Rapporteur
- Mme Chiara Osbat, Banque Centrale Européenne

Contents

1	General Introduction	1
1.1	Monetary macroeconomic theory	4
1.2	Empirical Monetary Macroeconomics	6
1.2.1	Monetary policy rules	7
1.2.2	Inflation dynamics	13
2	Introduction Générale	20
3	Taylor Rules and Liquidity in Financial Markets	41
3.1	Data	45
3.2	Empirical Results	49
3.2.1	Specifications	50
3.2.2	Full sample analysis	51
3.2.3	Exogenous breaks	53
3.3	Diagnostics on Structural Breaks and Markov Switching	55
3.3.1	Markov Switching estimation	57
3.4	Conclusion	61
3.A	Data Sources and Transformations	63
3.B	Additional plots	65
3.C	Empirical Appendix	69
3.C.1	Output gap	69
3.C.2	Full sample regression: residuals	69
3.C.3	GMM estimates	70
3.C.4	CUSUM test plots	70
4	A simple model with liquidity	71
4.1	A model with liquidity	74
4.1.1	Consumer	74
4.1.2	Firms	76
4.1.3	Monetary authority and market clearing	77
4.1.4	Linearised model	79
4.2	Calibration and IRFs	81
4.2.1	Liquidity with an accommodative central bank	84
4.3	Dissecting simulated inflation dynamics	87
4.4	The 2008 crisis: severe liquidity shortage	91

4.5	Conclusion	93
4.A	Model appendix	95
4.A.1	Obtaining the system of equations	95
4.B	Loglinearisation	96
4.C	CIA setup	99
4.D	Sensitivity analyses	105
4.D.1	Semi-elasticities for bonds and cash	105
4.E	How γ affects inflation dynamics	109
4.E.1	Misspecification	111
4.F	Full lags structure	112
4.G	Extreme policies	113
5	Inflation Persistence	115
5.1	Data and tools	119
5.2	Autoregressive analyses	122
5.2.1	Optimal lags selection	124
5.3	A Bayesian estimation of inertia	127
5.4	RNN-LSTM approach to persistence	130
5.4.1	LSTM forecasting	132
5.4.2	LSTM setup	133
5.4.3	Full sample: forecasting inflation and its persistence	134
5.4.4	Regressions on LSTM	134
5.4.5	Rolling LSTMs	138
5.5	Conclusion	141
5.A	A primer on the RNN-LSTM framework	142
5.B	Year-on-year series	148
5.C	Isolating oil and commodities' inflation from headline	149
5.D	Robustness with varying window width	152
5.E	Draw distributions	156
5.F	LSTM data and forecasts	157
5.F.1	LSTM predictions on non-overlapping subsamples	157
5.G	LSTM analyses	159
5.G.1	OLS regressions on decades	159
5.H	Trend estimates	162
5.H.1	Frequentist application	163
5.H.2	LSTM output: full sample	164
5.H.3	LSTM output: subsample analysis	165
5.H.4	LSTM output: rolling window	166
6	Conclusion	167
6.1	Policy Rules and Liquidity	167
6.2	Optimal Liquidity Management	168
6.3	Inflation Persistence	169
6.3.1	Macroeconometrics with Deep-Learning	170
6.3.2	Inflation and Structural Change	171

Acknowledgements

This doctoral dissertation is a partial snapshot of several years of work. Despite bearing a single name, it benefited from countless discussions, experimentations, fruitful errors, rushes and periods of doubtful pondering. Research springs from dissatisfaction with the current state of affairs and ends with – often – barely satisfying results: models could be generalised, extended, or revisited; empirical evidence could be corroborated with more robustness checks, alternative data, more sophisticated methods.

This is in part what pushes researchers farther, and has been the case for these chapters. This depiction of the research endeavour might look superficially unattractive, but in its depths it is threaded with long silver linings of joy and accomplishment, infrequent yet utterly rewarding. In largest part, for me at least, these linings are the wonderful people I met over the years, who are worth every hardship and deserve credit for all the joys that punctuated my Ph.D. journey, both scientific and personal. Without them, I would have hardly anything to say or write. These lines pay a due but insufficient credit to them, whom I am privileged to have met along the way.

My first *grazie* goes to my supervisor, Fabrizio Coricelli, who since the Masters' continuously stimulated my thinking, posed rightful questions, supported my work and myself. His guidance and nurture helped me becoming the researcher I am today, and will influence me for many years – and papers – to come. His supervision has been fundamental in the development of this dissertation: I often found myself in blind corners or lost in intricate, ambitious plans. Fabrizio helped me stay on track and steered my efforts, bringing me back on more fruitful and reasonable paths. His economic intuition and sharp questioning gave me motivation during these years and still inspires me to engage in research. I definitely learnt a lot, yet not enough.

I also want to thank Jean-Bernard Chatelain, whose suggestions, especially at the very beginning of my work, greatly improved the quality of these chapters. He was instrumental in the submission of the first chapter, I wish to thank him for the time he took to suggest modifications and improvements over the first, oversized version. His assistance during my Ph.D. journey eased many of the typical obstacles candidates encounter.

I want to thank the members of my thesis committee. I am grateful to Raoul Minetti and Roman Horváth for providing suggestions, comments, and help in improving the earlier version of these chapters and for accepting to referee my dissertation. Their suggestions greatly

improved the quality of the chapters and sparked the reworking and extension of some parts. Roman Horvát's comments pushed me to think again on some of the results of the first chapter, which eventually took me to start working on a whole follow-up paper, which complements and improves its predecessor. Raoul Minetti provided insightful suggestions on the second and third chapters: in particular, his comments are behind most of the second chapter's appendix on sensitivity and LIA form of the main model. Also, I thank him for the support he gave me during the last, unpredictable job market. Anne-Laure Delatte and Chiara Osbat joined recently the committee and I wish to thank them, too. RAing for Anne-Laure Delatte and Jean Imbs' project on the effects of ECB's non-conventional policies gave me the opportunity to learn a great deal on bank-level data and how to exploit them effectively in research.

Also, I greatly benefited from endless and constant support from PSE and Paris 1 administrative staff. In particular, Veronique Guillotin, Julia Piquemal, Jean-Christophe Fouillé, Loic Sorel, and more recently Christelle Gauvrit helped me immensely at every administrative step, therefore I wish to thank them for their support, beyond their official duties. I also am grateful to Sylvie Lambert, head of the doctoral program at PSE, for her comprehension and moral support during some phases of this journey.

I am greatly indebted to the PSE IT staff, Stephane and Radja, who assisted me in my coding and programming adventures, to the extent that they tolerated my 70% occupancy rate for the server. Thanks to both of you, for your endless patience.

My gratitude extends also to faculty members of PSE: Nuno Coimbra, Pablo Winant, Katheline Schubert, Agnes Bénassy-Quéré, Angelo Secchi, Jean Imbs, Luigi Moretti, Ariell Reshef, in particular helped me mature as a researcher and teacher.

There is a large community of fellow Ph.D.s whom I am indebted to. I shared three offices and several lunches, coffees, breaks, nights out, seminars and, eventually, friendship in its faceted forms. Therefore, in a somewhat scattered order, a huge thank you goes to Alexia, Anastasia, Anna, Brendan, Cem, Elisa, Emmanuel, Ezgi, Giulia, Hélène, Lennart, Luis, Matthieu, Medhi, Nhuung, Normann, Rudy, Sandra, Sarah, Sebastien, Stefaniya, Umberto, Yassine, Zeinab, and I am almost surely forgetting someone.

Alessandro, Can, Maria Chiara, Hector Ahuitzotl, Jaime chiquito, Sophia were the gang I frequently gravitated to, I owe them so many good moments and memories and therefore wish to thank them deeply. The time we spent together definitely enriched my journey, and I cannot wait to meet all of you again very soon!

Before venturing in the Ph.D. track, I met a number of amazing people, with whom I spent and shared one of the most defining parts of my academic career. During the ETE M2 forging years Antoine, Badis, Barbara, Christelle, Christina, Daniel, Kevin, Ludovic, Noe, Quentin, and Tanguy turned out to be not only great and inspiring individuals, but most importantly dear friends. We ploughed together through tough times and duly celebrated every step, I miss all of you!

In fact, my Parisian time started even earlier than then, in a literally random way. Most of the people I met back then went on for their own adventures, but I hold some of them dear to heart. Alice, Priscilla, Rachele, Maria Lucia, Andrea, Tiziano, I am deeply grateful to be your friend!

I wish to thank also my old friends back in Italy, an endless source of good laughs and discussions: Arianna, Bene, Bred, Cqz, Elena, Fede, Gatto, Gio, Jimmy, Lele, Matz, Meri, Scapi, somehow it always feels like we spent just a couple of days away.

I also owe a big thank you to Laura, who travelled with me through many, many unexpected corners and turns of my recent experience. Her help assisted me to keep myself together, yet multiple.

While naming names, I was blessed to cross the path of a relevant number of exceptional human beings. They are the most precious and rarest wealth I cumulated, the one that fuelled my efforts, especially over the last two, pandemic years.

I owe a lot of gratitude to Ilya, my hero in many dimensions. You made all the hours we shared in that office a real pleasure, and I cannot realise how much inspiration you infused me. Thank you for every serious discussion, every laughter, every coding tip, every nerdy meme, every musical discovery and especially for tolerating my awful Russian impression!

A long time ago I met Federica and she was kind enough to invite a random Erasmus student to meet some other QEM people. That day I did not know she would become one of my dearest friends, with whom I shared some dense moments of my recent life. I want to thank you for all the smiles and laughs you put on my face, for all the times you took me out of the office and brought me to enjoy some real, nutritious life. You really nurtured me while becoming the best mathematical economist I know, grazie mille Fede!

Irene, you blond, coffee-scented wind that blows right after lunch, you are a true force of nature. I admire your rigour, discipline and pragmatism, but I enjoy even more the time we spent (and will spend) together. You taught and inspired me to deal with hard times, but more importantly to deeply enjoy good times. Thank you for bringing light and lightness even during rainy days, for springing life, energy, and enthusiasm at every step. I owe you so much that words cannot tell, grazie infinite Ire!

Jaime, through this roller-coaster you have always been at my side, literally: you pushed me up every time I felt weak, at times taking care of me. I cannot describe the bond we have come to have, but I feel it is unique in its kind. I want to thank you for being my personal champion, reminding me of how good and satisfying life can be. And that some music and a little dance can change your day, aunque soy un tronco bailando. Especially over the last two years, Chiara and you (and Oliver!) were the bedrock of my daily life, I cannot find proper words to express how grateful I am to both of you for, basically, everything. Mil gracias jefe, grazie Chiara!

Marco, despite the first time you called me Francesco, something must have clicked. I feel just more than friendship, a sort of brotherhood that challenges your globetrottery habits. I wish to

thank you from the bottom of my heart for just everything, for believing in myself more than I do, for the brightest discussions and most mindless jokes. You can somehow handle my rough (very rough) edges and bring joy to everyone's life, mine included. You are an outstanding, thoughtful researcher, a sincere soul, just thank you, for all the moments we shared. There are simply not enough words to express my gratitude, grazie bro!

Finally, Titti: for a long time you have been the emotional propeller that pushed forwards my days. I wish to thank you for all the energy you poured and invested, indirectly and directly: you are a unique creature.

Needless to say, my whole family was fundamental all along, since I left Italy in 2013. My parents and sister supported me to an extent that I still cannot fully grasp, with unconditional love that travelled all the way from Trebaseleghe to Paris. Grazie, di tutto, di cuore.

Ζῆνα δέ τις προφρόνως ἐπινίχια κλάζων
τεύξεται φρενῶν τὸ πᾶν,
τὸν φρονεῖν βροτοὺς ὁδῶ
σαντα, τὸν πάθει μάθος
θέντα κυρίως ἔχειν.

But Zeus – if any, heart and soul, that name
Shouting the triumph-praise – proclaim,
Complete in judgment shall that man be found.
Zeus, who leads onward mortals to be wise,
Appoints that suffering masterfully teach.

Aeschilus - Agamemnon

Chapter 1

General Introduction

Inflation and prices are central topics in economics and pervasive in everyone's life experience. Their study produced rich streams of approaches, methods, frameworks, and even schools of thought. Since the timeless auctioneer in Walras (1874), economists ponder on price determinants and which forces drive their change. Inflation, indeed, is still a debated matter: research on the determining forces, as well as the intimate link with agents' expectations, spans several fields, well beyond macroeconomics. central banks' fundamental mandate in most advanced economies is to steer inflation so as to ensure stability in the macroeconomy. Over time, these institutions switched from managing the growth of monetary aggregates to setting interest rates in order to influence the path of inflation and curtail adverse economic shocks. In particular for the US, the change of regime was particularly apparent and consequential, since it is deemed to have set off the Great Moderation period, characterised by steady output growth, low and stable inflation, moderate fluctuations. The 2008 Global Financial Crisis, originated from the financial sector of the economy, pushed central banks to deploy additional tools on top of interest rates, and revived the role of liquidity. The SARS-CoV-19 shock, more recently, further reinforced the role of the latter.

The academic and policy consensus that formed around monetary issues hinged on a handful of central tenets. Besides central bank independence from fiscal policy, monetary policy narrowed down to adjusting a short term reference rate according to expected inflation in order to hit a predetermined inflation target. This institutional setup, usually, takes the form of a feedback rule, owing to the advantages of rule-based policies. Inflation expectations formed by economic agents, as it appears, play a crucial role in conducting monetary policy and in determining actual inflation, typically by some form of Phillips curve. Specifically, if firms and consumers expect future inflation to be high, they will internalise their belief into today's decisions and factually bring about higher rates of inflation, all else being equal. To correct inflation expectations and prevent said self-fulfilling spirals, most monetary models advocate for an active policy stance for the central bank, which requires adjusting interest rates more than proportionally to expected price change. To the empirical test, this setup proved broadly consistent: especially for the US, estimates of

a passive feedback rule for the Federal Reserve Bank correspond to high inflation years. This framework is often embodied in the so-called New Keynesian monetary models, which combine real shocks and nominal frictions to explain short-run fluctuations and draw policy prescriptions. As Figure 1.1 shows, inflation rates across advanced economies reduced volatility and report a remarkable convergence to low levels. This trend appears to start around the 1980s, with the adoption of inflation rate targeting and after the oil crisis.

The Zero Interest Rate Policies deployed during the 2008 recession, though, provided a compelling counterexample to the effectiveness of active policies and questioned the theoretical apparatus of monetary models. New Keynesian models mostly overlooked the financial side of the economy, while also implying perverse effects of policy rates pegged at zero for long stretches of time. Furthermore, major central banks turned to liquidity management and deployed numerous lines and facilities to address liquidity shortages and shield financial assets from fire-sales. These policies, broadly classified as unconventional or Quantitative Easing policies, moved on uncharted territory in 2008, but were once again put in place in 2020.

Looking at historical inflation, the combination of liquidity policies and irresponsive policy rates did not trigger self-fulfilling equilibria, but rather a twin missing deflation and subsequent reflation, which were predicted by the NK theoretical framework. Indeed, inflation displayed a rather smooth behaviour, in line with previous decades. Notwithstanding such stability, though, over the Great Moderation decades the inflation process became paradoxically stabler and less predictable.

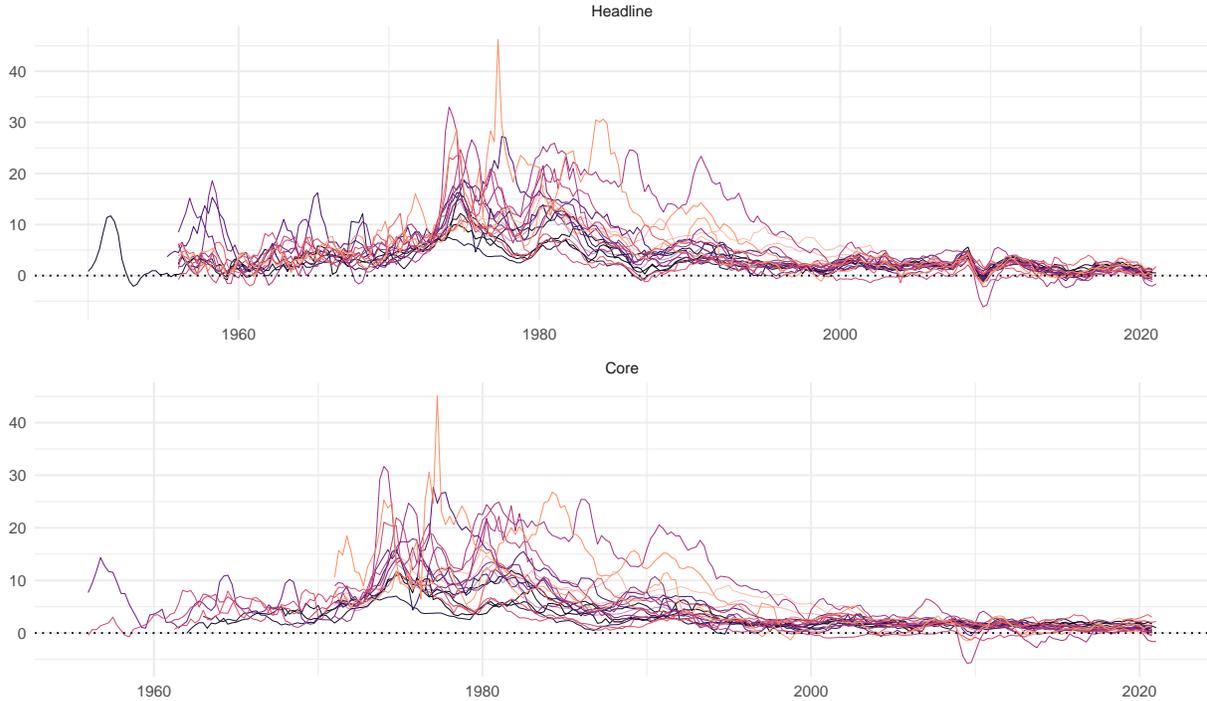
This set of facts motivates much research work and puzzles policymakers: What is the influence of monetary policy on inflation dynamics? Do central banks consider financial liquidity in their policy decisions? Do liquid assets affect in any significant way the limits of policy effectiveness?

The present dissertation revolves around the influence of central banks decisions on inflation, and how financial liquidity interacts with both factors. Its chapters expose the linkages of monetary policy, liquidity, and inflation from both empirical and theoretical angles.

The first chapter, titled "Taylor Rules and Financial Liquidity," revisits the historical evidence on the Fed's policy stance and reports on the factors entering the decision process of the US Federal Reserve Bank. Among these factors, financial liquidity turns out to be a consistent predictor of interest rate setting, especially in periods of financial distress and economic headwinds. On top of significant instability around the feedback rule of the Fed, this chapter also offers a new perspective on the necessity of an active stance to ensure determinacy.

The second chapter, "A Simple Model with Liquidity," rationalises the role of liquidity in a standard, New Keynesian framework and explores its implications for economic stability. This extension relaxes the conditions for determinacy of the model, even with a passive monetary stance. Virtually, all other features of New Keynesian models are preserved. The intensity of the central

Figure 1.1 – Historical inflation rates



Source: OECD. Selected advanced economies, quarterly CPI inflation rates: Austria, Belgium, Canada, Denmark, Finland, France, Greece, Ireland, Italy, Japan, Netherlands, Norway, Portugal, Sweden, UK, US. Top: headline inflation; Bottom: excluding food and energy items. Dashed lines are averages for G7, G20, OECD members, European OECD members.

bank reaction to expected inflation, though, affects the persistence in the model variables: a more accommodative stance is reflected in a more persistent inflation series.

The final chapter, "Inflation Persistence," takes this implication to the data and extends the study of inflation dynamics to several series and new methods. In particular, after reviewing standard macroeconomic methods, this chapter applies deep-learning tools to gauge how inflation persistence varied over time. The main pattern, consistent across econometric methods and inflation measures, is that inflation persistence follows a hump-shaped profile: it peaks in the mid-90s and decreases since. Thus, inflation is currently similar to a white noise process, bearing little dependence on past realisations.

To correctly place these chapters, the following sections provide a bird's-eye view of the existing literature and collocate the main contributions of this dissertation.

1.1 Monetary macroeconomic theory

The theoretical landscape in monetary macroeconomics is rich and burgeoning. Currently, New Keynesian Dynamic Stochastic General Equilibrium models are the reference tools to think about business cycles and test alternative policies. The origin of NKDSGE models dates back to Real Business Cycle models (Kydland and Prescott, 1982; Long and Plosser, 1983) and, farther back, to the growth model introduced by Solow (1956) and extended by Lucas (1972, 1975, 1977), although the focus is admittedly on short-run horizons. They rely on microfoundations to overcome the Lucas (1976) critique, and feature intertemporal constrained optimal choices under varying degrees of uncertainty.

In a nutshell, NKDSGE models add frictions and inefficiencies to the core structure of RBCs, so to sidestep monetary neutrality and carve a role for monetary policy. In their simplest form, NKDSGEs summarise aggregate supply and demand in two linearised, dynamic, and forward-looking equations. The supply side incorporates a Phillips curve that depends on inflation expectations and current economic activity. Aggregate demand is condensed in an Euler equation, which relates current consumption to interest rates and future consumption. Finally, if the fiscal side of government is assumed away, the model is completed with a third equation that sets the path for interest rates, in essence a representation for the central bank. Most often, the stylised economy is a cashless limit case, or otherwise a money demand is introduced with real balances entering the utility function.

This framework presents an appealing flexibility but also a number of predictions at odds with empirical evidence. Especially when the policy rate is pegged (eg, at the Zero Lower Bound, ZLB), NKDSGE models overpower the current effects of policy changes that are distant in the future. For example, the current effect of an announced policy rate cut increases with the distance into the future. Moreover, and key to correctly place the contribution of this dissertation, the relation governing the path of interest rates turns out to be crucial in disciplining the model characterisation.

These paradoxes are often addressed by reworking the rational expectations hypothesis: by assuming rational inattention (Sims, 2010), bounded rationality (Gabaix, 2016), or positing an explicit cognitive process (Garcia-Schmidt and Woodford, 2019). The result is that current decisions are much less responsive to changes occurring far into the future. The dynamics of an otherwise standard NKDSGE radically change and in particular, the model is not subject to indeterminacy at the zero lower bound.

In a feedback rule such as that often postulated, how intensively the interest rate reacts to expected inflation is a key factor for the equilibrium determinacy. If the central bank levers the policy rate less than one-to-one with respect to inflation expectations, the economy displays nominal and real indeterminacy. This implies that any level of prices and output is consistent with the model, and hence negligible shocks can lead to explosive paths. In this light, the ZLB episodes in

the recent past proved damaging for the workhorse model, and motivated numerous attempts to fix such puzzles. Liquidity factors were included in this set of fixes, but seldom included in simple and compact settings.

A complementary approach addresses the core structure of NKDSGEs and the role of money and liquidity. As previously pointed out, NKDSGEs build on a RBC kernel, and hold on to the real dimension of economic forces (Kocherlakota, 2016). Thus, studies like Canzoneri et al. (2011), Diba and Loisel (2021), and Michailat and Saez (2021) introduce money or bonds in various forms to circumvent a set of puzzles or paradoxes implied by NKDSGEs. In this vein, particularly relevant are the recent contributions by Calvo (2016), who puts at centre stage in the macroeconomic models issues related to money and liquidity.

The chapter 4 titled "A simple model with liquidity" is part of such an effort to amend NKDSGEs, but has the advantage of keeping close contact with the workhorse model while proposing minor deviations. Such deviations, though, suffice to relax the requirement of an active and aggressive central bank: the model displays nominal and real determinacy even when the interest rate is nearly irresponsive to inflation expectations. The chapter offers a systematic study on the interplay between liquid assets and monetary policy from a reduced-form standpoint. As cash and liquid bonds (or equivalently bank account deposits) provide transactional services, and hence utility, their inclusion in the utility function of the representative agent is a straightforward option to model different forms of liquidity.

The central bank, in contrast to the workhorse NKDSGEs, operates with two policy instruments. First, it sets the total level of nominal liquidity, defined as the sum of liquid bonds and money in circulation. Second, it adjusts the interest rate on the liquid bond in reaction to expected inflation, rather than the nominal interest rate on the illiquid bond that serves intertemporal smoothing.

Crucially, the resulting set of equilibrium conditions is modified in two ways: the Euler equation is augmented with a money wedge that alters its usual form; then, total liquidity demand also allocates between money and liquid bonds, both in real terms. Inflation, besides imparting movements in interest rates, also influences total real liquidity. This feedback loop does not compromise the behaviour of the economy, which reacts to technological and monetary policy shocks in line with the standard NKDSGE.

Relaxing the requirement of an active feedback rule allow exploring how aggregates are affected by liquidity factors and a potentially destabilising monetary policy. Specifically, the model can be used to rerun a liquidity dry-up similar to that of the 2008 GFC, when several assets suffered liquidity losses. It turns out that an active central bank, swiftly adjusting the policy rate, can significantly speed up the convergence of output, trading off lower levels of real liquidity and stranded liquid bonds. Symmetrically, a passive monetary stance prolongs the contraction of output and also struggles to revive inflation, while preserving real liquidity and bond holdings.

Passive monetary policy stance also affects aggregate dynamics, as the chapter illustrates. First off, the mere inclusion of liquidity induces a higher level of inertia in the observed aggregates. This inertia is still sizeable in comparison with much richer, detailed NKDSGEs that include trend inflation or sophisticated financial frictions and nominal indexation. Focusing on inflation, the model shows by means of stochastic simulations that a less responsive central bank induces more lag-dependency, as the liquidity dry-up exercise suggests.

The chapter also provides an alternative setup, where a mix of liquid assets is used to purchase consumption goods, in the vein of cash-in-advance models. Such alternative channel does not alter the fundamental insights of the reduced-form model, but rather reinforces the role of liquidity, especially for it significantly affects monetary policy.

All in all, this chapter offers an account of the relevance of liquidity conditions in the economy and its implications for policy. As the toolkit of unconventional policies shows, policymakers have clear considerations for financial matters in general, and liquidity provision in particular. The model, keeping the core structure of the baseline NKDSGEs, allows for promising extensions: many of the additional blocks developed for existing NKDSGEs can be included in its setup in a relatively easy fashion.

This is hardly possible in other approaches that abandon fundamental assumptions that have characterised models in the past 40 years, such as rational expectations. This does not imply that new ways for modelling expectation formation or other modifications of behavioural assumptions of the traditional model are not worth exploring. Focusing on the monetary-financial side of the models and introducing concepts, such as liquidity, which are well-founded in microeconomic and financial theories (Eggertsson and Krugman, 2012; Holmström and Tirole, 2011). The model discussed in the chapter lends itself as a bedrock for further extensions and refinement, particularly for analysing fiscal policy and banking.

Nevertheless, the most appealing way forwards sits in a full endogenisation of total nominal liquidity management by the central bank. This would provide monetary policy with a full-fledged second instrument, replicating more closely the actual conduction of QEs in recent years.

In light of the considerations presented above, hitting the ZLB in 2008 and 2020 while also implementing liquidity policies provided extremely valuable data points to bring to the empirical test the predictions of NKDSGEs. The next section overviews the most relevant contributions on the empirical side of monetary macroeconomics, and places the first and third chapters in such literature.

1.2 Empirical Monetary Macroeconomics

The tools traditionally used to study macroeconomic empirical questions are drawn from the time series toolkit. This includes reduced-form or Structural Vector Autoregressive (SVAR) models, as

well as less recent Dynamic Simultaneous Equations models. Restricting the focus to monetary issues, (S)VARs are the reference tools, since they require simple but stringent assumptions, and NKDSGEs can often be recast in SVARs form. In the last two decades, in particular, Bayesian methods further pushed the frontier of empirical investigation, providing powerful tools to estimate rather complex models with fairly transparent assumptions.

With respect to monetary policy, it is useful to distinguish two approaches: limited information estimation and full information one. The latter typically involves a system of structural relations and is informed by a theoretical apparatus – either on the restrictions (VARs) or on the mutual influence of shocks and observable aggregates. The former, instead, zooms in on a particular macroeconomic relation and assesses its reduced-form validity in isolation from other forces, although with possibly complex modelling assumptions.

Taylor (1993) monetary policy rules and Phillips (1958) curves are especially salient examples of the limited information approach: both relations are intimately intertwined and are crucial components of the modern monetary framework.

The next sections offer a broad overview of the current state of affairs for monetary policy rules and Phillips curves in relation to the first and third chapters of this dissertation.

1.2.1 Monetary policy rules

Feedback rules are currently the first-line instrument for monetary policy. Therefore, correctly identifying how central banks operate in the economic environment is essential for policy predictability and transparency. In the same vein, correct estimates of the monetary policy rule are decisive in evaluating how effectively central banks keep inflation near its target.

Eyeballing actual data as pictured by Figure 1.2, it is possible to appreciate the different phases the US economy traversed since World War II. The figure presents time series for the Federal Funds Rate, headline Consumer Price Index, and an ex-post measure of economic activity, the output gap. Until the early 1970s, inflation was relatively stable under 5% and not particularly volatile. Similarly, the FFR followed a smooth upward path, following the apparent trend of inflation. In contrast, the dynamics for economic activity appears relatively more volatile, with large swings in negative and positive territory. Then, since the 1970s, inflation ramps on an upward, decade-long trend, closely tracked by the policy rate. The output gap still displays large swings, but mainly in negative territory, signalling an economy under its full potential.

A break seems to take place around the beginning of the 1980s: FFR shoots up, the output gap signals a recession, and a deflation takes place. Indeed, the inflation rate returns to previous levels with reduced volatility, while the FFR moves more smoothly and responds both to inflation and economic activity. The latter also shows decidedly less volatility than previous periods, with longer upward trends recovering from mild recessions. This period, often referred to as Great Moderation, indeed starkly contrasts with previous years: less volatile in general, with lower in-

flation. The switch corresponds roughly to the appointment of Volcker as Chairman of the Federal Reserve Bank, suggesting that he brought about a discrete policy break.

The later decades contrast with the Great Moderation, too, but along different dimensions. In particular, at the onset of the 2008 Global Financial Crisis, the Fed pushed the policy rate to the ZLB, which bound for several years afterwards. Yet, inflation did not spiral out of control as during the 1970s. The key difference in this late part of the sample is summarised by Figure 1.3: since mid-2008 the Fed inflated its balance sheet in order to inject unprecedented amounts of liquidity in the financial system. This large balance sheet expansion was composed of several distinct operations and facilities, but is usually referred to as Quantitative Easing policy. Interestingly, similar policies were deployed by many other major central banks – including the European Central Bank, Bank of England, Bank of Japan.

The effects of such policies, though, are not yet well understood, even if they signal two relevant facts: first, interest rate setting is not the only policy tool at central banks' disposal; second, liquidity, broadly speaking, is a relevant factor for inflation control and economic stability, at least in central banks' framework. As chapter 4 explored theoretically the role of liquidity in NKDS-GEs, chapter 3 takes an empirical approach to liquidity and monetary policy rule, in the narrow financial sense.

The modern literature on monetary policy rule estimation sprung from the seminal paper of Taylor (1993), which in a simple setting estimated the reaction function of the Fed. In that contribution, the specification takes to the empirical test the then-recent theoretical insights on policy credibility, the effectiveness of rule-based policies vis-à-vis discretion, and the focus on interest rates rather than money growth. Taylor postulates a feedback rule that implements a strong response of the Federal Funds Rate to the year-on-year inflation rate, calibrated to 1.5. Assuming a response to output of about .5 and a long-run target of 1% for inflation, this simple rule fit very well the actual behaviour of the Fed over the period 1983Q1-1992Q3. The result, making direct contact with the nascent research of NKDSGEs, sparked a lively debate on policy rules, so much so that monetary policy rules of this form are often referred to as Taylor rules. Alongside a rich theoretical discussion, Taylor (1993) kick-started a whole strand of empirical literature on the estimation of policy rules. The early focus of this research was to understand how policy regimes changed in the US, especially to ascertain the prevailing monetary regime during the high inflation period of the 1970s. Clarida, Gali, and Gertler (2000) is one of the main contributions to such literature, setting the ground for the methods employed and the ensuing narrative about policy stance. In a nutshell, estimating policy rules is a complex task because of the issue of intrinsic endogeneity, which biases standard estimates. When the error terms are correlated with the regressors, in fact, plain OLS asymptotic does not hold and estimated parameters are biased. To overcome these issues, Clarida, Gali, and Gertler (2000) deploy a GMM-IV approach, taking lags of endogenous variables as instruments for their expected values. Moreover, they pioneer a split-sample strategy,

Figure 1.2 – US Phillips Curve and Monetary Policy

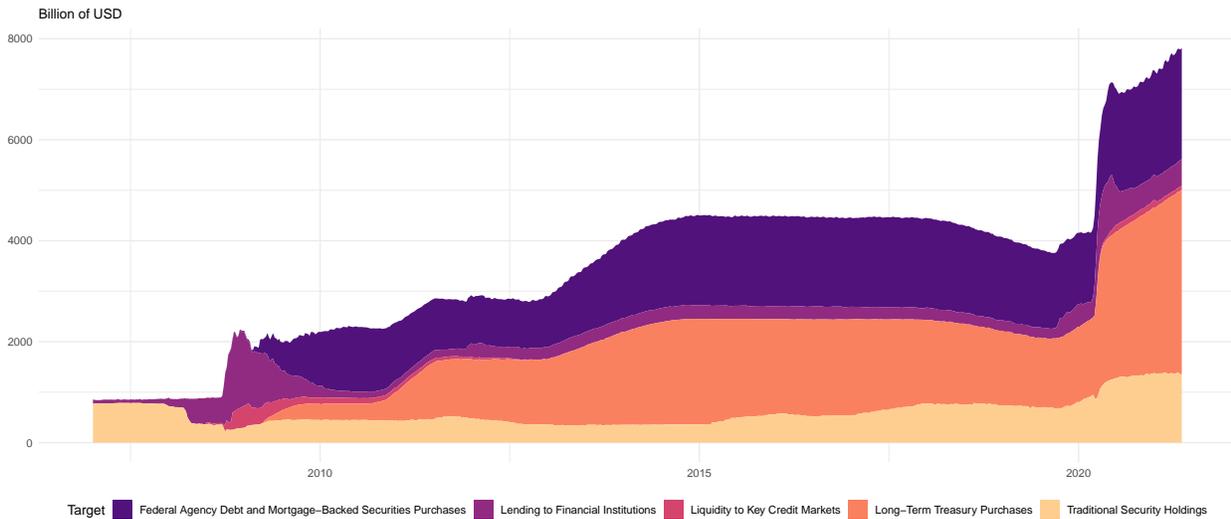


Federal Funds Rate, headline CPI, *ex post* output gap for the US economy. The output gap is computed as a percentage deviation of revised actual real output with respect to the Congressional Budget Office's estimates of potential real output. Sources: Federal Reserve Bank, CBO.

which consists in estimating the policy rule on two subsamples: the splitting point corresponds to Volcker's first chairmanship. The approach is palatable enough as it makes contact with conventional wisdom and anecdotal evidence at the time. The findings, in sum, are that the first period sees an accommodative central bank, corresponding to high inflation and great volatility. The key estimate for the reaction to expected inflation is well below one: drawing from the theoretical apparatus, this policy stance engenders sunspot equilibria and severe indeterminacy. Over the second subsample, initiated by Volcker's disinflation, the Fed is reported to strongly react to expected inflation, thus ruling out indeterminacy and effectively keeping inflation under control. This analysis, taken together, informs much of the established consensus on the conduct of monetary policy in the US; it also affords some criticisms. For instance, instrumenting expected inflation by its lagged values leads to weak instruments, since the inflation rate is often best described by a white noise. Moreover, two choices turned out particularly consequential: the use of revised data, and the – sensible but somewhat exogenous – choice of subsamples' timing.

Following research improved on both aspects. Reproducing as closely as possible the information set of the policy-maker at the time of each decision is at odds with revised data, which

Figure 1.3 – Fed’s Liquidity Facilities



Balance sheet of the Federal Reserve System, from 2007Q1. Classification of holdings into main grouping types: Federal Agency debt and MBS; Lending to Financial Institutions; Credit Market Liquidity; Long Term Treasury holdings; Traditional Holdings. Values in billions of current US dollars. Source: Cleveland Fed

undergo corrections as more precise information becomes available. In this sense, Orphanides (2001) improves on Clarida, Gali, and Gertler (2000) by collecting real-time data on inflation and output gap forecasts available to the Federal Open Market Committee at each meeting, drawing from the Greenbook database. The timing bias is sizeable in the estimates of policy rules, to the extent that a Taylor rule estimated over the 1987Q1-1993Q4 with real-time data and forecasts pictures a passive stance for the Fed – at least at forecast horizons shorter than three quarters. In a follow-up contribution, Orphanides (2004) replicates the split-sample exercise of Clarida, Gali, and Gertler (2000) employing only real-time data. The two regimes in place before and after Volcker’s appointment turn out to be more similar than expected. Estimates are broadly similar across the two regimes: both subsamples consistently produce estimates of proactive reaction to expected inflation, which downplays its stabilising power. The findings with real-time data actually point to excessive responsiveness to output during the 1970s, hinting at overconfident policy-makers about smoothing fluctuations of the business cycle. These contributions relied essentially on OLS or Non-Linear Least Squares (NLLS), without instrumenting for expected inflation but rather using actual forecasts from the FOMC staff. The use of real-time data, though, does not solve entirely the endogeneity issue, as forecasts are usually conditional on the future path for the policy rate, which risks polluting the estimates.

The issue of a historically informed, but exogenous subsampling strategy has been tackled with a variety of econometric approaches. Among these, Time-Varying Parameters (TVP) meth-

ods yielded insightful results, together with adaptations of Markov Switching estimates. Boivin (2006) and Kim and Nelson (2006) present limited information results based on TVP estimates. The former depart from the GMM-IV apparatus to venture into dealing with TVP models with endogenous regressors and heteroskedasticity in the shocks. Summing up the main takeaways, they exploit standardised forecast errors in the IV step to do away with endogeneity and correct the bias it brings. The resulting estimates paint yet a more nuanced picture. There appear to be three distinct phases in the post-WWII Fed: the first phase from 1970 to 1975 the Fed's reaction to inflation was not statistically different from one; during the second phase, from about 1975 through the early 1980s, the reaction ticked up slightly, still not significantly different from one, though. By contrast, the final regime, which covers at least the 1980-2002 period, features a proactive stance for the Fed, but with increasing uncertainty around the point estimate. The latter is possibly due to the low volatility of inflation, which further weakens the power of lags as instruments.

Finally, Boivin, 2006 combines heteroskedasticity-robust TVP with real-time data, summarising the contributions of Kim and Nelson (2006) and Orphanides (2004). Contrary to Kim and Nelson, 2006, Boivin finds an active stance towards inflation at the beginning of the 1970s and from about 1982, with a steep decrease in between. The transition back to a proactive stance, in particular, does not look discrete, while uncertainty still increases from the 1990s onwards.

A later strand of literature took a similar approach, but relied on discrete regimes instead of smooth transitions. It is the case for Davig and Leeper (2006, 2011) and Murray, Nikolsko-Rzhevskyy, and Papell (2015). The former in particular consider jointly fiscal and monetary policy rules, focusing on active and passive Markov regimes for both. While the combination of fiscal and monetary rules nuances the general picture, they find evidence of an active stance over the periods 1979-1990, 1995-2001, and a passive stance during 1949-1979, 1990-1995, 2001 onwards, in part confirming the early results from Clarida, Gali, and Gertler (2000). At the aggregate level, each state results particularly persistent, despite the frequent switches.

Following this trend of endogenous policy switches, Murray, Nikolsko-Rzhevskyy, and Papell (2015) apply the Markov switching tooling developed by Hamilton (1989) to the sole Taylor rule. Mixing real-time data and regime switch they confirm in a limited information exercise the likely existence of multiple regimes within the usual periodisation. The distribution of such regimes contradicts the typical narrative built on Clarida, Gali, and Gertler (2000): Fed's stance was virtually always active during the episodes of high inflation. In particular, the Fed proactively counteracted expected inflation in the periods 1965Q4-1972Q4, 1975Q1-1979Q3, and 1985Q2 until the end of their sample.

Chapter 3 builds on this approach and augments the baseline Taylor rule specification to account for liquidity factors, which proved relevant in actual policymaking.

The inclusion of financial stress indexes was previously explored in the monetary policy literature, but the role of liquidity was overlooked until 2008. Some notable exceptions are Baxa, Horváth, and Vašíček (2013) and Istrefi, Odendahl, and Sestieri (2020), which relate financial stability concerns to policy decisions – testing a battery of financial indexes for the former; parsing information from official minutes for the latter. The set of unconventional policies included a great deal of liquidity management: one of the main concerns at the Fed was to unlock the financial markets and reboot trades. This concern is evident in the notes that introduced each policy innovation in the wake of the GFC: *AMLF*; *CPFF*; *MMIFF*; *PDCF*; *TAF*; *TSLF* notes all mention liquidity as the target for such policies. Extending the traditional Taylor rule to account for the role of financial markets essentially boils down to proxy the state of liquidity in the financial system.

Including liquidity is reasonable in light of policy developments, but its empirical counterpart is particularly thorny. Liquidity is a rather ephemeral feature of assets. For one, it hinges on the trade side one takes: an asset in high demand is easy to sell, therefore liquid; conversely, it is hard to buy, thus illiquid. A narrower operative definition of liquidity rests on the discount that is necessary to bear when transforming an asset into cash. To overcome such issues, the chapter proposes to capture a proximate measure of liquidity via spreads on largely traded, low risk assets, so as to pick a measure that is devoid of risk and endogeneity as much as possible. These measures are two spreads between extremely safe assets, ie US Treasury bills, and less safe ones. Precisely, the measures used in the estimates are spreads between BAA corporate bonds' average yield and quarterly returns on the Standard & Poor's 500 index. Maturity at issuance for the latter is a relevant factor in determining the premium over safer assets. To account for this timing factor, the spreads are computed on Treasury bills of comparable maturity, either ten years or three months.

The inclusion of such measures of liquidity in the policy rule, thus, captures the role of financial concerns in actual policymaking. The intuition builds partially on leaning-against-the-wind literature (Svensson, 2017b), but elicits the input set of the Fed with respect to liquidity concerns.

The main results of the chapter follow from estimating a Markov switching model with liquidity-augmented Taylor rules. Such empirical model has the advantage of relying solely on information contained in the data to endogenise regime changes. While the standard specification, including Greenbook forecasts for current output gap and expected inflation, finds virtually only proactive regimes, at odds with the established consensus. Liquidity measures, though, provide a different narrative: when the short-run liquidity proxy is included, there is evidence of both active and passive stances. Interestingly, the passive regime seems to prevail over the post-WWII sample, while the active one prevails for some short-lived periods, precisely the first half of the 1970s and for 1979-1982. Taking the long-run liquidity proxy, based on BAA bonds, the relevance of expected inflation is severely disputed, so much so that the Fed appear to react negatively to it – remarkably during the ZLB period. Across specifications, though, financial liquidity is consistently found as

a significant predictor for movements in the policy rate. Interestingly, passive regimes were likely in place during periods of tranquil inflation: this result casts doubts on the stringency of active monetary policy to tame inflation. The estimates also suggest an association between the probability of different parameter regimes and the direction of policy changes. Periods of increasing interest rates are associated with a high probability of being in a Volcker type regime (high coefficient on inflation in the policy reaction function). By contrast, periods of monetary easing through declining interest rates are associated with a higher probability of being in a less reactive regime.

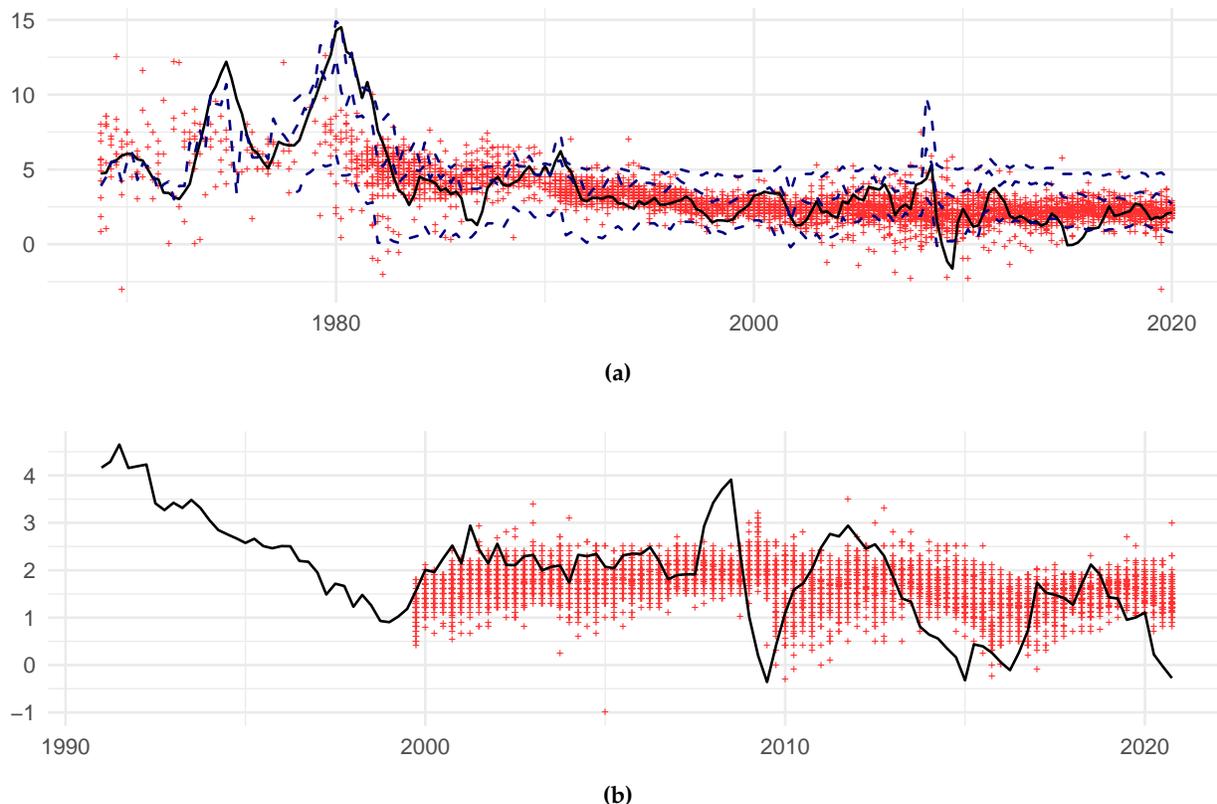
In an effort to summarise established empirical strategies, the chapter proposes also full- and split-sample analyses, which broadly corroborate the main result, on top of the relevance of financial liquidity for monetary policy-making. The full-sample analysis reveals that including either proxy of liquidity severely downsizes the coefficient on expected inflation: both specifications estimate a reaction to inflation slightly below one, much lower than the estimates for classic Taylor rule. In the split-sample exercise, the cuts follow the established empirical strategy, but add a third subsample covering the post-2008 periods. The latter part is naturally less informative because of data availability (Greenbook data are released with a five-year lag) and features little to zero variation in the policy rate. In line with Orphanides (2004), the standard specification finds an active stance during the pre-Volcker subsample, and a stronger active stance during the Great Moderation. The specifications that include financial proxies report weak responses to expected inflation in the pre-Volcker period, and, similarly, stronger responses from the 1980s onwards, although point estimates are significantly smaller than the standard specification.

1.2.2 Inflation dynamics

Akin to the Taylor rule, the Phillips curve is the key mechanism behind inflation dynamics: it relates current inflation with some measure of economic activity and, more elusively, with expected inflation. Notably, the expectations-augmented version of this relation is the channel for possible self-fulfilling expectations when the monetary policy stance is passive. More recently, the Phillips curve has mobilised a great deal of research efforts: has it disappeared? Has it flattened and how much? These questions spring from the unexpected behaviour of inflation in the last two decades: current price changes appear less connected to economic activity. The clearest example of this weakened connection is the missing deflation and reflation after the 2008 recession: before the 2020 pandemic, indeed, the US posted the longest expansion and pushed the unemployment rate below 3%, all with inflation in check. For the European economy, this fact is even more severe and puzzling (Ciccarelli and Osbat, 2017).

Moreover, the expectations channel does not provide guidance, but rather poses additional questions. As Figure 1.5a shows for the US, expectations from professional forecasters, regular consumers, and actual inflation follow significantly different paths. The same, for professional forecasters only, holds for inflation in the Euro Area, as shown in 1.5b.

Figure 1.4 – Inflation Expectations



Top panel: US CPI data, solid line is actual, year on year CPI; red dots are individual Survey of Professional Forecasters; blue dashed lines are top quartile, mean, bottom quartile from Survey of Consumers. Bottom panel: CPI and SPF data for the Euro Area. Forecasts are one year ahead CPI inflation rate, shifted to correspondig forecast date. Sources: FRED, Philadelphia Fed, European Central Bank, Michigan University.

Within the Phillips curve framework, therefore, observed inflation is influenced by current economic activity and future inflation. A simple forward iteration of this basic relation, thus, expresses current inflation as the present discounted sequence of deviations from the natural level of economic activity. That is, fluctuations around the so-called potential output, often expressed as the economy's frictionless and non-stochastic equilibrium. The straightforward implication is that, as long as shocks are zero mean and serially uncorrelated, inflation dynamics reflects the underlying dynamics of the business cycle and its determinants.

An additional challenge in estimating the New Keynesian Phillips Curve comes, somewhat paradoxically, from policy effectiveness. Indeed, if inflation expectations are exceptionally well-anchored and monetary policy offsets demand shocks, then observed inflation will appear to "dance to its own music," driven solely by transitory shocks.

These implications align well with the fact that inflation is notably difficult to predict with satisfying accuracy, and increasingly so (Stock and Watson, 2007).

However, the literature has put forward a host of econometric approaches to estimate the Phillips curve, including alternative specifications. It is useful to assort these approaches in two classes. First, use past observations to instrument inflation expectations in an IV setting; second, exploit existing surveys to gauge aggregate expectations. Each econometric strategy, though, tackles some issues but poses new challenges.

Instrumenting expected inflation with its past realisation is appealing because it matches the exclusion restriction. One of the first applications was proposed by Gali and Gertler (1999): they develop a hybrid Phillips curve with a share of backwards-looking firms and exploit forecast errors to replace actual one-period ahead inflation with its expectation. Expectations are then instrumented with lagged variables in a GMM-IV setting. Assuming rational expectations, this strategy meets the exclusion restriction but, in light of inflation predictability, only offers a weak instrument. The now standard inclusion of backwards-looking firms, though, is key to generate inertia and match the observed behaviour, although quantitatively dominated by the forward-looking force. Within this framework, a recent innovation exploits data revisions as external instruments for expected inflation, but with contestable empirical success (Mavroeidis, Plegborg-Moller, and Stock, 2014).

Survey-based estimates were pioneered by Roberts (1995) and recently surged to more general interest (Coibion, Gorodnichenko, and Kamdar, 2018) in studying inflation and expectations. The key idea is to use actual expectations – from professional forecasters, consumers, firms – to target aggregate expectations. While appealing, this direct approach poses subtle challenges with respect to rationality. Aggregate expectations are rational and model consistent, properties that cannot map directly to survey participants but are required for sound microfoundation. Moreover, deviations from rationality differ according to the surveyed population: consumers' bias is different from that of professional forecasters.

Regarding the various specifications, the most frequent strategy to match the data is to depart from the purely forward-looking formulation of the Phillips curve. This is usually obtained by assuming that a share of firms that cannot fully re-optimize will index their prices to past inflation. The intrinsic inertia that results from this departure usually improves the fit on US data. A further extension relaxes the zero-inflation steady state: when steady-state inflation is positive it corresponds to the long-run inflation level, and the Phillips curve is typically expressed in gaps from such trend inflation. Changes in such trend, though, are consequential for estimation and, crucially, for the dynamic properties of inflation, as trend inflation ultimately maps to inflation expectations anchoring.

The role of trend inflation has spurred a research debate that explored its implications for persistence and volatility of observed inflation, as well as its interplay with policy. Several contri-

butions highlighted how the main driving force of inflation dynamics comes from movements in its long-run trend. In a similar fashion, modelling trend inflation has relevant consequences for optimal policy.

Drawing from the comprehensive review of Ascari and Sbordone (2014), it is clear that trend dynamics do have an impact on observed inflation dynamics. The emerging finding is that observed inflation is highly influenced by dynamics in the underlying long-run trend, especially for volatility. Cogley, Primiceri, and Sargent (2008), Cogley and Sbordone (2009), and Stock and Watson (2007) offer empirical evidence on the interaction of trend inflation and persistence, which is fundamental for the design of sound monetary policy. Indeed, the necessary time for inflation to reach a new equilibrium after a shock is crucial for appropriate policy intervention. Similarly, proper measures of persistence provide central banks with correct trade-offs between inflation and activity: the impression of a low and unresponsive inflation could push a central bank to excessive easing, and vice versa.

The variations in US inflation persistence have been the subject of contributions of Fuhrer (2011) and Pivetta and Reis (2007), on which much of the third chapter is based on. Their focus is precisely the estimation of one particular aspect of inflation dynamics, precisely inflation. While Fuhrer (2011) spans limited and full information settings for forward- and backwards-looking Phillips curves, Pivetta and Reis (2007) takes a univariate approach on inflation. They build the analysis on a Bayesian autoregressive model for inflation and find little changes in persistence across several measures.

The third chapter, titled "Inflation Persistence," revisits much of the empirical toolkit presented in Fuhrer (2011) and Pivetta and Reis (2007), with two contributions. The analysis extends the array of inflation indexes for the estimates, so as to disentangle possible confounding factors like globalisation and commodity prices. The different baskets and goods measured by the Consumer Price Index, the Personal Consumption Expenditure index, and the Gross Domestic Product deflator offer a straightforward way to tell apart the forces influencing inflation dynamics. Further distinguishing between headline and core CPI and PCE is similarly informative to isolate the pass-through of (often volatile) commodity prices to observed headline inflation. In a similar way, the data used also cover a longer period of time, inclusive of the 2008 global recession. This last data portion is particularly instructive, since it includes the prolonged period of ZLB and the following rate hikes. In the same vein as Pivetta and Reis (2007), the chapter represents inflation as a univariate autoregressive process, in isolation from fiscal and monetary policy, economic activity, and agents' expectations. A summary of the results is presented in 1.5.

The chapter starts off with a simple, frequentist analysis of inflation persistence. The goal is to obtain first-pass evidence of variations in inflation inertia, which are indeed present across the five measures used. In particular, measuring persistence as the first autoregressive coefficient or as the sum of all relevant lags coefficients reveals a hump-shaped profile for inflation inertia. This

evidence motivates further investigation into inflation behaviour, with more sophisticated tools. Indeed, the autoregressive structure imparts statistical discipline to the estimates, but is limited in scope: it makes limited use of the information present in the data or the subsample it is presented with. Yet, inflation dynamics are possibly affected by long-run movements in the trend or discrete, policy-related switches.

Resorting to Bayesian methods relaxes a good deal of such constraints. Drawing on the statistical setup proposed by Pivetta and Reis (2007), the chapter takes a policy-maker perspective and, at any point in time, subsumes past information on inflation into an autoregressive process with (possibly) unit-root coefficients. From there, it explores future paths for inflation that are consistent with the current state. Such paths allow for the computation of persistence at every period: more precisely, the Bayesian structure provides a distribution for inflation inertia. The distribution is particularly useful, as it combines data-informed beliefs up to present and uncertainty about future developments. Specifically, this approach allows for explosive paths for inflation, those more relevant to policy-makers concerned with avoiding high inflation episodes. The results broadly fall in line with the frequentist exercise, with some caveats. While the hump-shaped profile is corroborated, the transition is sensibly smoother. This is due in part to the random walk transition that autoregressive parameters follow. The degree of uncertainty, though, ticks up considerably towards the end of the sample, reflecting relatively low levels of inertia and unpredictability.

For both exercises, the profile for persistence is found during the first five years of the 1990s, at the beginning of the Great Moderation. The timing is suggestive of forces other than policy, trade, or commodity volatility. Policy hardly operates with such long lags, nor agents' learning can take a decade to be fully incorporated. Global price pressures could have been operating already, but China did not enter the international trade stage with full WTO rights before 2001. Similarly, oil's and other commodities prices were relatively stable, if not for a one-off spike around the first Gulf War.

In order to further relax the structure imposed on the data generating process, the chapter harnesses the flexibility of deep learning methods, which constitutes the key contribution of this work. The edge of this class of methods lies in the wide array of non-linearities that they can approximate. Therefore, deep learning tools can handle fairly complex data structures with ease, departing from the linear autoregressive structure embedded in the frequentist and Bayesian exercises. Among the several statistical learning, the chapter presents Long-Short Term Memory (LSTM) models, from the Recursive Neural Network class. Typically employed for natural language processing, LSTMs are explicitly designed to gauge complex structures in the data they handle. In several applications, they consistently outperform sophisticated forecasting tools, especially when predicting inflation on longer horizons (Almosova and Andresen, 2019). This feature is particularly compelling for time series analysis in general, and for inflation persistence in particular. Despite relatively short series, it is reasonable to assume that the process behind inflation

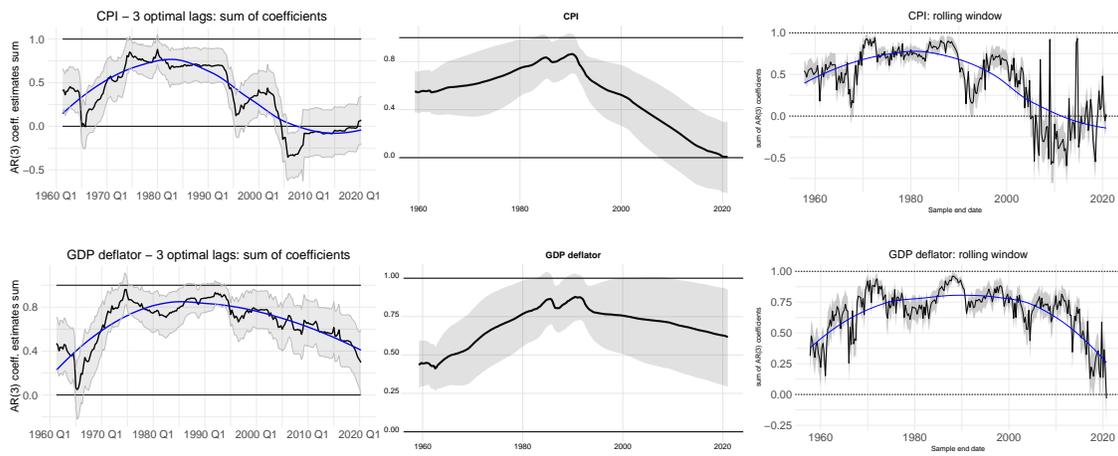
displays breaks, switches, slow-moving trends, or combinations thereof. Thus, methods that can jointly account for these factors are welcome and can shed light on inflation dynamics. As a proof of concept, the third chapter adapts LSTMs to the study of inflation persistence, complementing the more traditional techniques.

The empirical strategy, in this case, is a mixture of the previous two approaches: first, the full sample is used to train the LSTMs, which then is used to forecast inflation for the following ten years. Secondly, the sample is split into non overlapping subsamples of ten years each, used for training; subsequently, the trained networks predict the path for inflation over the following decade. This split-sample strategy isolates features of the inflation process that are peculiar to a given historical period, thus extending the original subsample with synthetic (but informed) data points. Third, a 56-quarter rolling window trains several LSTMs on actual data. Again, before moving to the next data point, the trained model forecasts the following decade of inflation, so as to augment the sample with historically consistent artificial data.

Finally, persistence is measured on the joint time series: this provides a measure of how inertia varied over time, together with reasonable confidence intervals. In sum, the results corroborate the findings from more traditional approaches: inflation inertia peaked in the mid-90s, then steadily decreased, imparting an almost-white-noise behaviour to observed inflation. The timing is consistent with other methods, too: the persistence reversal seems unrelated to policy, trade, or commodities. Remarkably, the deep-learning approach reveals a great deal of uncertainty over the most recent period, since point estimates are particularly unstable starting from the 2008 Global Financial Crisis.

This exercise demonstrates that deep-learning methods can effectively complement more standard macroeconomic tools. Going forward, LSTMs can be generalised to take a VAR-like structure, overcoming at once the issues of multifaceted non-linearity and the curse of dimensionality.

Figure 1.5 – Inflation Persistence



Summary plots for selected series, all methods. Top row is headline CPI, bottom row is GDP deflator. Left column: results for $AR(k^*)$ sum of coefficients; middle column: Bayesian estimates with 95% credibility intervals; left column: rolling LSTM, sum of coefficients.

Introduction Générale

L'inflation et les prix sont des sujets centraux en économie et omniprésents dans l'expérience de vie de chacun. Leur étude a donné lieu à de riches courants d'approches, de méthodes, et même d'écoles de pensée. Depuis le commissaire-priseur de WALRAS (1874), les économistes s'interrogent sur les déterminants des prix et les forces qui les font évoluer. L'inflation, en effet, fait toujours l'objet de débat : la recherche sur ses déterminants, ainsi que le lien avec les anticipations des agents, s'étend à plusieurs domaines, bien au-delà de la macroéconomie. Dans la plupart des économies avancées, le mandat fondamental des banques centrales est de contrôler l'inflation de manière à assurer la stabilité de la macroéconomie. Au fil du temps, ces institutions sont passées de la gestion de la croissance des agrégats monétaires à la fixation des taux d'intérêt, afin d'influencer la trajectoire de l'inflation et de limiter les chocs économiques négatifs. Pour les États-Unis, ce changement de régime a été particulièrement apparent et conséquent, puisqu'il est réputé avoir déclenché la période de la Grande Modération, caractérisée par une croissance régulière de la production, une inflation faible et stable, des fluctuations modérées. La crise de 2008, issue du secteur financier de l'économie, a poussé les banques centrales à déployer des outils supplémentaires en plus des taux d'intérêt, et a ravivé le rôle de la liquidité. Le choc du SRAS-CoV-19, plus récemment, a encore renforcé le rôle de cette dernière.

Le consensus académique et politique qui s'est formé autour des questions monétaires s'est articulé autour d'un ensemble de principes centraux. Outre l'indépendance de la banque centrale vis-à-vis de la politique budgétaire, la politique monétaire s'est contentée d'ajuster le taux de référence à court terme en fonction de l'inflation attendue, afin d'atteindre un objectif d'inflation prédéterminé. Ce dispositif institutionnel prend généralement la forme d'une règle de réaction, en raison des avantages des politiques fondées sur des règles. Les anticipations d'inflation formées par les agents économiques jouent un rôle crucial dans la conduite de la politique monétaire et dans la détermination de l'inflation réelle, généralement à travers une courbe de Phillips. Plus précisément, si les entreprises et les consommateurs s'attendent à ce que l'inflation future soit élevée, ils internaliseront leur conviction dans leurs décisions d'aujourd'hui et provoqueront de fait des taux d'inflation plus élevés, toutes choses étant égales par ailleurs. Pour corriger les anticipations d'inflation et empêcher ces spirales auto-réalisatrices, la plupart des modèles monétaires préconisent une position de politique active pour la banque centrale, ce qui nécessite d'ajuster

les taux d'intérêt plus que proportionnellement à la variation attendue des prix. À l'épreuve empirique, cette configuration s'est avérée globalement efficace : en particulier pour les États-Unis, les estimations d'une règle de réaction passive pour la Federal Reserve Bank correspondent à des années de fortes inflations. Ce cadre est souvent incarné par les modèles monétaires dits néokeynésiens, qui combinent chocs réels et frictions nominales pour expliquer les fluctuations à court terme et en tirer des prescriptions politiques. Comme le montre Figure 2.1, les taux d'inflation dans les économies avancées ont réduit leur volatilité et font état d'une convergence remarquable vers de faibles niveaux. Cette tendance semble débiter autour des années 1980, avec l'adoption du ciblage des taux d'inflation et après la crise pétrolière.

Les politiques de taux d'intérêt à zéro déployées pendant la récession de 2008 ont toutefois fourni un contre-exemple convaincant de l'efficacité des politiques actives et remis en question le paradigme théorique des modèles monétaires. Les modèles néokeynésiens ont majoritairement négligé l'aspect financier de l'économie, tout en impliquant des effets pervers des taux directeurs fixés à zéro pendant de longues périodes. De plus, les principales banques centrales se sont tournées vers la gestion de la liquidité et ont déployé de nombreuses mesures pour faire face aux pénuries d'actifs liquides et protéger les actifs financiers des ventes massives. Ces politiques, classées en gros comme des politiques non conventionnelles ou d'assouplissement quantitatif, ont exploré un territoire inconnu en 2008, mais ont été à nouveau mises en place en 2020.

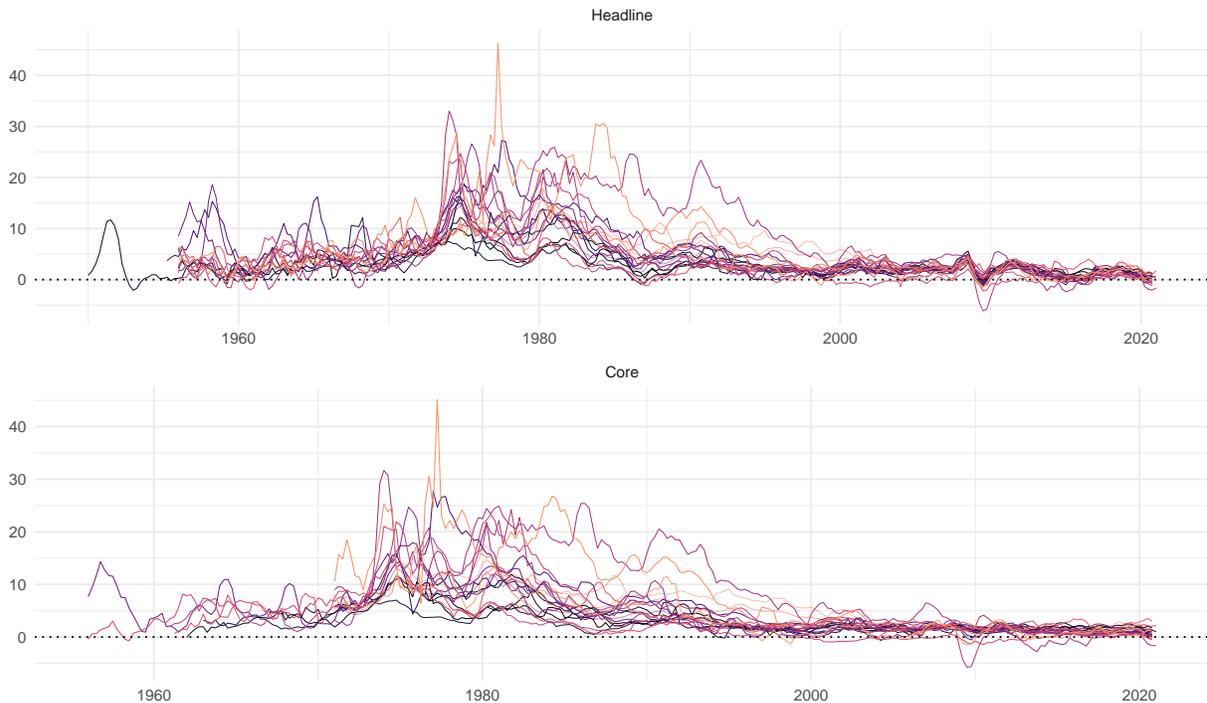
Si l'on considère l'inflation historique, la combinaison de politiques de liquidité et de taux directeurs à zéro n'a pas déclenché d'équilibres autoréalisateurs, mais plutôt une déflation et une reflation manquées, qui étaient prévues par le cadre théorique néokeynésien. En effet, l'inflation a affiché un comportement plutôt lisse, comme les décennies précédentes. Malgré cette stabilité, au cours des décennies de la Grande Modération, le processus d'inflation est devenu paradoxalement plus stable et moins prévisible.

Cet ensemble de faits motive de nombreux travaux de recherche et rend les décideurs politiques perplexes : quelle est l'influence de la politique monétaire sur la dynamique de l'inflation ? Les banques centrales tiennent-elles compte de la liquidité financière dans leurs décisions politiques ? Les actifs liquides affectent-ils de manière significative l'efficacité de la politique monétaire ?

Cette thèse s'articule autour de l'influence des décisions des banques centrales sur l'inflation, et de la manière dont la liquidité financière interagit avec ces deux facteurs. Ses chapitres exposent les liens entre la politique monétaire, la liquidité et l'inflation sous des angles à la fois empiriques et théoriques.

Le premier chapitre, intitulé "Règles de Taylor et liquidité financière", revient sur la narration historique de la politique de la Fed et rend compte des facteurs qui entrent dans le processus de décision de la Federal Reserve Bank américaine. Parmi ces facteurs, la liquidité financière s'avère être un fort prédictor des mouvements des taux d'intérêt, notamment en période de détresse financière et de crise économique. En plus de l'instabilité significative autour de la règle de réaction

FIGURE 1.6 – Taux d’inflation historiques



Source : OCDE. Sélection d'économies avancées, taux d'inflation trimestriels de l'IPC : Autriche, Belgique, Canada, Danemark, Finlande, France, Grèce, Irlande, Italie, Japon, Norvège, Pays-Bas, Portugal, Suède, Royaume-Uni, États-Unis. En haut : inflation globale; en bas : hors alimentation et énergie. Les lignes pointillées sont des moyennes pour les membres du G7, du G20, de l'OCDE et les membres européens de l'OCDE.

de la Fed, ce chapitre offre également une nouvelle perspective sur la réelle nécessité d'une position active pour assurer la détermination.

Le deuxième chapitre, "Un modèle simple avec liquidité", rationalise le rôle de la liquidité dans un cadre néokeynésien standard, et explore ses implications pour la stabilité économique. Cette extension assouplit les conditions de déterminabilité du modèle, même avec une position monétaire passive. De fait, toutes les autres caractéristiques des modèles néo-keynésiens sont préservées. L'intensité de la réaction de la banque centrale à l'inflation attendue affecte toutefois la persistance des variables du modèle : une position plus accommodante se traduit par une série d'inflation plus persistante.

Le dernier chapitre, "Persistance de l'inflation", vérifie cette implication empiriquement et étend l'étude de la dynamique de l'inflation à plusieurs séries et à de nouvelles méthodes. En particulier, après avoir passé en revue les méthodes macroéconométriques standard, ce chapitre applique des outils d'apprentissage statistique pour évaluer comment la persistance de l'inflation a varié dans le temps. La principale tendance, cohérente entre les méthodes économétriques et les

mesures de l'inflation, est que la persistance de l'inflation suit un profil descendante : elle atteint un pic au milieu des années 90 et diminue depuis. Ainsi, l'inflation s'apparente actuellement à un processus de bruit blanc, qui dépend peu des réalisations passées.

Afin de situer correctement ces chapitres, les sections suivantes offrent une vue d'ensemble de la littérature existante et situent les principales contributions de cette thèse.

Théorie macroéconomique monétaire

Le paysage théorique de la macroéconomie monétaire est riche et en plein essor. Actuellement, les modèles néokeynésiens d'équilibre général dynamique et stochastique (NKDSGE) sont les outils de référence pour la réflexion sur les cycles économiques et tester des politiques alternatives. L'origine de ces modèles remonte aux modèles de cycle économique réel (RBC) (KYDLAND et PRESCOTT, 1982; LONG et PLOSSER, 1983) et, plus loin, au modèle de croissance introduit par SOLOW (1956) et développé par LUCAS (1972, 1975, 1977), bien que l'accent soit certes mis sur le court terme. Ils s'appuient sur des microfondations pour surmonter la critique LUCAS (1976), et présentent des choix optimaux intertemporels contraints sous divers degrés d'incertitude.

En un mot, les modèles NKDSGE ajoutent des frictions et des inefficacités à la structure de base des RBC, de manière à contourner la neutralité monétaire et à conférer un rôle à la politique monétaire. Dans leur forme la plus simple, les NKDSGE résument l'offre et la demande globales en deux équations linéarisées, dynamiques et anticipatives. L'offre prend la forme d'une courbe de Phillips qui dépend des anticipations d'inflation et de l'activité économique actuelle. La demande agrégée est condensée dans une équation d'Euler, qui relie la consommation actuelle aux taux d'intérêt et à la consommation future. Enfin, si on abstrait du côté fiscal du gouvernement, le modèle est complété par une troisième équation qui fixe la trajectoire des taux d'intérêt, essentiellement une représentation de la banque centrale. Le plus souvent, l'économie stylisée est un cas limite sans monnaie, ou sinon une demande de monnaie est introduite avec des soldes réels entrant dans la fonction d'utilité.

Ce cadre présente une flexibilité attrayante mais aussi un certain nombre de prédictions en désaccord avec les faits empiriques. En particulier, lorsque le taux directeur est fixe (par exemple, à sa limite inférieure zéro, ZLB), les modèles NKDSGE prédisent des effets actuels très forts des changements de politique qui sont éloignés dans le futur. Par exemple, l'effet aujourd'hui d'une réduction annoncée du taux directeur augmente avec la distance dans le futur, quand ce changement aura effectivement lieu. De plus, et c'est la clé pour placer correctement la contribution de cette thèse, la relation régissant la trajectoire des taux d'intérêt s'avère être fondamentale pour discipliner la caractérisation du modèle.

Ces paradoxes sont souvent abordés en remaniant l'hypothèse des anticipations rationnelles : en supposant une inattention rationnelle (SIMS, 2010), une rationalité limitée (GABAIX, 2016), ou

en postulant un processus cognitif explicite (GARCIA-SCHMIDT et WOODFORD, 2019). Le résultat est que les décisions actuelles sont beaucoup moins réactives aux changements qui se produisent loin dans le futur. La dynamique d'un NKDSGE autrement standard change radicalement et, en particulier, le modèle n'est pas sujet à l'indétermination si la politique monétaire est à sa limite inférieure.

Dans une règle de réaction telle que celle souvent postulée, l'intensité de la réaction du taux d'intérêt à l'inflation attendue est un facteur clé pour que l'équilibre soit unique et stable. Si la banque centrale fait varier le taux directeur de moins de un pour un par rapport aux anticipations d'inflation, l'économie présente une indétermination nominale et réelle. Cela implique que tout niveau de prix et de production est compatible avec le modèle, et donc que des chocs négligeables et transitoires peuvent conduire à des équilibres explosifs. Dans cette optique, les épisodes de ZLB dans un passé récent se sont avérés des forts contre-exemples pour le modèle de base, et ont motivé de nombreuses tentatives pour résoudre ces énigmes. Les facteurs de liquidité ont été inclus dans ces efforts, mais rarement dans des cadres simples et compacts.

Une approche complémentaire aborde la structure de base des NKDSGE et le rôle de la monnaie et de la liquidité. Comme nous l'avons souligné précédemment, les NKDSGEs s'appuient sur un noyau RBC, et tiennent à la dimension réelle des forces économiques (KOCHERLAKOTA, 2016). Ainsi, des études comme CANZONERI et al. (2011), DIBA et LOISEL (2021) et MICHAILLAT et SAEZ (2021) introduisent de la monnaie ou des obligations sous diverses formes pour contourner un ensemble de paradoxes impliqués par les NKDSGEs. Dans ce sens, les contributions récentes de CALVO (2016), qui place au centre des modèles macroéconomiques les questions liées à la monnaie et à la liquidité, sont particulièrement importantes.

Le chapitre 4 intitulé "A simple model with liquidity" fait partie d'un tel effort de modification des NKDSGE, mais présente l'avantage de rester en contact étroit avec le modèle standard, en proposant que des déviations mineures. Ces déviations suffisent toutefois à assouplir l'exigence d'une banque centrale active et agressive : le modèle affiche une déterminabilité nominale et réelle même lorsque le taux d'intérêt est presque insensible aux anticipations d'inflation. Le chapitre propose une étude systématique de l'interaction entre les actifs liquides et la politique monétaire d'un point de vue de la forme réduite. Étant donné que la liquidité et les obligations liquides (ou de manière équivalente les dépôts sur les comptes bancaires) fournissent des services transactionnels, et donc de l'utilité, leur inclusion dans la fonction d'utilité de l'agent représentatif est une option simple pour modéliser différentes formes de liquidité.

La banque centrale dans ce modèle, contrairement aux NKDSGEs, opère avec deux instruments de politique. Premièrement, elle fixe le niveau total de liquidité nominale, définie comme la somme des obligations liquides et de la monnaie en circulation. Deuxièmement, elle ajuste le taux d'intérêt sur l'obligation liquide en réaction à l'inflation attendue, plutôt que le taux d'intérêt nominal sur l'obligation illiquide, qui sert au lissage intertemporel.

De façon cruciale, l'ensemble des conditions d'équilibre qui en résulte est modifié de deux façons : l'équation d'Euler est augmentée d'un coin monétaire qui modifie sa forme habituelle ; ensuite, la demande totale de liquidité se répartit entre la monnaie et les obligations liquides, toutes deux en termes réels. L'inflation, en plus d'imprimer des mouvements aux taux d'intérêt, influence également la liquidité réelle totale. Cette boucle ne compromet pas le comportement de l'économie, qui réagit aux chocs technologiques et de politique monétaire conformément au NKDSGE standard.

Le relâchement de l'exigence d'une règle de réaction active permet d'explorer comment les agrégats sont affectés par les facteurs de liquidité et une politique monétaire potentiellement déstabilisante. Plus précisément, le modèle peut être utilisé pour reproduire une crise de liquidité similaire à celle de 2008, quand plusieurs actifs ont subi des pertes de liquidité. Il s'avère qu'une banque centrale active, ajustant rapidement le taux directeur, peut accélérer considérablement la convergence de la production, en échangeant des niveaux plus faibles de liquidité réelle et des obligations liquides bloquées. Symétriquement, une position monétaire passive prolonge la contraction de la production et peine également à relancer l'inflation, tout en préservant la liquidité réelle des obligations.

L'orientation passive de la politique monétaire affecte également la dynamique globale, comme l'illustre le chapitre. Tout d'abord, la simple prise en compte de la liquidité induit un niveau plus élevé d'inertie dans les agrégats observés. Cette inertie est encore considérable par rapport à des NKDSGE beaucoup plus riches et détaillés, qui incluent l'inflation tendancielle ou des frictions financières sophistiquées et une indexation nominale. En se concentrant sur l'inflation, le modèle montre, au moyen de simulations stochastiques, qu'une banque centrale moins réactive induit une plus forte autocorrélation, comme le suggère l'exercice d'assèchement des liquidités.

Le chapitre présente également une configuration alternative, où un mélange d'actifs liquides est utilisé pour acheter des biens de consommation, dans la tradition des modèles cash-in-advance. Ce canal alternatif ne modifie pas les idées fondamentales du modèle de forme réduite, mais renforce plutôt le rôle de la liquidité, notamment parce qu'elle affecte de manière significative la politique monétaire.

Dans l'ensemble, ce chapitre offre un compte rendu sur l'importance des conditions de liquidité dans l'économie et ses implications pour la politique. Comme le montrent les politiques non conventionnelles, les décideurs ont des considérations claires pour les questions financières en général, et l'approvisionnement en liquidités en particulier. Le modèle, qui conserve la structure de base des NKDSGEs de référence, permet des extensions prometteuses : de nombreux blocs supplémentaires développés pour les NKDSGEs existants peuvent être inclus dans sa configuration de manière relativement facile.

Cela n'est guère possible dans d'autres approches qui abandonnent les hypothèses fondamentales qui ont caractérisé les modèles au cours des 40 dernières années, comme les anticipations

rationnelles. Cela ne signifie pas que de nouvelles façons de modéliser la formation des anticipations ou d'autres modifications des hypothèses comportementales du modèle traditionnel ne méritent pas d'être explorées. En se concentrant sur l'aspect monétaire et financier des modèles et en introduisant des concepts, tels que la liquidité, qui sont bien fondés dans les théories microéconomiques et financières (EGGERTSSON et KRUGMAN, 2012 ; HOLMSTRÖM et TIROLE, 2011). Le modèle présenté dans ce chapitre se prête à des extensions et des raffinements ultérieurs, notamment pour l'analyse de la politique budgétaire et des banques.

Néanmoins, la voie la plus attrayante pour l'avenir réside dans une endogénéisation complète de la gestion de la liquidité nominale totale par la banque centrale. Cela fournirait à la politique monétaire un second instrument, reproduisant plus fidèlement la conduction réelle des QEs ces dernières années.

Vues les considérations présentées, le fait d'atteindre la ZLB en 2008 et 2020 tout en mettant également en œuvre des politiques de liquidité a fourni des points de données extrêmement précieux pour mettre à l'épreuve empirique les prédictions des NKDGE.

La section suivante donne un aperçu des contributions les plus importantes sur le plan empirique de la macroéconomie monétaire, et situe les premier et troisième chapitres dans cette littérature.

Macroéconomie Monétaire Empirique

Les outils traditionnellement utilisés pour étudier les questions empiriques macroéconomiques sont tirés du milieu des séries temporelles. Cela inclut les modèles de forme réduite ou les modèles vectoriels autorégressifs structurels (SVAR), ainsi que les modèles moins récents d'équations simultanées dynamiques. En se limitant aux questions monétaires, les (S)VAR sont les outils de référence, car ils nécessitent des hypothèses simples mais strictes, et les NKDSGE peuvent souvent être reformulés sous forme de SVAR. Au cours des deux dernières décennies, en particulier, les méthodes bayésiennes ont encore repoussé la frontière de l'investigation empirique, en fournissant des outils puissants pour estimer des modèles plutôt complexes avec des hypothèses assez transparentes.

En ce qui concerne la politique monétaire, il est utile de distinguer deux approches : l'estimation à information limitée et celle à information complète. La seconde fait généralement intervenir un système de relations structurelles et s'appuie sur un appareil théorique – soit sur les restrictions (VAR), soit sur l'influence mutuelle des chocs et des agrégats observables. La première, au contraire, se concentre sur une relation macroéconomique particulière et évalue sa validité de forme réduite, en l'isolant des autres forces, bien qu'avec des hypothèses de modélisation éventuellement complexes.

Les règles de politique monétaire comme celle de TAYLOR (1993) et les courbes dites de PHILLIPS (1958) sont des exemples particulièrement marquants de l'approche fondée sur l'information limitée : ces deux relations sont intimement liées et constituent des éléments cruciaux du cadre monétaire moderne.

Les sections suivantes offrent un large aperçu de l'état actuel des règles de politique monétaire et des courbes de Phillips, en relation avec les premier et troisième chapitres de cette thèse.

Règles de politique monétaire

Les règles de réaction constituent actuellement l'instrument de première ligne de la politique monétaire. Par conséquent, l'identification correcte du fonctionnement des banques centrales dans l'environnement économique est essentielle pour la prévisibilité et la transparence de la politique. Dans le même ordre d'idées, des estimations correctes de la règle de politique monétaire sont déterminantes pour évaluer l'efficacité avec laquelle les banques centrales maintiennent l'inflation près de son objectif.

En observant les données telles qu'illustrées par Figure 2.2, il est possible d'apprécier les différentes phases traversées par l'économie américaine depuis la Seconde Guerre mondiale. La figure présente les séries pour le taux des fonds fédéraux, l'indice global des prix à la consommation et une mesure ex-post de l'activité économique, l'écart de production. Jusqu'au début des années 1970, l'inflation était relativement stable sous 5% et pas particulièrement volatile. De même, le FFR a suivi une trajectoire ascendante régulière, suivant la tendance apparente de l'inflation. En revanche, la dynamique de l'activité économique apparaît relativement plus volatile, avec de fortes oscillations en territoire négatif et positif. Ensuite, depuis les années 1970, l'inflation suit une tendance à la hausse sur une décennie, suivie de près par le taux directeur. L'écart de production affiche toujours de grandes variations, mais principalement en territoire négatif, signalant une économie en dessous de son plein potentiel.

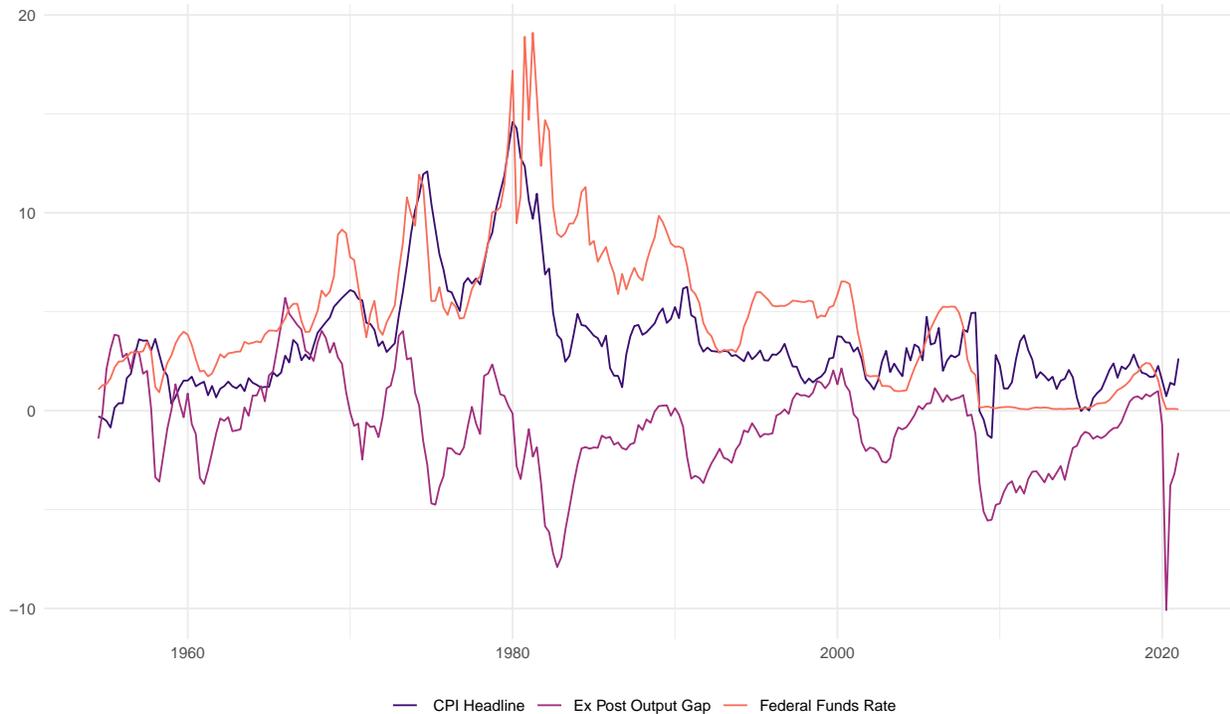
Une rupture semble se produire au début des années 1980 : le FFR s'envole, l'écart de production signale une récession, et une déflation a lieu. En effet, le taux d'inflation revient aux niveaux précédents avec une volatilité réduite, tandis que le FFR évolue de manière plus régulière et réagit à la fois à l'inflation et à l'activité économique. Ce dernier présente également une volatilité nettement inférieure à celle des périodes précédentes, avec des tendances à la hausse plus longues se remettant de légères récessions. Cette période, souvent qualifiée de Grande Modération, contraste en effet fortement avec les années précédentes : moins volatile en général, avec une inflation plus faible. Le changement correspond à peu près à la nomination de Volcker en tant que président de la Fed, ce qui suggère qu'il a provoqué une rupture du régime de politique.

Les dernières décennies contrastent également avec la Grande Modération, mais pour des aspects différentes. En particulier, au début de la crise de 2008, la Fed a poussé le taux directeur jusqu'à zéro, qui a été la borne pendant plusieurs années par la suite. Pourtant, l'inflation n'est pas

devenue incontrôlable comme dans les années 1970. La différence essentielle dans cette dernière partie de l'échantillon est résumée par Figure 2.3 : depuis mi-2008, la Fed a gonflé son bilan afin d'injecter des quantités sans précédent de liquidités dans le système financier. Cette vaste expansion du bilan était composée de plusieurs opérations distinctes, mais elle est généralement désignée sous le nom de politique d'assouplissement quantitatif. Il est intéressant de noter que des politiques similaires ont été déployées par de nombreuses autres grandes banques centrales - notamment la Banque Centrale Européenne, la Banque d'Angleterre et la Banque du Japon.

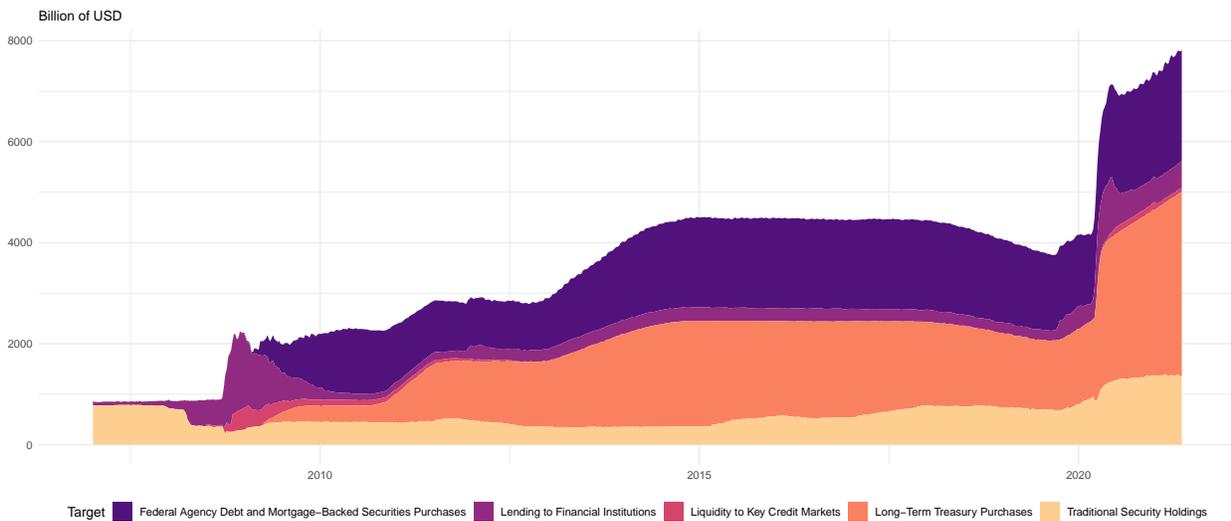
Les effets de ces politiques ne sont cependant pas encore bien compris, même s'elles signalent deux faits importants : premièrement, la fixation des taux d'intérêt n'est pas le seul outil politique à la disposition des banques centrales; deuxièmement, la liquidité, au sens large, est un facteur important pour le contrôle de l'inflation et la stabilité économique, du moins dans le cadre des banques centrales. Comme le chapitre 4 a exploré théoriquement le rôle de la liquidité dans les NKDSGEs, le chapitre 3 adopte une approche empirique à l'étude de la liquidité et de la règle de politique monétaire, au sens financier.

FIGURE 1.7 – Courbe de Phillips américaine et politique monétaire



Taux des fonds fédéraux, IPC global, *ex post* écart de production pour l'économie américaine. L'écart de production est calculé comme un écart en pourcentage de la production réelle révisée par rapport aux estimations de la production réelle potentielle du Congressional Budget Office. Sources : Federal Reserve Bank, CBO.

FIGURE 1.8 – Les mesures de liquidité de la Fed



Bilan du Federal Reserve System, depuis 2007T1. Classification des avoirs en principaux types de groupements : Dette de l'agence fédérale et MBS; Prêts aux institutions financières; Liquidité du marché du crédit; Avoirs du Trésor à long terme; Avoirs traditionnels. Valeurs en milliards de dollars courants. Source : Fed de Cleveland

La littérature moderne sur l'estimation des règles de politique monétaire est fleurie à partir de l'article de TAYLOR (1993), qui, dans un cadre simple, estimait la fonction de réaction de la Fed. Dans cette contribution, la spécification soumet à un test empirique les idées théoriques alors récentes sur la crédibilité de la politique, l'efficacité des politiques fondées sur des règles par rapport au pouvoir discrétionnaire, et l'accent mis sur les taux d'intérêt plutôt que sur la croissance monétaire. En effet, Taylor postule une règle de réaction qui met en œuvre une forte réponse du taux des fonds fédéraux au taux d'inflation en glissement annuel, calibré à 1,5. En supposant une réponse à la production d'environ .5 et une cible à long terme de 1% pour l'inflation, cette règle simple explique très bien au comportement réel de la Fed sur la période 1983T1-1992T3. Le résultat, en prise directe avec la recherche naissante des NKDSGE, a suscité un débat animé sur les règles de politique monétaire, à tel point que les règles de politique monétaire de ce type sont souvent appelées règles de Taylor. Parallèlement à une riche discussion théorique, TAYLOR (1993) a inspiré une branche de la littérature empirique sur l'estimation des règles de politique monétaire. L'objectif initial de cette recherche était de comprendre comment les régimes de politique changeaient aux États-Unis, en particulier pour déterminer le régime monétaire dominant pendant la période de forte inflation des années 1970. CLARIDA, GALI et GERTLER (2000) est l'une des principales contributions à cette littérature, posant les bases des méthodes employées et du consensus qui en découle sur l'orientation de la politique. En un mot, l'estimation des règles de politique est une tâche complexe en raison de l'endogénéité intrinsèque, qui biaise les estima-

tions standard. En effet, lorsque les termes d'erreur sont corrélés avec les régresseurs, l'asymptotique des MCO simples ne tient pas et les paramètres estimés sont biaisés. Pour surmonter ces problèmes, CLARIDA, GALI et GERTLER (2000) déploient une approche GMM-IV, prenant les valeurs passées des variables endogènes comme instruments pour leurs valeurs attendues. En outre, ils inaugurent une stratégie d'échantillonnage, qui consiste à estimer la règle de politique sur deux sous-échantillons : le point de séparation correspond à la première présidence de Volcker. L'approche est assez acceptable car elle entre en contact avec le discours conventionnel et les anecdotes de l'époque. Les conclusions, en résumé, sont que la première période voit une banque centrale accommodante, correspondant à une inflation élevée et une grande volatilité. L'estimation clé de la réaction à l'inflation attendue est bien inférieure à un : en s'appuyant sur l'appareil théorique, cette orientation de la politique engendre des équilibres autoréalisateurs et une indétermination sévère. Sur le second sous-échantillon, initié par la désinflation de Volcker, la Fed réagirait fortement à l'inflation attendue, excluant ainsi l'indétermination et maintenant effectivement l'inflation sous contrôle. Cette analyse, prise dans son ensemble, éclaire une grande partie du consensus établi sur la conduite de la politique monétaire aux États-Unis ; elle permet également de formuler certaines critiques. Par exemple, l'instrumentation de l'inflation attendue par ses valeurs décalées conduit à des instruments faibles, puisque le taux d'inflation est souvent mieux décrit par un bruit blanc. En outre, deux choix se sont avérés particulièrement critiques : l'utilisation de données révisées et le choix exogène des sous-échantillons.

Les recherches suivantes ont permis d'améliorer ces deux aspects. Reproduire aussi fidèlement que possible l'ensemble des informations dont dispose le décideur au moment de chaque décision est en contradiction avec les données révisées, qui subissent des corrections au fur et à mesure que des informations plus précises sont disponibles. En ce sens, ORPHANIDES (2001) améliore CLARIDA, GALI et GERTLER (2000) en collectant des données en temps réel sur les prévisions d'inflation et sur l'écart de production dont dispose le Federal Open Market Committee à chaque réunion, depuis la base de données du Greenbook. Le biais temporel est important dans les estimations des règles de politique, dans la mesure où une règle de Taylor estimée sur la période 1987T1-1993T4 avec des données et des prévisions en temps réel indique une position passive pour la Fed – au moins à des horizons de prévision inférieurs à trois trimestres. Dans une contribution complémentaire, ORPHANIDES (2004) reproduit l'exercice sur les sous-échantillons de CLARIDA, GALI et GERTLER (2000), en utilisant uniquement des données en temps réel. Les deux régimes en place avant et après la nomination de Volcker s'avèrent plus similaires que prévu. Les estimations sont largement similaires entre les deux régimes : les deux sous-échantillons produisent systématiquement des estimations de réaction proactive à l'inflation attendue, ce qui minimise son pouvoir stabilisateur. Les résultats avec des données en temps réel indiquent en fait une réactivité excessive à la production au cours des années 1970, ce qui laisse supposer que les décideurs politiques étaient peut-être trop confiants quant au lissage des fluctuations du cycle économique.

Ces contributions s'appuyaient essentiellement sur les MCO ou les moindres carrés non linéaires (MCNL), sans instrumenter l'inflation attendue, mais plutôt en utilisant les prévisions réelles du personnel du FOMC. L'utilisation de données en temps réel ne résout cependant pas entièrement le problème de l'endogénéité, car les prévisions sont généralement conditionnelles à la trajectoire future du taux directeur, ce qui risque de biaiser les estimations.

La question d'une stratégie de sous-échantillonnage basé sur faits historiques, mais exogène, a été abordée avec une variété d'approches économétriques. Parmi celles-ci, les méthodes Time Varying Parameters (TVP), ainsi que des adaptations des estimations Markov Switching, ont donné des résultats intéressants. BOIVIN (2006) et KIM et NELSON (2006) présentent des résultats à information limitée basés sur des estimations TVP. Les premiers s'écartent de l'appareil GMM-IV pour s'aventurer dans le traitement des modèles TVP avec des régresseurs endogènes et hétéroscédasticité dans les chocs. Pour résumer les principales conclusions, ils exploitent les erreurs de prévision standardisées dans l'étape IV pour éliminer l'endogénéité et corriger le biais qu'elle entraîne. Les estimations qui en résultent brossent un tableau encore plus nuancé. Il semble y avoir trois phases distinctes dans le régime de la Fed après la Seconde Guerre mondiale : durant la première phase, de 1970 à 1975, la réaction de la Fed à l'inflation n'était pas statistiquement différente de un ; pendant la deuxième phase, de 1975 environ jusqu'au début des années 1980, la réaction a légèrement augmenté, mais n'était toujours pas significativement différente de un. En revanche, le dernier régime, qui couvre au moins la période 1980-2002, se caractérise par une position proactive de la Fed, mais avec une incertitude croissante autour de l'estimation ponctuelle. Ce dernier point est peut-être dû à la faible volatilité de l'inflation, qui affaiblit encore plus le pouvoir des retards en tant qu'instruments.

Enfin, BOIVIN (2006) combine la TVP robuste à l'hétéroscédasticité avec des données en temps réel, résumant les contributions de KIM et NELSON (2006) et ORPHANIDES (2004). Contrairement à KIM et NELSON (2006), Boivin trouve une position active vis-à-vis de l'inflation au début des années 1970 et à partir de 1982 environ, avec une forte diminution entre les deux. La transition vers une position proactive, en particulier, ne semble pas discrète, tandis que l'incertitude augmente encore à partir des années 1990.

Un courant ultérieur de la littérature a adopté une approche similaire, mais s'est appuyé sur des changements de régime discrets au lieu de transitions progressives. C'est le cas de DAVIG et LEEPER (2006, 2011) et MURRAY, NIKOLSKO-RZHEVSKYY et PAPELL (2015). Ces premiers, en particulier, considèrent conjointement les règles de politique budgétaire et monétaire, en se concentrant sur les régimes Markoviens actifs et passifs pour les deux. Bien que la combinaison des règles budgétaires et monétaires nuance le tableau général, ils trouvent des preuves d'une position active sur les périodes 1979-1990, 1995-2001, et d'une position passive pendant les périodes 1949-1979, 1990-1995, 2001 et suivantes, confirmant en partie les premiers résultats de CLARIDA,

GALI et GERTLER (2000). Au niveau agrégé, chaque régime est particulièrement persistant, malgré les changements fréquents.

Suivant cette tendance de changements de politique endogènes, MURRAY, NIKOLSKO-RZHEVSKYY et PAPELL (2015) appliquent le modèle de changement de régime Markovien développé par HAMILTON (1989) à la seule règle de Taylor. En mélangeant les données en temps réel et le changement de régime, ils confirment dans un exercice à information limitée l'existence probable de régimes multiples au sein de la périodisation habituelle. La distribution de ces régimes est en contradiction avec la narration construite sur CLARIDA, GALI et GERTLER (2000) : l'attitude de la Fed a été pratiquement toujours active pendant les épisodes de forte inflation. En particulier, la Fed a contrôlé de manière proactive l'inflation attendue au cours des périodes 1965T4-1972T4, 1975T1-1979T3, et 1985T2 jusqu'à la fin de leur échantillon.

Le chapitre 3 s'appuie sur cette approche et augmente la spécification de base de la règle de Taylor pour tenir compte des facteurs de liquidité, qui se sont avérés importants dans la prise de décision de politique monétaire.

L'inclusion d'indices de stress financier a déjà été explorée dans la littérature sur la politique monétaire, mais le rôle de la liquidité a été négligé jusqu'en 2008. Quelques exceptions notables sont BAXA, HORVÁTH et VAŠÍČEK (2013) et ISTREFI, ODENDAHL et SESTIERI (2020), qui établissent un lien entre les préoccupations en matière de stabilité financière et les décisions de politique monétaire – en testant une batterie d'indices financiers pour le premier ; en analysant les informations des procès-verbaux officiels pour le second. L'ensemble des politiques non conventionnelles comprenait une grande part de gestion de la liquidité : l'une des principales préoccupations de la Fed était de débloquer les marchés financiers et de relancer les transactions. Cette préoccupation est évidente dans les notes qui introduisent chaque innovation politique à la suite de la crise de 2008 : les notes *AMLF* ; *CPFF* ; *MMIFF* ; *PDCF* ; *TAF* ; *TSLF* mentionnent toutes la liquidité comme cible de ces politiques. L'extension de la règle de Taylor traditionnelle pour prendre en compte le rôle des marchés financiers se résume essentiellement à donner une approximation de l'état de la liquidité dans le système financier.

L'inclusion de la liquidité est raisonnable vus les développements des politiques, mais sa contrepartie empirique est particulièrement épineuse. La liquidité est une caractéristique plutôt éphémère des actifs. D'une part, elle dépend du côté dont on la considère : un actif très demandé est facile à vendre, donc liquide ; à l'inverse, il est difficile à acheter, donc illiquide. Une définition opérationnelle plus étroite de la liquidité repose sur la décote qu'il faut supporter pour transformer un actif en espèces. Pour surmonter ces problèmes, le chapitre propose de capturer une mesure proche de la liquidité par le biais d'écarts sur des actifs largement échangés et à faible risque, afin de choisir une mesure dépourvue de risque et d'endogénéité autant que possible. Ces mesures sont deux spreads entre des actifs extrêmement sûrs, à savoir les titres du Trésor américain, et des actifs moins sûrs. Plus précisément, les mesures utilisées dans les estimations sont les écarts entre

le rendement moyen des obligations d'entreprises BAA et les rendements trimestriels de l'indice Standard & Poor's 500. L'échéance à l'émission pour ces dernières est un facteur important pour déterminer la prime par rapport aux actifs plus sûrs. Pour tenir compte de ce facteur temporel, les écarts sont calculés sur des titres du Trésor d'échéance comparable, soit dix ans ou trois mois.

L'inclusion de ces mesures de liquidité dans la règle de politique monétaire permet donc de saisir le rôle des préoccupations financières dans l'élaboration de la politique. L'intuition s'appuie en partie sur la littérature sur le *leaning-against-the-wind* (SVENSSON, 2017b), mais fait appel à l'ensemble des données de la Fed en ce qui concerne les préoccupations en matière de liquidité.

Les principaux résultats du chapitre découlent de l'estimation d'un modèle changement de régime Markovien, avec des règles de Taylor augmentées par les mesures de liquidité. Ce modèle empirique présente l'avantage de s'appuyer uniquement sur les informations contenues dans les données pour endogénéiser les changements de régime. Alors que la spécification standard, incluant les prévisions du Greenbook pour l'écart de production actuel et l'inflation attendue, ne trouve pratiquement que des régimes proactifs, en désaccord avec le consensus établi. Les mesures de liquidité, cependant, fournissent un récit différent : lorsque l'indicateur de liquidité à court terme est inclus, il y a des preuves de régimes actifs et passifs. Il est intéressant de noter que le régime passif semble prévaloir sur l'échantillon postérieur à la Seconde Guerre mondiale, tandis que le régime actif prévaut pour certaines périodes de courte durée, précisément la première moitié des années 1970 et pour la période 1979-1982. Si l'on prend l'indicateur de liquidité à long terme, basé sur les obligations BAA, l'importance de l'inflation attendue est sévèrement contestée, à tel point que la Fed semble y réagir négativement – remarquablement pendant la période du ZLB. Cependant, dans l'ensemble des spécifications, la liquidité financière est systématiquement considérée comme un prédicteur significatif des mouvements du taux directeur. Il est intéressant de noter que les régimes passifs étaient probablement en place pendant les périodes d'inflation sous contrôle : ce résultat jette des doutes sur la rigueur de la politique monétaire active pour dompter l'inflation. Les estimations suggèrent également une association entre la probabilité de différents régimes de paramètres et la direction des changements de politique. Les périodes de hausse des taux d'intérêt sont associées à une probabilité élevée de se trouver dans un régime de type Volcker (coefficient élevé de l'inflation dans la fonction de réaction de la politique). En revanche, les périodes d'assouplissement monétaire par une baisse des taux d'intérêt sont associées à une probabilité plus élevée d'être dans un régime moins réactif.

Dans le but de résumer les stratégies empiriques établies, le chapitre propose également des analyses en échantillon complet et en sous-échantillons, qui corroborent largement le résultat principal, en plus de l'importance de la liquidité financière pour l'élaboration de la politique monétaire. L'analyse en échantillon complet révèle que l'inclusion de l'un ou l'autre proxy de la liquidité réduit sévèrement le coefficient de l'inflation attendue : les deux spécifications estiment une réaction à l'inflation légèrement inférieure à un, bien plus faible que les estimations de la règle

de Taylor classique. Dans l'exercice par sous-échantillons, les répartitions suivent la stratégie empirique établie, mais ajoutent un troisième sous-échantillon couvrant les périodes post-2008. Cette dernière partie est naturellement moins informative en raison de la disponibilité des données (les données du Greenbook sont publiées avec un décalage de cinq ans) et présente une variation faible ou nulle du taux directeur. Conformément à ORPHANIDES (2004), la spécification standard trouve une position active pendant le sous-échantillon pré-Volcker, et une position active plus forte pendant la Grande Modération. Les spécifications qui incluent des proxys financiers rapportent des réponses faibles à l'inflation attendue dans la période pré-Volcker, et, de la même manière, des réponses plus fortes à partir des années 1980, bien que les estimations ponctuelles soient significativement plus petites que la spécification standard.

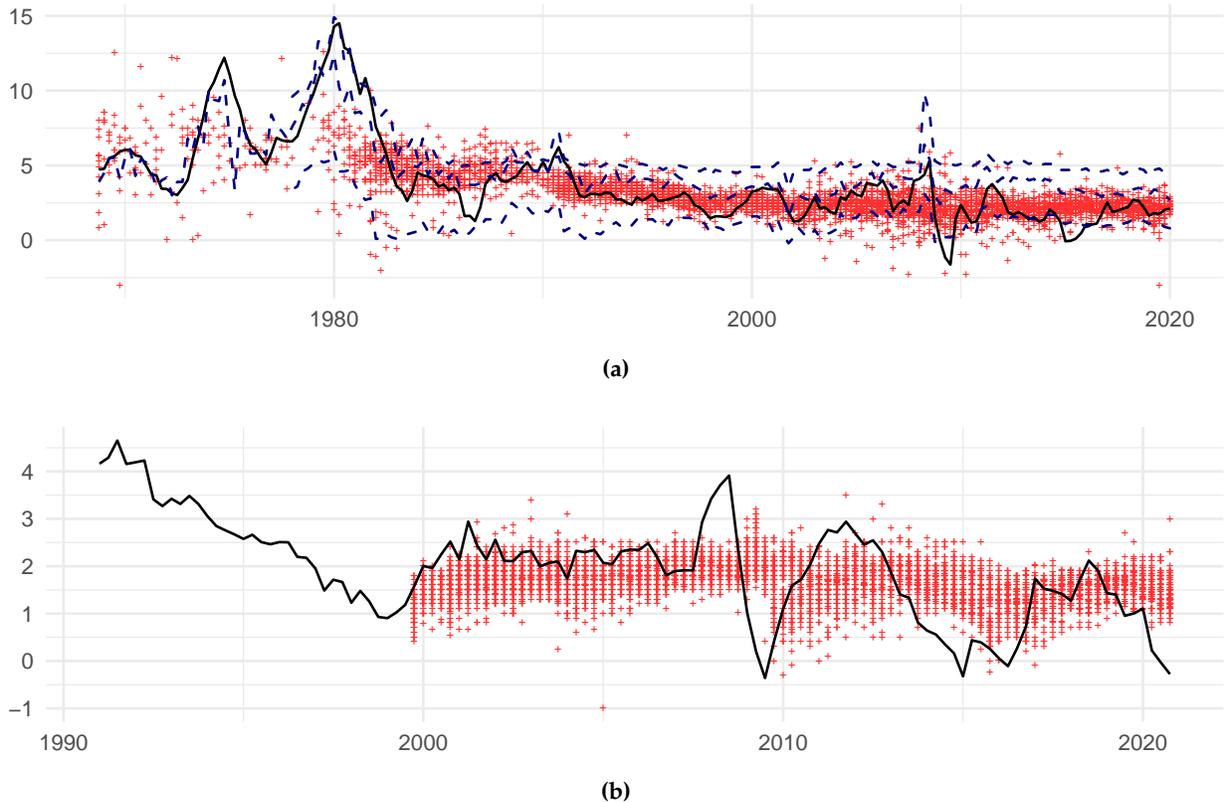
Dynamique de l'inflation

Tout comme la règle de Taylor pour la politique monétaire, la courbe de Phillips est le mécanisme clé de la dynamique de l'inflation : elle met en relation l'inflation actuelle avec une certaine mesure de l'activité économique et, de manière plus subtile, avec l'inflation attendue. Notamment, la version augmentée des anticipations de cette relation est le canal des éventuelles anticipations autoréalisatrices lorsque l'orientation de la politique monétaire est passive. Plus récemment, la courbe de Phillips a mobilisé de nombreux efforts de recherche : a-t-elle disparu ? S'est-elle aplatie et de combien ? Ces questions découlent du comportement inattendu de l'inflation au cours des deux dernières décennies : les variations des prix courants semblent moins liées à l'activité économique. L'exemple le plus clair de cette connexion affaiblie est la déflation et la reflation manquantes après la récession de 2008 : avant la pandémie de 2020, en effet, les États-Unis ont affiché la plus longue expansion et ont fait passer le taux de chômage en dessous de 3%, le tout avec une inflation maîtrisée. C'est le cas pour l'économie européenne, également (CICCARELLI et OSBAT, 2017).

En outre, le canal des anticipations ne fournit pas d'indications, mais pose plutôt des questions supplémentaires. Comme le montre la 2.5a pour les États-Unis, les anticipations des prévisionnistes professionnels, celles des consommateurs ordinaires et l'inflation réelle suivent des chemins sensiblement différents. La même chose, pour les prévisionnistes professionnels uniquement, s'applique à l'inflation dans la zone euro, comme le montre la 2.5b.

Dans le cadre de la courbe de Phillips, l'inflation observée est donc influencée par l'activité économique actuelle et l'inflation future. Une simple itération vers l'avant de cette relation de base exprime donc l'inflation actuelle comme la séquence actualisée des déviations du niveau naturel de l'activité économique. C'est-à-dire les fluctuations autour de ce que l'on appelle la production potentielle, souvent exprimée comme l'équilibre sans friction et non stochastique de l'économie. L'implication directe est que, tant que les chocs sont de moyenne nulle et sériellement

FIGURE 1.9 – Attentes d’inflation



Panneau supérieur : Données de l'IPC américain, la ligne continue correspond à l'IPC réel, en glissement annuel; les points rouges correspondent à l'enquête individuelle sur les prévisionnistes professionnels; les lignes pointillées bleues correspondent au quartile supérieur, à la moyenne et au quartile inférieur de l'enquête sur les consommateurs. Panneau inférieur : Données IPC et SPF pour la zone euro. Les prévisions sont le taux d'inflation IPC à un an d'avance, décalé à la date de prévision correspondante. Sources : FRED, Fed de Philadelphie, Banque centrale européenne, Université du Michigan.

non corrélés, la dynamique de l'inflation reflète la dynamique sous-jacente du cycle économique et de ses déterminants.

Un défi supplémentaire dans l'estimation de la courbe de Phillips néokeynésienne provient, quelque peu paradoxalement, de l'efficacité des politiques. En effet, si les anticipations d'inflation sont exceptionnellement bien ancrées et que la politique monétaire compense les chocs d'offre, alors l'inflation observée semblera "danser à sa propre musique", mue uniquement par des chocs transitoires.

Ces implications s'accordent bien avec le fait que l'inflation est notoirement difficile à prévoir avec une précision satisfaisante, et de plus en plus difficile à prévoir (STOCK et WATSON, 2007).

Cependant, la littérature a mis en avant une quantité d'approches économétriques pour estimer la courbe de Phillips, y compris des spécifications alternatives. Il est utile de classer ces

approches en deux catégories. Premièrement, utiliser les observations passées pour instrumenter les anticipations d'inflation dans un cadre IV ; deuxièmement, exploiter les enquêtes existantes pour jauger les anticipations globales. Chaque stratégie économétrique, cependant, aborde certains problèmes mais pose de nouveaux défis.

L'instrumentation de l'inflation attendue avec sa réalisation passée est attrayante car elle correspond à la restriction d'exclusion. Une des premières applications a été proposée par GALI et GERTLER (1999) : ils développent une courbe de Phillips hybride avec une part d'entreprises rétrospectives et exploitent les erreurs de prévision pour remplacer l'inflation anticipée. Les anticipations sont ensuite instrumentées avec des variables retardées dans un cadre GMM-IV. En supposant des anticipations rationnelles, cette stratégie répond à la restriction d'exclusion mais, vue la faible prévisibilité de l'inflation, elle n'offre qu'un instrument faible. Cependant, l'inclusion désormais standard des entreprises rétrospectives est essentielle pour générer l'inertie et correspondre au comportement observé, bien que quantitativement dominé par la force prospective. Dans ce cadre, une innovation récente exploite les révisions de données comme instruments externes pour l'inflation attendue, mais avec un succès empirique contestable (MAVROEIDIS, PLEGBORG-MOLLER et STOCK, 2014).

Les estimations fondées sur des enquêtes ont été lancées par ROBERTS (1995) et ont récemment suscité un intérêt plus général (COIBION, GORODNICHENKO et KAMDAR, 2018) pour l'étude de l'inflation et des anticipations. L'idée clé est d'utiliser les anticipations réelles – des prévisionnistes professionnels, des consommateurs, des entreprises – pour cibler les anticipations globales. Bien que séduisante, cette approche directe pose des défis subtils en matière de rationalité. Les anticipations agrégées sont rationnelles et cohérentes avec le modèle, des propriétés qui ne peuvent pas être transposées directement aux participants à l'enquête mais qui sont requises pour une microfondation solide. En outre, les écarts par rapport à la rationalité diffèrent selon la population interrogée : le biais des consommateurs est différent de celui des prévisionnistes professionnels.

En ce qui concerne les diverses spécifications, la stratégie la plus fréquente pour faire correspondre les données consiste à s'écarter de la formulation purement prospective de la courbe de Phillips. Celle-ci est généralement obtenue en supposant qu'une partie des entreprises qui ne peuvent pas se réoptimiser complètement indexeront leurs prix sur l'inflation passée. L'inertie intrinsèque qui résulte de cet écart améliore généralement le fit sur les données américaines. Une autre extension assouplit l'hypothèse d'état stationnaire sans inflation : lorsque l'inflation à l'état stationnaire est positive, elle correspond au niveau d'inflation à long terme, et la courbe de Phillips est généralement exprimée en écarts par rapport à cette tendance d'inflation. Les changements de cette tendance ont toutefois des conséquences importantes sur l'estimation et, surtout, sur les propriétés dynamiques de l'inflation, car l'inflation tendancielle correspond en fin de compte à l'ancrage des anticipations d'inflation.

Le rôle de l'inflation tendancielle a suscité un débat intense qui a exploré les implications pour la persistance et la volatilité de l'inflation observée, ainsi que son interaction avec la politique. Plusieurs contributions ont souligné comment la principale force motrice de la dynamique de l'inflation provient des mouvements de sa tendance à long terme. De la même manière, la modélisation de l'inflation tendancielle a des conséquences importantes pour une politique optimale.

En s'appuyant sur la revue de ASCARI et SBORDONE (2014), il est clair que la dynamique des tendances a un impact sur la dynamique de l'inflation observée. La conclusion qui se dégage est que l'inflation observée est fortement influencée par la dynamique de la tendance sous-jacente à long terme, en particulier pour la volatilité. COGLEY, PRIMICERI et SARGENT (2008), COGLEY et SBORDONE (2009) et STOCK et WATSON (2007) offrent des preuves empiriques de l'interaction entre l'inflation tendancielle et la persistance, ce qui est fondamental pour la conception d'une politique monétaire efficace. En effet, le temps nécessaire à l'inflation pour atteindre un nouvel équilibre après un choc est crucial pour une intervention appropriée de la politique monétaire. De même, des mesures appropriées de la persistance fournissent aux banques centrales des compromis corrects entre l'inflation et l'activité : l'impression d'une inflation faible et peu réactive pourrait pousser une banque centrale à un assouplissement excessif, et vice versa.

Les variations de la persistance de l'inflation américaine ont fait l'objet de contributions de FUHRER (2011) et PIVETTA et REIS (2007), sur lesquelles se base une grande partie du troisième chapitre. Leur objectif est précisément l'estimation d'un aspect particulier de la dynamique de l'inflation, à savoir la persistance. Alors que FUHRER (2011) couvre les paramètres d'information limitée et complète pour les courbes de Phillips anticipative et rétrospectives, PIVETTA et REIS (2007) adopte une approche univariée sur l'inflation. Ils construisent l'analyse sur un modèle autorégressif bayésien pour l'inflation et trouvent peu de changements dans la persistance pour plusieurs mesures.

Le troisième chapitre, intitulé "Persistance de l'inflation", reprend une grande partie de outils empirique présentée dans FUHRER (2011) et PIVETTA et REIS (2007), avec deux contributions. L'analyse élargit l'éventail des indices d'inflation pour les estimations, afin de démêler les éventuels facteurs de confusion comme la mondialisation et les prix de l'énergie et des produits alimentaires. Les différents paniers et biens mesurés par l'indice des prix à la consommation, l'indice des dépenses de consommation personnelle et le déflateur du produit intérieur brut offrent un moyen simple de distinguer les forces qui influencent la dynamique de l'inflation. Une distinction plus subtile entre l'IPC headline et l'IPC core est tout aussi instructive pour isoler la transmission des prix des produits alimentaires et de l'énergie (souvent volatils) à l'inflation globale observée. De la même manière, les données utilisées couvrent également une période plus longue, incluant la récession mondiale de 2008. Cette dernière partie des données est particulièrement instructive, car elle inclut la période prolongée de ZLB et les hausses de taux qui ont suivi. Dans le même socle que PIVETTA et REIS (2007), le chapitre représente l'inflation comme un processus autorégressif

univarié, isolé de la politique budgétaire et monétaire, de l'activité économique et des anticipations des agents. Un résumé des résultats est présenté dans la 2.5.

Le chapitre débute par une analyse simple et fréquentiste de la persistance de l'inflation. L'objectif est d'obtenir une première preuve des variations de l'inertie de l'inflation, qui sont effectivement présentes dans les cinq mesures utilisées. En particulier, la mesure de la persistance en tant que premier coefficient autorégressif ou en tant que somme de tous les coefficients sur les retards révèle un profil concave pour l'inertie de l'inflation. Cette évidence motive une étude plus approfondie du comportement de l'inflation, avec des méthodes plus sophistiquées. En effet, la structure autorégressive confère une forme statistique rigide aux estimations, mais sa portée est limitée : elle n'exploite que peu les informations présentes dans les données ou les sous-échantillons qui lui sont présentés. Pourtant, la dynamique de l'inflation peut être affectée par des mouvements à long terme de la tendance ou par des changements discrets, liés à la politique monétaire ou fiscale.

Le recours aux méthodes bayésiennes relâche une bonne partie de ces contraintes. S'inspirant de la formalisation statistique proposée par PIVETTA et REIS (2007), le chapitre adopte la perspective du décideur politique et, à tout moment, subsume les informations passées sur l'inflation dans un processus autorégressif avec des coefficients possiblement explosifs. À partir de là, la méthode explore les trajectoires futures de l'inflation qui sont cohérentes avec l'état actuel. De tels chemins permettent de calculer la persistance à chaque période : plus précisément, la structure bayésienne fournit une distribution pour l'inertie de l'inflation. Cette distribution est particulièrement utile, car elle combine les croyances informées par les données jusqu'à présent et l'incertitude quant aux développements futurs. Plus précisément, cette approche permet des trajectoires explosives pour l'inflation, celles qui sont plus importantes pour les décideurs politiques soucieux d'éviter les épisodes de forte inflation ou déflation. Les résultats sont globalement conformes à l'exercice fréquentiste, avec quelques distinctions. Si le profil concave est confirmé, la transition est sensiblement plus lisse. Cela est dû en partie à la transition de marche aléatoire que suivent les paramètres autorégressifs. Cependant, le degré d'incertitude augmente considérablement vers la fin de l'échantillon, ce qui reflète des niveaux relativement faibles d'inertie et une forte imprévisibilité.

Pour les deux exercices, le profil de persistance se trouve pendant les cinq premières années des années 1990, au début de la Grande Modération. Le moment suggère des forces autres que la politique monétaire, le commerce international ou la volatilité des matières premières. La politique ne fonctionne guère avec des décalages aussi longs, ni l'apprentissage des agents peut prendre une décennie pour être pleinement intégré. Les pressions mondiales sur les prix auraient pu être déjà à l'œuvre, mais la Chine n'est pas entrée sur la scène du commerce international avec tous les droits de l'OMC avant 2001. De même, les prix du pétrole et d'autres matières premières étaient relativement stables, si ce n'est pour une hausse transitoire autour de la première guerre du Golfe.

Afin d'assouplir davantage la structure imposée au processus de génération des données, le chapitre exploite la flexibilité des méthodes d'apprentissage statistique les plus récents (réseaux

neuronaux profonds), ce qui constitue la contribution clé de ce travail. L'avantage de cette classe de méthodes réside dans le large éventail de non-linéarités qu'elles peuvent approximer. Par conséquent, les réseaux neuronaux peuvent traiter des structures de données assez complexes avec facilité, en s'écartant de la structure linéaire autorégressive intégrée dans les exercices fréquentistes et bayésiens. Parmi les différents réseaux, le chapitre présente les modèles LSTM (Long-Short Term Memory), de la classe des réseaux neuronaux récurrents. Généralement employés pour le traitement du langage naturel, les LSTM sont explicitement conçus pour jauger les structures complexes des données qu'ils traitent. Dans plusieurs applications, ils surpassent systématiquement les outils de prévision traditionnels, notamment lorsqu'il s'agit de prévoir l'inflation sur des horizons plus longs (ALMOSOVA et ANDRESEN, 2019). Cette caractéristique est particulièrement convaincante pour l'analyse des séries temporelles en général, et pour la persistance de l'inflation en particulier. Malgré des séries relativement courtes, il est raisonnable de supposer que le processus à l'origine de l'inflation présente des ruptures, des commutations, des tendances à évolution lente, ou des combinaisons de ces éléments. Ainsi, les méthodes qui peuvent tenir compte conjointement de ces facteurs peuvent éclairer la dynamique de l'inflation. Comme preuve de concept, le troisième chapitre adapte les LSTM à l'étude de la persistance de l'inflation, en complément des techniques plus traditionnelles.

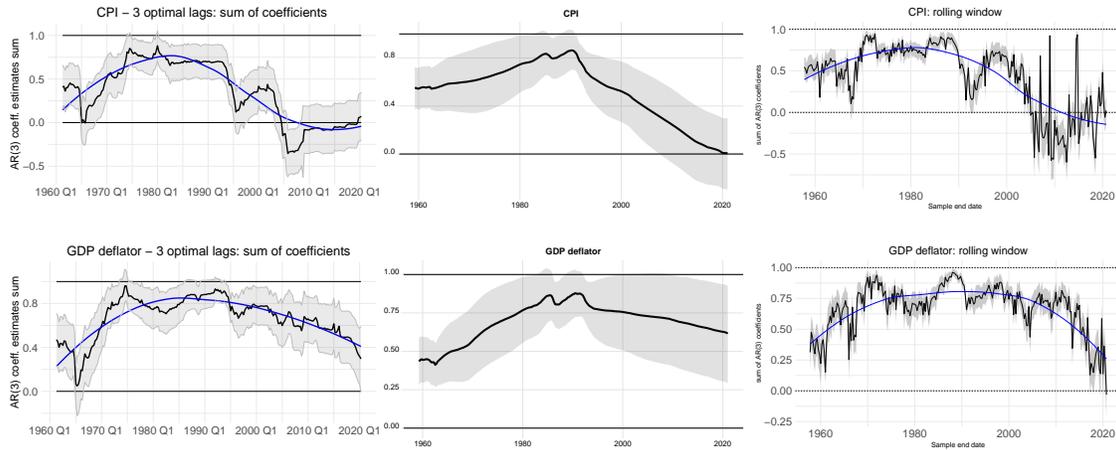
La stratégie empirique dans ce cas est un mélange des deux approches précédentes : premièrement, l'échantillon complet est utilisé pour entraîner les LSTM, qui sont ensuite utilisés pour prévoir l'inflation pour les dix années suivantes. Deuxièmement, l'échantillon est divisé en sous-échantillons non chevauchants de dix ans chacun, utilisés pour l'apprentissage des réseaux ; ensuite, les réseaux entraînés prédisent la trajectoire de l'inflation au cours de la décennie suivante. Cette stratégie de sous-échantillonnage isole les caractéristiques du processus d'inflation qui sont propres à une période historique donnée, étendant ainsi le sous-échantillon original avec des points de données synthétiques mais informés sur la période précédente. Troisièmement, une fenêtre glissante de cinquante-six trimestres est utilisée pour entraîner plusieurs LSTM sur des données réelles. Là encore, avant de passer au point de données suivant, le modèle entraîné prévoit l'inflation de la décennie suivante, afin d'augmenter l'échantillon avec des données artificielles historiquement cohérentes.

Enfin, la persistance est mesurée sur les séries observées et prédites conjointes : cela fournit une mesure de la variation de l'inertie dans le temps, ainsi que des intervalles de confiance raisonnables. En résumé, les résultats corroborent les conclusions d'approches plus traditionnelles : l'inertie de l'inflation a atteint un pic au milieu des années 90, puis a régulièrement diminué, conférant un comportement de quasi-bruit blanc à l'inflation observée. La temporalité est également cohérente avec d'autres méthodes : le renversement de la persistance ne semble pas lié à la politique, au commerce ou aux matières premières. De manière remarquable, l'approche d'apprentissage statistique révèle une grande incertitude sur la période la plus récente, puisque les

estimations ponctuelles sont particulièrement instables à partir de la crise financière mondiale de 2008.

Cet exercice démontre que les méthodes d'apprentissage profond peuvent compléter efficacement les outils macroéconométriques plus standard. À l'avenir, les LSTM peuvent être généralisés pour prendre une structure de type VAR, surmontant à la fois les problèmes de non-linéarité à multiples facettes et le problème de la dimensionnalité.

FIGURE 1.10 – Persistance de l'inflation



Plots récapitulatifs pour les séries sélectionnées, toutes méthodes confondues. La ligne supérieure correspond à l'IPC global, la ligne inférieure au déflateur du PIB. Colonne de gauche : résultats pour la somme des coefficients $AR(k^*)$; colonne du milieu : Estimations bayésiennes avec intervalles de crédibilité de 95% ; colonne de droite : rolling LSTM, somme des coefficients.

Chapter 2

Taylor Rules and Liquidity in Financial Markets

Before the 2008 Global Financial Crisis, macroeconomists underestimated the role of the financial sector in stabilisation policies, especially monetary policies. The dominant framework in theoretical and empirical macroeconomics, the New Keynesian Dynamic Stochastic General Equilibrium (NKDSGE), largely neglects the financial sector (Gali, 2015; Walsh, 2003; Woodford, 2003). In this framework, the scope of central banks policies is limited to actively adjust a single policy interest rate in response to expected inflation to stabilise inflation and to smooth out deviations of output from its potential, following Taylor (1993)-type rules. Specifically, central banks were supposed to react more than one to one to expected inflation, in line with the so-called Taylor Principle. The Global Financial Crisis (GFC) shattered the above consensus and brought to the center stage the role of the financial sector and liquidity.

This Chapter investigates whether the US Federal Reserve Bank consistently acted in line with the Taylor rule. To account for the policy measures and debates that followed the 2008 GFC, we extend the analysis by testing an alternative specification that includes financial variables.

The common characterisation of the evolution of monetary policy regimes in the US runs as follows. In the 1970s, high inflation rate cohabited with an accommodative monetary stance. The beginning of the 1980s marks a structural break, with the shift to a non-accommodating policy that characterised the Volcker period. Such period of *monetary dominance* lasted until the GFC. Indeed, the GFC brought another structural break, with the policy interest rate hitting the Zero Lower Bound (ZLB) during the periods 2008Q4-2015Q4 and 2020Q2-present. In these periods, the policy interest rate did not react at all to swings in expected inflation, thus conflicting with the Taylor Principle. Monetary policy instead worked through non standard measures, addressing liquidity shortages¹ (broadly defined Quantitative Easing, QE). Before then, liquidity of financial

¹AMLF (2016), CPFF (2016), MMIFF (2016), PDCF (2016), TAF (2016), and TSLF (2016) programs descriptions all heavily stress liquidity concerns behind the implementation of each monetary intervention.

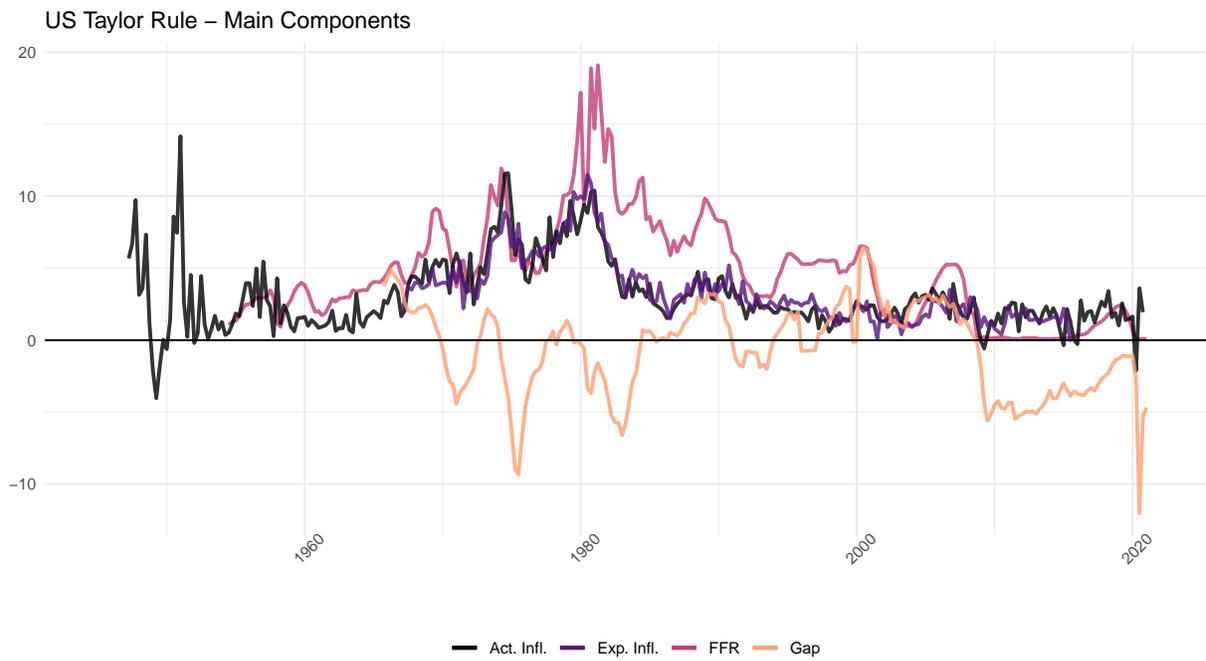
assets, including quantitative measures of money and credit, had no relevance in monetary policy conduct. Their consideration was even discouraged because of the difficult identification of bubbles or excessive risk-taking (Bernanke, 2002; Greenspan, 2002).

In this chapter, we challenge such description of monetary policy regimes. The traditional specification of the monetary policy reaction function does not fully account for the Fed's behaviour over the post-WWII period. We augment a standard specification of the Taylor rule with financial liquidity proxies to test whether the Fed took into account such factors in its policy-making over the past fifty years. In contrast with previous studies, when we consider a standard Taylor rule with real-time data, we do not find deviations from the Taylor Principle even in periods of high inflation. More importantly, the Fed consistently deviates from the Taylor Principle when financial liquidity is included in the Taylor rule. Our measures of financial liquidity rely on the spread between risk-free liquid assets and less liquid assets. When we include proxies for liquidity, we find regimes in which the Fed lowers the weight on expected inflation, while giving more weight to financial liquidity. Moreover, we report instability in the policy rule and therefore investigate the presence of multiple regimes.

Our contribution is two-fold. First, we find mixed evidence on the consistency between actual policies and the Taylor principle. In particular, we do not find strong evidence of a clear-cut switch to monetary dominance in correspondence with Volcker's appointment. Second, we find evidence of a role for liquidity in the Fed's decision set, as proxied by spreads on financial assets. In some episodes, financial liquidity considerations turn out to be more relevant than expected inflation in predicting the policy rate. In contrast with predictions of the NKDSGE models, this policy does not disanchor inflation expectations. To the best of our knowledge, our contribution is the first empirical attempt to assess the properties of a Taylor rule in an environment in which assets liquidity *strictu sensu* has a role. Our interest focuses on reproducing as closely as possible the information set available to the Fed at the time of the decision and include liquidity proxies that capture the dynamics in the financial market and its liquidity.

This inquiry is motivated by the unconventional tools employed by major central banks in the aftermath of the GFC. In December 2008, upon hitting the ZLB on the Federal Funds Rate, the Fed rolled out a growing set of QE policies, *de facto* liquidity facilities. These policies aimed at restoring liquidity and reviving several segments of the financial market, frozen by the recent financial turmoil. Thus, when conventional interest rate setting turned into a loose cannon, liquidity management was elected as primary policy intervention, alongside with forward guidance. On the basis of such policy innovations, we investigate whether Taylor-type rules can effectively accommodate financial and liquidity considerations, as suggested by leaning-against-the-wind literature (Svensson, 2017b) and, most importantly, the large body of post-GFC contributions (Calvo, 2016; Diba and Loisel, 2021; Hall and Reis, 2016).

Figure 2.1 – US Macroeconomy



Variables employed in the estimates. Historically revised GDP deflator (Act. Infl.); Greenbook one-quarter GDP deflator forecast (Exp. Infl.); Federal Funds Rate (FFR); Greenbook estimated output gap (Gap).

Literature overview

We contribute and relate to a vast literature. First, the literature on exogenous regime definition is relevant for our study (Clarida, Gali, and Gertler, 2000). Second, our focus on real-time data is inspired by Orphanides (2001, 2004). In studying how monetary policy rules evolved over time we consider the insights from Boivin (2006), Canova and Sala (2009), Primiceri (2005), and Sullivan (2016), and finally follow Davig and Leeper (2006, 2011), Murray, Nikolsko-Rzhevskyy, and Papell (2015), and Sims and Zha (2006) in adopting a Markov Switching model to let the data speak as freely as possible under bare-bone restrictions. State-dependent policy rules have been theoretically explored as well: estimated DSGEs like Bianchi (2013) model the Fed as switching between hawkish (strongly anti-inflationary) and doveish (weakly anti-inflationary) regimes and study the changes in regimes since WWII and their interplay with agents beliefs. Lhuissier (2018) and Lhuissier and Tripier (2019) employ Bayesian Markov-Switching methods to study volatility and uncertainty changes and their interplay with credit and financial frictions.

Closely related to our study is Leveigue (2002), who analyses the relevance of a set of financial indexes in US, German, and Japanese monetary policy, but finds no role for share indexes on monetary policy. A standard definition of liquidity also enters the forecasting models for the US inflation in D'Agostino and Surico (2009): money growth in the G7 countries does actually predict well inflation on the medium horizon. Regarding the specifications we utilise to look at the data, we also consider the contributions on leaning against the wind policies (LAW, see Svensson (2017a,b) as comprehensive references), although this avenue is outside the scope of this work. Several studies focused on broad measures of financial stress and financial stability. Notably, Baxa, Horváth, and Vašíček (2013) investigate the effects of financial stress on policymaking and correlate rates easing around stock and bank stress. Symmetrically, Gertler, Horváth, and Jonášová (2020) document how ECB communication affects comovements in financial markets. On the other hand, we relate to works that analyse official reports and documents to assess whether the Fed takes into account financial factors. Oet and Lyytinen (2017) and Peek, Rosengren, and Tootell (2016) find that discussing financial stability in FOMC meetings affects policy decisions, Wischnewsky, Jansen, and Neuenkirch (2019) reinforces this evidence analysing congressional hearings of Fed Chairmen. Lastly, Istrefi, Odendahl, and Sestieri (2020) analyse Fed's Governors and Presidents speeches finding that financial concerns are associated with more accommodative stances.

Finally, there is a large body of research in finance that focuses on liquidity and its effects on stocks' and bonds' yields. We mainly refer to Chen, Lesmond, and Wei (2007), Houweling, Mentink, and Vorst (2005), and Lin, Wang, and Wu (2011) who document how liquidity is indeed priced into corporate bond and shares yields in the US and in the European markets. Moreover, Chordia, Sarkar, and Subrahmanyam (2005) provide vector autoregressive evidence on the positive relation between liquidity, financial volatility, and monetary expansions: they focus on US Treasuries and stocks and find correlation between innovations in net banks' reserves and as-

sets' liquidity, besides mutual influences between stocks and Treasury bills. Finally, Dick-Nielsen, Feldhütter, and Lando (2012) analyse how liquidity differentials affect yield spreads in corporate bonds trading at several risk notches: they find that liquidity dries up in crises, especially for sub-AAA bonds, like those indexed in our study.

The remainder of the chapter is structured as follows. Section 3.1 presents the data set we compile; Section 3.2 presents the specifications we estimate, with results for the full sample analysis and the exogenous sub-sampling ones; Section 3.3 presents stability tests on the specifications that motivate a Markov Switching model estimation. Section 5.5 concludes the Chapter.

2.1 Data

To investigate the conduct of US monetary policy, we build a database of the most relevant time series on the US economy aggregates. Such database includes historical and real-time data at the macroeconomic level, as well as statistics from specific microeconomic data to account for expectations. We collect data on inflation, interest rates, real economic activity, monetary aggregates, government debt and deficit, financial market indicators, and finally (measures of) expectations of these variables. In addition, we also collect and test two prominent synthetic shadow rates produced by Wu and Xia (2016) and Krippner (2015).²

For each of these aggregates, we collect a set of more specific measures that differ in the exact definition or computation: the clearest examples are the GDP deflator, the CPI, and the CPE for inflation, or the capacity utilisation, lay-off rate, unemployment, and Fed's own calculations for the output gap.³

To obtain real-time data and distributions on agents' expectations, we source micro-data from the Greenbook data-set on forecasts and expectations, as well as the University of Michigan Survey of Consumers. The former includes the Survey of the Professional Forecasters, which provides expectations on economic aggregates at several time horizons. While the UM Survey of Consumer focuses on a representative sample of the US population, the Philadelphia Fed's SPF surveys a rolling panel of forecasters belonging to a number of industries (financial services at large, academics, manufacturing) and collects forecasts on a wide range of economic indicators. The forecast horizon is heterogeneous and ranges from last quarter to two years ahead, with longer horizons presenting less frequent forecasts.

For financial variables, we employ two measures: the quarterly returns of the S&P 500 index and the weighted average return of BAA corporate bonds. We elaborate on these series to obtain

²The related results are not included as they do not add any further insight to our analysis. They are used as robustness checks, though.

³A side contribution of this paper is to provide a comprehensive and harmonised database as well as the tools to maintain, fine-tune, and customise it. This database is freely available on the [author's website](#). The maintenance programs are available in [this Git repository](#).

reasonable proxies for the liquidity conditions in the economy. The advantage of these proxies is that both are computed on publicly available and reasonably fine-grained: data on individual bond issuances or share- and bond-level trading information allow for deeper analyses and more precise measurement of liquidity premia, but are only available under commercial license.

The vast majority of the series are retrieved from the Federal Reserves of St. Louis and Philadelphia: a complete list is provided in the Appendix, Table (3.7) alongside with plots of the time series.

For a subset of the data, we perform some transformations and manipulations to isolate precise information. We briefly introduce the transformed data below.

Output gap The output gap is an elusive measure of economic slack (Andrle, 2013). Besides the idea of a frictionless economy operating at full capacity, it is especially challenging to gauge the current absolute level of economic activity. In an effort to comprehensively capture the state of the US economy at the time of the policy decision, we collect several different measures of output slack, among which two are worth detailing. First, official output gap nowcast: this extrapolation uses data from the Congressional Budget Office’s Greenbook database on the real-time estimates on the GDP level and implements the methodology mentioned in Murray, Nikolsko-Rzhevskyy, and Papell (2015). For each available date t , we regress the time series against a quadratic time trend and finally take the residual of the latest available data point, ε_t , as output gap observation for date t . We label the resulting time series as *real time output gap*.⁴ It is closer to the signal policymakers receive at the time of the decision. Throughout the analysis, this latter is our preferred indicator of economic slack, covering a reasonably long period and not differing substantially from other measures.

Secondly, we compute the percentage difference between installed capacity and actual GDP, both provided by St. Louis Fed. We call this series *ex-post output gap* since it relies on historical, revised data, not necessarily those available to policymakers at the time of their policy meetings.

The database also presents series on lay-offs, deviations from natural unemployment level and other related proxies for economic cycles.⁵

Inflation and expectations Regarding the measures of inflation, we include revised time series like the indexes of GDP deflator, the Consumer Price Index (CPI), and the Personal Consumption Expenditure (PCE). For the last two, we also include two versions excluding food and energy prices from the reference basket, the core CPI and PCE. The headline PCE is the explicit inflation objective for the Federal Open Market Committee, but limited availability prevents a full-fledged

⁴Further details are presented in the Data Appendix (3.C.1).

⁵Our choice of slack indices are informed on several contributions, mainly Benhabib, Schmitt-Grohe, and Uribe (2001), Bilbiie and Straub (2013), Boivin (2006), Cochrane (2011), Gali and Gertler (1999), Gali, Gertler, and Lopez-Salido (2001), and Orphanides (2001, 2004).

analysis in our study. Thus, throughout the analysis we employ CBO's Greenbook one-quarter ahead forecasts as measures of expected inflation.

The Greenbook database from the Philadelphia Fed provides information on last, current and future expected values for three of the aforementioned indexes – CPI, PCE, and GDP deflator. In particular, forecasts are available up to eight quarters ahead of the survey time t . The same database also offers now-casts on t and $t - 1$ value of these variables. These expectations are part of the information set of the policy-maker at the time of decision and thus represent a reliable way to gauge the policy rule. These data offer the opportunity to test monetary policy with different horizons of expectations, which in itself represents an interesting exercise to assess effects of forward guidance and expectations anchoring.

Liquidity proxies and financial indexes After the financial market collapse that triggered the 2008 GFC, liquidity gained momentum as research topic (alongside with safe assets and thus risk) especially following massive injections carried out by the Federal Reserve.⁶

We employ the financial condition of the economy to infer underlying liquidity. Financial market prices embody a great deal of different information, so the risk of picking up the wrong signal or incurring in endogeneity is high. Considering these threats we compute our indicators as premia over safe assets of comparable maturity, which are subject to "fire-purchases" in times of uncertainty or financial turmoil. In this respect, the finance literature offers robust ways of capturing liquidity variations in isolation from risk factors (see Schestag, Schuster, and Uhrig-Homburg, 2016, for a measurement survey): these methods rely heavily on commercial information at the transaction level over relatively short periods of time, but the main findings are in line with our results, eg Chordia, Sarkar, and Subrahmanyam (2005).

On this ground, we proxy liquidity in financial markets with two spreads. The first one is the difference between the Standard & Poor's 500 quarterly returns and the 3-month Treasury Bill, the second is the gap between the weighted average yield of BAA corporate bonds (as provided by Moody's) and the US 10-year Treasury note. The motivation for using BAA bonds instead of AAA hinges on the flight-to-quality factor that affects top-rated bonds and Treasuries alike: Dick-Nielsen, Feldhütter, and Lando (2012) find evidence for such factor over the 2005-2009 period, thus our choice to focus on the BAA notch has more chances to pick up liquidity movements.

Safety and *liquidity* of any given asset are intimately linked: safety concerns the discrepancy between face value and gross return *at maturity*, while liquidity characterises the difference between face value and *realised selling price*. Thus, the main difference between these concepts boils down to the timing difference. It follows that comparing any asset with its safest equivalent of appropriate maturity isolates its liquidity properties, at least theoretically.

⁶See among others Caballero, Farhi, and Gourinchas (2016, 2017), Canzoneri and Diba (2005), Canzoneri et al. (2008a,b), Del Negro et al. (2017), and Hall and Reis (2016) for summaries and examples.

With real data, liquidity and safety are more arduous to tell apart in a clear-cut way, especially at a macroeconomic scale. We thus proxy liquidity with spreads that isolate as much as possible the liquidity component of assets. Comparing safe assets with widely traded and riskier ones helps to minimise the confounding of risk and liquidity, especially since we consider aggregate indexes rather than individual assets. This last factor is relevant as diversification forces operate already to drive down the risk component of our observables. Indexes naturally incorporate a diversification factor that helps with the idiosyncratic sources of uncertainty, still not undoing entirely the macroeconomic risk.

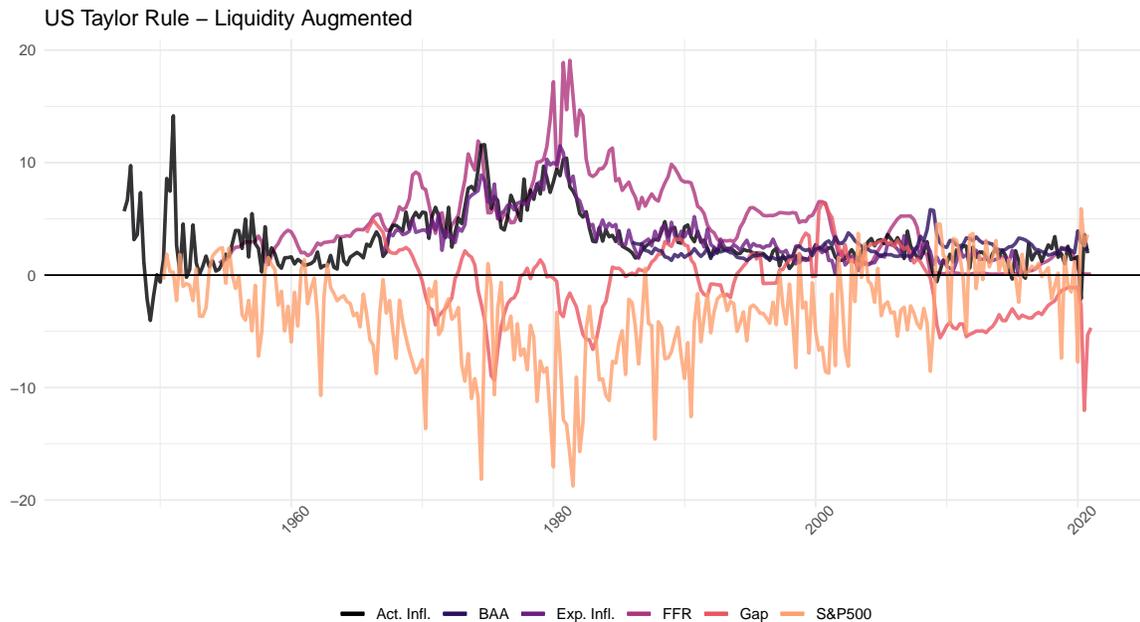
The intuition goes as follows: prior to a recession or slowdown in economic activity, publicly-traded assets are fully liquid and smoothly traded. When uncertainty kicks in or expectations turn pessimistic, these assets become second choice to safer assets with high liquidity. Agents on the market readjust their holdings to shield from possible effects of the incoming downturn. As they sell these second-choice assets and shelter with Treasuries and AAA bonds, selling prices of stocks and BAA bonds decrease (therefore pushing up expected return), whilst safe assets prices increase (and return plummets, possibly turning negative, too) as they maintain their attractiveness and ease of liquidation. In this process, selling parties are willing to bear losses with respect to face value upon transaction, while stomaching increased prices to buy safer assets. Referring back to the liquidity definition above, this dynamics mirrors a liquidity dry-up for second choice assets that percolates into relative returns. Therefore, the spread between the former and the latter factors reflects variation in liquidity of assets triggered by movements in market expectations.

We include these spreads in the decision rule of the central bank in order to test whether policymakers are also attentive to financial and liquidity conditions in the economy when setting their policies. This approach is well settled especially in the leaning-against-the-wind literature, where financial conditions, however defined, are related to monetary policy (Baxa, Horváth, and Vašíček, 2013; Svensson, 2017b)

Throughout the rest of the chapter we use the GDP deflator as the measure of inflation, since it maps closely the price dynamics emerging from the US economy. Moreover, its time series is the longest and affords estimation in the pre-Volcker period. To reproduce the information set of the Fed, we use our real-time output gap measure. Both series strongly correlate with the alternatives and bear little differences in trends.

For completeness, Table (3.8) in the Appendix presents correlations among selected variables in our data set over the period from 1986Q1 to 2013Q4, when all series overlap. The whole set of series used in this chapter is collected in Fig.(3.2).

Figure 2.2 – Liquidity-Augmented Taylor Rule Variables for the US



Main variables analysed with financial liquidity measures, BAA and S&P500 spreads (dot-dashed and long-dashed)

2.2 Empirical Results

There is a broad consensus on the empirical validity of the Taylor rule, the more so since Volcker’s chairmanship. At the beginning of his term he induced a switch in policy from a regime of indeterminacy (accommodative policy) to one of determinacy (aggressive policy).⁷ We revise the most common methods to estimate a Taylor rule, with the inclusion of financial liquidity proxies, before proposing a Markov Switching model to account for multiple regimes. Closer to our take to endogenise policy changes is Murray, Nikolsko-Rzhevskyy, and Papell (2015), where Hamilton (1989) algorithm for Markov processes estimation is applied to monetary policy rules. They explore the two-state case, finding two periods of indeterminate policy.

The point estimates from Clarida, Gali, and Gertler (2000) and Murray, Nikolsko-Rzhevskyy, and Papell (2015) are summarised in the tables below; in both tables μ represents the long-term inflation target, γ captures the (over) reaction to expected inflation as in $(1 + \gamma) E_t \pi_{t+1}$, ω repre-

⁷One of the first attempts to verify such break in policy is Clarida, Gali, and Gertler (2000), who exogenously divide their sample in two periods. The Fed was passive during the first part of the sample, whereas it reacted aggressively after Volcker chairmanship. Boivin (2006) estimates a Taylor rule with drifting parameters over the full sample of real-time data. Inflation response was weak in the second half of the 1970s, but strong in the rest of the sample. The response to real activity decreased significantly after the 1970s. In the same vein, Primiceri (2005) employs a Bayesian SVAR to show that the Fed complied with the Taylor Principle even prior to Volcker. Bilbiie and Straub, 2013, though, use an early TANK estimated model to highlight how financial market participation was key in setting off indeterminate equilibria.

sents the reaction to output gap fluctuations, and finally ρ estimates the smoothing factor for the federal funds rate.

Table 2.1 – Taylor Rule estimates

Exogenous break	μ	γ	ω	ρ	Markov State	μ	γ	ω	ρ
<i>Pre – Volcker</i>	4.24	-.17	.27	.68	S ₁	9.44	-.3	.46	.49
	(1.09)	(.07)	(.08)	(.05)		(3.07)	(.34)	(.26)	(.15)
<i>Post – Volcker</i>	3.58	1.15	.93	.79	S ₂	.59	.85	.58	.8
	(.5)	(.4)	(.42)	(.04)		(.66)	(.23)	(.09)	(.05)

(a) Clarida, Gali, and Gertler (2000)

(b) Murray, Nikolsko-Rzhevskyy, and Papell (2015)

We first study the stability of the parameters of the Taylor rule in the post-WWII period. We propose a set of estimates of the decision rule followed by the monetary authority. After analysing the full sample with OLS estimators,⁸ we exogenously split the sample into three sub-samples and compare the parameters. Then, we investigate possible structural breaks over the full sample via stability diagnosis. Third, we let the sub-sampling be endogenous with a Markov Switching estimation for two possible states. Throughout these steps, we also include liquidity proxies in the monetary policy rule, to assess its relevance in the Fed’s input set.

2.2.1 Specifications

We first assess the robustness of the traditional specification in comparison to the alternatives with liquidity proxies. Our interest lies particularly in the stability of the parameters over different methods, sample cuts, and with the inclusion of financial conditions in the Fed’s information set.⁹ This analysis solely focuses on the interest-setting rule, while leaving aside direct management of liquidity in the economy. Optimal central bank balance sheet policy is nevertheless of paramount relevance and deserves further research effort. A first step in this direction is developed in chapter (4), where we propose a setup that includes liquid assets and sets the stage for liquidity policies by the central bank.

Throughout this Section, we will estimate equation (3.1): r is the effective federal funds rate, π_{t+h} is h -period ahead expected inflation (mapped to forecasts, up to 8-quarter), \hat{y} is output gap in percentage deviation, and X collects any additional variables used in the study, as detailed in Section (3.1).

$$r_t = (1 - \rho) \left[\mu + (1 + \gamma) E_t \pi_{t+h} + \omega \hat{y}_t + \mathbf{X}'_t \boldsymbol{\beta} \right] + \rho r_{t-1} + \varepsilon_t \quad (2.1)$$

⁸Carvalho, Nechio, and Tristao (2018) provide motivation for using OLS in estimating monetary policy rules and quantify the bias of such method compared to GMM or IV.

⁹We also employ the Generalised Method of Moments over the same specifications as a way to circumvent endogeneity issues: these results are presented in the Appendix (3.C.3).

In equation (3.1), we assume that the central bank smooths its policy decision putting a weight $0 < \rho < 1$ on past interest rate level. Therefore, we need to recover estimates and confidence intervals from the estimated ρ . Moreover, we allow for inflation targeting including an intercept μ . ε_t captures the exogenous shock the monetary authority impulses to the rate path.

Our parameters of interest are β_i and γ . The sign and the magnitude of the former will tell how relevant other factors are for the central bank; on the other hand, γ will shed light on the robustness of the Taylor Principle. Established consensus points to a value close to $\gamma = .8$ following the onset of the Great Moderation and the inflation conquest carried out by Volcker.

Before exposing the results, we flesh out the specifications we estimate throughout this Section and briefly motivate their utilisation. Other specifications, although conceptually appealing, do not necessarily add interesting insights or fall outside the scope of investigating the role of financial liquidity.

Spec. I $X = 0$: the standard specification as in Taylor (1993) and many other works. We employ one period ahead forecasts of GDP deflator as expected inflation and real-time gap for output slack. Considerably, we employ uniquely real time data so to track as closely as possible the information set available to the central bank.

Spec. II $X = \text{BAA}$: we introduce in this specification a proxy for financial liquidity distress, namely the spread between BAA corporate bonds and 10 years Treasury bonds. This specification captures long term liquidity in the economy and thus accounts for the Fed's concern of longer run financial stability.

Spec. III $X = \text{S\&P500}$: we test a second proxy for liquidity with this specification. We exploit quarterly returns on the stock market to obtain a spread with 3-month Treasury Bills. This spread captures shorter term concerns in financial liquidity, such as those that triggered the GFC. This specification is prone to picking up also solvability concerns.

2.2.2 Full sample analysis

The first step is to estimate our specifications on the full sample, overlooking the possibility of structural breaks or fluctuations in the parameters. The period covered varies with the series included in the specification, from a maximal of about 185 observations to a minimum of slightly less than 115.¹⁰ Assuming parameter instability of any form and correct specification, this approach produces simple averages over the possible parameter values. In this sense, regimes that are more frequently in place will be more represented in the final estimate. Table (3.2) summarises the results of OLS estimates on the full sample for the different specifications.

¹⁰Greenbook real-time data, in particular, are released to the public with a 5-year lag with respect to the estimates or forecasts and thus presents data until 2013. Moreover, data on BAA spreads are available only from 1986Q1.

Table 2.2 – Full sample estimates

	Specification		
	I	II	III
μ	.1327 (.7766)	8.1884*** (1.6854)	.3355 (.5268)
$E_t(\text{GDP defl.})_{t+1}$	1.5811*** (.2736)	.9522* (.4673)	.9528*** (.1883)
Real time \hat{y}	.48** (.1439)	.2496+ (.1437)	.2361* (.0998)
$BAA_{spr.}$		-2.9673*** (.4939)	
$SP_{spr.}$			-.4349*** (.0951)
FFR_{t-1}	.7782*** (.0369)	.8643*** (.0289)	.6877*** (.0402)
Obs.	188	112	188
R^2	.8949	.9733	.9052
BIC	628	157	612

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1; SE in parentheses. Observations up to 2013 Q2.

These results are interesting in a number of aspects. First, the sample encompasses a variety of regimes: from the pre-Volcker era to the ZLB period, with the Great Moderation dwarfing other regimes in terms of observations. Therefore, it blends together different rules and behaviours with diverse weights: the Great Moderation accounts for roughly 55% of the sample, the high-inflation period for about 30%, and finally only 15% of the sample covers the most recent times.

Second, the addition of liquidity in the policy rule of the central bank marks a consistent violation of the Taylor Principle, with γ estimates below 1 for Specifications II and III. The inclusion of liquidity proxies significantly lowers the weight on inflation expectations. This finding is even more surprising when considering that the BAA spread series starts in 1986Q1, at the end of Volcker's mandate. These results suggest that financial conditions enter the decision set of the central bank with the expected sign. While the point estimates signal violations of the Taylor principle, standard errors are in the neighbourhood of 1: nevertheless, our results for Spec. II and III are in sharp contrast with traditional results, as those listed in Tab.(3.2a) and (3.2b). This first evidence begs for further and more refined inquiry into the role of financial liquidity in the Fed's decision set.

As mentioned before, the length of the sample blends multiple regimes and stances. In this light it is not surprising to find substantial parameters instability, as Table (3.2) presents. Under the hypothesis of μ as inflation target, its estimates greatly vary according to the sample and the specification. The same, with less variability, applies for ω and ρ , the output gap weight and the

smoothing factor, respectively. To address this instability issue, we exogenously split the sample into three sub-samples, assuming different regimes.¹¹

2.2.3 Exogenous breaks

The estimates above might result from a variety of underlying regimes, either smooth (Boivin, 2006; Primiceri, 2005) or discrete (Murray, Nikolsko-Rzhevskyy, and Papell, 2015). To account for such regime change, a straightforward approach is to look for relevant historical events that mark a discontinuity and estimate the policy rule before and after such dates. This approach follows the seminal work of Clarida, Gali, and Gertler (2000), and maps historical events to discrete switches in the monetary policy stance. It boils down to splitting the sample into three sub-samples: post-WWII period, the Great Moderation, and the GFC. The latter in particular encompasses the QE liquidity injections in correspondence of the Zero Lower Bound period. This last characteristic distinguishes the first from the third sub-sample. In fact, in the postwar period the federal funds rate averaged around 5%, seldom below 4%. The downside is that at date this sub-sample presents a small number of observations, which makes inference rather heroic. We split the sample to obtain three subsets that map relevant monetary and economic events:

- (i) pre-Volcker regime [-:1981Q4],
- (ii) the Great Moderation [1982Q1:2007Q2],
- (iii) and finally the GFC [2007Q3:-].

Historically, the period covered by sub-sample (i) reports high inflation and federal funds rate, as well as volatile cyclical fluctuations. According to Clarida, Gali, and Gertler (2000), among others, the Fed carried out an accommodative monetary policy in that period, following inflation instead of aggressively responding to its expectations. Hence, we expect to see values close to those presented in Table (3.2a). Unfortunately, data availability severely limits the estimation of Spec. III, which is excluded from the first subsample.

The second chunk of data covers the inflation conquest and the steady, sustained growth that followed, with mild recessions and inflation in check. Supposedly, these conditions were brought about by a central bank eventually fighting back inflation aggressively, adjusting the FFR more than one-to-one with respect to expected inflation.

The third period starts right before the GFC. Data are still scarce: to date, we have about 10 years of quarterly data with hardly enough variation, mainly because of the FFR hitting the zero lower bound and hovering in its neighbourhood. Therefore, the estimates here shall be considered

¹¹Appendix (3.C.2) offers the residuals plot for the full sample regression. Eye-balling these plots provides sufficient motives to carry out additional analyses on model instability.

cum grano salis: statistical significance is hardly found with so few observations, but point estimate might be qualitatively informative.

Table 2.3 – Exogenous splits, three samples

Sample	Spec.	FFR_t					Obs.	R^2	BIC	
		μ	$E_t(\text{GDP defl.})_{t+1}$	Real time \hat{y}	BAA spr.	SP spr.				FFR_{t-1}
(i)	I	-.2145 (1.6457)	1.4885*** (.362)	.5266* (.2207)			.5985*** (.0882)	60	.7664	258
	II	.1106 (.9632)	.8069*** (.216)	.2226+ (.1324)		-.4938*** (.1131)	.4022*** (.0889)	60	.8234	244
(ii)	I	-.7987 (1.4227)	2.1619** (.6898)	.3646+ (.2179)			.8495*** (.0404)	102	.9352	238
	II	9.79*** (2.18)	1.1145* (.5147)	.6558** (.2027)	-4.215*** (.7594)		.8642*** (.0299)	86	.9606	120
	III	-.2352 (.9721)	1.4082** (.4705)	.1749 (.1503)		-.3954** (.1188)	.7892*** (.0425)	102	.9413	232
(iii)	I	-48.64+ (26.01)	3.6837 (8.22)	-8.984+ (4.37)			.9846*** (.1015)	26	.9351	29
	II	2.7847 (.347)	.1705 (.5859)	.0925 (.4117)	-1.1005* (.4948)		.8004*** (.1249)	26	.945	27
	III	-10.43 (9.185)	1.058 (2.621)	-1.662 (1.675)		.7416 (.549)	.9526*** (.1024)	26	.9375	30

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '+ ' 0.1; SE in parentheses. (i) runs from the earliest available observation to 1981Q4; (ii) runs from 1981Q4 to 2007Q2; (iii) goes from 2007Q3 to the latest observation available, currently 2013Q2, as some data are published with a five years lag.

Skimming through Table (3.3) it is interesting to compare the regimes in place. Although underlying heterogeneity restricts significantly the econometric robustness of such exercise, a number of regularities emerges.

In particular, γ estimates are highly volatile, both across and within sub-samples. Contrary to the consensus (but in line with Boivin (2006) and Primiceri (2005)), the Fed broadly complied with the Taylor Principle, although Spec. II and III present values statistically closer to 0 for γ . In the second period, in line with the narrative of a committed and credible central bank, γ is consistent with the Taylor Principle in all specifications, but presents a significant heterogeneity in the estimates. The final period is morphed by the extremely sluggishness of the federal funds rate sticking at the ZLB. For all specifications, the only statistically significant factor is the past policy rate level. Strikingly, inflation expectations seem to disappear from the relevant set for monetary policy-making, as its estimates are definitely not significant. At a first pass, for this subsample, a Taylor-type rule appears irrelevant, unless a more refined analysis includes monetary aggregates and possibly a role for forward guidance.

Likewise, output gap measures seem less relevant in the wake of the GFC than in previous periods, when ω estimates take expected values and signs without much volatility. The implicit

inflation target, μ , displays an erratic behaviour and provide a less than convincing picture over the periods.

Focusing on cross samples comparison, the BAA spread holds robust estimates over time and across specifications, like the S&P500 spread, pointing to our intuition that financial conditions are key. When liquidity dries up because of financial or real turmoil (and hence spreads increase) the monetary authority puts in place accommodating policies by decreasing the reference interest rate. On top of that, there is a noticeable increase in the federal funds rate persistence over time. The first sub-sample present a quite volatile policy rate, whilst sub-sample (ii) reports significant increases in ρ . The most recent sub-sample (iii), with severely scarce observations, includes a key policy rate that barely moves, with other variables displaying more variability. This explains why in all regressions the most significant variable is the lagged interest rate, while all other variables present odd estimates. Nevertheless, some results are suggestive of fundamental parameters instability, consistently with the Fed switching to QE policies (*de facto* balance sheet expansions to deploy liquidity facilities).

These early results point towards an inconsistent behaviour of the central bank – if we assume its only decision function takes the form of a strictly parametrised Taylor rule and exclude other tools altogether. The great deal of volatility in the estimates across sub-samples motivates a deeper investigation into the stability of the parameters. This instability likely mirrors a changing policy stance for the monetary authority, as encapsulated by the Taylor rule specifications. The next step, therefore, is to diagnose possible sources of structural breaks or parameter variability and to address this with proper, flexible tools.

2.3 Diagnostics on Structural Breaks and Markov Switching

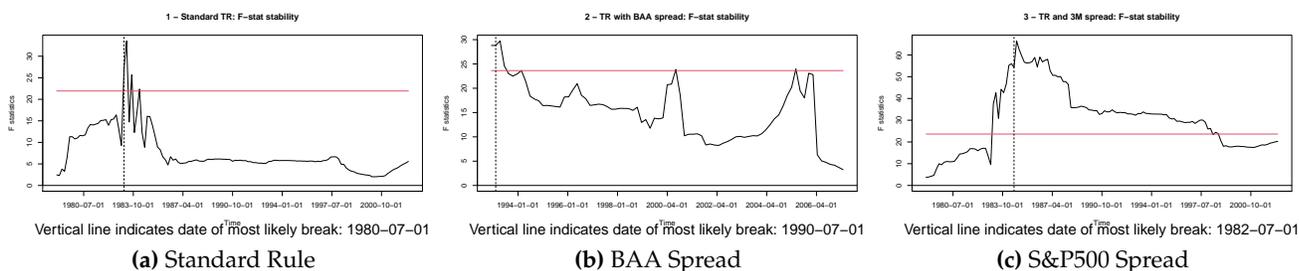
Sub-sampling according to historical events is a straightforward strategy, but it does not provide a robust statistical motivation. Exogenous subsampling is an appealing shortcut, but mutes off the endogenous source of regime changes. A more complete positive approach would explain these switches with the current state of the economy, agents' expectations, and their information. To address this shortcoming, we exploit a number of tests to diagnose the stability of the estimated parameters. Compared to exogenous sub-sampling, this approach is more data-driven, as it makes use of the information contained in the sample to check for breaks and eventually propose the most likely break date(s). We run these diagnostics on the full sample to identify precise dates. In line with the established consensus and based on previous evidence, we expect to find breaks in correspondence of the Volcker chairmanship and the GFC.

We take the models estimated on the full sample and run first a simple CUSUM diagnostic test, then a Chow (1960) test. The latter in particular is flexible enough to provide an optimal sample segmentation based on parameters stability. Evaluating on the Bayes Information Criterion, the

Chow test can produce the most likely date of break, provided that we require only one single date; or it can provide the most likely number of breaks in the sample. We exploit these properties in our analysis.

As a straightforward check, CUSUM tests do not report significant fluctuations in the empirical process, meaning that the cumulative sum of the residuals eventually levels off to zero without significant erratic deviations. It is nevertheless interesting to remark how persistent over time the deviations are and how consistently the CUSUM statistics builds up towards the end of the sample. Plots of CUSUM diagnostic are presented in Appendix (3.C.4). The F-Test derived from Chow (1960) points in another direction, though, as Fig. (3.3) shows. The output of the test actually reports multiple breaks along the sample, some of which occur with unexpected timing.

Figure 2.3 – Structural Breaks



F-statistic plots for specifications I to III. Dashed line indicates the statistics value, solid horizontal line marks the significance area at 95%. Individual captions offers most likely date for a singular structural break in the specification pointed by the vertical dashed line.

When only the most likely date break is requested, two out of three specifications report it around two years into Volcker’s Chairmanship (specifications I and III, excluding the BAA spread, which starts in 1986), decidedly in line with the established consensus. The 3-month S&P500 spread specification, among those analysed, report F-statistics hovering above the threshold for substantial periods. The results of optimal segmentation of the sample pave the way to the Markov switching estimation below as it motivates further inquiry into the existence of multiple policy regimes.

Table (3.4) summarises the analysis on optimal segmentation: most of the specifications likely involve more than one structural break. On top of Volcker’s regime change, one would reasonably expect that the mix of ZLB and unconventional policies would be sufficient to mark an additional break. Moreover, a striking result is that in two cases (specifications I and III) the end of Volcker’s two terms is also a candidate point for a structural discontinuity.¹² These results suggest that there has been a structural break when Volcker left the Chair, with Greenspan chairmanship possibly

¹²Volcker was nominated July 25th, 1979, sworn shortly after August 6th, and left the Chair in August 11th, 1987. In our quarterly data set it translates in 1979Q3:1987Q3. This last insight deserves more documentary research effort, since it could signal that it was actually Greenspan to introduce a Taylor-type reaction function in the Fed decision process.

Table 2.4 – Optimal segmentation and break dates

	Specification		
	I – Standard Rule	II – BAA Spread	III – S&P500 Spread
Sing. break	1980Q3	1990Q3	1982Q3
N. of breaks	2	3	2
1 st Date	1980Q3	1989Q4	1980Q3
2 nd Date	1987Q3	2000Q4	1987Q3
3 rd Date	-	2008Q3	-

The first row presents the most likely break admitting only a single one. Third to fifth rows present break dates when up to 5 breaks are allowed.

introducing an additional new regime – somewhat a *new normal*. Hence, this structural break analysis finds suggestive evidence of two or more monetary policy regimes, among which the Federal Reserve might possibly switch back and forth depending on a variety of indicators. This state-dependent rule includes economic activity factors as well as financial ones. Therefore, it is worth pursuing additional insights into these structural breaks with adequate techniques.

2.3.1 Markov Switching estimation

We further unconstrain the data via a Markov Switching model: we adopt Hamilton (1989, 1994) and Murray, Nikolsko-Rzhevskyy, and Papell (2015) approach to our extended sample and only assume it comprises k discrete states. We restrict our analysis to $k = 2$, in line with the discussion on the determinacy or indeterminacy regimes at the beginning and at the end of our sample. As aforementioned, pre-Volcker and post-GFC periods yield deeper insight on the functioning of the Federal Reserve monetary policy conduct away from the Great Moderation “steady state.” Therefore, eq.(3.1) takes now the form

$$r_t = (1 - \rho^{(S)}) \left[\mu^{(S)} + (1 + \gamma^{(S)}) E_t \pi_{t+h} + \omega^{(S)} \hat{y}_t + \mathbf{X}'_t \boldsymbol{\beta}^{(S)} \right] + \rho^{(S)} r_{t-1} + \varepsilon_t^{(S)} \quad (2.2)$$

$$\text{var} \left(\varepsilon_t^{(S)} \right) = \sigma^{(S)} \quad \text{with } S = k \in \{1, \dots, K\}$$

Hamilton (1989) provides the algorithm to estimate our specifications with k states, generating also transition matrices and smoothed probabilities to pick the prevailing regime in any date t . For every specification, we allow for the variation of every parameter and the error variance: in k different states, all parameters are freely estimated, with no constraint posed by other states' estimates.¹³

¹³Alternatively, any subset of parameters can be optionally estimated across all regimes, so its estimates are invariant to the prevailing regime.

Table 2.5 – Estimates Markov Switching model

Spec.	State	FFR_t					FFR_{t-1}	Adj.R ²	BIC
		μ	$E_t(\text{GDP defl.})_{t+1}$	Real time \hat{y}	BAA spr.	SP spr.			
I	S ₁	-.7759 (.2025)	1.693*** (.4816)	.3554 (.329)			.6128*** (.1118)	.7853	518
	S ₂	-.4052 (1.199)	2.012*** (.4131)	.6329** (.2274)			.9208*** (.0253)		
II	S ₁	2.74* (-)	-.7505+ (-)	.0564 (-)	-.3956 (-)		1.099*** (.0234)	.9987	142
	S ₂	11.01*** (1.264)	.3199 (.3338)	.442** (.1367)	-4.066*** (.3692)		.8385*** (.0234)		
III	S ₁	1.45 (1.322)	.5299*** (.1596)	-.2855+ (.1458)		-.6061*** (.0938)	-.0642 (.1416)	.8505	518
	S ₂	-.8855 (.9915)	1.561*** (.3917)	.4535* (.2005)		-.2201 (.1712)	.8982*** (.0286)		

$k = 2$ states. Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '+ ' 0.1; SE in parentheses. (-) stands for non-convertible SE: significance is hence derived from the main regression.

Table (3.5) presents estimates for the two-state Markov switching model. In contrast to the results of Murray, Nikolsko-Rzhevskyy, and Papell (2015), Specification I – mirroring those of the cited work – finds two states in compliance with the Taylor Principle.¹⁴ This traditional Taylor rule presents two distinct states that differ in terms of reaction *intensity*. S_2 estimates present larger coefficients in absolute values for virtually all variables. Focusing on γ estimates, both states comply to the Taylor principle, pointing towards a *more aggressive* reaction in the second state. This evidence contrasts the established consensus, but rather corroborate Orphanides (2001, 2004) findings.

Interestingly, S_1 is more likely to be in place during downturns, with a lower persistence of the policy rate. This last piece of evidence suggests that the Fed might react asymmetrically to evolving economic conditions: aggressively, in light of downturns; cautiously, when the recovery materialises. With respect to the timing, it appears to precede economic turmoil periods: it covers roughly the years of the oil shock, the high inflation that followed, the dot-com crash, and the early stages of the GFC. It is suggestive to connect this regime with a cautious approach of the Federal Reserve: as soon as slowdown headwinds build up, the policy intervention intensity needs to adjust accordingly. By contrast, when risk, uncertainty, and sources of slowdown weaken, the central bank acts with decision to steer the economy, possibly in a strategic way that gains room for further rate leveraging.

¹⁴Most likely this discrepancy arises from our longer sample and the slightly different method employed in the estimation.

Turning to the financial specifications (II and III), the picture is less clear but insightful. Estimates for γ display both regimes of determinacy and indeterminacy. Across these two specifications estimates for the weight of the spread are rather consistent with previous results: to a higher relevance of the financial condition corresponds a lower reaction to expected inflation and real activity (summarised in S_1 for both Specifications), putting in place an accommodative policy regime. Again, the interest rate persistence is lower in the indeterminate regime. The prevailing period for the accommodative monetary policy corresponds roughly to Volcker’s first term and the first half of the 1970s: this evidence challenges the established narrative that sees Volcker putting in place an aggressive regime, engineer a recession, and trigger the Great Moderation and credibility build-up. Looking at the time dimension of these regimes, the picture is complementary to the one drawn from the traditional Taylor rule. In both cases, the indeterminate regime is in place at times of economic and financial distress. One difference is worth noticing, though: in Specification III the ZLB period is described by an active monetary policy regime, while the opposite is found for Spec. II. This difference depends on the fact that bonds are more exposed to the liquidity risk, as well as incorporate a quantity of default risk. This latter factor is less present in the most capitalised companies on the financial market¹⁵ and might drive the result.

The panels collected in Fig. (3.4) depict the prevailing state along the sample for the estimates of Table (3.5). We also propose the transition matrices for the two estimated states. In general, every state appears to be an attractor: virtually all states will persist to the following period with a probability greater than .8.

Table 2.6 – Transition matrices

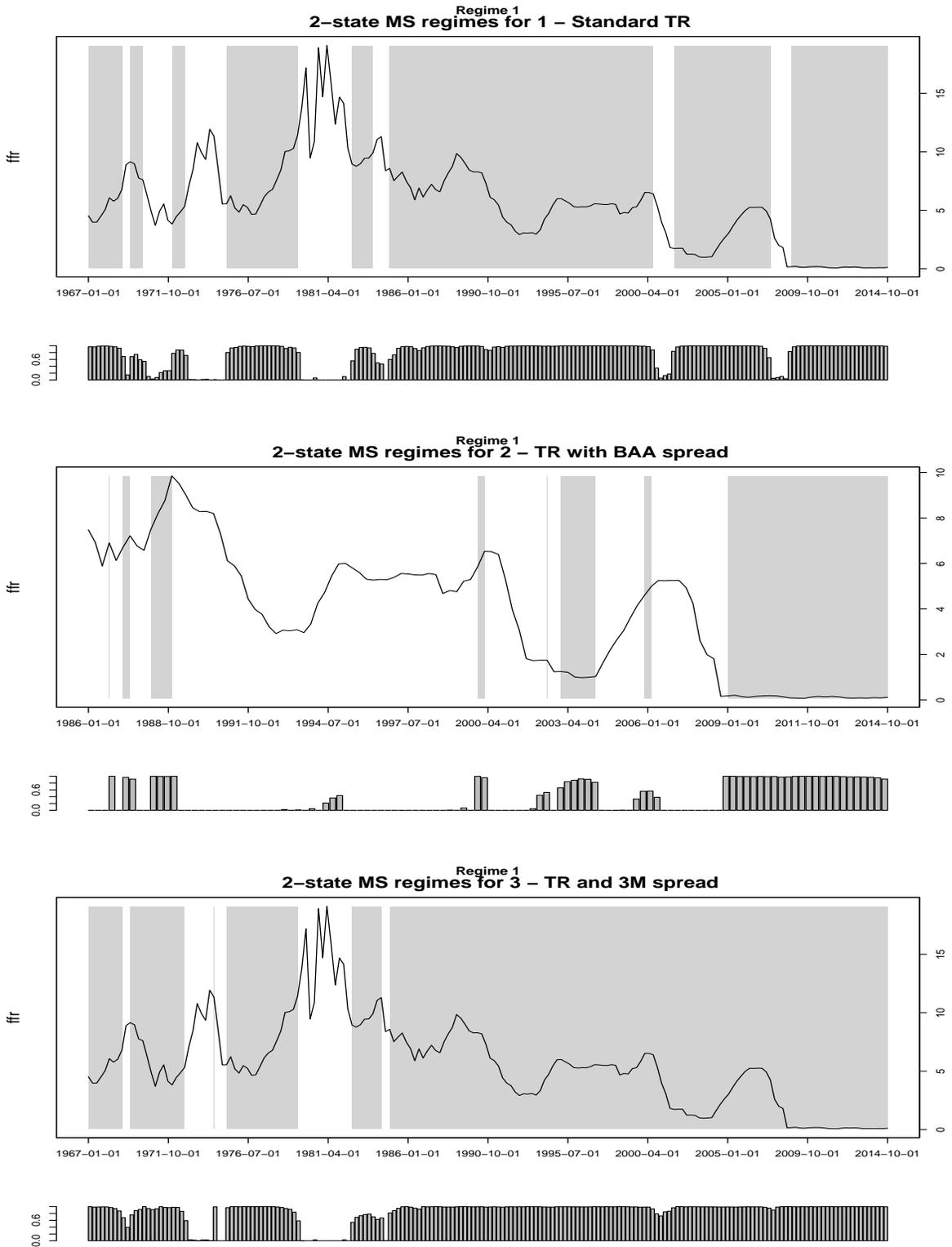
I – Standard Rule		II – BAA Spread		III – S&P500				
	S_1	S_2		S_1	S_2			
S_1	.945	.171	S_1	.9086	.2399	S_1	.8451	.0511
S_2	.055	.829	S_2	.0914	.7601	S_2	.1549	.9489

Transition probabilities for two states Markov process. Columns are current state, hence conditional on it next state is one of the rows.

The overall evidence points towards a fundamental instability in parameters of the Taylor rule, entailing periods of violation of the TP and a consistent role of financial liquidity. We retrieved periods with – theoretically – destabilising monetary rules and inflation under control at the same time, once liquidity is included in the decision set of the monetary authority. These results suggest that the Taylor Principle might not play a fundamental role in anchoring the inflation expectations: agents could – and did, according to our findings – form expectations about the future paths of prices that are non-degenerate even when the central bank deviates from the prescriptions of the

¹⁵While it is true that during the GFC some systemic banks underwent actual bailouts, it is disputable for the rest of the companies listed in the S&P 500.

Figure 2.4 – Markov Smoothed Probabilities



Markov States: in the above panels, shaded areas correspond to S_1 prevailing over S_2 , bottom panel show smoothed probabilities for S_1 .

New Keynesian workhorse. Other factors are therefore at play in anchoring inflation expectations: while the influence of central bank actions remains relevant for the determination of the inflation behaviour, the Taylor Principle appear to be less than granitic.

In our proposed specifications for the monetary rules, proxies for liquidity and for financial conditions have a sizeable and robust role across different methodologies. As soon as liquidity dries up, financial conditions worsen, the Federal Reserve Bank acts and reacts lowering the reference rate. This finding is consequential since these violations of the Taylor Principle are not accompanied by degenerate behaviour of inflation, at least in the most recent cases.

2.4 Conclusion

In this chapter, we collect and aggregate an important number of data on the US economy, both real-time and revised series. We use this extensive database to empirically investigate the stability of the monetary policy rule in the US postwar period and to test whether the Federal Reserve Bank considers financial markets liquidity in its policy decisions.

We study the robustness of the standard Taylor rule derived from NKDSGE models' restrictions and embedded within several central banks' decision set. A standard rule reveals compliance – with varying intensity – of the Fed to the Taylor principle over the whole period considered, contrary to the established consensus that identifies the anchoring of the Taylor Principle in the post-Volcker regime. We also provide statistical evidence for the presence of structural breaks in the policy rule. Furthermore, the inclusion of liquidity spreads reveals that the Fed also takes into account financial conditions in its interest rate setting. This inclusion produces violations of the Taylor Principle that nevertheless do not disanchor inflation nor set aggregates spiralling.

Estimating two-state Markov Switching models, we find multiple monetary policy regimes. When considering financial markets liquidity, the Fed reacts less than one-to-one to expected inflation in periods of looming economic uncertainty and financial tension. In such cases, it also weighs more the liquidity conditions in the economy. Conversely, it switches back to a standard Taylor rule in tranquil times. These results challenge the narrative of a passive US central bank until the regime switch associated with Volcker's chairmanship.

All in all, across estimates we find evidence of generalised parameters instability. Our findings shed new light on the functioning of the US monetary authority and on its information set. The potential effects of such instability on the dynamics of inflation and of other macroeconomic aggregates are an important subject for future research. The model presented in chapter 4 proposes a first attempt in such direction: the simple addition of a liquid asset relaxes the Taylor Principle requirement, but also adds one more lever to the central bank's toolkit, which is direct liquidity control. While these results revolve around the interest-setting leg of the monetary policy, they pave the way for more research on the whole tool-kit of central banking. Ideally, a characteri-

sation of the central bank decision mapping would include monetary aggregates and liquidity management alongside with an interest rate setting rule, so to describe the whole set of tools that a monetary authority can leverage to respond to economic fluctuations.

2.A Data Sources and Transformations

Table 2.7 – Data sources details for the US

Variable	Source	Mnemonics	Type, Transformation
FFR	1	FEDFUNDS	Rate, A
Deflator	2	gPGDP	Rate, A
CPI	2	gPCPI	Rate, A
Core	2	gPCPIX	Rate, A
Realtime y-gap	2	ROUTPUT	\$ bls, B
Expost y-gap	1	GDPPOT; GDPC1	Percentage Deviation, C
Unemployment	1	UNRATE	Rate
Layoff Rate	1	ICSA; PAYEMS	Rate
Empl. Fluctuations	1	NROU; UNRATE	deviation, C
BBA Spread	1	BAA10Y	Rate, E
S&P500 Spread	3	^GSPC	Rate, E
US Deficit	1	M318501Q027NBEA; GDP	share to GDP
Debt to GDP	1	GFDEGDQ188S	Share to GDP
Debt Level	1	GFDEBTN	\$ mln
Debt Gross Growth	1	GFDEBTN	Rate, D
Debt Held by FED	1	FDHBFBRBN	\$ bls
Debt Share Held by FED	1	FDHBFBRBN; GFDEBTN	share
SPF:CPI	2	SPFCPI.i	Unbal. Panel
SPF:CORECPI Rate	2	SPFCORECPI.i	Unbal. Panel
SPF:PCE Rate	2	SPFPCE.i	Unbal. Panel
SPF:COREPCE	2	SPFCOREPCE.i	Unbal. Panel
SPF:BBA spr	2		Unbal. Panel
Hist. CPI	1	CPIAUCSL	Rate, D
Hist. Deflator	1	GDPDEF	Rate, D
Hist. PCE Core	1	PCEPILFE	Rate, D
Hist. CPI Core	1	CPILFESL	Rate, D

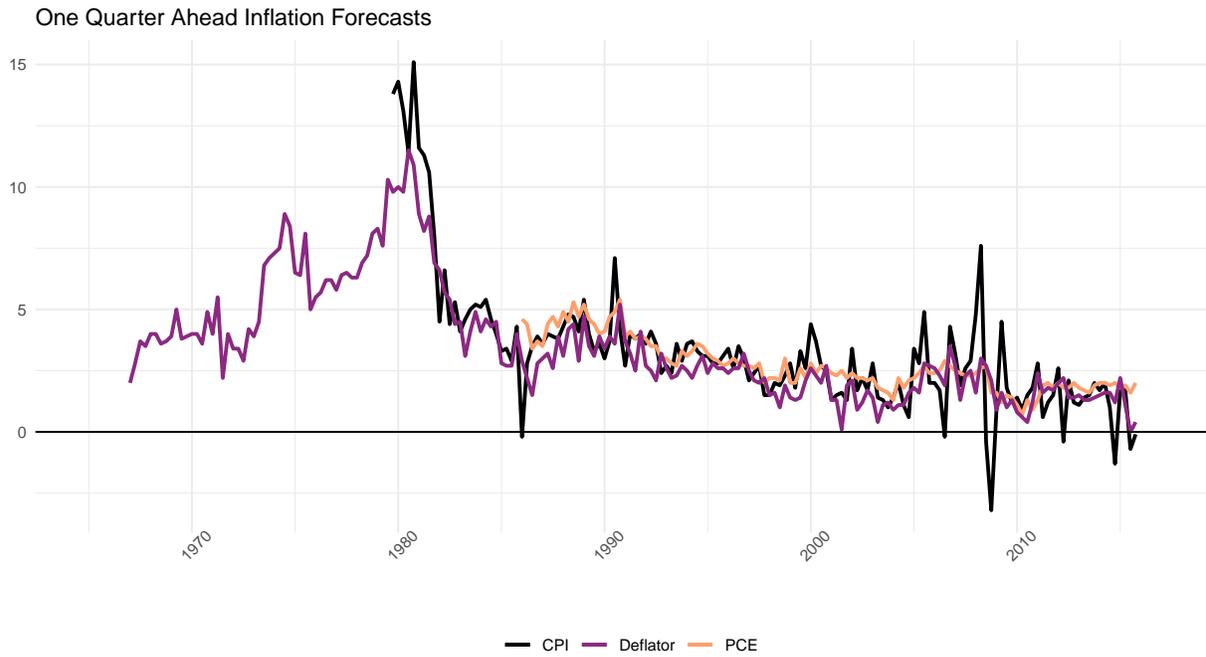
Sources: (1) St. Louis Fed's FRED database; (2) Philadelphia Fed's database; (3) Yahoo! Finance. Transformation: (A) each quarter's last observation; (B) residual of quadratic time trend OLS using observations up to time t ; (C) deviations from full potential (output) or long run estimates (employment); (D) annualised log difference $\hat{x}_t = 400 \times (\ln(x_t) - \ln(x_{t-1}))$; (E) spread with 3-month or 10-year Treasury Bill.

Table 2.8 – Correlations

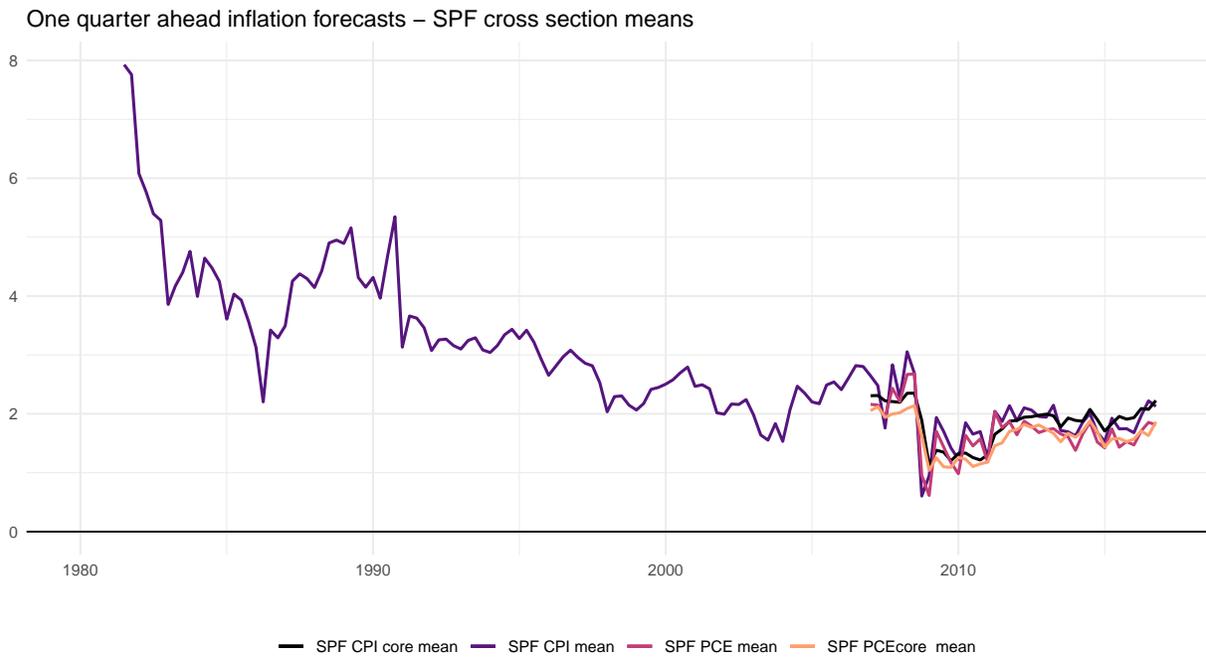
	r_t	r_{t-1}	$E_t\pi_t^{defl}$	$E_t\pi_{t+1}^{defl}$	$E_t\pi_t^{cpi}$	$E_t\pi_{t+1}^{cpi}$	$E_t\pi_t^{pce}$	$E_t\pi_{t+1}^{pce}$	$E_t\hat{y}_t$	\hat{y}_t	\hat{u}_t	SP	BAA
r_t	1	0.982	0.547	0.693	0.393	0.518	0.729	0.822	0.616	0.614	-0.735	-0.622	-0.623
r_{t-1}	0.982	1	0.549	0.682	0.356	0.484	0.735	0.817	0.612	0.575	-0.723	-0.629	-0.552
$E_t\pi_t^{defl}$	0.547	0.549	1	0.543	0.476	0.206	0.667	0.629	0.304	0.145	-0.311	-0.499	-0.311
$E_t\pi_{t+1}^{defl}$	0.693	0.682	0.543	1	0.410	0.571	0.727	0.813	0.218	0.192	-0.347	-0.406	-0.431
$E_t\pi_t^{cpi}$	0.393	0.356	0.476	0.410	1	0.493	0.500	0.431	0.223	0.223	-0.260	-0.179	-0.471
$E_t\pi_{t+1}^{cpi}$	0.518	0.484	0.206	0.571	0.493	1	0.469	0.582	0.214	0.204	-0.314	-0.292	-0.414
$E_t\pi_t^{pce}$	0.729	0.735	0.667	0.727	0.500	0.469	1	0.861	0.295	0.153	-0.366	-0.466	-0.399
$E_t\pi_{t+1}^{pce}$	0.822	0.817	0.629	0.813	0.431	0.582	0.861	1	0.357	0.222	-0.449	-0.533	-0.461
$E_t\hat{y}_t$	0.616	0.612	0.304	0.218	0.223	0.204	0.295	0.357	1	0.736	-0.881	-0.551	-0.355
\hat{y}_t	0.614	0.575	0.145	0.192	0.223	0.204	0.153	0.222	0.736	1	-0.889	-0.393	-0.557
\hat{u}_t	-0.735	-0.723	-0.311	-0.347	-0.260	-0.314	-0.366	-0.449	-0.881	-0.889	1	0.543	0.533
SP	-0.622	-0.629	-0.499	-0.406	-0.179	-0.292	-0.466	-0.533	-0.551	-0.393	0.543	1	0.235
BAA	-0.623	-0.552	-0.311	-0.431	-0.471	-0.414	-0.399	-0.461	-0.355	-0.557	0.533	0.235	1

Notes: r is the FFR, $E_t\pi^i$ is now- or one period ahead forecast for inflation measure i , $E_t\hat{y}_t$ is real time output gap, \hat{y}_t is capacity utilization, \hat{u}_t is fluctuations around the natural unemployment rate, SP and BAA are the spreads.

2.B Additional plots

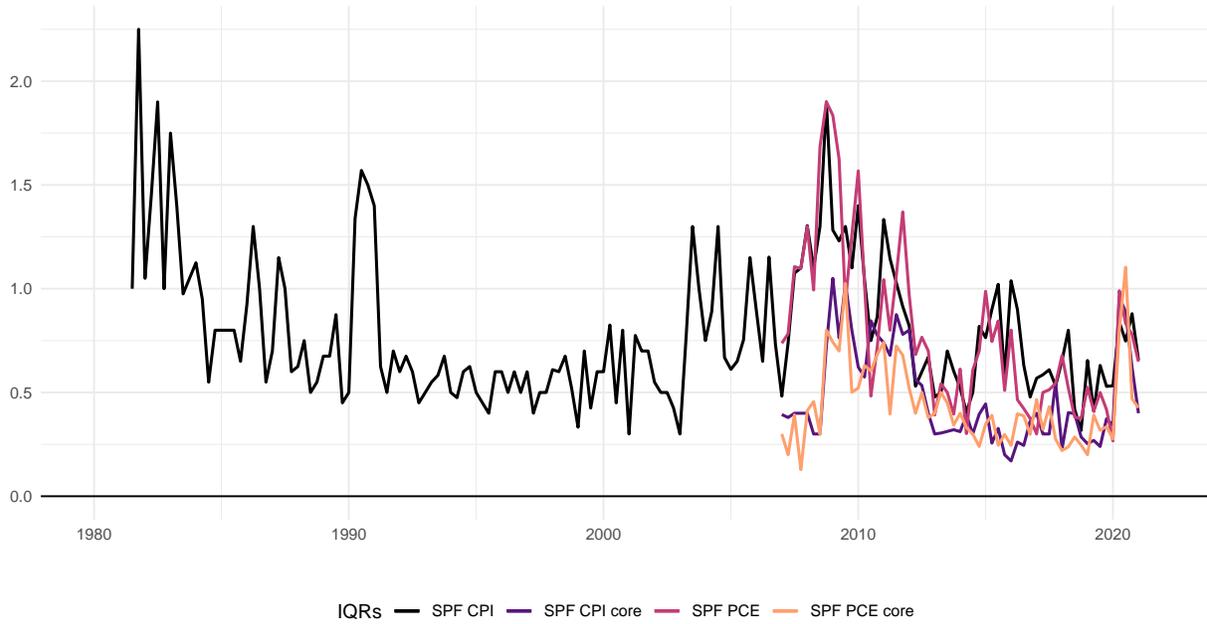


One-period ahead forecasts for CPI, PCE, GDP deflator; CBO's Greenbook database.



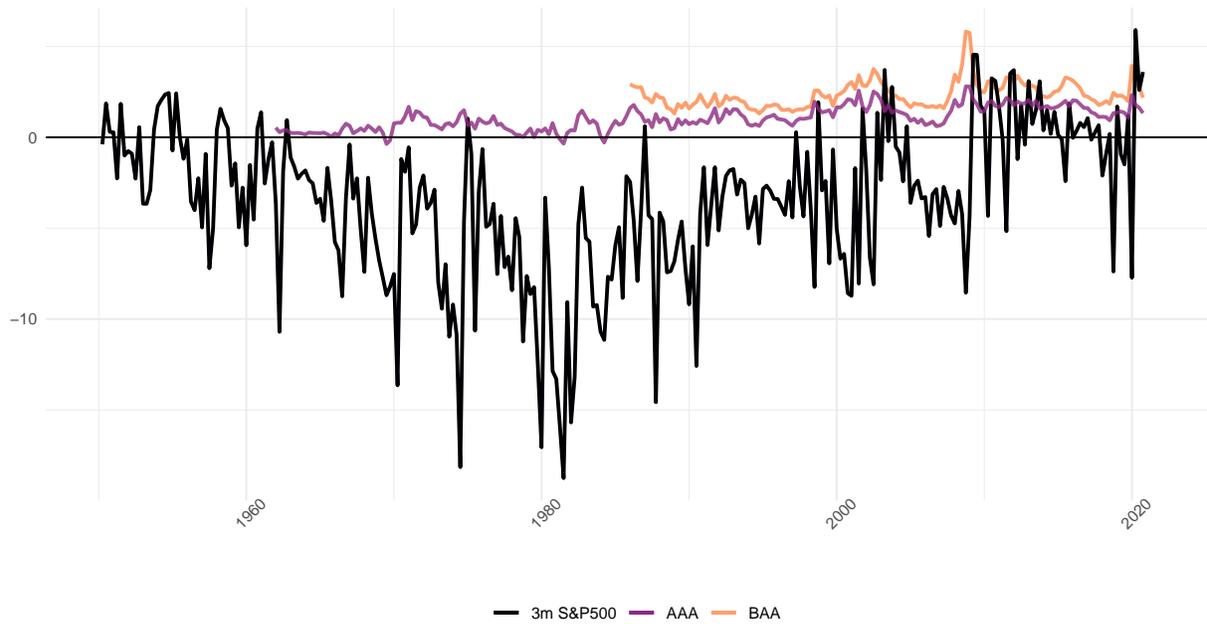
Mean and median from Survey of Professional Forecasters.

One Quarter Ahead Inflation Forecasts – SPF Cross Section IQR



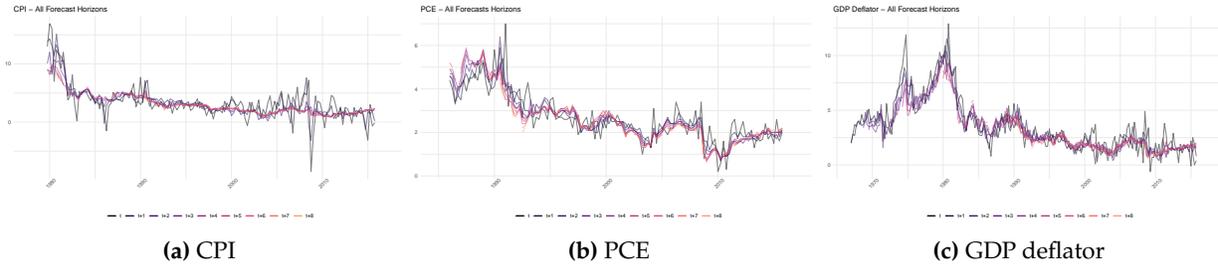
Interquartile range in inflation forecasts from the Survey of Professional Forecasters.

Liquidity Spreads – Financial Liquidity

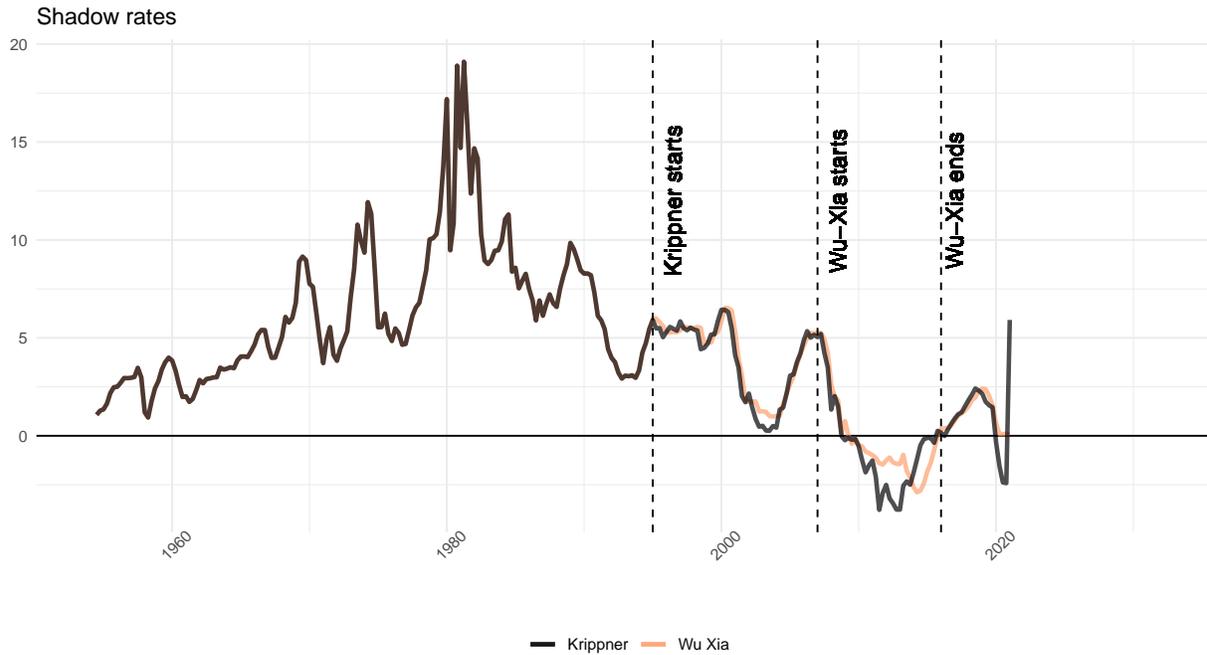


Spreads between AAA, BAA bond indexes and 10 Treasury Bill yield; spread between S&P 500 quarterly returns and 3-month Treasury Bill yield.

Figure 2.5 – Nowcasts and Forecasts

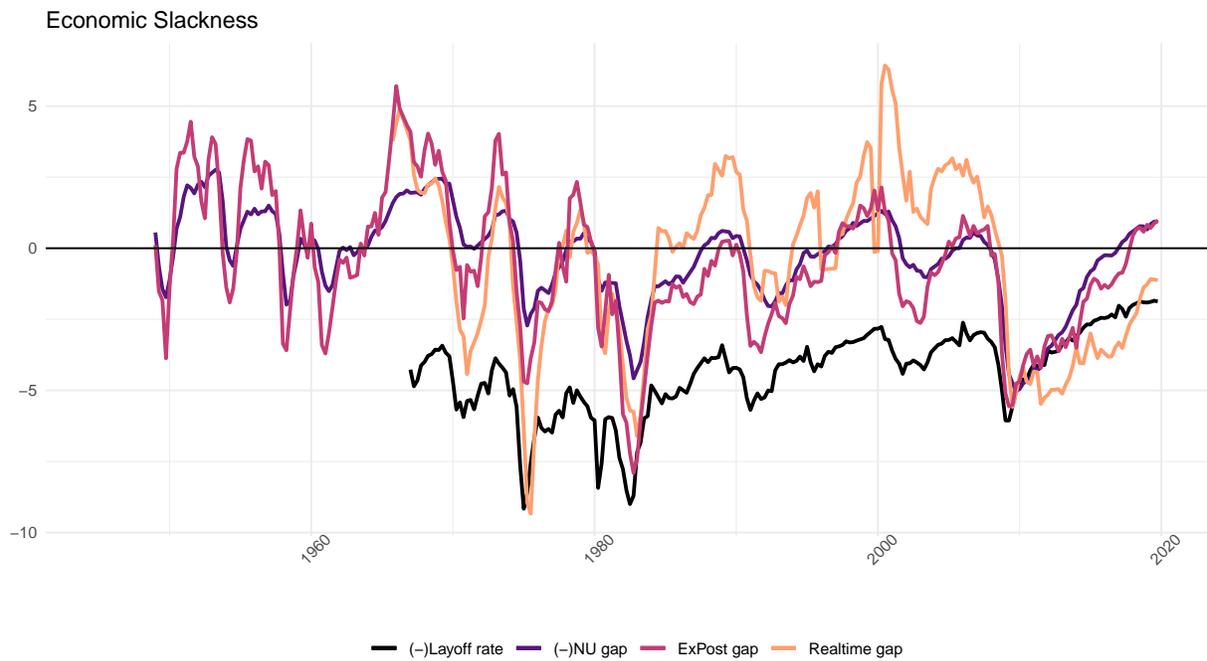


CPI, PCE, GDP deflator nowcasts and forecasts up to two years



Krippner (2015) and Wu and Xia (2016) shadow rates for the federal funds rate.

Figure 2.6 – Measures of output gap.



2.C Empirical Appendix

2.C.1 Output gap

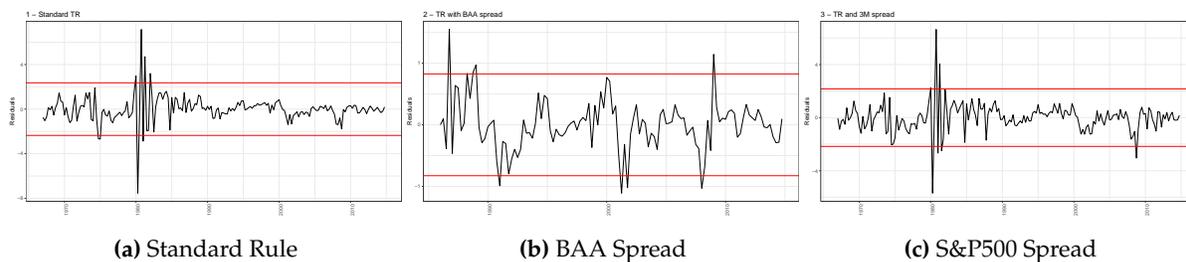
We detail the econometric derivation of the real-time output gap from the Greenbook data set. This data set offers data and forecasts prepared by the Federal Reserve System staff for the Federal Open Market Committee meetings, where policy rate decisions are discussed. To gauge the output gap at any date t , we select all observations up to time t , run a simple quadratic time trend and keep the last observation residual. This term is then normalised as percentage deviation from the computed trend and used as output gap observation. The algorithm is the following:

- 1 set a number j of observations sufficiently high to compute precise estimates
- 2 for all t in $[j, T]$ run the following routine
 - i* sub-select observations in the $[j, t]$ interval
 - ii* estimate $y_s = \alpha_1 s + \alpha_2 s^2 + \epsilon_s, s = j, j + 1, \dots, t$; recover fitted values \hat{y}_s
 - iii* compute $gap_t = \left(\frac{y_t - \hat{y}_t}{y_t} \right) \times 100 = \frac{\hat{\epsilon}_t}{y_t} \times 100$; keep the latest observation/nowcast prior to the FOMC meeting
- 3 stack and date all gap_t to construct the *real time output gap* variable

2.C.2 Full sample regression: residuals

Fig.(3.7) plots the residuals generated from the regression on the full sample. As OLS sort of averages over the full sample, sudden and ample fluctuations in the residuals point to observations where the model underperforms. This occurs typically in the late '70s, late '80, and around the GFC period.

Figure 2.7 – OLS Residuals



Residuals plot for the three specifications. Dashed line depicts single residuals as time series, solid horizontal lines contour the 2-SDs area around zero, the expected – and empirical – residuals mean.

2.C.3 GMM estimates

Table 2.9 – Full Sample GMM

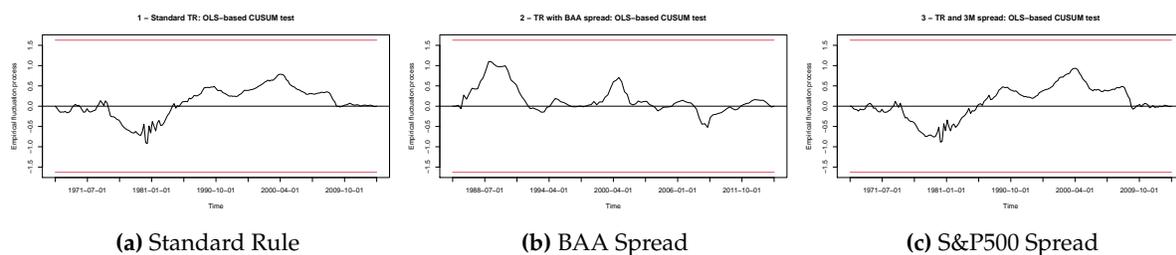
	Specification		
	I – Standard Rule	II – BAA Spread	III – S&P500 Spread
μ	-.9766 (1.68)	12.24* (4.944)	.5633 (1.7384)
$E_t(\text{GDP defl.})_{t+1}$	1.8765* (.8457)	.6913 (.7656)	.8099 (.8209)
Real time \hat{y}	.8096* (.3305)	.3479 (.2919)	.614* (.2858)
BAA		-3.4725*** (.8153)	
S&P500			-.4231 (.3903)
FFR_{t-1}	.907*** (.0332)	.9022** (.0222)	.8966*** (.0416)
Obs.	188	112	188

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '+' 0.1; SE in parentheses. Observations up to 2013 Q2.

2.C.4 CUSUM test plots

This section presents the CUSUM plots for the three specifications of the Taylor rule studied in Section (3.2). The dashed line marks the cumulative sum of residuals, whilst the horizontal solid lines define the significance areas, in which the sum of the residuals signals a change in the underlying data generating process.

Figure 2.8 – CUMSUM plots



Dashed line depicts cumulative sum of residuals, solid horizontal lines mark the significance area for changes in the underlying DGP.

Chapter 3

A simple model with liquidity

In the wake of the 2008 Global Financial Crisis liquidity and finance have risen to the central stage of macroeconomics. As financial and interbank markets froze and policy interest rates fell to zero, major central banks injected unprecedented amounts of liquidity (Quantitative Easing measures, QE) to revive financial markets. Fed's Chairman Ben Bernanke famously commented on the ex-ante effectiveness of Quantitative Easing policies stating that "The problem with QE is [that] it works in practice, but it doesn't work in theory." Indeed, monetary policy frameworks at the time were ill-equipped to deal with such unconventional interventions. Besides lacking a thorough characterisation of the effects of financial turmoil, workhorse models implied a number of puzzles at the Zero Lower Bound (ZLB): implausibly large current effects of small (announced) policy interventions in the future, counter-intuitive effects of increased price flexibility, and sunspot equilibria for output and inflation as the policy rate becomes irresponsive to expectations. The last implication became salient when the ZLB became binding and central banks switched to liquidity management.

This chapter contributes to the study of the role of liquidity in monetary policy in general, and how it affects the restrictions on interest rate rules in particular. We introduce liquid bonds and total liquidity in an otherwise standard New Keynesian monetary model. We equip the monetary authority with a simple rule that fixes total liquidity and sets the interest rate on liquid assets. Expanding the set of bonds, differentiated by their degree of liquidity, permits a more realistic characterisation of interest rate policy, with the policy rate associated to a liquid bond, rather than the risk-free non-liquid bond typical of standard models. Moreover, with liquid and illiquid assets, the monetary instruments of the central bank are both an interest rate and a quantitative measure of total liquidity.

The first result is that, in contrast with the traditional Taylor principle, this framework yields a stable equilibrium as long as the central bank responds positively to expected inflation, regardless of the intensity of the interest rate adjustment. Secondly, we study how inflation dynamics are affected by monetary policy stances. We find that less reactive central banks impart a higher degree

of persistence to realised inflation, while active policy rules reduce volatility and persistence in simulated inflation.

This model thus provides a simple framework to analyse how the introduction of liquidity affects the response of the economy under a supposedly indeterminate regime. We compare this liquidity model to the baseline standard model that is at the core of the New Keynesian Dynamic Stochastic General Equilibrium (NKDSGE) class, with nominal frictions, technological shocks, and rational expectations. The inclusion of liquidity and a simpler monetary policy rule are sufficient to broadly match and reproduce the behaviour of the NKDSGE baseline model. This approach has several attractive features. First, it allows us to move in known territory, making thus possible to compare the properties of the model with the standard NKDSGE model. Second, it can be easily implemented in existing theoretical structures with negligible adjustments, making possible a direct test against other policy-relevant models. Moreover, such a common ground allows further extensions to the analysis of fiscal policy and banking.

Through stochastic simulations, we analyse how the degree of policy accommodation (summarised by the coefficient on inflation in the interest rate rule) directly affects inflation dynamics. We find that an accommodative monetary policy induces more persistence in the inflation process, making it harder to reach the inflation target.

Finally, in close connection to the 2008 crisis, we experiment with a large drop in liquidity and analyse how different policy regimes affect economic responses.

Related literature

This chapter studies the role and consequences of liquidity in monetary policy and formulates a setup encompassing both aggressive and passive monetary policy interest setting rules. This theoretical model builds on Calvo (2016) and also relates in spirit with Michailat and Saez (2015, 2021), Benhabib, Schmitt-Grohe, and Uribe (2001), and Canzoneri and Diba (2005) and Canzoneri et al. (2008a,b).

These latter contributions in particular extensively study the role of liquid (government) bonds in providing transaction services. As shown in those works, this specification is equivalent to one in which rather than liquid government bonds, the distinction is made between base money and bank deposits, which pay an interest rate, which in turn represents the policy rate. In this different specification, total liquidity is given by broad money (base money plus deposits, usually defined as M1). Canzoneri et al. (2011) take a similar approach in showing that money, liquid bonds, and transaction frictions help in ruling out controversial sunspot equilibria even with passive monetary policy. A similar framework, developed in Canzoneri and Diba, 2005, draws comparable conclusions with a more detailed role for fiscal policy. Our results are similar to those achieved in Canzoneri and Diba (2005) and Canzoneri et al. (2011), but arise from a simpler framework that abstracts from the presence of transaction frictions and fiscal policy. A series of contributions (Ben-

habib, Schmitt-Grohe, and Uribe, 2001; Benhabib, Schmitt-Grohé, and Uribe, 2002; Schmitt-Grohe, Benhabib, and Uribe, 2001; Schmitt-Grohe and Uribe, 2000, 2007) investigates the determinacy properties of monetary policy rules with a framework similar to ours. A common feature of those works is to assume that also firms use money for transactions, and that the central bank reacts only to inflation. We adopt a similar simplified feedback rule for the central bank but we restrict liquidity only to the consumer block of the model. Schmitt-Grohe and Uribe (2007), moreover, show how indeed feedback rules that include output are dominated by simpler rules in terms of welfare.

We elaborate the Calvo (2016) model by fully specifying the supply side of the economy and studying its behaviour in discrete time. This model is a stripped-down version of the framework presented in Calvo and Végh (1996) and Calvo and Végh (1990a,b), which instead are applications to exchange rates in a small open economy. Throughout the chapter, our aim is to focus on a small deviation from the standard NKDSGE framework that constitutes the core of modern fluctuations theory in macroeconomics (Gali, 2015; Walsh, 2003; Woodford, 2003). Diba and Loisel (2021) also propose a simple departure from the baseline NK model, featuring a money-in-utility extension and a pegged policy rate, that solves part of the puzzles.

A foundational approach to the role of money in a search setup is developed and discussed in a series of contributions such as Aiyagari and Wallace (1992), Aiyagari, Wallace, and Wright (1996), and Kiyotaki and Wright (1991). The main focus there is to provide a role for fiat money as medium of exchange. Our take is simpler and assumes away the utility provided by liquid assets and cash, in the interest of studying how they affect monetary policy. Our take is simpler, as we concentrate on reduced-form models in which money or liquidity entered directly the utility function or it enters through liquidity in advance constraints. A general overview of the salient approaches to model liquidity is provided in Lagos, Rocheteau, and Wright (2017), with a focus on money liquidity.

This paper is also related to studies on conditions for determinacy in several vintages of New Keynesian models. Benhabib, Schmitt-Grohe, and Uribe (2001) include money in the production function and studies under which fiscal and monetary policy regimes the model displays indeterminacy. Relatedly, Leeper (1991) studies how fiscal and monetary policy interact and establishes parametric regions for determinacy: fiscal dominance can ensure determinacy over a passive monetary policy. Along the same lines of fiscal interaction, Cochrane (2020) further develops a price level theory grounded on debt and taxes, working around the limitations and shortcomings of the monetary NKDSGEs. In this model, we abstract from fiscal considerations and focus exclusively on the liquidity factor. Likewise, we also avoid considerations on market frictions and distributional consequences of liquidity shocks, which are discussed more broadly in Iacopetta and Minetti (2019): we stick to a representative agent framework and focus on the consequences of liquidity for determinacy.

Similarly, our simplified model abstracts from a refined banking sector to emphasise the role of liquidity. Recent works find that the credit and bank lending channel affect the transmission of monetary policy, especially when credit is directed to the housing sector (Iacoviello and Minetti, 2003, 2008): the inclusion of such channel, though, constitutes a promising avenue for future research.

The remainder of the paper is structured as follows: Section 4.1 introduces liquidity in the NKDSGE framework; Section 4.2 compares the liquidity model with a baseline one; Section 4.3 analyses how monetary rules influence inflation dynamics; Section 4.4 studies the effects of a severe liquidity shock. Finally, Section 4.5 concludes.

3.1 A model with liquidity

We introduce liquid bonds in an otherwise standard, small scale NKDSGE model. This extension is motivated by the utility they provide in being liquid assets, which provide utility services. Introducing liquidity services directly in the agent's utility function is a reduced-form approach, but is an informative exercise to explore the implications of such extensions. Reasonable micro-foundations call for more complete modelling of the financial block of our stylised economy. The main idea is to model agents' liquidity portfolio allocation between plain money and liquid assets, interchangeably bonds or bank deposits. Including banks, firms' financial position, financial frictions is typically carried out in a structural way following the footsteps of Bernanke and Gertler (1989), Bernanke, Gertler, and Gilchrist (1996), and Kiyotaki and Moore (1997). Since our interest lays in understanding the consequences of liquidity rather than its deeper motives, at present we opt for the simplicity of a transparent reduced form.

In the same vein, we approach the policy modelling. While Gertler and Karadi (2011) set out to directly model unconventional monetary policies, we restrict the scope of monetary policy with a simplified rule and the assumption of a fixed amount of nominal circulating assets. We thus postulate a simple feedback rule that relates the interest rate paid on liquid assets to expected inflation. Such simplified rules have been extensively explored for their debated determinacy properties, as is shown in Benhabib, Schmitt-Grohe, and Uribe (2001), Benhabib, Schmitt-Grohé, and Uribe (2002), Schmitt-Grohe, Benhabib, and Uribe (2001), and Schmitt-Grohe and Uribe (2000, 2007).

3.1.1 Consumer

The major variation from workhorse models is limited to the structure of flow-utility and overall assets portfolio of the consumer. We postulate an economy with an infinitely-lived representative agent, who works, consumes, and holds money alongside two types of assets. Therefore, total wealth is allotted between a liquid bond B , cash M , and an illiquid bond X . This latter serves

the only purpose of smoothing consumption over time and allocate wealth intertemporally. All assets mature after one period. We assume the consumer is willing to hold B and M because they provide transaction services, and therefore utility. In addition, B bonds pay a nominal interest rate s , X ones pay nominal interest rate i . Cash holdings pay no interest and are carried on to the next period, suffering inflation erosion.

While inserting money balances in the utility function is not new (Sidrauski, 1969), adding bond holdings in the utility function is less common. In this respect, our model is close to Michailat and Saez (2021), although the resulting Euler equation at the steady-state is not modified by bonds but rather by money holdings. In fact, it is completely homomorphic to introduce a cash-in-advance (or rather, liquidity-in-advance) constraint, as in Calvo and Vegh (1996).¹

Under these assumptions, the utility maximization problem of the consumer takes the following form:

$$\begin{aligned} & \max_{\{c_s, M_s, b_s, N_s\}_{s=0}^{\infty}} E_t \left[\sum_{s=0}^{\infty} \beta^s (u(c_{t+s}) + h(b_{t+s}) + v(M_{t+s}/P_{t+s}) - g(N_{t+s})) \right] \\ & \text{s.t. } C_t + M_t + X_t + B_t = W_t N_t + (1 + s_{t-1}) B_{t-1} + (1 + i_{t-1}) X_{t-1} + M_{t-1} \end{aligned} \quad (3.1)$$

Where we assume additively separable (dis)utilities for consumption c , cash m , liquid bonds b , and hours worked N . Moreover, c is the result of aggregating a measure one of differentiated goods via a Dixit-Stiglitz aggregator with constant elasticity of substitution θ . This also implies that P is the price index of the underlying goods.

The intertemporal budget constraint summarises expenditures, allocations and income sources: interests promised at $t - 1$, carried-on money, and labour income.

It is useful to reformulate the budget constraint in terms of total wealth and real quantities before deriving the system of first-order conditions. Therefore, let $D = X + M + B$ be the total wealth held by the consumer. Replacing $X = D - M - B$ in the budget constraint and appropriately dividing through P gives the following real budget constraint.

$$c_t + d_t = w_t N_t - \frac{i_{t-1} - s_{t-1}}{1 + \pi_t} b_{t-1} + (1 + r_{t-1}) d_{t-1} - \frac{i_{t-1}}{1 + \pi_t} m_{t-1} \quad (3.2)$$

Where $\pi = \frac{P_t}{P_{t-1}} - 1$ is the inflation rate and lower-case indicates real quantities. Rewriting the constraint in such forms highlights the opportunity costs of holding bonds b and cash m : assuming a positive spread $i > s$, the consumer gives up the additional interest paid by illiquid bonds for every additional unit of b the consumer holds.

¹Appendix 4.C proposes an alternative setup with a Liquidity-In-Advance which yields very similar quantitative results.

Similarly, holding cash entails giving up entirely on the nominal interest rate i . These two opportunity costs are offset by the marginal utilities provided by holding such assets and will determine holdings and allocations at the equilibrium.²

With this reformulation, the FOCs system implies the equilibrium conditions in (4.3):

$$\begin{aligned} u'(c_t) &= E_t [\beta (1 + r_t) u'(c_{t+1})] & h'(b_t) &= E_t \left[\beta \lambda_{t+1} \frac{i_t - s_t}{1 + \pi_{t+1}} \right] \\ \frac{g'(N_t)}{u'(c_t)} &= w_t & v'(m_t) &= E_t \left[\beta \lambda_{t+1} \frac{i_t}{1 + \pi_{t+1}} \right] \end{aligned} \quad (3.3)$$

where λ_t is the Lagrangian multiplier, the first equation is the usual Euler equation for intertemporal consumption, the second equates marginal cost and benefits of work, and the last two equations govern the allocation decision between liquid bonds b and real money balances m . In particular, the latter condition discounts money holdings by the real interest rate, that is the loss from inflation; the equilibrium condition on bonds, instead, regulates liquid asset holdings on the spread between nominal interest rate i and yield of such bonds, both in real terms. For the purpose of this chapter, we will assume a non-negative spread between i and s , such that the consumer holds simultaneously cash and bonds at every period.

3.1.2 Firms

The production side of this economy is straightforward and assumes a measure one of infinitesimal firms, indexed by $j \in [0, 1]$. Each firm is embedded with a technology that employs only labour, so that the production function is

$$Y_{jt} = A_t N_{jt}^a. \quad (3.4)$$

The term A captures the stochastic productivity of the economy and follows a simple AR(1) process, $a \in (0, 1)$ represents the decreasing returns to scale, and N_{jt} the individual employment of each firm. As the consumption good results from the CES aggregator, every firm j faces a demand schedule (4.5), relative to aggregate production, with θ being the elasticity of substitution and P_{jt} the firm's price.

$$Y_{jt} = \left(\frac{P_{jt}}{P_t} \right)^{-\theta} Y_t \quad (3.5)$$

²This reduced form specification is presented to highlight the effects of liquidity on monetary policy. A Liquidity-in-Advance version is deferred to Appendix (4.C).

We assume nominal rigidity à la Calvo (1983):³ every period, a $(1 - \alpha)$ portion of firms is given the chance to update their price, while the remaining share will stick to previously set prices. This entices a forward-looking behaviour in firms when they optimise their expected discounted profits, as they take into account the duration of their price. Firm j 's marginal cost is $\mathcal{MC}_{jt} = \frac{W_t/P_t}{aY_{jt}/N_{jt}}$, whence the expected discounted profits in eq.(4.6).

$$\begin{aligned} \max_{P_j^*} \quad & E_t \left[\sum_{s=0}^{\infty} \alpha^s Q_{t,t+s} \left(P_j^* Y_{jt+s} - \mathcal{MC}_{jt+s} Y_{jt+s} \right) \right] \\ \text{s.t.} \quad & Y_{jt+s} = \left(\frac{P_{jt+s}}{P_{t+s}} \right)^{-\theta} Y_{t+s} \end{aligned} \quad (3.6)$$

Where α is the Calvo pricing parameter, $Q_{t,t+s}$ is the stochastic discount factor between periods t and $t + s$ the consumer, who owns all firms, uses to weight future profits, and P_j^* is the optimal price chosen by the firm. Factoring in the constraint and solving the program with a symmetry argument gives two results. First, firms price with a constant markup over the marginal cost, and second that the optimal price is specified as a function of expected future marginal costs, price index levels and economic activity, as shown in eq.(4.7). This same equation also presents the inflation dynamic as an autoregressive process of order one, depending on past prevailing prices and current updated prices.

$$\frac{P^*}{P_t} = \frac{\theta}{\theta - 1} \frac{E_t \sum_{s=0}^{\infty} (\beta\alpha)^s Y_{t+s} \left(\frac{P_{t+s}}{P_t} \right)^{\theta} \mathcal{MC}_{t+s}}{E_t \sum_{s=0}^{\infty} (\beta\alpha)^s Y_{t+s} \left(\frac{P_{t+s}}{P_t} \right)^{\theta-1}} \quad P_t^{1-\theta} = (1 - \alpha) P_t^{*1-\theta} + \alpha P_{t-1}^{1-\theta} \quad (3.7)$$

3.1.3 Monetary authority and market clearing

We also depart from the standard setup in characterising the central bank, its policy, and total liquidity management. We assume the existence of a monetary authority that operates in two ways in the economy. First, it sets the total amount of liquidity in circulation, namely eq.(4.8). Importantly, the central bank does not directly determine the allocation between cash and liquid bonds, but only the sum of the two, irrespectively of the portfolio composition. This setup provides the baseline building block for further extensions of this framework. Equipping monetary policy with quantitative easing tools equates to provide the central bank with a balance sheet and the channels to adjust the total nominal liquidity Z_t . For the sake of simplicity, we will also assume a fixed nominal amount of circulating liquidity, so that

³The precise modelling of the nominal rigidity is inconsequential, as the main novelties are in the consumers' side of the model. Hence, quadratic adjustment costs as in Rotemberg (1982) might be added without loss of fundamental insights.

$$Z_t = B_t + M_t = \bar{Z} \quad \forall t \quad (3.8)$$

A fixed amount of liquidity implies no issuance of liquid bonds: ideally, to capture the effects of the central bank's quantitative easing and balance sheet policies, B would be issued by either a government (Cochrane, 2020), banks (Gertler and Karadi, 2011), firms (Ravenna and Walsh, 2006; Surico, 2008), or combinations thereof. The same principle applies to money, which is injected or withdrawn to seamlessly counteract the movements in liquid bonds trade. In this light, B closely resembles bank deposits, for both can be readily exchanged for cash. One can think of Z as a fixed money supply (M2), although a full characterisation would have B issued by the government in the form of liquid, risk-free treasury bonds. Along the same lines, the central bank would influence the circulation of these bonds to manage total liquidity and, more importantly, to coordinate with fiscal policy.⁴

The central bank also sets the policy interest rate s_t paid on liquid bonds following a constant rule. It is useful to think in this context to the link between the Federal fund rate and the interest rate on the shortest-maturity Treasury Bill. In detail, the rule responds solely to inflation expectations and does ignore any level of economic slack contrary to more usual Taylor feedback rules:

$$\exp(s_t) = \exp(E_t \gamma \pi_{t+1}). \quad (3.9)$$

This skeletal structure is clearly a simplified Taylor rule, but the values for γ are crucial. An accommodative central bank, with $0 < \gamma \leq 1$, does not necessarily trigger unstable sunspot equilibria in our model, but rather stabilises the economy despite the accommodative stance.⁵ Such a rule can easily be extended to a full-fledged Taylor Rule including real slack without impairing or modifying the final results. As an actual example, this specification of the Taylor Rule relates to that of the European central bank, which in its mandate contemplates explicitly only price stability and not employment or economic slack.

The rationale behind this specification is that, keeping under control the liquidity in circulation, the Central bank assures that inflation follows a specified path. This particular specification of the monetary policy links the inflation rate and the return yield on liquid bonds. Therefore, the central bank in our model levers the liquidity allocation via interest rate setting over a fixed amount of Z_t .

⁴Notable examples are the CARES Act in the US, or the combined SURE and Next Generation EU in Europe, both in response to Covid-19 crisis. These were preceded by ECB's APP-PSPP and Fed's Open Market Operations as measures for the 2008 crisis.

⁵Although it is worth recalling that a central bank complying with the Taylor principle actually *destabilizes* the economy (Benhabib, Schmitt-Grohe, and Uribe, 2001).

Last, taking eq.(4.8) and dividing through by the prices level, one can obtain the values for liquidity allocation in real terms, as well as a backwards-looking expression for real liquidity depending on current inflation.

$$z_t = m_t + b_t \iff z_t = \frac{z_{t-1}}{1 + \pi_t} \quad (3.10)$$

This last equation, with the market clearing condition $C_t = Y_t$, closes the model.

3.1.4 Linearised model

In this section we briefly present the system of equations resulting from loglinearising eqs. (4.3), (4.7), (4.9), and (4.10) around a zero-inflation steady state. To this end, we assume precise functional forms for the utility functions, namely CRRA

$$\begin{aligned} u(c) &= \frac{c^{1-\sigma}}{1-\sigma} & v(m) &= \frac{m^\psi}{\psi} \\ h(b) &= \frac{b^\phi}{\phi} & g(N) &= \chi \frac{N^{1+\eta}}{1+\eta} \end{aligned} \quad (3.11)$$

With $1 \geq \phi > \psi > 0$, which implies that the consumer is more sensitive to bonds rather than money, in line with everyday financial decisions. The other functional forms assumed are consistent with more traditional exercises and well settled in the NKDSGE literature. This shared ground highlights how consequential liquid assets can be, once modelled as a complement to cash, especially on the policy rule.

The loglinearised model consists of a system of linear equations whose properties can be easily and extensively studied. We perform a comparison with the simple 3-equation model presented in Gali (2015), for example.⁶

$$\hat{y}_t = \frac{1-\psi}{\sigma} \hat{m}_t + E_t \hat{y}_{t+1} + \frac{1}{\sigma} E_t \pi_{t+1} \quad (3.12)$$

$$\pi_t = \beta E_t \pi_{t+1} + \kappa (\hat{y}_t - \hat{y}_t^f) \quad (3.13)$$

$$(1-\phi) \left(1 + m^{\frac{\phi-\psi}{1-\psi}}\right) \hat{z}_t = \left[\frac{1-\psi}{m^2} + m^{\frac{\phi-\psi}{1-\psi}} (1-\psi) \right] \hat{m}_t + \left(\beta y^{-\sigma} m^{1-\psi} \right) s_t \quad (3.14)$$

⁶A more detailed walk-through for obtaining this system is provided in the Appendix (4.A.1)

$$\hat{z}_t = \hat{z}_{t-1} - \pi_t \quad (3.15)$$

$$s_t = \gamma E_t \pi_{t+1} \quad (3.16)$$

Eq.(4.13) is the Phillips curve of this economy, in line with more classical models. Eq.(4.12) is the Euler equation augmented with the real money balances, which affect positively the contemporaneous output gap. This wedge subsumes the transaction utility of liquidity: positive values for \hat{m} entail higher output levels. As the Phillips Curve (4.13) relates economic slack to inflation, the wedge also affects this latter, and eventually its expected value.

Eq.(4.14) summarises the money demand function (depending on liquid bonds interest rate s , total real liquidity z). This relation shows how, indeed, the central bank affects asset allocation for a given level of total liquidity, leveraging the policy rate s_t .

Eq.(4.15) captures intertemporal changes in total real liquidity, and finally eq.(4.16) represents the monetary policy rule.

We adopt the convention that \hat{x} is the percentage deviation of x from its steady-state, while all lower-case, unhatted, and time-independent variables are steady-state values. Moreover, we use y^f for frictionless output, so that $(\hat{y}_t - \hat{y}_t^f)$ is the output gap.

Finally, some exogenous variables and parameters are grouped as follows:

$$\begin{aligned} \kappa &= \left[\frac{(1-\alpha)(1-\alpha\beta)a}{\alpha(a+\theta(1-a))} \right] \left[\frac{1+\eta+a(\sigma-1)}{a} \right] & m &= \left(\frac{y^\sigma}{1-\beta} \right)^{\frac{1}{1-a}} \\ \hat{y}_t^f &= \frac{\eta+1}{1+\eta+a(\sigma-1)} \hat{A}_t & y &= \frac{\eta+1}{1+\eta+a(\sigma-1)} \end{aligned} \quad (3.17)$$

As this approximation of the full model can be easily simulated, we exploit it to check for which calibration sets the model generates a unique and stable equilibrium.

To fully specify the model process, we introduce two standard shocks to perturb the model around its steady state to display convergence dynamics. Both shocks follow a stationary autoregressive process of order one:

$$\begin{aligned} \hat{A}_t &= (1-\rho_A) \bar{A} + \rho_A \hat{A}_{t-1} + \epsilon_t^A & \epsilon_t^A &\sim N(0, \sigma_A) \\ v_t &= \rho_v v_{t-1} + \epsilon_t^v & \epsilon_t^v &\sim N(0, \sigma_v) \end{aligned}$$

In addition to the real, technological disturbance, we add a policy shock impulsed by the central bank. The former induces a change in the total factor productivity A from its steady-state value, set to 1. The latter is a shock to the monetary rule detailed in eq.(4.9). These two shocks

allow the comparison with the aforementioned standard models, so to perform a horse race and check the consistency of our augmented model. Assessing whether our setup replicates the classic reactions of well-known models is a first paramount consistency check to validate its structure and internal workings before further analysing policy stance implications.

Both shocks are simply added to their respective equations, the production function and the monetary policy rule, and throughout the whole simulations, we calibrate their persistence parameters ρ to the same value.

3.2 Calibration and IRFs

Calibration and simulation is a usual exercise and there is a large literature to inform our choice of parameters' values.

We calibrate the model to the values presented in Table (4.1), taking the most common values used in the literature. The novel restriction involved in the model concerns exponents of bonds and real balances utility functions. As of the values chosen for semi-elasticities ψ and ϕ , values from the literature are scarce, thus we conservatively underweight liquid bonds with respect to money holdings. For the "Taylor Principle" parameter of our policy rule, we explore two candidate values. The first is supposed to violate the Blanchard and Kahn (1980) condition and thus produce an unstable solution, whilst the second is the one most commonly found in both empirical studies and theoretical exercises.⁷ These two values are also what we find in 3 for a variety of regimes that include or exclude financial liquidity.

The remaining calibrated parameters are fairly standard. Price duration is chosen to obtain four quarters on average Woodford (2003). The parameter governing returns to scale matches the labour share usually observed in the data. Intratemporal elasticity offers several values and affects firms' markup, we conservatively pick a low value from Bilbiie, Ghironi, and Melitz (2019), used to target micro evidence. For intertemporal elasticity of substitution, we pick a fairly high value in comparison to what Havranek et al. (2015) find in a meta-analysis of roughly 170 studies. If anything, this high value penalises the wedge present in the Euler equation (4.12), but equally affects the comparison with NKDSGE models.

Two facts emerge from this calibration exercise. First, our calibrated model generates a unique, stable equilibrium for both values of γ . Thus, sunspot equilibria are ruled out even if the central bank reacts passively to the inflation expectations of the economy.

Second, our model behaves as expected once compared with the 3-equation NKDSGE counterpart.⁸

⁷Appendix 4.D offers some values for semi-elasticities ψ and ϕ . These changes mainly affect persistence and do not dramatically affect the quantitative results, nor do they impart the qualitative profiles of IRFs. Similarly, Appendix 4.E explores the implications of varying values of γ on inflation persistence.

⁸Gali (2015), chapter III, version with interest rate rule.

Table 3.1 – Calibration for model simulations

Parameter	Descr.	Value
a	ret. to scale	.6
β	discount rate	.975
σ	intertemp. el. of subst.	5
θ	intratemp. el. of subst.	3.8
α	price duration	.75
η	Frisch elast.	1
χ	labour disutility	1
ρ_A	persistence, TFP shock	.65
ρ_v	persistence, MP shock	.65
ψ	bond el.	.02
ϕ	money el.	.65
γ	Passive Rule	.5
	Active Rule	1.8

To illustrate this point, we first compare side-by-side the effects of a technological shock (Fig.(4.1)), then compare the effects of a monetary policy shock under two regimes for our model (Fig.(4.2) and Fig.(4.4)).

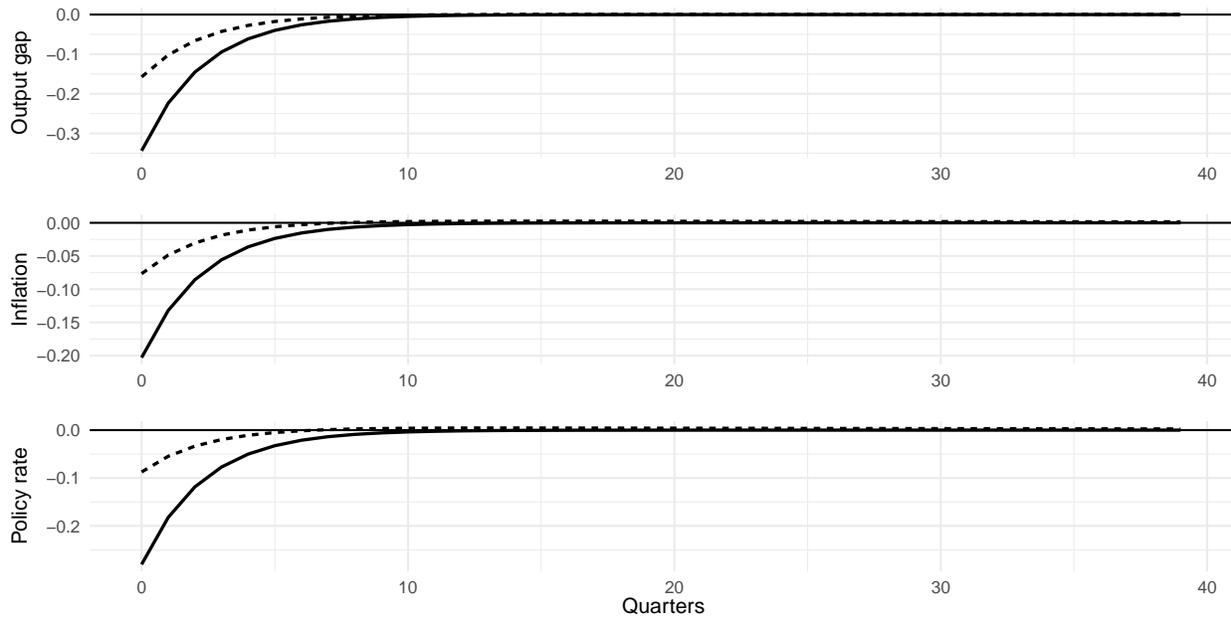
Both shocks are present in the baseline versions of the models in exam, which allow for a meaningful comparison. In order to carry out the comparison on equal ground, the two models are calibrated with the very same values for the common parameters. In fact, under the same calibration, the two models show the same behaviour in terms of reactions to shock, as it is possible to see in the next figures.

Technological shock

Fig. (4.1) depicts the reactions to a one-standard-deviation, positive shock to total factor productivity, \hat{A}_t . It produces the same response in our model and in the standard NKDSGE one. This first result is not surprising, because of the very same structure of the production side of the two stylised economies.

We focus on three common endogenous aggregates to evaluate the matching between the models. Following a TFP shock, in both cases the output gap turns negative, inflation falls, and the reference interest rate falls in response. While this reaction is common across models, magnitudes mark little differences: output and inflation fall more in the NKDSGE model, whereas our proposal implies less movement in these aggregates: for inflation and output gap, the response in our

Figure 3.1 – TFP Shock – IRFs



Liquidity model (dashed, $\gamma = 1.8$) and Gali (2015) (solid): impulse response functions following a one standard-deviation technology shock. Values in percentages.

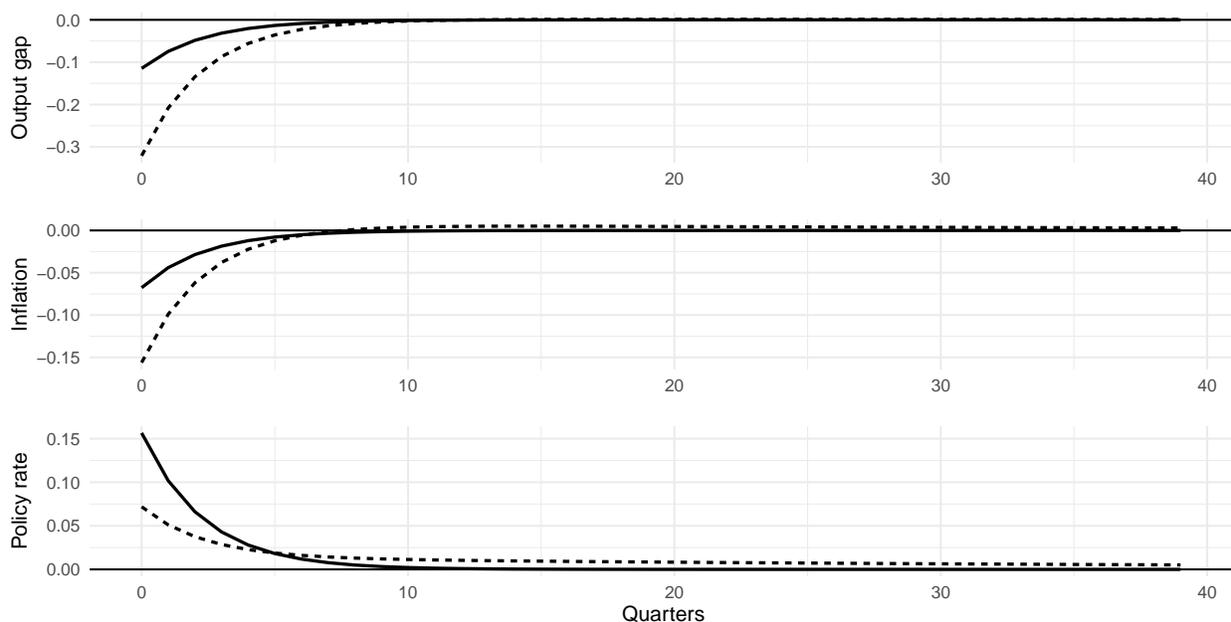
model is broadly half on impact. This milder response in our model is driven by the stripped-down version of our monetary policy rule, which prevents output gap movements to percolate into the variations in s . An alternative, concurrent explanation for the dampened propagation works through the two complete financial markets included in our model, where the TFP shock is dissipated smoothly through two complete financial markets.

Monetary policy shock

When the modelled economies are hit by a monetary policy shock and the liquidity model complies with the Taylor Principle, they generate the IRFs pictured in Fig.(4.2). The IRFs produce the same profiles in both cases, but the liquidity models report higher volatility and greater impact on output gap and inflation. In both cases, the magnitude is roughly thrice that of the baseline NKDSGE, while direction and adjustment correspond closely. The skeletal policy rule engrained in the liquidity model abstracts from the economic activity level, thus the central bank is not facing trade-offs between economic activity and inflation and focuses solely on the latter. This monetary policy rule, moreover, impedes the feedback loops between policy rate, inflation and economic activity, designing a different propagation mechanism: adjustments in the market for money and liquid bonds affect consumption and, then, the supply side of the economy.

Focusing on the policy rate, our model responds roughly twice less than the NKDSGE model to the very same shock: the propagation mechanism works mainly through the consumer's financial

Figure 3.2 – Interest Rate Shock – IRFs



Liquidity model (dashed, $\gamma = 1.8$) and Gali (2015) (solid): impulse response functions following a 1% key rate hike, annualised. Values in percentages.

position in the model with liquidity. In addition, our monetary shock is well more persistent in the interest rate, converging back to zero only when inflation levels off, too. The absence of economic activity in the rule makes that the policy rate essentially mirrors the inflation profile. In comparison to the baseline model, the liquidity one presents a slower convergence path to the steady-state. This model, therefore, entails a higher degree of persistence in its design, even with an active central bank.

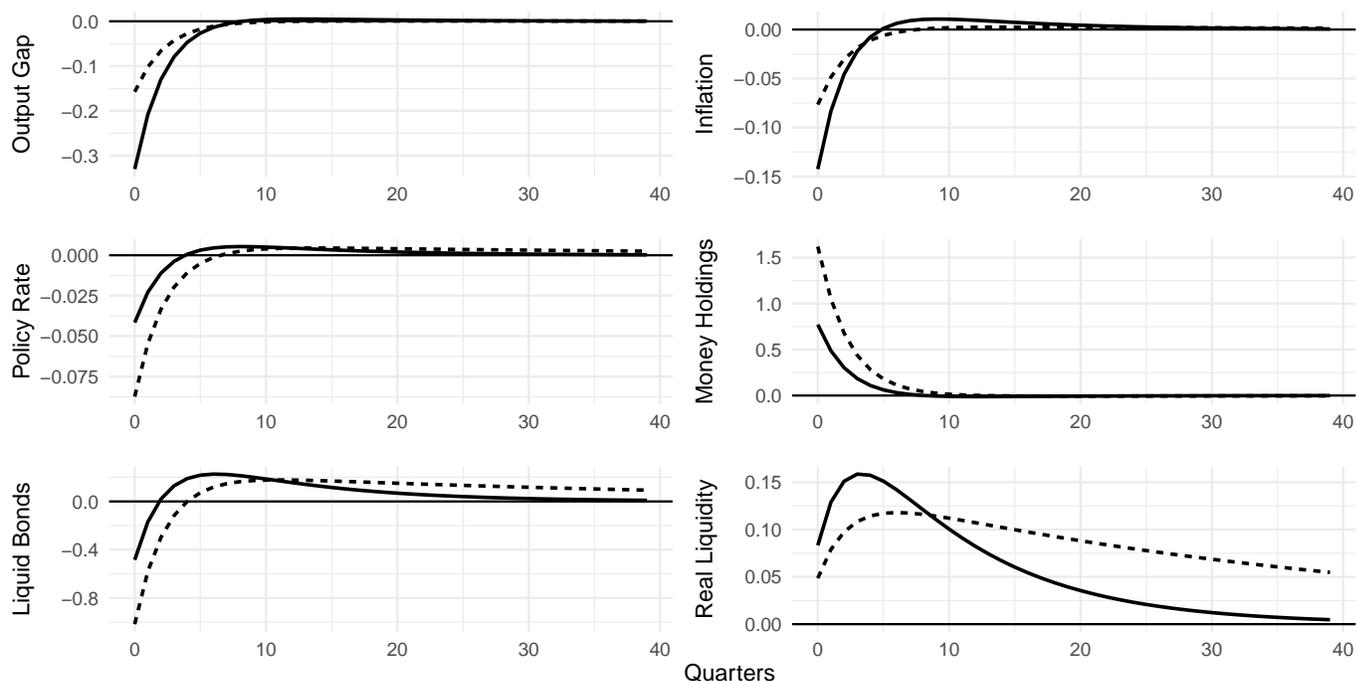
A more thorough analysis of the effects of monetary policy stance on inertia will be the focus of Section 4.3.

3.2.1 Liquidity with an accommodative central bank

After checking that our model consistently matches the response of the workhorse model, we turn now to showing how the liquidity model reacts to the previous shocks under different policy stances. The focus of this exercise is on the passive regime, where we set $\gamma = .5$ and analyse the behaviour of all endogenous aggregates in comparison to the active policy stance. We also include the dynamics of liquid bond holdings, recovered from the FOCs, to complete the picture.

We start with the IRFs generated by a positive TFP shock, collected in Fig.(4.3). In the aggressive regime, output, inflation, and the policy rate s fall. Interestingly, s and π slightly overshoot before converging to zero from above after about five quarters. Inflation and policy rate dynamics

Figure 3.3 – TFP Shock IRFs – Both Regimes



Impulse Response functions for our model: with accommodative (solid) and aggressive (dashed) reaction to expected inflation after a productivity shock (one standard deviation) – all endogenous variables, values in percentages.

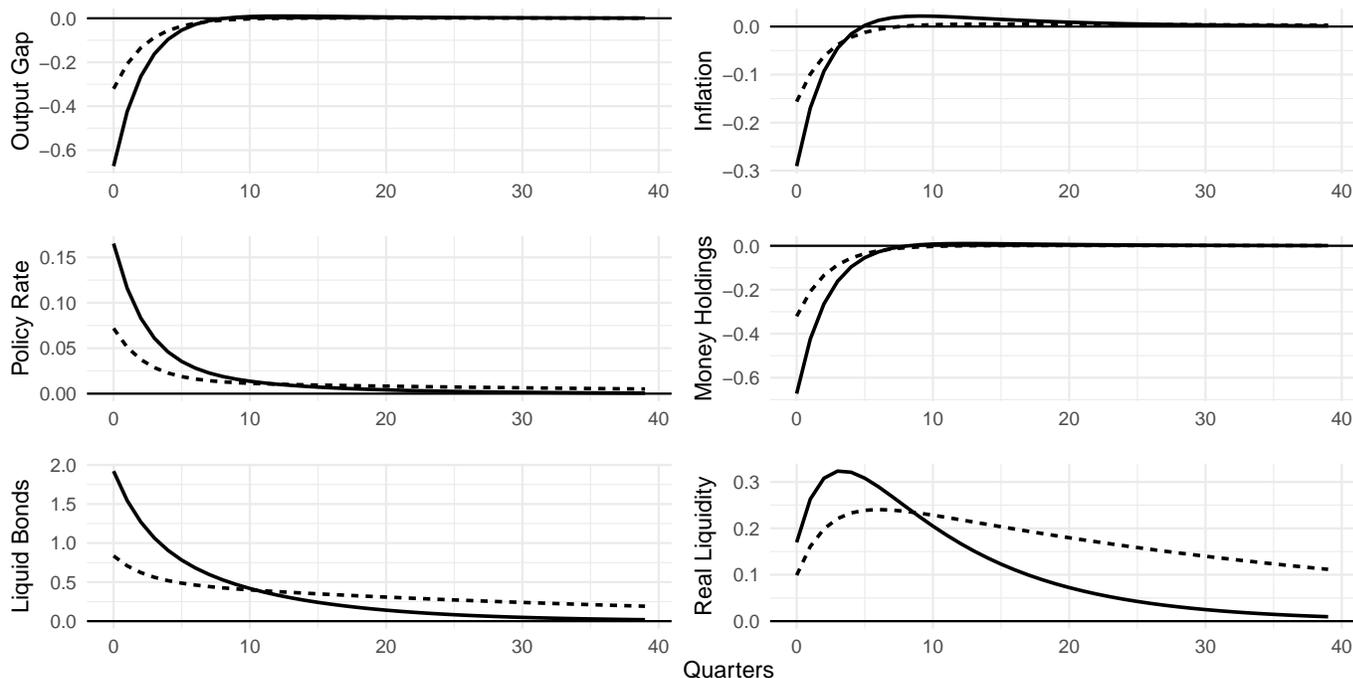
map almost one to one to the asset allocation and real liquidity. This behaviour is reflected in real liquidity and its components: in comparison to the active stance, the technological shock doubles the impact on total real liquidity, the interest rate drop triggers a reallocation away from bonds towards cash m due to deflation, until s overshoots and subsequently offsets the overall effect on liquidity reshuffling.

All in all, the main discrepancy between the two regimes is in the asset block: the inflation path influences dramatically and on a shorter period the amount of real liquidity when the central bank is passive. Conversely, the path back to the steady-state is significantly quicker in the passive regime.

Figure 4.4 summarises the response to a monetary policy shock that raises the interest rate s on bonds of 1%. First off, in both policy regimes output moves in the same direction and is affected similarly. Secondly, inflation falls as expected and after 15 quarters the shock is fully absorbed, without unexpected evolutions of the prices. This last result is particularly relevant as one would expect erratic paths instead of a well-determined dynamic for prices under a passive monetary policy regime.

In close connection to the findings of chapter 3, accounting for the effects of financial liquidity relaxes the stringency of the Taylor Principle. The relevance of financial markets is also shown by

Figure 3.4 – Interest Rate Shock IRFs – Both Regimes



Impulse Response functions for our model: with accommodative (solid) and aggressive (dashed) reaction to expected inflation after a monetary policy shock (annualised 1% policy rate hike) – all endogenous variables, values in percentages.

Bilbiie and Straub (2013), which provide a theoretical framework to analyse how financial market participation impacts the policy requirements for determinacy.

On the total liquidity side, two forces are at work: the inflation effect and arbitrage towards the more remunerative asset. On impact, inflation decreases: this path influences real liquidity as future inflation will be higher than today, increasing current liquidity until inflation overshoots its steady-state level and turns slightly positive. This happens over a relatively short period. When inflation turns back positive (although extremely close to zero), real liquidity peaks and decreases smoothly.

In this process, the agent adjusts its portfolio of assets profiting from the increased return on the liquid asset, b . In this respect, the IRF for m mirrors that of b , with a reallocation away from money holdings to liquid bonds. The asset flows following a positive productivity shock are thus flipped when the monetary authority raises the policy rate.

Under the passive monetary policy regime the economy experiences different magnitudes in the aggregates, but qualitatively similar profiles. When the central bank is accommodative, output, inflation, and the key policy rate double their impact change, and the latter becomes less persistent, overall. This last outcome falls in line with Primiceri (2005), relating to a passive pre-Volcker Fed, and sheds light on how monetary policy stances can affect aggregate dynamics in

general, and inflation dynamics in particular. Taking this implication to the data would imply that prior to the Great Moderation period inflation was reporting higher degrees of inertia, whilst it would decrease afterwards. chapter 5 reports indeed decreasing levels of inertia, although the timing does not univocally point to policy switches.

We analyse the effects of different monetary policy regimes on inflation dynamics in Section (4.3).

3.3 Dissecting simulated inflation dynamics

As discussed, the monetary policy stance could influence the dynamics of some endogenous aggregates. We are particularly interested in inflation, as typically a passive central bank enables sunspot, degenerate paths for prices. While this inquiry relies on stylised economies, chapter 5 delves into an empirical analysis of inflation persistence in the US.

The motivating question we ask is, thus, whether a passive central bank produces unstable inflation dynamics. In this framework, it is straightforward to generate abundant time series and hence conduct some ex-post econometric exploration. To offer a more comprehensive comparison, we simulate and analyse two other DSGE models, namely Ascari and Sbordone (2014) and Smets and Wouters (2007). The former study how trend inflation affects aggregate dynamics and policy in a generalised New Keynesian model, while the latter builds a rich environment with nominal frictions and indexation for wages and prices, investment adjustment cost, and a large number of shocks and outperform VARs in short term forecasting.

We generate for each model 500 thousand observations, or 125 thousand years of simulated history: this should assure convergence of the estimators and tight confidence intervals. By the very structure of the models, the Data Generating Processes of these series is a linear system shocked by AR(1) normal innovations: one should not be surprised that the data generated are also Gaussian. Comparability across models is meaningful because of the close parametrisation: as in the previous Sections, all common blocks are calibrated to the same values. All remaining parameters are either those chosen in the original paper or those estimated. Crucially, every model is also perturbed with the very same sequence of monetary and technological shocks, so to ideally elicit differences in the transmission and propagation mechanisms.

For our interest is chiefly on inflation dynamics, we limit our interest to the global autoregressive properties of the inflation series, which is generated by a single, stable, and well behaved DGP for each case. The statistical framework of reference is an autoregressive one, with varying orders. As we calibrate the persistence of all shocks to $\rho = .65$, we expect to find values in this neighbourhood, the more so since all shocks are identical and other exogenous disturbances are muted. Any difference between the simulated series is due to the propagation mechanism and, most importantly, the monetary response function. Moreover, when comparing our model in its

two policy regimes, we will be able to pick up the different inflation dynamics enticed by the passive monetary policy stance.

Our analysis starts with an $AR(5)$: our model complying to the Taylor Principle, the same violating it, and the workhorse NKDSGEs. Secondly, we set an upper bound on the lags to 120, and pick the optimal lag number minimising the Bayesian Information Criterion. Table (4.2) presents the results for an $AR(5)$.

Table 3.2 – Estimates on simulated data

	Simulated Inflation				
	(1)	(2)	(3)	(4)	(5)
Const.	-.0001 (.0003)	-.0002 (.0005)	-.0001 (.0002)	-.0002 (.0003)	-.0004 (.0005)
1 st lag	.648*** (.001)	.597*** (.001)	.640*** (.001)	.652*** (.001)	.966*** (.001)
2 nd lag	.001 (.002)	-.017*** (.002)	-.005*** (.002)	-.001 (.002)	-.004** (.002)
3 rd lag	.00003 (.002)	-.014*** (.002)	-.003** (.002)	.0004 (.002)	-.004** (.002)
4 th lag	-.001 (.002)	-.013*** (.002)	-.003** (.002)	.0001 (.002)	-.005** (.002)
5 th lag	.002 (.001)	-.029*** (.001)	-.009*** (.001)	.001 (.001)	-.025*** (.001)
Adjusted R ²	.421	.338	.402	.424	.875

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

$AR(5)$ estimates on simulated data from five models: (1) Gali (2015), (2) liquidity model $\gamma = .5$, (3) liquidity model $\gamma = 1.8$, (4) Ascari and Sbordone (2014), (5) Smets and Wouters (2007). Only technological and monetary policy shocks are allowed, each model is simulated for 500000 periods, after discarding the first 100000 iterations. All shocks are set to have zero mean, equal variance, and are iid. Second and third columns present estimates for our model with liquidity, complying to the Taylor Principle and violating it, respectively.

Looking at the coefficients on the first lag, we see our expectations confirmed, as all coefficients are tightly close to the calibrated parameter, no constant is statistically different from zero, and significance decreases after the first lag for classic NKDSGE models. A notable exception is the Smets and Wouters (2007) model, which reports five significant lags with relatively large coefficients. The most remarkable feature of such model, though, is that it displays a quasi unit-root in inflation, as the first lag is statistically close, but different, than one.

Interestingly, as we depart from the NKDSGE models to analyse an accommodating central bank, the coefficient on the first lag moves away from the calibrated value, downwards (column (2) in Table 4.2). This result corroborates the consensus that a passive monetary authority has less command over the inflation path and fails at taming its dynamics back on target. The flip side of this latter aspect is that, when the first lag becomes less relevant, previous ones acquire more weight. Overall, hence, inflation seems to become more persistent when central banks do not follow an aggressive Taylor rule.

On the other hand, the magnitude of the significant coefficients of the two calibrations of the liquidity model is comparatively small. While for an active central bank (model (3)) the autoregressive coefficients quickly approach zero, for a passive one these still become smaller but remain roughly ten times bigger than those of the other models, and still significant.

These results point to a role for monetary policy stance in substantially influencing the inflation dynamics, along the same lines traced by Cogley, Primiceri, and Sargent (2008), Cogley and Sargent (2002, 2005), and Primiceri (2005).

To carry out a deeper analysis into inflation dynamics, we analyse a crude measure of persistence, that is to compare the number of optimal lags for an $AR(k)$ process. This procedure finds that the optimal number of lags for our model with liquidity and Taylor principle compliance (model (3) in Tab.(4.3)) is around 70, whilst for the version parametrised in accordance to the Taylor Principle (model (2), *ibid*) it is around 50, for the standard NKDSGEs it is merely 2. Table (4.3) offers more details on this result.

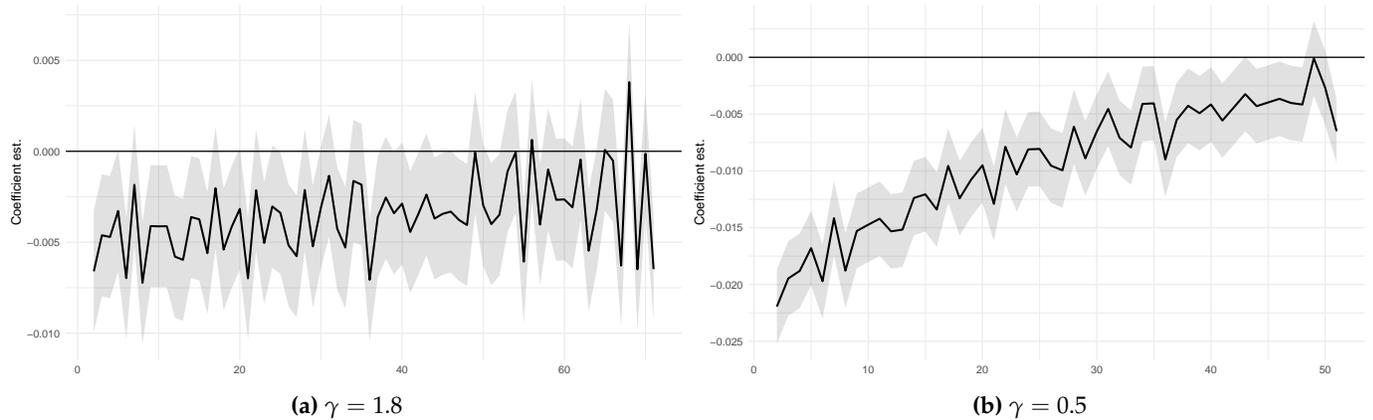
Table 3.3 – Optimal lags

Models	Opt. lags	Sign. lags	adj. R ²	BIC
(1) Gali (2015)	2	50%	.421	−280676.9
(2) Liq. $\gamma = .5$	51	80%	.348	−413028.9
(3) Liq. $\gamma = 1.8$	71	32%	.401	−1033066
(4) Ascari and Sbordone (2014)	2	50%	.421	−422629.6
(5) Smets and Wouters (2007)	13	38%	.719	−486558.7

Optimal lags for AR process of inflation for (1) baseline NKDSGE (Gali, 2015), (2) liquidity model violating the Taylor Principle ($\gamma = 0.5$), (3) liquidity model complying to it ($\gamma = 1.8$), a model with time-varying trend inflation (Ascari and Sbordone, 2014), and (5) a medium scale workhorse DSGE (Smets and Wouters, 2007). Optimal lags are those minimising the BIC. All models are fed the same sequence of shocks of the same variance, generating 500000 quarterly observations.

An optimal lag number does not imply that *all* regressor lags are significant. It only implies that all significant lags are part of the regression, likely a small subset of lags explains the largest share of variation. With this in mind, our model of liquidity with an accommodative monetary

Figure 3.5 – $AR(k^*)$ Lags



Autoregressive estimates on optimally selected lags. Left panel 4.6a liquidity model with aggressive central bank ($\gamma = 1.8$). Right panel 4.6b liquidity model with accommodative policy ($\gamma = 0.5$). Estimated coefficients in solid, bands are twice the estimated standard errors. Note: first lag excluded from the plot for scale readability.

authority shows that today’s inflation depends on a long sequence of lags. The number of optimal lags increases when we let the central bank respond aggressively to expected inflation.

On the other hand, the baseline NKDSGE model produces a process with extremely short memory, likewise for the model of trend inflation. In line with the result from Tab.(4.2), the medium-scale Smets and Wouters (2007) model displays longer endogenous lags, due to numerous frictions and feedback engrained in this rich model.

Panel 4.6a plots the point estimates and confidence bands⁹ for the 71 optimal lags of the liquidity model with an active policy rule. Of all lags included, only 21 are strongly significant ($\pm 30\%$); interestingly, these are largely negative in sign and relatively small in magnitude. This last evidence might suggest that in this policy regime inflation depends negligibly on past realisations, being constantly nudged into a well-determined path.

Panel 4.6b plots the same information for an accommodative policy rule. What is striking is the length of lags deemed relevant and the share of significant ones ($\pm 80\%$), as opposed to the aggressive policy rule. As remarked in Table 4.2, coefficients are greater and all significant at 1% up to the 33rd. These features point toward a higher persistence in inflation, something compatible with a central bank that, for instance, targets monetary aggregates (prior to Volcker’s chairmanship) or finds itself short of conventional monetary tools (QE at the ZLB), as it has been for long periods recently. Nonetheless, this setting does not imply necessarily sunspot equilibria or spiralling aggregates, the system eventually converges back to the steady-state even when the monetary policy stance is accommodative.

⁹For the sake of readability, we do not plot the first lag. All-inclusive plots are in Appendix 4.F.

This result offers a framework to analyse inflation dynamics in light of diverse monetary policy regimes. Higher levels of persistence – as long as it is measured as the number of significant lags – are expected in periods like the US high inflation of the '70s, when indeed prices skyrocketed out of control. Conversely, tranquil periods results from aggressive monetary stances, like the Great Moderation. The inclusion of liquidity is extremely helpful to systematise the 2008 crisis, while also accounting for the inactive interest rate policy, as shown in chapter 3. For this latter case, though, other factors need to be considered, like the level of liquidity and the wide range of unconventional policies put in place, from which we abstracted in this chapter.

3.4 The 2008 crisis: severe liquidity shortage

This model lends itself to an interesting experiment: although nominal liquidity is assumed to be constant at \bar{Z} and real liquidity z moves with the inflation rate, we hit z with a negative shock and study the behaviour of our model, as in eq.(4.18). Although not orthodox, this is a practical shortcut: neglecting where that missing liquidity goes *physically* – and in which proportion money and bonds are affected – lets us focus on the dynamics of convergence to the steady-state. This negative shock affects solely the current level of real liquidity \hat{z}_t , leaving inflation untouched. The adjustment then has to run primarily through the remaining contemporaneous relations (4.12) and (4.14). In eq. (4.14), for given inflation expectations $E_t\pi_{t+1}$, the policy interest rate does not yet move, so that the sudden drop in real liquidity has to be offset by money holdings \hat{m}_t . This adjustment then passes on to the Euler equation eq. (4.12) through the wedge: current output gap moves accordingly and then transmits to current inflation via the Phillips Curve (4.13).

One could think of this experiment as a sudden drop in the combined liquidity of bonds and money, closely related to the liquidity dry up triggering the 2008 Global Financial Crisis. In the aftermath of such recession, major central banks rapidly hit the zero lower bound on policy interest rates, thus in effect adopting an accommodative monetary policy stance, equivalent to that admitted by the liquidity model. In this respect, analysing how aggregates react to a sudden liquidity dry up sheds light on how different interest rate rules interact and affect the overall dynamics.

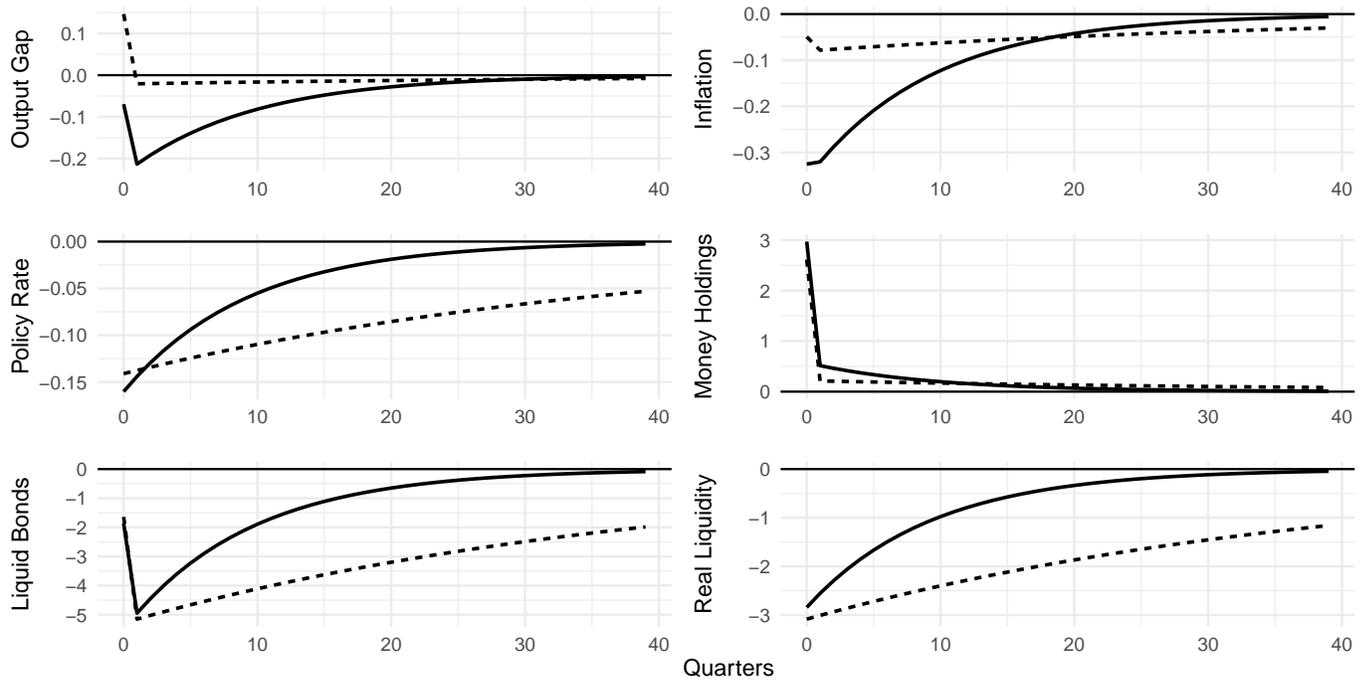
We produce IRFs and discuss their economic interpretations under the two regimes of monetary policy.

$$\hat{z}_t - \epsilon_t^z = \hat{z}_{t-1} - \pi_t \quad (3.18)$$

Following an abrupt and violent liquidity dry-up the model shows a general reaction broadly independent from the behaviour of the monetary authority, but substantial differences emerge in magnitude and reversion to steady-state levels.

The shock impacts at first money demand \hat{m} , for a given policy rate s , which spikes up. Conversely, the representative agent disinvests from liquid bonds, proportionally more than the miss-

Figure 3.6 – Liquidity Shock



Real liquidity shock: aggressive ($\gamma = 1.8$, dashed) vs accommodative ($\gamma = .5$, solid) policy rules, values in percentages.

ing liquidity. This results from the preference for money with respect to bonds that were assumed in the calibration.

At this stage, the Phillips curve (4.13) and the IS equation (4.12) propagate the shock to the rest of the economy. The money wedge in the Euler equation (4.12) transmits the shock to current output that spikes as well on impact. Under a passive monetary policy, this translates into a limited effect of the shock on impact, and a degradation in the following quarters. The Phillips curve then squares the expected inflation with current π and a widened output gap. Inflation expectations subsequently drive the monetary policy decisions which translate into two distinct paths for realised inflation.

The money term in the IS curve is the telling point between the two regimes, together with the path for liquid bonds, b .

Under both regimes, money converges back to the steady-state relatively quickly, showing that real liquidity is deeply intertwined with liquid bonds. Most notably, under an accommodative central bank real liquidity recovers rapidly, thanks to a deflation that accelerates the recovery of z but impedes a quick rebound in output. Interestingly, when the money authority conducts an active policy, the inflation path – contained deflation with slow recovery – turns into persistently low rates, well beyond the case of a passive central bank.

The sharp difference in the reaction of the two regimes lies in the severity of the impact and the duration of the recovery. Inflation and output, in particular, show starkly different behaviours: when the central bank has a passive stance, a liquidity dry-up triggers a deep recession with a painfully long recovery (more than thirty quarters); the same applies to inflation, too. An active monetary stance against a liquidity shock tames the damage and facilitates the recovery, somewhat containing the effects within the financial sector of the economy.

To sum up, and combining these results with previous information, this exercise suggests that central banks complying with the Taylor Principle have a firmer control on contagion when a liquidity crisis hits. According to our stylised model, in fact, active monetary policy helps to contain and limit the damage to the sole financial sector of the economy, with reduced impact and consequence on real activity. The stark difference in the set of IRFs, clearly, lies in the fall of the output gap: it widens under passive monetary policy while it marks a mild recession under active policy, although the duration is similar. As we already remarked, recovery speeds are substantially different.

Stretching our model to draw policy implications, one can think back to the Zero Interest Rate Policies (ZIRP) that major central banks deployed both in the 2008 GFC and in the current Covid19 crisis. Although hazardous, such an aggressive reaction could be strong enough to facilitate a speedy recovery, as it has been the case for the US after the GFC.

3.5 Conclusion

Since the Global Financial Crisis, liquidity has gained a central role in the general macroeconomic discussion, and in the debate on monetary policy in particular. It has been the main concern for major central banks in engaging in unconventional policies.

In this chapter, we present a parsimonious framework with minor departures from the core new Keynesian model of monetary policy and derive relevant results for central banks' mandate of price stabilisation.

First and foremost, we show that the simple addition of a liquid asset – and the consequent modification of the intertemporal Euler equation – pins down a solution with an accommodative, stripped-down policy rate rule. The latter needs only to *positively* correlate the policy rate and expected inflation to rule out degenerate, multiple equilibria. These latter sunspot equilibria would arise in the baseline New Keynesian Dynamic Stochastic General Equilibrium model when the central bank does not react to inflation aggressively.

We compare the responses of our liquidity model with those of the baseline NKDSGE and confirm that they broadly match for technological and monetary policy shocks, under identical, common calibration. We then study how our model responds to such shocks when the central

bank under-reacts to expected inflation. We find no evidence of degenerate behaviour for the model aggregates, contrary to the predictions of the baseline NKDSGE.

In fact, all aggregates, and inflation especially, broadly react in the same way under the two regimes: we signal, though, a change in the dynamics, rather than direction. We find that an accommodative central bank generates more persistent inflation. We test such hypothesis on simulated data, including two other workhorse models from the monetary field.

We find that the inclusion of liquidity and liquid bonds generates overall more persistence in inflation. Within the sets of calibration parameters, inflation displays more persistence when the monetary policy stance is passive, in line with evidence from the US high inflation period.

Going forward, more sophisticated versions of this model may be easily developed, as it accommodates additional layers of complexity. For example, one might want to include sticky wages, capital stock, financial blocks in the spirit of the financial accelerator, or occasionally binding constraints like a properly modelled Zero Lower Bound. These theoretical devices have been developed as modules for the basic NKDSGE model, with which our model shares the core features. The first, natural extension for this model would be the relaxation of the fixed total liquidity in nominal terms, \bar{Z} . This would provide the monetary authority an additional tool to carry out its mandate and, most importantly, a framework to study liquidity management.

3.A Model appendix

3.A.1 Obtaining the system of equations

We combine the equations presented in the body of the paper so to obtain the system of equations that will be later loglinearised and fed to Dynare for simulations.

The first target is the augmented Euler equation. The version proposed in the paper includes real money balances on top of the usual terms. Take the first intertemporal FOC from the consumer utility maximisation program as the starting point:

$$u'(c_t) = E_t [\beta (1 + r_t) u'(c_{t+1})]$$

recall the Fisher equation $(1 + r_t) (1 + \pi_{t+1}) = (1 + i_t)$ and replace r_t

$$\begin{aligned} u'(c_t) &= E_t \left[\beta u'(c_{t+1}) \frac{1 + i_t}{1 + \pi_{t+1}} \right] \\ &= E_t \left[\frac{\beta u'(c_{t+1})}{1 + \pi_{t+1}} + \frac{\beta u'(c_{t+1}) i_t}{1 + \pi_{t+1}} \right] \end{aligned}$$

Now recall the condition on marginal utility of real money balances, $v'(m_t)$, and employ it to define the nominal interest rate, i_t :

$$\begin{aligned} v'(m_t) &= E_t \left[\beta \lambda_{t+1} \frac{i_t}{1 + \pi_{t+1}} \right] \\ i_t &= v'(m_t) E_t \left[\frac{1 + \pi_{t+1}}{\beta \lambda_{t+1}} \right] \\ \text{moreover } \lambda_{t+1} &= u'(c_{t+1}) \end{aligned}$$

Turning back to the Euler equation, plug the nominal interest rate in the relation just recovered:

$$u'(c_t) = E_t \left[\frac{\beta u'(c_{t+1})}{1 + \pi_{t+1}} + \frac{\beta u'(c_{t+1})}{\cancel{1 + \pi_{t+1}}} v'(m_t) \frac{1 + \pi_{t+1}}{\beta u'(c_{t+1})} \right]$$

which rearranges in

$$= E_t \left[\beta u'(c_{t+1}) \left(\frac{1}{1 + \pi_{t+1}} + \frac{v'(m_t)}{\beta u'(c_{t+1})} \right) \right].$$

This last expression is then loglinearised to obtain Equation 4.12.

To condense the money equation, start with the last two relations in (4.3):

$$v'(m_t) = E_t \left[\beta \lambda_{t+1} \frac{i_t}{1 + \pi_{t+1}} \right]$$

$$i_t = v'(m_t) E_t \left[\frac{1 + \pi_{t+1}}{\beta \lambda_{t+1}} \right]$$

plug this result into $h'(b_t)$

$$h'(b_t) = E_t \left[\beta \lambda_{t+1} \frac{i_t - s_t}{1 + \pi_{t+1}} \right]$$

$$\Rightarrow h'(b_t) = E_t \left[\frac{\beta \lambda_{t+1}}{1 + \pi_{t+1}} v'(m_t) \frac{1 + \pi_{t+1}}{\beta \lambda_{t+1}} - \frac{\beta \lambda_{t+1} s_t}{1 + \pi_{t+1}} \right]$$

exploit the fact that $\lambda_{t+1} = u'(c_{t+1})$ and $b_t = z_t - m_t$ to obtain

$$h'(z_t - m_t) - v'(m_t) = E_t \left[-\frac{\beta s_t u'(c_{t+1})}{1 + \pi_{t+1}} \right]$$

Setting this equation to its steady-state and using a first Taylor approximation generates equation (4.14) in the text.

The backward dependence of real liquidity (4.10) results as follows:

$$Z_t = \bar{Z} = M_t + B_t$$

$$\frac{\bar{Z}}{P_{t-1}} = \frac{(m_t + b_t) P_t}{P_{t-1}}$$

$$\Rightarrow z_t = \frac{z_{t-1}}{1 + \pi_t}$$

Concerning the Phillips curve, it remains unchanged from traditional New Keynesians models and derives from the use of equations (4.7). The output gap it includes results from the comparison to the flexible prices version of the model. Other relations do not need further manipulation.

3.B Loglinearisation

The model is loglinearised around a zero-inflation steady state as in the early New Keynesian models. This assumption yields the remaining variable values in the long run and without shocks. We employ directly the functional forms from (4.11).

Steady state values

Euler equation / IS curve at the steady state:

$$\begin{aligned}u'(c_t) &= E_t [\beta (1 + r_t) u'(c_{t+1})] \\c^{-\sigma} &= m^{\psi-1} + \beta \frac{c^{-\sigma}}{1 + \pi} \\m^{1-\psi} &= \frac{c^\sigma}{1 - \beta}\end{aligned}$$

Output level: $y = y^f$.

Money demand and policy rate:

$$\begin{aligned}s &= \gamma\pi = 0 \\h'(z - m) - v'(m) &= \left[-\frac{\beta s u'(c)}{1 + \pi} \right] \\(z - m)^{\phi-1} - m^{\psi-1} &= -\frac{\beta \overbrace{s}^{=0} c^{-\sigma}}{1 + \underbrace{\pi}_{=0}} = 0 \\(z - m)^{\phi-1} &= m^{\psi-1} \\z - m &= m^{\frac{\psi-1}{\phi-1}} \\z &= m + m^{\frac{1-\psi}{1-\phi}} = m \left(1 + m^{\frac{\phi-\psi}{1-\phi}} \right)\end{aligned}$$

Remarkably, if $\psi = \phi$, so that the agent is indifferent between money and liquid bonds, $z = 2m$ as in the log preferences case.

Linearised system

Linearised liquidity law of motion:

$$\begin{aligned}z_t &= \frac{z_{t-1}}{1 + \pi_t} \\ \Rightarrow \hat{z}_t &= \hat{z}_{t-1} - \pi_t\end{aligned}$$

This equation does not pin down the steady-state value for z , since it results from the sum of real money balances, m , and liquid bonds, b .

$$\hat{b}_t = \frac{z}{z - m} \hat{z}_t - \frac{m}{z - m} \hat{m}_t$$

Linearised Euler equation:

$$c_t^{-\sigma} = E_t \left[\beta c_{t+1}^{-\sigma} \left(\frac{1}{1 + \pi_{t+1}} + \frac{m_t^{\psi-1}}{\beta c_{t+1}^{-\sigma}} \right) \right]$$

taking logs and first derivatives - drop E_t for convenience - and rearrange:

$$\begin{aligned} -\sigma \hat{c}_t &= (\psi - 1) \hat{m}_t + (-\sigma \hat{c}_{t+1} - \pi_{t+1}) \\ \hat{c}_t &= \frac{1 - \psi}{\sigma} \hat{m}_t + \hat{c}_{t+1} + \frac{1}{\sigma} \pi_{t+1} \end{aligned}$$

Linearised money demand:

$$\begin{aligned} h'(z_t - m_t) - v'(m_t) &= E_t \left[-\frac{\beta s_t u'(c_{t+1})}{1 + \pi_{t+1}} \right] \\ (z_t - m_t)^{\phi-1} - m_t^{\psi-1} &= E_t \left[-\frac{\beta s_t c_{t+1}^{-\sigma}}{1 + \pi_{t+1}} \right] \\ (1 + \pi_{t+1}) (z_t - m_t)^{\phi-1} &= (1 + \pi_{t+1}) m_t^{\psi-1} - \beta c_{t+1}^{-\sigma} s_t \end{aligned}$$

Focus first on the left-hand side of the above equation and drop the steady state terms – which will cancel out eventually:

$$\ln(1 + \pi_{t+1}) + (\psi - 1) \ln(z_t - m_t) \Rightarrow \pi_{t+1} + (\psi - 1) \left[\frac{z}{z - m} \hat{z}_t - \frac{m}{z - m} \hat{m}_t \right]$$

Now replace the steady state value for z found in the money demand above and plug in the previous equation, factoring out the common terms, we obtain

$$\begin{aligned} \pi_{t+1} + (\psi - 1) \left[m^{\frac{1-\psi}{1-\phi}} (z \hat{z}_t - m \hat{m}_t) \right] \\ \pi_{t+1} + (\psi - 1) \left[\hat{z}_t + (\hat{z}_t - \hat{m}_t) m^{\frac{\psi-\phi}{1-\phi}} \right] \end{aligned}$$

Turning to the right-hand side of the previous equation, we know that at the steady-state and in logs it equals $\ln(m^{\psi-1})$, as the term involving s collapses to 0. Furthermore, we can break down the linearisation into two chunks, the first yielding simply

$$\pi_{t+1} + \frac{\psi - 1}{m^2} \hat{m}_t$$

and the second one not more difficult. Namely, linearising with respect to c yields 0, as it contains $s = \gamma\pi = 0$ at the steady-state; approximation for s gives

$$-\beta \frac{c^{-\sigma}}{m^{\psi-1}} s_t$$

Gathering all pieces together gives, after slight rearrangements:

$$\left(\frac{1-\psi}{m^2}\right) \hat{m}_t + \beta c^{-\sigma} m^{1-\psi} s_t - (1-\phi) \left[\hat{z}_t + (\hat{z}_t - \hat{m}_t) m^{\frac{\psi-\phi}{1-\phi}} \right] = 0$$

From which we can recover eq.(4.14) in the text.

3.C CIA setup

The model presented in the main body follows the reduced-form strategy that was introduced by Sidrauski (1969), which is usually referred to as Money in the Utility function, MIU. It is a straightforward way to have agents hold cash balances or other assets, but it lacks a proper microfoundation. Cash in Advance (CIA) models partly overcome this lack of microfoundation: instead of providing direct utility, money holdings are used to purchase the desired amount of consumption. The CIA approach was proposed by Clower (1967) and further developed by Grandmont and Younes (1972) and Lucas (1980).

Compared to MIU models, therefore, CIA models add one more inequality constraint: agents allocate money to purchase the desired level of consumption, either within the same period or with a lag. This timing difference is key. Lucas (1982) develops a deterministic CIA model where assets are allocated at the beginning of the period, in line with the consumption decision. Svensson (1985), conversely, constrains agents to allocate assets in advance, before shocks and adjustments take place. This latter approach yields more informative dynamics, since agents might over (under) accumulate money balances with respect to their future desired level of consumption. Finally, Cooley and Hansen (1989, 1991) build on Lucas and Stokey (1987) and propose a RBC model with CIA and uncertainty. The setup proposed here is akin to this latter, as we embed in the NKDSGE structure a Liquidity in Advance (LIA) constraint. In this exercise, we draw from Calvo and Végh (1990a,b), who study a similar economy under a different setup.

The LIA constraint hinges on a liquidity production function $Q(M, B)$ that combines enough bonds and cash to purchase the desired level of consumption. Liquidity in-house production function Q might take several forms (Calvo and Végh (1990a,b) lay out some restrictions on these forms): to grant flexibility and generality, assume Q is increasing and concave in both assets. For the sake of generality, timing will be detailed momentarily. Under these assumptions, the LIA constraint takes the following form:

$$C_t \leq Q_t(M_{t-j}, B_{t-j}) \tag{3.19}$$

With this additional constraint, the representative agent now solves the following program:

$$\begin{aligned} \max_{c_t, N_t} E_t \sum \beta^t (u(c_t) - g(N_t)) \\ \text{st: } C_t + B_t + M_t + X_t = W_t N_t + (1 + i_{t-1}) X_{t-1} + (1 + s_{t-1}) B_{t-1} + M_{t-1} \\ C_t \leq Q(M_{t-j}, B_{t-j}) \end{aligned} \quad (3.20)$$

Notice that rewriting the budget constraint in terms of total (real) assets highlights the nominal (real) opportunity costs incurred when holding positive quantities of money and bonds. Let $D_t = X_t + B_t + M_t$, where D is total assets. One can replace X in the flow constraint and turn to real quantities:

$$\begin{aligned} C_t + B_t + M_t + X_t &= W_t N_t + (1 + i_{t-1}) X_{t-1} + (1 + s_{t-1}) B_{t-1} + M_{t-1} \\ C_t + D_t &= W_t N_t + (1 + i_{t-1}) D_{t-1} + (i_{t-1} - s_{t-1}) B_{t-1} - i_{t-1} M_{t-1} \\ c_t + d_t &= w_t N_t + (1 + r_{t-1}) d_{t-1} - \frac{i_{t-1} - s_{t-1}}{1 + \pi_t} b_{t-1} - \frac{i_{t-1}}{1 + \pi_t} m_{t-1} \end{aligned} \quad (3.21)$$

Where $(1 + i_t) / (1 + E_t \pi_{t+1}) = (1 + r_t)$ follows from Fischer equation and is the real interest rate. As of the liquidity in advance constraint in real terms, timing is relevant:

$$c_t = \begin{cases} Q(m_t, b_t) & \text{for } j = 0 \\ \frac{Q(m_{t-1}, b_{t-1})}{1 + \pi_t} & \text{for } j = 1 \end{cases} \quad (3.22)$$

When allocating one additional unit of income to liquid bonds, the consumer will receive interest s next period, but gives up i on the purely illiquid bond. Therefore the opportunity cost of holding such bonds is $i - s$. Similarly, money yields no interest so that the agent gives up entirely to i . The elasticity of substitution q , together with the liquidity-in-advance constraint will force agents to hold non-negative amounts of m and b .

Lagrangian reads

$$\mathcal{L} = \sum \beta^t \left[\begin{aligned} &[u(c_t) - g(N_t)] + \\ &\lambda_t \left(w_t N_t + (1 + r_{t-1}) d_{t-1} - \frac{i_{t-1} - s_{t-1}}{1 + \pi_t} b_{t-1} - \frac{i_{t-1}}{1 + \pi_t} m_{t-1} - c_t - d_t \right) + \\ &\mu_t \left(\frac{Q(m_{t-1}, b_{t-1})}{1 + \pi_t} - c_t \right) \end{aligned} \right] \quad (3.23)$$

FOCs for consumption, labour supply, and intertemporal allocation are independent of the timing of the liquidity constraint and fairly standard:

$$\begin{aligned}
u'_t - \lambda_t - \mu_t &= 0 \\
-g'_t + \lambda_t W_t &= 0 \\
-\lambda_t + \beta \lambda_{t+1} (1 + r_t) &= 0
\end{aligned} \tag{3.24}$$

FOCs for money holdings and bonds, instead, slightly differ depending on the timing of assets allocation. The $j = 1$ is the most interesting case: with uncertain productivity and monetary policy shock, agents will rationally choose their portfolio composition, level of consumption, labour supply. For instance, for a productivity shock, they wish to increase consumption but face a binding liquidity in advance constraint and therefore do not adjust as they wish.

For $j = 1$, FOCs for cash and bonds read:

$$\frac{\mu_{t+1} Q_t^M}{1 + \pi_{t+1}} = \left(\frac{i_t}{1 + \pi_{t+1}} \right) \lambda_{t+1} \quad \frac{\mu_{t+1} Q_t^B}{1 + \pi_{t+1}} = \left(\frac{i_t - s_t}{1 + \pi_{t+1}} \right) \lambda_{t+1} \tag{3.25}$$

Where λ and μ are the multipliers associated with the budget and liquidity in advance constraint, respectively. Taking the ratio of the two equations provides the relative allocation between cash and bonds. The functional form of Q is now required, and we assume a Cobb-Douglas one with elasticity α :

$$\begin{aligned}
\frac{Q_t^B}{Q_t^M} &= \frac{i_t - s_t}{i_t} \\
\frac{m_t}{b_t} &= \frac{\alpha}{1 - \alpha} \frac{i_t - s_t}{i_t} \quad \forall t
\end{aligned} \tag{3.26}$$

The last equation allows for recovering demand for bonds b (money m) as a function of interest rates and money holdings m (bonds b) for each time t . Indeed, since the one-period-ahead LIA constraint in real terms reads $c_t = (1 + \pi_t)^{-1} m_{t-1}^\alpha b_{t-1}^{1-\alpha}$ it obtains that, for b and m given from $t - 1$:

$$\begin{aligned}
b_{t-1} &= \left(\frac{1 - \alpha}{\alpha} \right) \left(\frac{i_{t-1}}{i_{t-1} - s_{t-1}} \right) m_{t-1} & m_{t-1} &= \left(\frac{\alpha}{1 - \alpha} \right) \left(\frac{i_{t-1} - s_{t-1}}{i_{t-1}} \right) b_{t-1} \\
c_t(m_{t-1}) &= m_{t-1}^\alpha \left[\left(\frac{1 - \alpha}{\alpha} \right) \left(\frac{i_{t-1}}{i_{t-1} - s_{t-1}} \right) m_{t-1} \right]^{1-\alpha} & c_t(b_{t-1}) &= b_{t-1}^{1-\alpha} \left[\left(\frac{\alpha}{1 - \alpha} \right) \left(\frac{i_{t-1} - s_{t-1}}{i_{t-1}} \right) b_{t-1} \right]^\alpha \\
&= \left[\frac{1 - \alpha}{\alpha} \frac{i_{t-1}}{i_{t-1} - s_{t-1}} \right]^{1-\alpha} m_{t-1} & &= \left[\frac{\alpha}{1 - \alpha} \frac{i_{t-1} - s_{t-1}}{i_{t-1}} \right]^\alpha b_{t-1} \\
m_{t-1}(c_t) &= \left[\frac{\alpha}{1 - \alpha} \frac{i_{t-1} - s_{t-1}}{i_{t-1}} \right]^{1-\alpha} c_t & b_{t-1}(c_t) &= \left[\frac{\alpha}{1 - \alpha} \frac{i_{t-1} - s_{t-1}}{i_{t-1}} \right]^{-\alpha} c_t
\end{aligned} \tag{3.27}$$

Whence it follows that:

$$b_{t-1} = \left(\frac{\alpha}{1-\alpha} \frac{i_{t-1} - s_{t-1}}{i_{t-1}} \right)^{-\alpha} c_t (1 + \pi_t); \quad m_{t-1} = \left(\frac{\alpha}{1-\alpha} \frac{i_{t-1} - s_{t-1}}{i_{t-1}} \right)^{1-\alpha} c_t (1 + \pi_t); \quad (3.28)$$

Taking $z_t = m_t + b_t, \forall t$ in combination with $z_t = z_{t-1} / (1 + \pi_t)$, it finally obtains:

$$\begin{aligned} z_{t-1} &= \left[\left(\frac{\alpha}{1-\alpha} \frac{i_{t-1} - s_{t-1}}{i_{t-1}} \right)^{1-\alpha} + \left(\frac{\alpha}{1-\alpha} \frac{i_{t-1} - s_{t-1}}{i_{t-1}} \right)^{\alpha} \right] c_t (1 + \pi_t) \\ \frac{z_t}{c_t} &= \left[\left(\frac{\alpha}{1-\alpha} \frac{i_{t-1} - s_{t-1}}{i_{t-1}} \right)^{1-\alpha} + \left(\frac{\alpha}{1-\alpha} \frac{i_{t-1} - s_{t-1}}{i_{t-1}} \right)^{\alpha} \right] \\ \frac{c_t}{z_t} &= \frac{\left(\frac{\alpha}{1-\alpha} \frac{i_{t-1} - s_{t-1}}{i_{t-1}} \right)^{\alpha}}{\left(1 + \frac{\alpha}{1-\alpha} \frac{i_{t-1} - s_{t-1}}{i_{t-1}} \right)} = h(i_{t-1}, s_{t-1}) \end{aligned} \quad (3.29)$$

Which determines the level of real consumption relative to real liquidity as a function of past interest rates i and s .

The zero-inflation steady state, which implies that $s = 0$, pins down the ratio:

$$\frac{c_{ss}}{z_{ss}} = \frac{\left(\frac{\alpha}{1-\alpha} \right)^{\alpha}}{\left(1 + \frac{\alpha}{1-\alpha} \right)} = \alpha^{\alpha} (1 - \alpha)^{1-\alpha} \quad (3.30)$$

which pins down the amount of real consumption over total real liquidity.

The LIA constraint also affects the wedge between marginal utility and λ_t , wealth shadow price. In fact, marginal consumption from the first line of Equation 4.24 equates the sum of the two Lagrangian multipliers. To factor out μ , we exploit the FOCs for the LIA constraint. Indeed, from the first order condition on cash

$$\mu_{t+1} = \frac{i_t \lambda_{t+1}}{Q_t^M(m_t, b_t)} \quad \forall t \quad (3.31)$$

Where indeed $\frac{i_t}{Q_t^M(m_t, b_t)}$ is a wedge introduced by the LIA constraint that increases the marginal cost of consumption above unity. As noted in Calvo and Végh (1990a, footnote 18), this wedge appears because one extra unit of consumption requires one extra unit of liquidity $\frac{1}{Q_t^M}$, which in cash terms costs i_t : giving up on nominal interest.

Using this in the marginal utility FOC, with appropriate timing, one obtains

$$u'_t = \left(1 + \frac{\overbrace{i_{t-1}}^{=v_{t-1}}}{Q^M(m_{t-1}, b_{t-1})} \right) \lambda_t \quad (3.32)$$

$$\Rightarrow u'_t = (1 + i_t) \beta \frac{1 + v_{t-1}}{1 + v_t} u'_{t+1}$$

Simpler liquidity production

The Cobb-Douglas functional form for liquidity production determines c/z as $h(i_{t-1}, s_{t-1})$ but does not offer an analytical expression for h^{-1} . To further simplify the analysis, let us assume that Q is a linear liquidity production function. In the most general case one has:

$$Q = \phi m + \psi b \quad (3.33)$$

With (ϕ, ψ) strictly positive. Then Equation 4.26 boils down to a simpler linear relation

$$\frac{\psi}{\phi} = \frac{i - s}{i} \Rightarrow i = \frac{\phi}{\phi - \psi} s \quad (3.34)$$

In such a setting, ensuring a positive spread between nominal interest rate i and the policy rate on liquid bonds s equates to assuming $\phi > \psi > 0$, mirroring the restriction in the body of the chapter. This assumption entails a higher, constant marginal liquidity productivity of cash with respect to liquid bonds (or bank deposits). This relation pins down the nominal interest rate as a proportional function of the policy interest rate. Transmission of policy shocks is therefore immediate.

To further simplify the setup, one could assume that liquidity is an affine combination of cash and bonds so that $\phi = \alpha$ and $\psi = 1 - \alpha$. This implies that

$$i = \alpha s \quad (3.35)$$

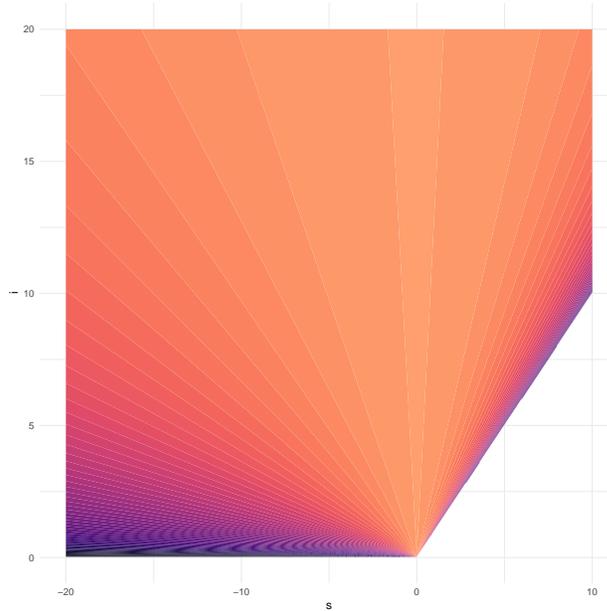
which can be easily plugged into other expressions – s is set by the forward-looking central bank. Thus

$$v_t = \frac{i_t}{Q^M} = \frac{\alpha s_t}{\alpha} = s_t \quad (3.36)$$

$$u'_t = (1 + i_t) \beta \frac{1 + s_{t-1}}{1 + s_t} u'_{t+1}$$

The resulting Euler/IS equation is therefore modified: in this setting, past and current policy rates affect the intertemporal allocation of consumption. Akin to the money wedge in the main

Figure 3.7 – h with Cobb-Douglas liquidity function



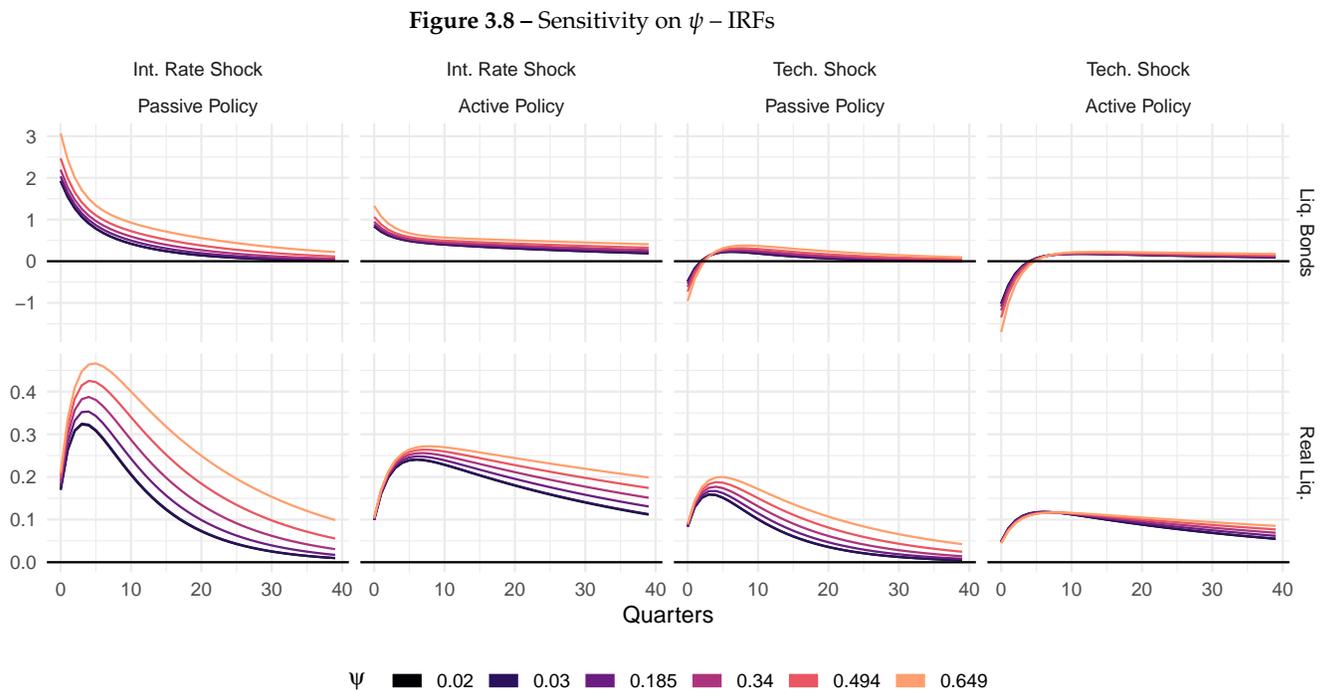
Numerical surface for $c/z = h(i, s)$ as defined in Equation 4.29, with $i > s$ and $\alpha = 1/3$. Lighter area, higher values of h .

model, the wedge hinges on liquidity, although through a different channel. In the MIU form, money holdings are indirectly affected by changes in the liquid bonds' interest rate: the higher such rate, the lower the money holdings, the lower current consumption for given values of other aggregates. In the LIA for, by contrast, the effect is direct: all else equal, higher current s (hence, higher allocation towards liquid bonds) depresses current consumption.

3.D Sensitivity analyses

3.D.1 Semi-elasticities for bonds and cash

To assess how our results depend on some particular values in the calibration of the parameters, this section briefly explores a sensitivity exercise for some values of ψ and ϕ , which are the most exotic plug-in from the main model. These parameters are also particularly hard to estimate in an empirical setting, with very scattered references. We select the more affected variables for each parameter, liquid bond holdings b , money holdings m , and real liquidity z . Other variables report negligible variations in responses to shocks in magnitude, persistence, and overall profile, but are available in the companion online repository.



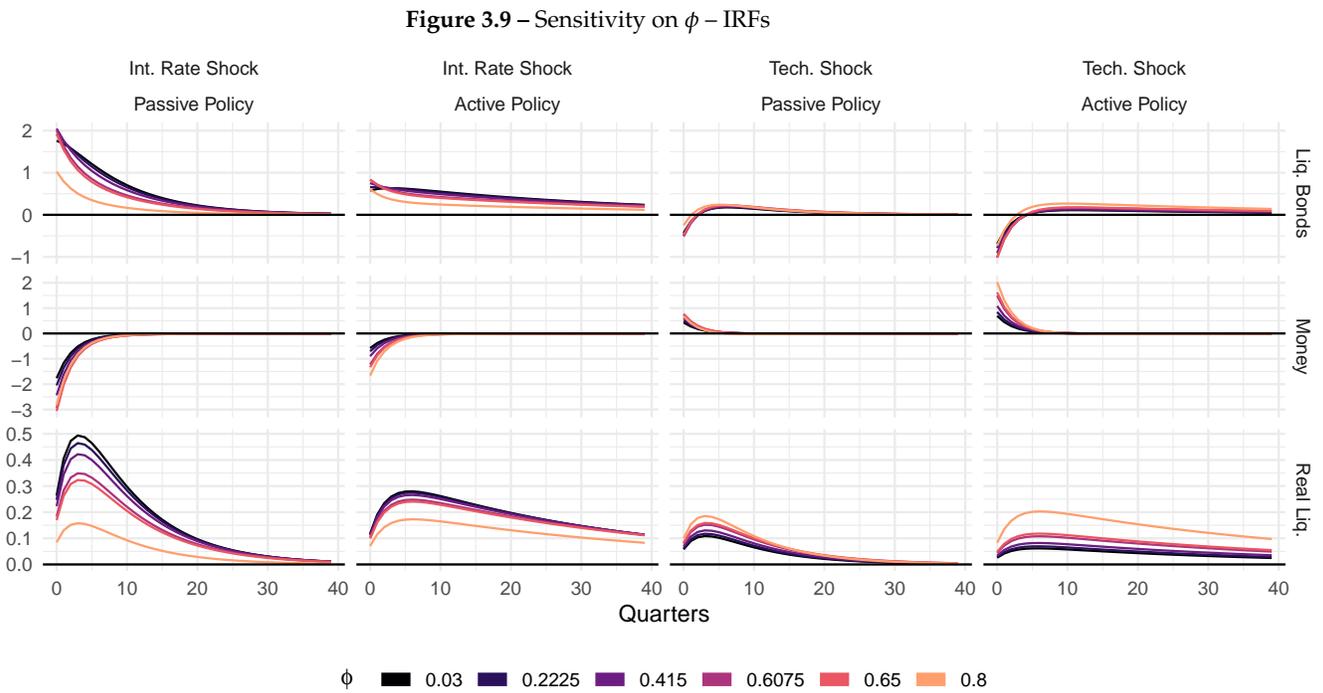
Liquidity model, IRFs for bonds b (top row) and real liquidity z (bottom row), under several values for bond holdings semi-elasticity, $0 < \psi < \phi = .65$. Monetary shock is a 1% key rate hike, annualised; technology shock is one standard-deviation. All values in percentages.

While the overall IRFs profile does not vary much over the values, higher semi-elasticity in bond holdings mainly affects the magnitude and persistence of the shock. Case in point is the monetary policy shock under a passive regime, where a higher value of ψ monotonically amplifies adjustments in real liquidity, driven by increased action in bond holdings. The same line of reasoning applies, to a reduced scale, to other combinations of shocks and regimes.

In turn, technological shocks display for both regimes a slight variation in reaction: impact variation inversely depends on ψ , but convergence to steady-state values takes longer. These changes are nevertheless limited in magnitude, so their effect is eventually negligible with respect to overall dynamics.

All in all, the especially low value used in the body of the paper is conservative and reduces considerably the persistence properties ingrained in our model of liquidity. Higher values would further increase the persistence that our model generates in comparison to other setups.

Turning to the semi-elasticity for money holdings, ϕ , Figure 4.9 plots an ensemble of IRFs for several candidate values. In such case, ψ is fixed at 0.02 as in the main analysis and ϕ takes increasing values. As ϕ affects mainly m and z , less so b , we restrict our analysis to these variables.

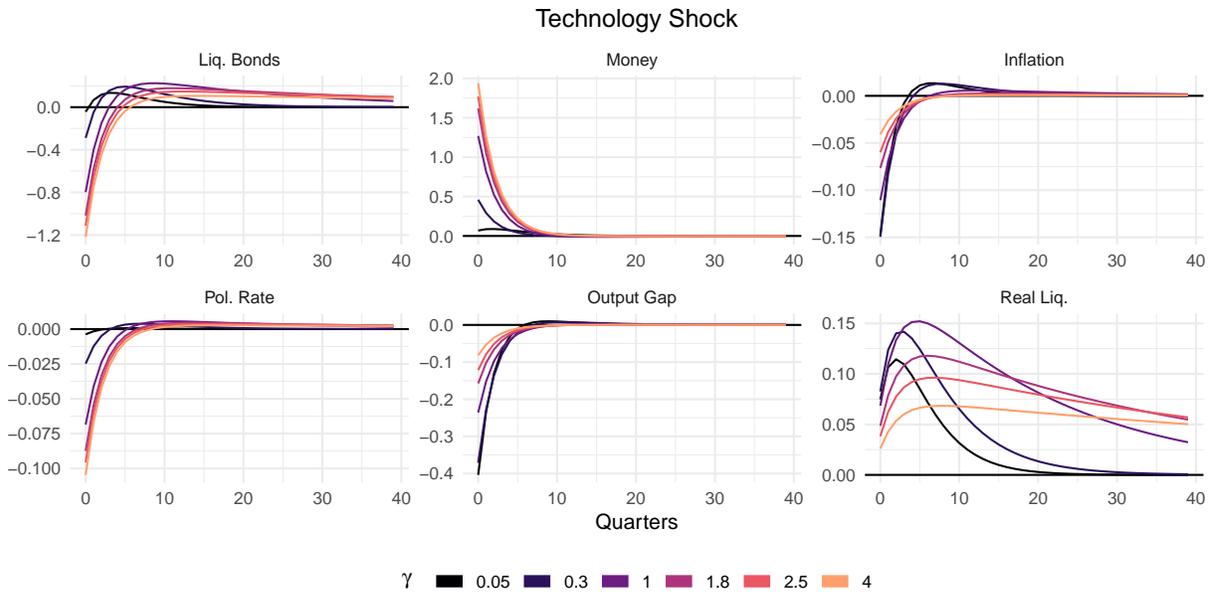


Liquidity model, IRFs for bonds b (top row), money holdings m (middle row), and real liquidity z (bottom row), under several values for money holdings semi-elasticity, ϕ , with $\psi = .02$. Monetary shock is a 1% key rate hike, annualised; technology shock is one standard-deviation. All values in percentages.

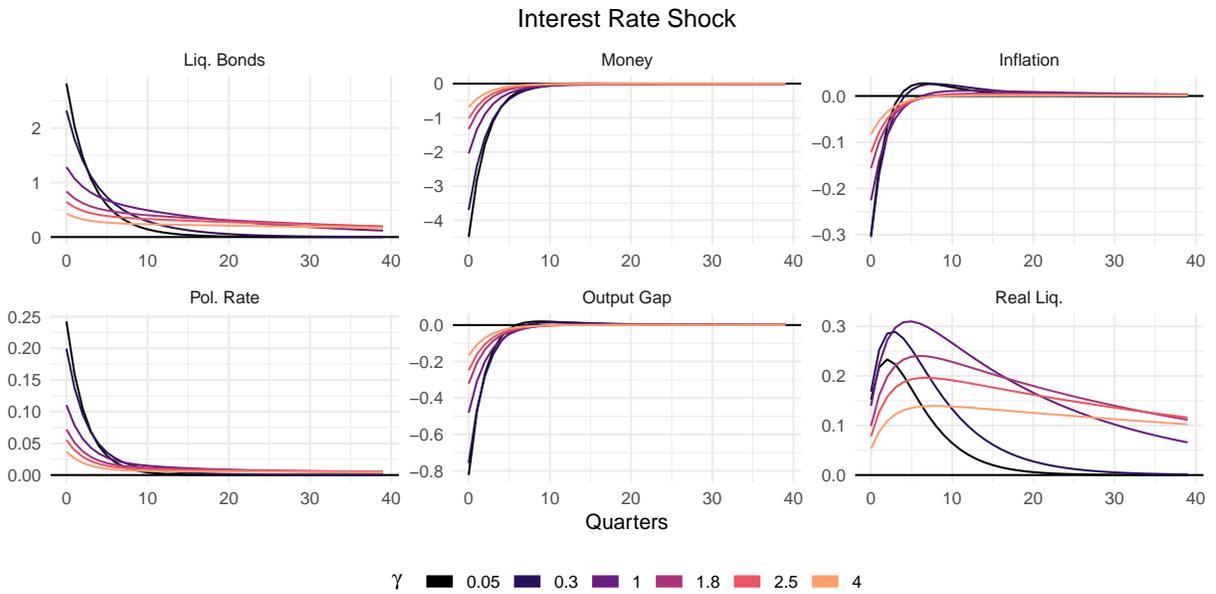
For both technology and interest rate shocks, higher semi-elasticity ϕ monotonically translates in higher responses on impact and little to no effect in terms of convergence speed. On the other hand, for liquid bonds, ϕ affects the response profile, especially for the monetary shock. In both regimes when ϕ is close to ψ IRFs take a smooth declining profile, whereas as ϕ grows away the reaction on impact decreases and the overall profile becomes more convex, speeding up convergence to the steady-state.

Turning to the effects on real liquidity z , the shock's nature determines the changes. Following a monetary shock, irrespective of the policy regime, higher values of ϕ imply less volatile movements in z and a quick convergence. The opposite holds following a technological shock, again for both policy regimes, although an active monetary policy seems to impart higher persistence in real liquidity after a TFP shock.

Figure 3.10 – Sensitivity on γ – IRFs



(a)

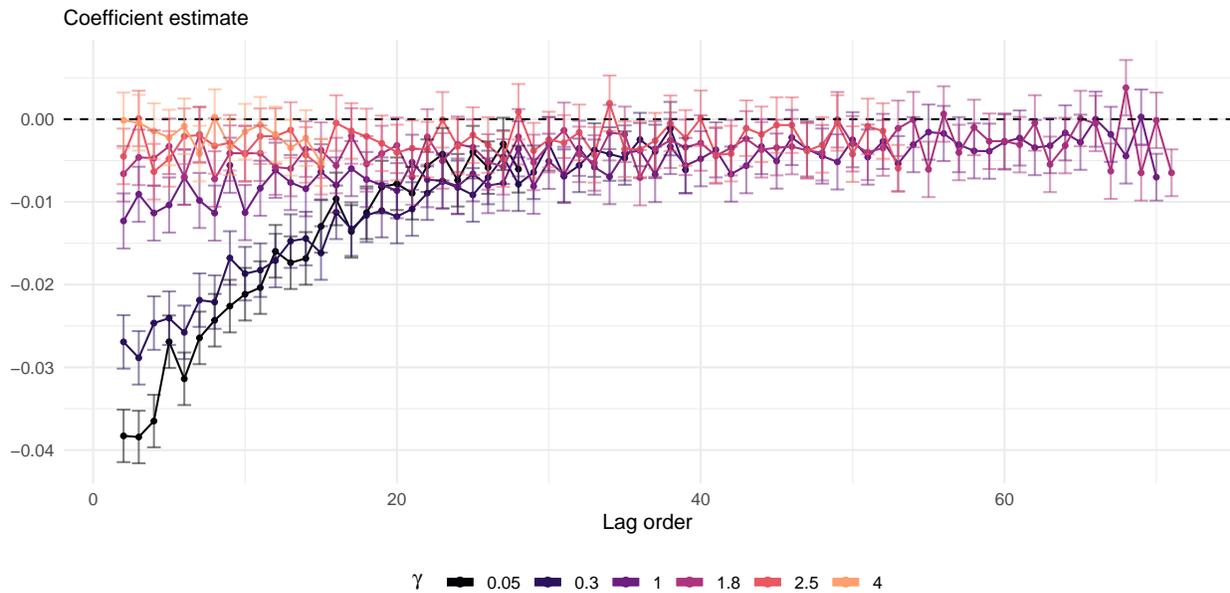


(b)

Full set of IRFs for several values of γ , subsuming the monetary policy stance.

3.E How γ affects inflation dynamics

Figure 3.11 – Sensitivity on $\gamma - AR(k^*)$ on Inflation



Coefficient estimates for autoregressive model on simulated inflation, optimal lag order. Error bars are twice estimated standard deviations around the point value. The first lag is excluded in the figure for scale: it is overall in the neighbourhood of the shock persistence, .65.

In the main body of the chapter, we restrict our attention to two values for γ , while this Section experiments with a larger set of values to assess how it affects the model dynamics. As shown in the chapter, the model we propose allows for passive monetary policy, which indeed influences inflation dynamics.

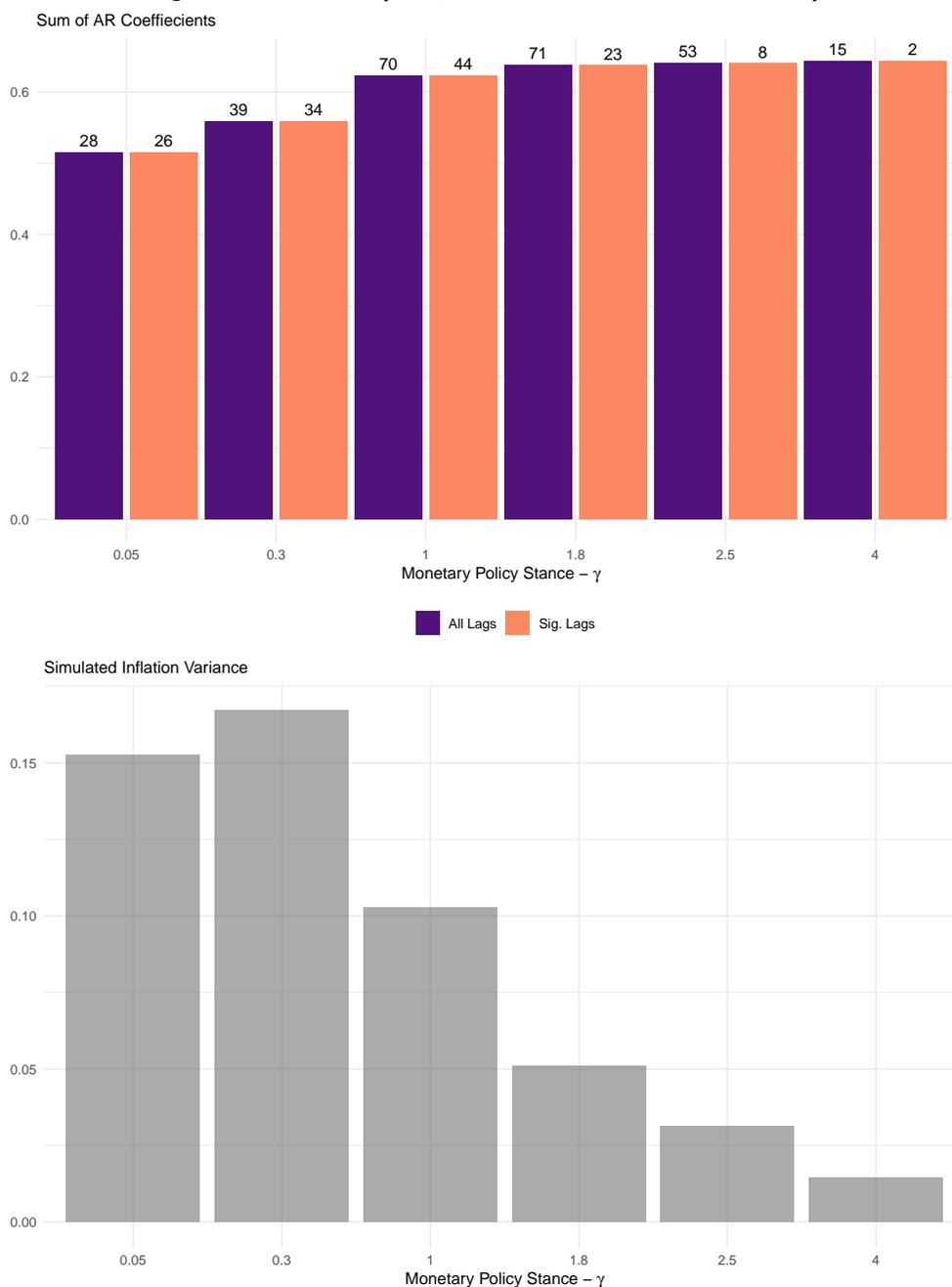
Consistently with the analysis in the main body, we compute optimal lags for the simulated series: as shown in Figure 4.11, when the central bank under-reacts to expected inflation in its rate leveraging, actual inflation depends more heavily on past values and readjusts more slowly. The number of significant lags and their magnitude is sensibly higher than the case of an extremely aggressive response to inflation (ie, $\gamma = 4$).

Figure 4.12 shows that inflation in itself exhibits lower persistence, but most importantly a high variance – which drives the results on persistence as measured with the sum of an AR coefficients. The tension originates from the effectiveness (or lack thereof) of the central bank to steer inflation back to its zero steady-state after a shock.¹⁰

The mechanism is the following: current high inflation π_t drives down real liquidity \hat{z}_t , which relates to liquidity allocation (eq. 4.14 in the main text). As the central bank does not keep suffi-

¹⁰The same logic holds for the baseline NKDSGE model, see below Section 4.G.

Figure 3.12 – Sensitivity on γ – Inflation Persistence and Volatility



Results based on simulations. Values for π persistence. Top: Sum of optimal autoregressive coefficients; Bottom: variance for simulated π .

ciently up with a raise in s_t , the liquid bond interest rate, money holdings \hat{m}_t need to compensate; reallocating away from cash pushes down s_t to clear the asset market. In turn, degradation of money holdings and liquid bonds' yield fosters demand, triggering inflationary pressure in ex-

pectations. As expected inflation $E_t\pi_{t+1}$ increases, the central bank raises the interest rate s_t by γ , ie less than proportionally, thus not enough to absorb into liquid bonds the asset reshuffling, which feeds again back into demand.

3.E.1 Misspecification

From a statistical standpoint, we can illustrate the relation between measured persistence, observed variance, and model misspecification. Assuming an econometrician who observes the series and wants to measure persistence, she fits a simple first-order autoregressive model $\pi_t = \rho\pi_{t-1} + \varepsilon_t$. The error in such specification is correlated with observed inflation via the transmission mechanism that is imparted by the model, namely the Phillips curve and the feedback through the central bank reaction. While the structural shocks impulsing the dynamics are the same sequence across simulation rounds, residuals also capture the channels of asset holding and total liquidity.

To see formally the link between measured persistence, observed variance, and model misspecification, recall that the variance of a first order autoregressive process is $V(\pi) = \frac{\sigma_\varepsilon^2}{1-\rho^2}$, so that expressing ρ leads to

$$\rho = \sqrt{1 - \frac{\sigma_\varepsilon^2}{V(\pi)}}. \quad (3.37)$$

Provided the fundamental shocks are the same across simulations and have thus identical variance σ_ε^2 , a more volatile series influenced by varying transmissions mechanisms will report a higher estimated persistence, ρ .

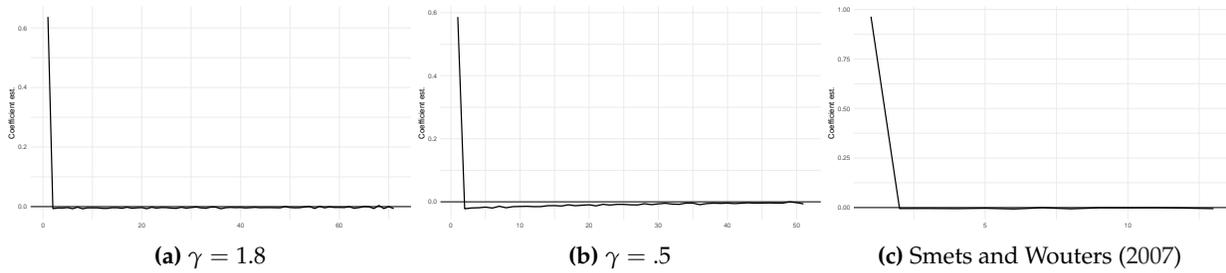
This relation is less transparent when the model to fit comprises more lags and the econometrician measures persistence summing over the autoregressive coefficients. Then, the estimated model imparts the following relationship, where L is the usual lag operator and $\Psi(L)$ is the associated characteristic polynomial:

$$\begin{aligned} \pi_t &= \mu + \sum_{i=1}^p \rho_i \pi_{t-i} + \varepsilon_t \\ \left(1 - \sum_{i=1}^p \rho_i L^i\right) \pi_t &= \Psi(L) \pi_t = \varepsilon_t \end{aligned} \quad (3.38)$$

Switching over to the roots of Ψ while keeping the relation with ρ coefficients is algebraically cumbersome for $p > 2$, the more so when the econometrician measures persistence summing over the estimated ρ 's.

3.F Full lags structure

Figure 3.13 – Autoregressive process lags



$AR(k^*)$ model, with k^* optimally selected lags. Estimated coefficients in solid, bands are twice the estimated standard errors. Full lags, only the intercept is omitted.

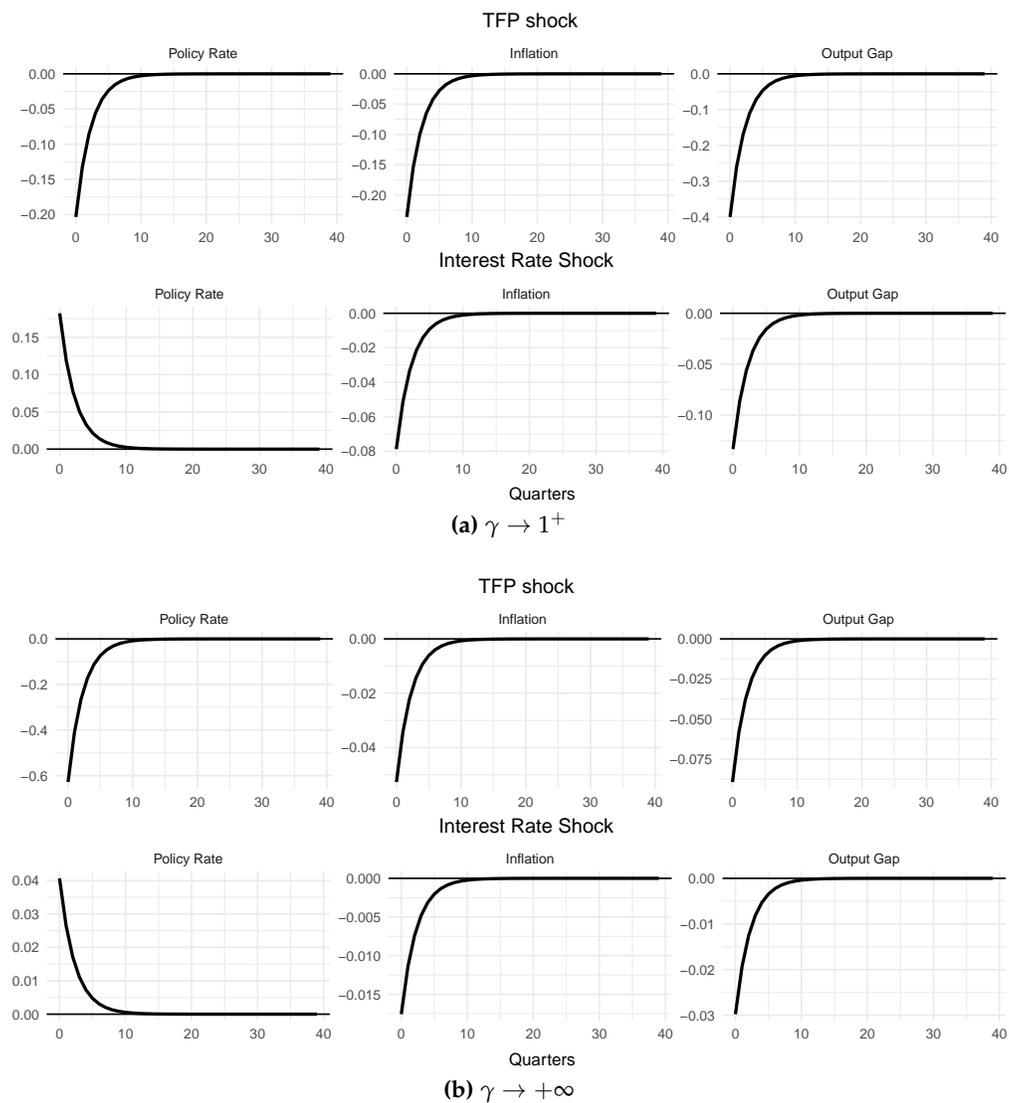
3.G Extreme policies

This section investigates what are the consequences of two extreme calibrations for the textbook NKDSGE models.

We let the reaction to expected inflation be negligibly close to its determinacy limit, such that $\gamma \rightarrow 1^+$, and then test for extremely high levels of reaction. The IRFs for these two cases are presented in the Figures below.

Interestingly, while calibrations close to indeterminacy do not display dramatic changes but only a higher effect on impact, an extremely reactive central bank can substantially suppress a sizeable share of volatility in the economy. Indeed, it appears to curb inflation and economic gap in a much more effective way when a fundamental productivity shock hits the economy – at the cost of considerable volatility in the policy rate, though. On the monetary policy shock, the decrease in volatility is even sharper.

Figure 3.14 – Extreme Policies in baseline NKDSGE



Baseline NKDSGE (Gali, 2015) with extreme monetary policies. Top panel: very low reaction to expected inflation, $\gamma \rightarrow 1^+$. Bottom panel: extremely aggressive reaction to expected inflation. Values in percentages.

Chapter 4

Inflation Persistence

Inflation is one of the main topics in macroeconomics. Its origin, dynamic behaviour, and control have sparked immense strands of research, from microeconomics to forecasting. In the last decade, inflation was – and still is – part of a lively discussion on monetary policy.

During the 2010's decade, inflation has been unexpectedly low and stable in advanced economies. This low trend and mild volatility are even more baffling in light of the large swings in economic activity, commodities price, monetary and fiscal policies. This recent dynamics of inflation is puzzling both if one compares it with historical data and if one looks at the predictions of most macroeconomic theories. From a historical perspective, during the last two decades inflation has become at the same time harder and easier to predict (Stock and Watson, 2007): significantly less volatile than in the post-war period and yet well modelled by a white noise rather than more structured models. Furthermore, according to conventional theories based on simple Taylor rules, the new, unconventional tools adopted by central banks in response to the 2008 Global Financial Crisis could have had small or zero effect or have generated inflation spiralling out of control, as in the late '70s (see, for example, Taylor (2014)). In contrast to these predictions, the US economy posted its longest expansion since WWII, until COVID-19 upended it, in a context of moderate inflation.

This chapter presents a wide-ranging empirical analysis on the dynamics of inflation mainly based on reduced-form models. We focus on the univariate properties of the inflation series since WWII, abstracting from an analysis of its determinants. Over this period, there were different phases characterised by various stages of structural change, varying degrees of trade openness, and different regimes for macroeconomic policies. To capture and exploit all information present in the data, we use five measures of price change covering consumption and production of goods and services in the US economy. Our contribution to the existing literature relates both to the sample and indicators of inflation, and on the methodology adopted. First, we study inflation persistence using longer time periods and several measures. Second, we implement relatively

recent methodologies, including artificial intelligence, which have not yet been fully exploited in the analysis of inflation.

Our analysis starts from simple, but reliable, autoregressive models that impose a structural straightjacket to the data. We then move to more flexible tools, such as a Bayesian state-space autoregressive analysis and the model-free deep learning approach. While Bayesian tools are largely common in macroeconomic analyses, machine learning is still at an exploratory stage, although their use in macroeconomic analysis is rapidly growing (Athey and Imbens, 2019; Varian, 2014), especially because of their excellent forecasting performances (Almosova and Andresen, 2019; Makridakis, Spiliotis, and Assimakopoulos, 2018). In this chapter we exploit the capabilities of these tools in order to identify with finer granularity the non-linear properties of inflation. Once the neural network is presented and trained with data, we use its forecasting properties to generate additional data points and assess more precisely how persistence has changed over time. We restrict our attention to inflation persistence and its dynamic changes. A broad definition of persistence relates to inertia, which is the property of an object to not deviate from its past dynamics in absence of external shocks. A highly persistent time series posting a 5% growth rate will likely move in such value's neighbourhood if nothing affects it. On the other hand, when the inertial series is hit by a shock, it will slowly incorporate and dissipate the shock over time. Similarly, weakly persistent series will display more variability and shocks will be depleted relatively quickly. An intuitive implication of inertia is predictability, which goes hand in hand with persistence.

The question of inflation persistence is particularly relevant for monetary policy. Assessing the sensitivity of inflation to changes is crucial for central banks when planning policy interventions: how aggressive should the intervention be to undo an inflationary spiral? Or symmetrically, what is the optimal timing for rate increases during a recovery or expansion? As monetary policy operates through lags, how much will it take for a shock to be transmitted to observed inflation? The degree of persistence also influences the tradeoff between inflation and economic activity: if inflation is persistent and far from the target, it will require corresponding larger output gaps. Likewise, the analysis of persistence provides information on how the sectoral structure of the economy, technology, international finance and trade affect the country's inflation. Analysing inflation persistence also sheds light on the process of price formation: is persistence stable? If not, how does it vary over time and why? Is there a mutual influence with measures of output persistence and volatility? Moreover, as shown by the large body of studies on the Taylor Rule parametrisation and its changes over time, it is still unclear whether these changes affect in any measurable manner the dynamics of inflation (see chapter (4)).

Understanding the dynamic properties of inflation improves the decision making of central banks in two crucial ways: before the policy decision, an extended information set helps to calibrate the intervention; and after the policy decision it helps to evaluate its effectiveness. Central

banks need to assess whether the sources of movements in inflation are inherited from deep, structural sources like price-setting strategies or, alternatively, whether are due to transient shocks to commodities prices. This is in turn useful also to evaluate the time lag between policy changes and changes in inflation, or the length of time needed to achieve the inflation target. For governments, the knowledge of inflation dynamics is important to design their fiscal policy. For example, with highly inertial inflation, a VAT increase will take several quarters to be fully transferred to final consumer prices. Moreover, if inflation, as in a traditional Phillips Curve, inherits its inertia from output, fiscal authorities might improve the global policy mix with the monetary authority.

Recent research highlights the interplay between trend inflation, inflation target, and persistence, see for example Cogley and Sbordone (2009), Kurozumi and Zandweghe (2019), and Stock and Watson (2007). To account for such interplay, throughout the chapter we control for trend inflation and for time-varying trends; nevertheless, such investigation is outside the scope of this study.

Literature overview

Inflation is a central theme in macroeconomics. We therefore contribute to a long and rich literature. Recent inflation dynamics in developed economies is the focus of Ciccarelli and Osbat (2017), Coibion, Gorodnichenko, and Kamdar (2018), and Miles et al. (2017), who apply diverse frameworks but overall report low and stable inflation rates since the 2000. The root causes for such dynamic behaviour are studied in three complementary strands of literature. A large number of studies focuses on the expectations in a Phillips Curve framework. A comprehensive overview of empirical strategies to estimate the effects of inflation expectations in the Phillips Curve is Mavroeidis, Plegborg-Moller, and Stock (2014), which emphasise the uncertainty and difficulties in precisely pinning down a robust specification. On the other hand, Coibion and Gorodnichenko (2015) and Coibion, Gorodnichenko, and Kamdar (2018) propose mechanisms of expectations formation to explore how these affect realised inflation.

An additional cause for inflation dynamics is found in the integration into global value chains, which ease the transmission of foreign shocks: Auer, Borio, and Filardo (2017) and Bianchi and Civelli (2015) fall in this line of research and find evidence of global inflation effects. These effects generally increase with openness but are stable over time. Along these lines, Jarociński and Bobeica (2017) augment a VAR with domestic and global factors to solve the twin puzzle of missing both disinflation and inflation in the Euro Area during the 2008 recession. They find that domestic factors counteracted global ones in the EA and can explain the missing inflation leg of the twin puzzles.

More generally, a third strand of literature has focused on the interplay of monetary policy regimes, inflation, and volatility shocks. Fernández-Villaverde, Guerrón-Quintana, and Rubio-Ramírez (2010, 2015) estimate DSGE models with stochastic volatility and evaluate the role of

shocks and policies in setting off the Great Moderation. Their findings point towards a minor role for policy in the steady dynamics of aggregates during the 1984-2007 period. Our approach is closely related to Pivetta and Reis (2007) and Fuhrer (2011): the former study inflation dynamics building on Cogley and Sargent (2002, 2005) with a flexible Bayesian approach. The latter offers a review of the state of the art in terms of measures, methods, and theories to evaluate inflation dynamics. The bottom line of both studies, though, is that inflation persistence is relatively stable in the post-WWII period, although both studies predate the 2008 recession and the ensuing policy innovations.

A series of interrelated studies investigated the dynamics of the inflation gap, which is the deviation from trend. Benati and Surico (2008) and Cogley, Primiceri, and Sargent (2008) estimate VAR models with a focus on predictability. Overall they find that US inflation has become less predictable and argue that this broadly corresponds to a more aggressive monetary stance or a change in the inflation target.

We extend the above analyses by using statistical learning tools such as machine learning, which are growing in empirical economic studies. Forecasting is one of the main uses of statistical learning tools. (Jung, Patnam, and Ter-Martirosyan, 2018; Kock and Teräsvirta, 2016; Makridakis, Spiliotis, and Assimakopoulos, 2018; McAdam and McNelis, 2005; Medeiros et al., 2019). In this respect, Nakamura (2005) employed plain neural networks to forecast inflation, while Almosova and Andresen (2019), in close relation to our work, compare recurrent neural networks against workhorse forecasting models to predict monthly inflation at several horizons. Researchers have previously ventured in adapting artificial intelligence to macroeconomic applications (Bajari et al., 2015; Chakraborty and Joseph, 2017; Giannone, Lenza, and Primiceri, 2018; Goulet Coulombe et al., 2019; Korobilis, 2018), econometrics (Athey, 2018; Athey and Imbens, 2015, 2019; Mullainathan and Spiess, 2017; Varian, 2014), or asset pricing (Gu, Kelly, and Xiu, 2020). Lastly, promising applications of sophisticated machine learning models have been proposed in computational economics, see for example Fernandez-Villaverde and Guerron-Quintana (2020), Fernández-Villaverde, Hurtado, and Nuño (2020), and Maliar, Maliar, and Winant (2019) who offer ML-based numerical solutions to DSGEs, or more generally Rackauckas et al. (2020) who incorporates ML for numerical solution of complex dynamic systems.

The rest of the chapter is organised as follows: Section (5.1) presents data, Section (5.2) overviews the empirical tools and resulting evidence from a plain autoregressive, frequentist approach, Section (5.3) describes the Bayesian take on persistence, Section (5.4) presents the results using the machine learning tools and results; finally Section (5.5) concludes.

4.1 Data and tools

While traditionally only one series is used in analysing inflation persistence, we consider three classes of inflation indexes for the US economy: the Consumer Price Index (CPI), the Personal Consumption Expenditure index (PCE), and finally the Gross Domestic Product Deflator. These three indexes are measured on different baskets of goods, hence discrepancies and deviations are due to the distinct subset of goods and services each index tracks. More precisely, the CPI mainly relates to consumers purchases, the PCE captures business sales, while the GDP deflator is measured on the goods and services produced domestically, abstracting from “imported” inflation. CPI and PCE also differ in the weights for each good and are available as “headline” and “core”, with the latter excluding volatile items like food, energy, and commodities. We cover almost entirely the post-WWII period, as series span 1948Q1:2020Q1 for CPI and GDP deflator, while PCE indexes start in 1960Q1. All series considered are historically revised to track as closely as possible the actual change of prices.¹

Researchers interested in monetary policy usually prefer PCE and the GDP deflator, as the former is the explicit target of the Federal Reserve Bank (Cogley and Sargent, 2005; Cogley and Sbordone, 2009), while forecasting and statistical analyses often rely on the CPI (Fuhrer, 2011; Pivetta and Reis, 2007). We cover the whole range of indexes to capture common trends in the dynamics of aggregate inflation.

All the series are sourced at a quarterly frequency from the FRED database of the Federal Reserve Bank of St. Louis. We take the raw level of the indexes and compute annualised quarter-on-quarter percentage change, to account for slow-moving trends in the data.² Throughout the analysis, our preferred measure is the GDP deflator index, since by design it tracks closely the variation in prices of goods and services produced and supplied within the US economy. This feature allows us to track more closely the underlying macro dynamics. Moreover, it also provides us with more observations, which can be used for checking the robustness of our results. Figure 5.1 plots all series that we will analyse.

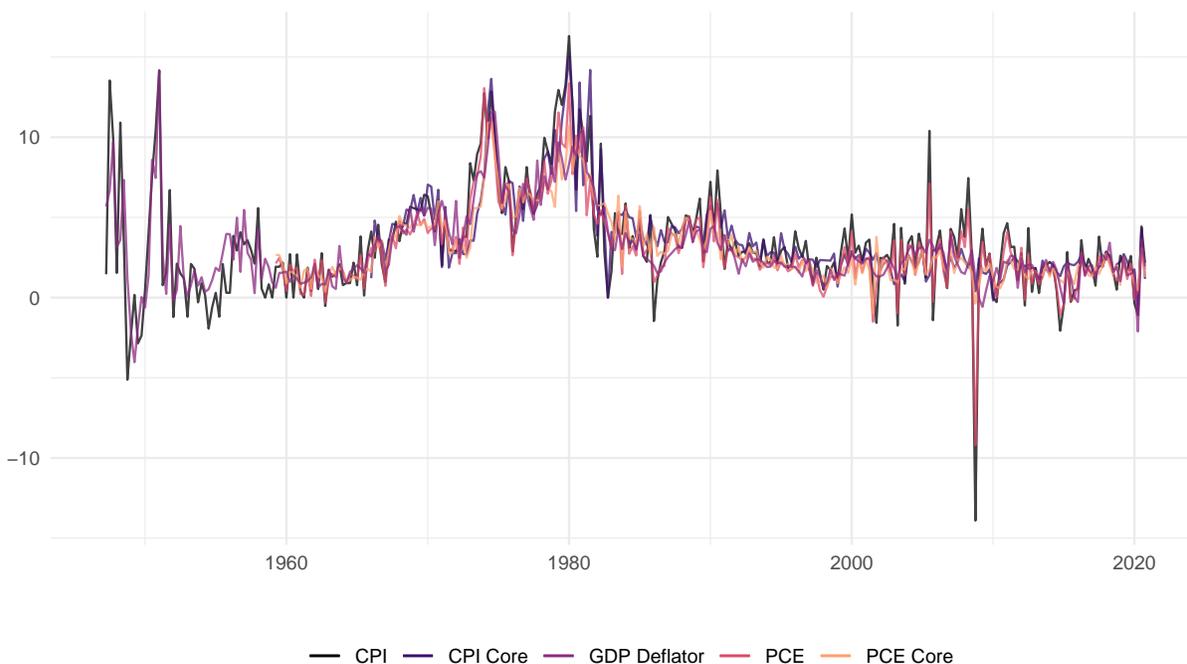
One stark fact emerges at a simple glimpse of the series. While for most observations headline and core inflation measures move hand in hand, they appear to diverge in volatility after 2000. To better tell apart these discrepancies, Fig.(5.2) compares headline and core measures for CPI and PCE over the whole sample and since 1990.

This difference in volatility is, at a first pass, due to commodities prices, which sharply fluctuated after 2000. The clearest example is the oil price, which posted threefold increases and

¹For Core CPI we drop the observations up until 1966, since those are interpolated from lower frequency data and thus carry very little signal to noise ratio.

²Quarter on quarter annualized percentage changes are left to the Appendix for comparison: overall trends do not vary significantly, whilst the levels change mildly.

Figure 4.1 – US Inflation Data



Inflation series: CPI, Consumer Price Index, including and excluding Food and Energy prices (starting 1947Q2, 1958Q4); PCE, Personal Consumption Expenditure index, including and excluding Food and Energy (1959Q2, 1959Q2); US Gross Domestic Product deflator (1947Q2). All series are computed as annualised log differences from previous quarter and end in 2020Q1.

contractions since early 2000s (Miles et al., 2017).³ Core inflation series are unaffected by these swings, as they exclude commodity prices. Therefore, analysing both headline and core series permits to isolate shocks, including their dynamic implications, arising from fluctuations in food and energy prices.

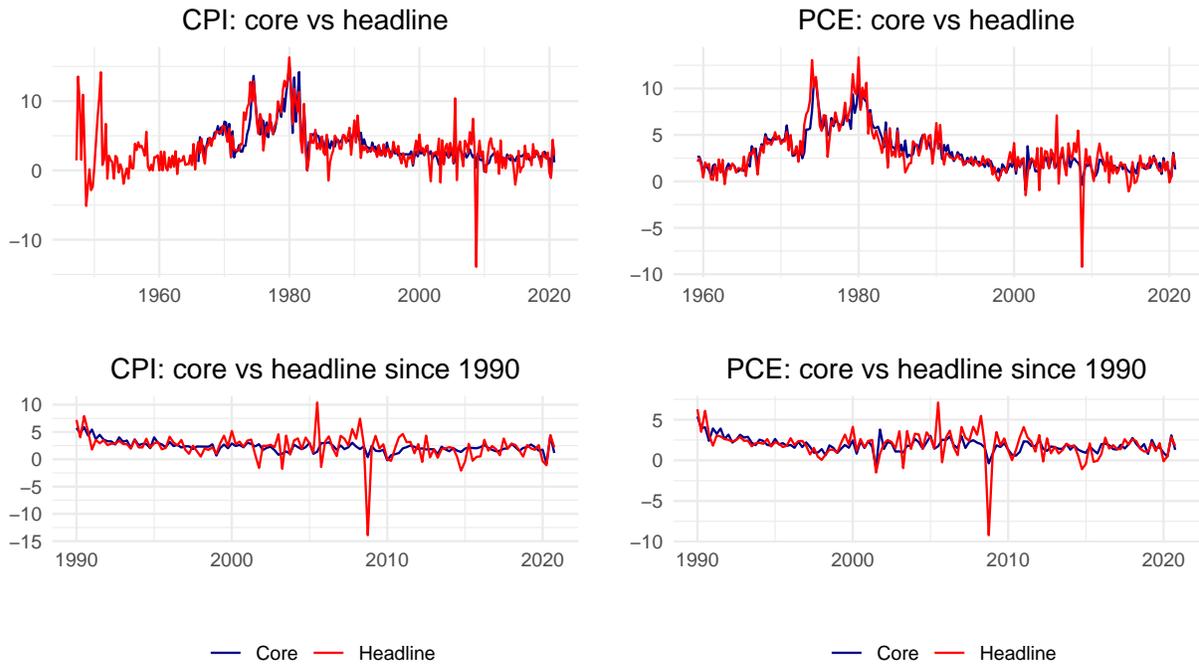
Tools and methods

Distinguishing core and headline inflation does not eliminate other sources of persistence in inflation, in particular the main channels working through forward-looking inflation expectations and the interactions with the level of economic activity, as postulated by several Phillips curve specifications:

$$\pi_t = \beta E_t (\pi_{t+i}) + \omega \hat{y}_t + \epsilon_t \quad (4.1)$$

³Fig.(5.10) in the Appendix plots price level and change for the West Texas Intermediate since mid-80s; Fig.(5.11) plots the same metrics for a global commodities index.

Figure 4.2 – Headline and Core Inflation



Top panel: full sample of headline and core series; bottom panel: headline and core series since 1990.

Thus, assuming errors are a zero-mean, drift-less iid process,⁴ inflation inertia is fully *inherited* from the dynamics of the output gap \hat{y} actualised at t . Several studies focus separately on expectations (Coibion and Gorodnichenko, 2015; Coibion, Gorodnichenko, and Ropele, 2019) or output gap (Mavroeidis, Plegborg-Moller, and Stock, 2014) to explain inflation persistence. The former find that market-implied, consumers', and professional forecasters' expectations drifted apart and match differently the actual value of inflation (Mankiw, Reis, and Wolfers, 2003; Trehan, 2015). In this chapter, we take a policymaker point of view and solely assess the transmission of shocks to actual inflation, disregarding effects on expectations or confounding sources such as movements in the output gap. This is clearly a reduced-form approach, but still informative on the underlying dynamics of inflation. Indeed, more general approaches are required to at least match what is observed in simple frameworks such ours.

We employ three classes of models to analyse inflation inertia, spanning increasing levels of refinement and complexity. The first step is a simple autoregressive approach with varying lags. Then, the same structure is extended in a Bayesian framework, and finally we fit a recurrent neural network borrowed from statistical learning. In all parts we measure inertia as the first order serial correlation or the sum of all autoregressive coefficients. These statistics convey enough informa-

⁴Potential measurement errors are alike, and omitted in the equation.

tion that allows us not only to judge inflation persistence but also to complement other common metrics, for example, the largest autoregressive root (LAR) or the half-life of a shock (both employed in Fuhler (2011) and Pivetta and Reis (2007)).⁵

4.2 Autoregressive analyses

As a first step to test whether inflation inertia has varied significantly over time, we estimate a simple $AR(1)$ model. To capture such variations, we estimate the $AR(1)$ model on a 56-quarter rolling window, in line with Fuhler (2011) and Pivetta and Reis (2007). Unstable estimates or large swings from such exercise would substantiate further analyses, aimed at decomposing the varying weight of past inflation on current price change. For this purpose, we consider inflation observations as drawn from the following process:

$$\pi_t = \beta_{0,t} + \beta_{1,t}\pi_{t-1} + \varepsilon_t \quad (4.2)$$

This barebone model represents the benchmark for our analysis. In such a framework β_0 represents the steady-state or trend inflation rate, while β_1 encapsulates any form of intrinsic inflation autocorrelation. We will primarily focus on $\beta_{1,t}$, without direct consideration of the intercept. Its consideration in our exercise would require the explicit modelling of trend inflation, which is out of the scope of this work.⁶ The error term ε , in this case, mops up new disturbances of any sorts, from expectations to technology, mark-up, demand shocks hitting inflation at time t .

In workhorse modern macro monetary models, the process generating inflation hinges mainly on expected inflation and disturbances in the current output gap: the Phillips curve, in this case, reads as in eq.(5.1) (Walsh, 2003; Woodford, 2003) and it is fully forward-looking with $i > 0$. The only sources of persistence are the serial correlation of shock to technology and the degree of price flexibility, both originating in the supply block of the economy subsumed in \hat{y} . This basic framework can be augmented by adding trend inflation and fluctuations around it.⁷ In empirical studies this extension helps to bridge the gap between observed and theoretical behaviour of inflation and links central banks' inflation targeting and short term fluctuations around such target.

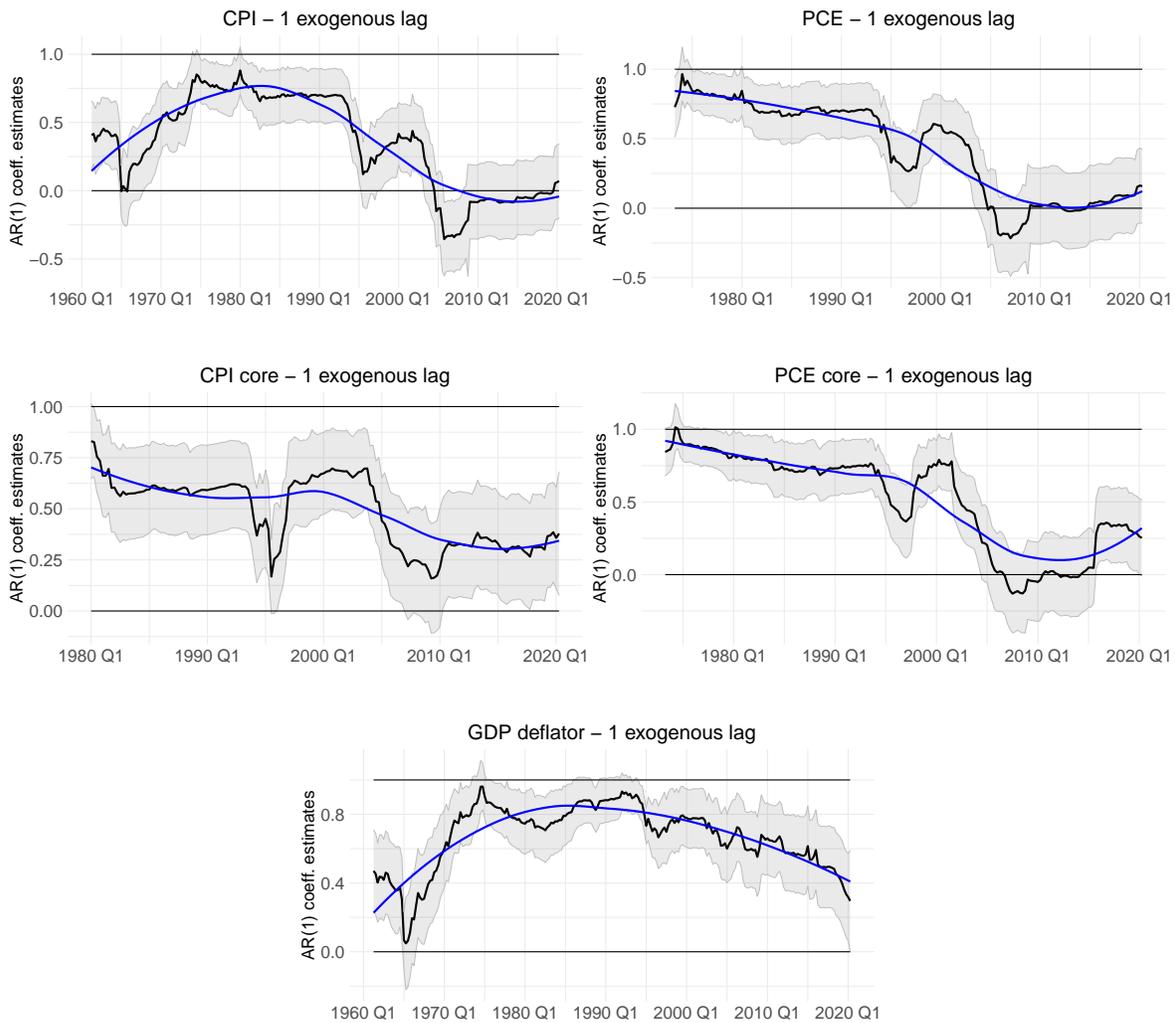
Fig.(5.3) collects the estimates from the $AR(1)$ process previously described, using the GDP deflator, CPI, and PCE indexes.

⁵Ideally, though, one would further add measures that are not easily summarised in one scalar, such as the variation of the autocorrelation function over time or the decomposition of permanent and transitory shocks' variances as presented in Stock and Watson (2007).

⁶Nevertheless, Appendix 5.H complements with the full range of trend inflation estimates.

⁷Cogley, Primiceri, and Sargent (2008) and Cogley and Sbordone (2009) provide a framework for modelling and estimating time-varying trend inflation and the inflation gap, respectively. Cogley, Primiceri, and Sargent (2008) find that the inflation gap display little persistence, while Cogley and Sbordone (2009) show that allowing trend inflation does away with inflation indexation in NK models.

Figure 4.3 – AR(1) Persistence



AR(1) estimates for CPI, PCE, GDP deflator series, rolling window of 56 quarters (14 years). Black line plots $\beta_{1,t}$ point estimates over time, red lines are 5% confidence intervals, blue line is LOESS fit to polynomially smooth out point estimates, grey bands are its 5% confidence intervals. Headline series are on the top row, core are on the middle row.

At a first glance, some regularities emerge: for all series analysed report sensible variation in $\beta_{1,t}$, alongside a generalised downward trend. This trend peaks roughly in the mid-90s in all series considered before decreasing at varying speeds. A 56-quarter window implies that the estimates for this period are based on a subsample that just excluded observations from the early 80s, a period of structural change and monetary (unexpected) intervention, namely the Volcker policy shift. Interestingly, the mid-90s estimates show high levels of variation, in comparison with the rest of the estimates: values for $\beta_{1,t}$ drop significantly before climbing back on trend, common

to all five panes, likely reflecting the switch induced by Volcker. The switch takes the form of a debasing of inflation from the previous trend, thus erasing dependence on past realisations.

Consistently with the increased variability displayed by headline series (as opposed to core ones excluding food and energy) the CPI series (left column) display relevant differences between headline (top left) and core series (mid left). This pattern is less evident in the PCE series, which in turn display a similar profile over time both in trend and magnitude. In three cases out of five, zero is included in the confidence interval roughly from 2005, implying a white-noise process.

Overall, these estimates point to a decreased inflation inertia, with significant drops taking place since the 2000s. This is particularly stark for CPI and PCE series but less clear for the GDP deflator: the first autocorrelation coefficient for this latter series starts decreasing in the mid-1990s and five years later displays a mild acceleration, with sensibly more smoothness than other indexes.

Such widespread dynamics allows the exclusion of commodity prices as the root cause of decreased inflation inertia, but the timing of the switches hints at factors like international trade shocks (Autor, Dorn, and Hanson, 2016; Bianchi and Civelli, 2015). Specifically, increasing economic integration at a global scale might foster the transmission of international shocks into domestic inflation, as argued by Auer, Borio, and Filardo (2017). This argument is corroborated by the fact that the GDP deflator displays a slightly different, and slower, decrease, tracking more closely the US national production.

The evidence offered by this simple analysis begs further investigation on the behaviour of inflation and its persistence. A more refined approach within the frequentist domain consists in extracting more information from the inflation time series by using an optimally chosen number of lags. This approach is applied in the next section.

4.2.1 Optimal lags selection

A natural step forward consists in adding more lags to the model we estimate. This addition allows a better framing of inflation persistence, since longer lags can capture dependence on realisations farther in the past. Two issues arise when comparing multiple lags estimates, though. Firstly, it is not clear whether a process with two lags like $.7$ and $.2$ is more, equally, or less persistent than a process with three lags, like $.5$, $.4$, $.3$. To circumvent this issue, we sum over the coefficients, compounding together all estimates. In this way we can compare a measure of persistence independently of the number and magnitude of the single parameters. Secondly, this approach allows for heterogeneity in the lags number for each series. We exploit this feature and compute, on the whole sample, the number of lags that minimises the Bayesian Information Criterion. Formally, the assumed process for inflation is

$$\pi_t = \beta_{0,t} + \sum_{i=1}^{k^*} \beta_{i,t} \pi_{t-i} + \varepsilon_t \quad \text{with} \quad k^* = \operatorname{argmin}_{k \leq \bar{k}} BIC(k; 1, \dots, T) \quad (4.3)$$

where \bar{k} is set to 18 quarters as an upper bound to the number of admissible lags. Conversely, we measure inflation persistence as follows, as presented by Fuhrer (2011) and Pivetta and Reis (2007):

$$\rho(k^*) = \sum_{i=1}^{k^*} \hat{\beta}_i \quad \text{with} \quad \beta_i = \frac{E(\pi_t \pi_{t-i})}{V(\pi)} \quad (4.4)$$

where k^* is computed on all available observations, at this stage. We estimate the $AR(k^*)$ process with a rolling window in order to study how $\rho(k^*)$ evolves over time, using again a width of 56 observations. Tab.(5.1) presents the optimal lags obtained for each series.⁸

Table 4.1 – Optimal lags selection via BIC

GDP Defl.	CPI headline	CPI core	PCE headline	PCE core
3	3	3	3	2

Fig.(5.4) shows that the same, generalized downward trend in persistence is found even when more lags are included in the model for inflation dynamics. The values reported are all in the same neighbourhood, corroborating the evidence of initially high but decreasing persistence.

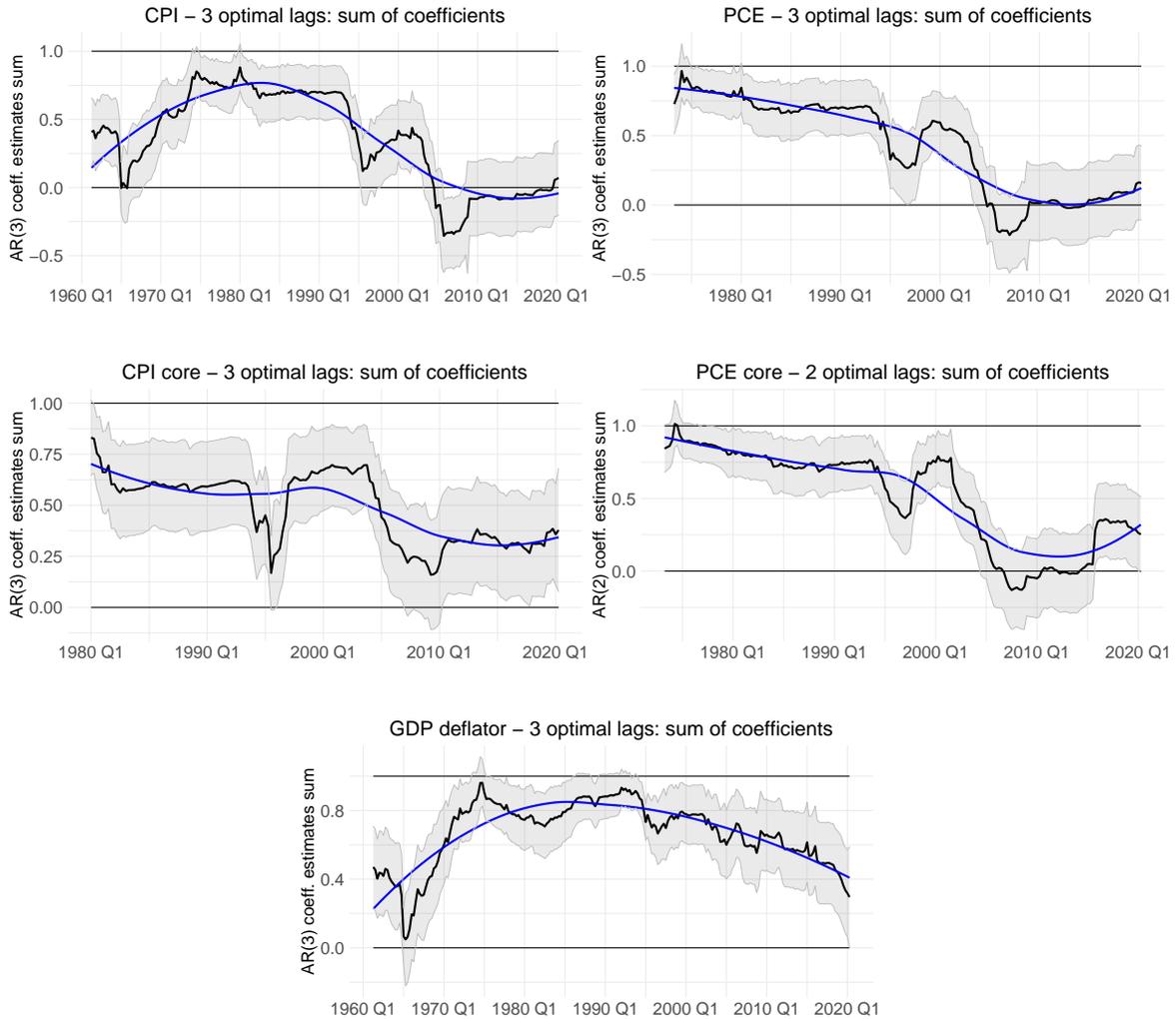
Individual profiles do not differ much from previous plots, with the GDP deflator showing the slowest downward trend in the series, possibly due to slow transformations taking place within the US economy’s composition and percolating onto prices. Although less stable than the previous case, the sum of autoregressive coefficients still reports a relevant drop around the mid-90s, when observations associated with Volcker’s initial period are phased out of the rolling window.

Again, a sharp decrease in inertia takes place around the year 2000 for CPI and PCE series (both core and headline). The sharp fall emerges from observations from the beginning of the Great Moderation, but for most series is followed by a modest rebound upwards. This last movement, though, does not fully offset the previous decrease and sets inertia on relatively lower values. The GDP deflator, consistently, follows a smoother, hump-shaped path, peaking during the 1970-90 period, with much less volatile point estimates and tighter error bands.

According to these slightly more refined analyses, inflation has become less persistent over the decades and has drastically accelerated this process during the last two decades. The pattern is consistent with two simple methods and is confirmed when we exclude volatile commodity prices, as in core series. The GDP deflator, which tracks more closely the economic activity in the

⁸We propose a similar table in the Appendix, including year-on-year growth rates, Tab.5.3. While year-on-year and annualised quarter-on-quarter series are computed on the same raw data and method, the former present significantly higher levels of autocorrelation: the optimal lags numbers are in all cases between 9 and 18.

Figure 4.4 – $AR(k^*)$ Persistence



$AR(k^*)$ estimates for $\rho(k^*)$, on a 56-quarter rolling window. Black line plots $\rho(k^*)$, grey bands are sums of SEs, blue line is LOESS polynomial fit to smooth out point estimates.

US economy, displays a much smoother dynamics of inertia. Appendix (5.D) presents results with varying windows widths, which do display the same pattern.

All in all, there is room for deeper investigation into the behaviour of inflation dynamics with refined methods. Next section uses a Bayesian approach to deepen the analysis and to exploit more of the information present in the data.

4.3 A Bayesian estimation of inertia

Adding a layer of sophistication to our inquiry, Bayesian methods helps in efficiently use data information, providing distributions of per-period measures of persistence. To this end, we adapt the approach illustrated in Pivetta and Reis (2007). We build upon this work operating on two margins: first, we use longer time series and, second, we run the estimations on five series rather than only one. Throughout this exercise, we set the lags to three, consistently with frequentist analyses outlined above.

The main framework of this section dates back to Cogley and Sargent (2002, 2005), subsequently extended in Pivetta and Reis (2007) to allow for degenerate, unit root draws. The assumed state-space model consists of the following components:

$$\begin{aligned}
 \pi_t &= \beta_{0,t} + \sum_{i=1}^3 \beta_{i,t} \pi_{t-i} + \varepsilon_t \\
 P(\beta_{t+1} | \beta_t, V) &\propto I(\beta_{t+1}) MVN(\beta_{t+1} | \beta_t, V) \\
 \implies \beta_{t+1} &= \beta_t + v_{t+1} \\
 \text{with } \text{var}(v) &= Q
 \end{aligned} \tag{4.5}$$

where the first equation is the measurement equation we also estimated in the previous section, the second line is the (hidden) state evolution, evolving as a multivariate normal distribution. β_t stacks all parameters at time t , $\beta_t = [\beta_{0,t}, \beta_{1,t}, \beta_{2,t}, \beta_{3,t}]'$. The state equation has the density of the parameters vector β depend on two components, an indicator function I that can be used to optionally exclude unit root draws and a multivariate normal density conditional on past draws of β , with *constant* covariance matrix V . The third line, implied by the Gaussian density, establishes that autoregressive parameters evolve over time as driftless random walks, potentially with unit roots.⁹

In this framework, β values are the model parameters, while the hyper-parameters are collected in the covariance matrix V . This latter gathers the co-variances of measurement and state equations:

$$V = \begin{bmatrix} \sigma_\varepsilon^2 & C' \\ C & Q \end{bmatrix} \tag{4.6}$$

⁹Ideally, further extensions accommodating for time-varying innovations would provide additional insights on the evolution of shocks and uncertainty in the economy. As a reference, see Bianchi (2013) and Lhuissier (2018), who develop estimated DSGE models with regime-switching uncertainty in volatility. Cogley and Sargent (2002) acknowledge such limitation in their work and tackle stochastic volatility in Cogley and Sargent (2005).

where $\text{var}(\varepsilon_t) = \sigma_\varepsilon^2$ is the variance of innovations in the measurement equation and Q is that for the state equation. C captures the covariance of measurement and state disturbances, set to zero.

To initialise β we use the first ten years of observations for each series, then the model is estimated on the remaining observations.¹⁰ These estimates are collected in $(\bar{\beta}, \bar{P}, \bar{V}, T_0)$, with $\bar{\beta}$ and \bar{P} being the OLS based mean and variance of a Gaussian distribution, and \bar{V}^{-1} and T_0 are scale matrix and degrees of freedom of a inverse-Wishart distribution, respectively. Therefore, the prior distribution on β_0 is a draw from the following joint prior

$$P(\beta_0, V) \propto I(\beta) \text{MVN}(\bar{\beta}, \bar{P}) \text{IW}(\bar{V}^{-1}, T_0) \quad (4.7)$$

After initialisation, the algorithm obtains draws covering the past posterior distribution of states and hyper-parameters. These are then used to compute conditional future paths for inflation and state. The final step computes persistence measures on these simulated paths. At each period t , conditional on past information, we simulate distributions for the next 120 periods.

To produce such simulations, we need to draw from the following posterior distribution:

$$P(\Pi^{t+1,t+h}, \beta^{t+1,t+h}, \beta^t, V | \Pi^t) \quad (4.8)$$

with Π^t collecting all observations until t . This posterior density can be separated into past and present beliefs and future uncertainty, conditional on time- t information, as follows:

$$P(\Pi^{t+1,t+h}, \beta^{t+1,t+h}, \beta^t, V | \Pi^t) = \underbrace{P(\beta^t, V | \Pi^t)}_{\text{beliefs on past and present}} \times \underbrace{P(\Pi^{t+1,t+h}, \beta^{t+1,t+h} | \beta^t, V, \Pi^t)}_{\text{future uncertainty}} \quad (4.9)$$

The first block can be sampled via a Gibbs sampler. Draws from this sampler are later used to simulate future trajectories conditional on data-informed beliefs up to time t . Additional details on the algorithm to sample from such posterior density are presented in Cogley and Sargent (2002) and Pivetta and Reis (2007).

We set 300 thousand total draws, with a burn-in of 150 thousand to deal with path dependency and to ensure convergence to the posterior. Therefore, we use in our computations of persistence 150 thousand actual draws. Future paths are simulated up to a 120 quarters horizon, equivalent to thirty years of synthetic history. In line with Pivetta and Reis (2007) but in contrast to Cogley

¹⁰Looking back at Fig.5.1, though, this step produces potentially biased hyperparameters, in light of the stark heterogeneity with the rest of the sample.

and Sargent (2002), we do not rule out explosive roots, so to report the complete, rather than the truncated, distribution of β s.

The results are plotted in Fig.(5.5) for all series in exam, together with upper and lower 5% credibility intervals.¹¹ The advantage of the Bayesian setup is to summarise all information on persistence at time t from both data and prior in order to extend the sample simulating a wealth of future, consistent paths. At any date t the state incorporates all information available in the past observations. Consistently with this information, the space of future realisations is duly explored far into the future and covering large swaths of the domain of the distributions. Therefore, it pins down more precisely the measurement of inertia at time t rather than relying on a rolling window approach with a fixed number of observations.

The output of this Bayesian exercise broadly corroborates the findings of simpler, frequentist approaches, with some significant departures. Overall, swings in inflation inertia are smoother and more gradual, unfolding over the whole length of the sample.

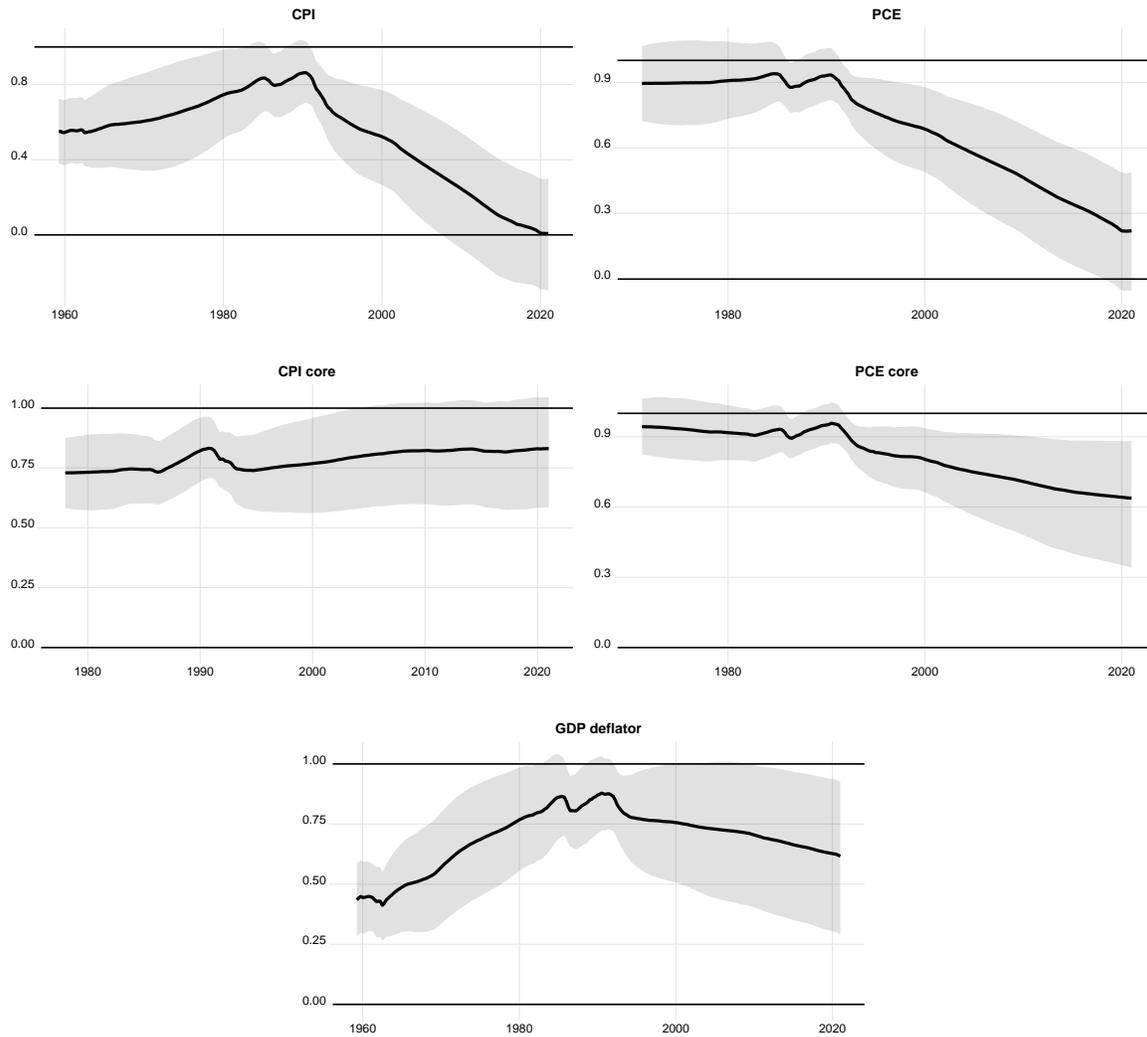
Scrutinising the general pattern, though, one can easily make contact with previous results: inflation inertia decreases from relatively high levels. Point estimates report generally higher values than Fig. (5.3) and 5.4, notwithstanding a common temporal path. The peak is generally reached in the mid-90s, which backdates slightly the onset of the decrease and somewhat weakens the international trade cause for such decline, as China officially entered WTO only in 2001 (Autor, Dorn, and Hanson, 2016). Moreover, this decline appears to be preceded by a phase of increase, as suggested in previous analyses. Core CPI persistence, though, stands out: it shows a rather stable path, if not slightly increasing.

A second common pattern is an increasing uncertainty around the median: in most series credibility intervals widen visibly toward the end of the period, with GDP deflator's and core series' intervals covering much of the unit space, in line with the higher unpredictability of inflation put forward by Stock and Watson (2007).

This unpredictability is even corroborated as credibility intervals reach zero for the headline CPI and PCE. This levelling down corresponds to the flattening displayed in Figs.(5.3) and (5.4), suggesting that after the Global Financial Crisis inflation is much closer to white noise than to an autoregressive process. In turn, it is necessary to couple such reduced form analysis with more interactions with other economic forces to fully unbundle the effects of economic slack, fiscal and monetary policies, international spillovers. A first sophisticated step towards such setup is carried out in Fernández-Villaverde, Guerrón-Quintana, and Rubio-Ramírez (2010, 2015), which estimate a rich DSGE with volatility shocks and study whether monetary policy switches had more effects than reduced variance on taming inflation.

¹¹Full draws distributions per quarter are deferred to Appendix (5.E).

Figure 4.5 – Bayesian Estimates



Sum of $AR(3)$ coefficients calculated on the simulated future paths for CPI, PCE (both with and without food and energy), GDP deflator. Black solid lines are the median values, shaded grey areas are 95% credibility intervals. Total draws 300000, burn-in 150000, thus 150000 final draws, per period t .

As a side note, an advantage of this Bayesian approach is the potential to extend this kernel to include more structured and informed models. In fact, Cogley, Primiceri, and Sargent (2008) and Cogley and Sargent (2005) do take this approach to structural models.

4.4 RNN-LSTM approach to persistence

Based on our reduced form perspective, the lag structure and possible non-linearities play a crucial role. For instance, inflation may display a slow-moving drift that affects realisations at sensibly

long horizons. To tackle this possibility, we borrow from a class of models that are precisely designed to handle long, short, and time-varying lags in a flexible and dynamic way.

Long Short-Term Memory models (LSTMs) are machine learning algorithms that exploit the structure and advantages of Recurrent Neural Networks (RNNs). This class of essentially non-parametric models has the advantage of effectively handling a very large set of functional forms under mild regularity requirements (Kidger and Lyons, 2019; Leshno et al., 1993; Tabuada and Gharesifard, 2020). The downside is the infamous black-box nature and the complexity of the inner mechanisms: the resulting estimated network can hardly provide intuitive insights on the connections and relations between data points. Conversely, though, machine learning models produce reliable results in terms of fit and forecasts.

The broader class of neural networks (NN) does not keep up well with persistence, as these models conserve little information about data with potentially long time dependencies. This issue is known as the “vanishing gradient” problem. It originates with the back-propagation algorithm, which, in a nutshell, is an efficient way to minimize a loss function evaluated on data samples by adjusting the NN parameters according to the values of the (chained) gradient. This, coupled with parameters being typically constrained within the $[-1; 1]$ interval, implies that deep networks sequentially multiply small adjustment values, quickly falling to zero. In this way, past information is lost.

Recurrent neural networks (RNN) overcome this shortcoming by explicitly carrying forward relevant information through a hidden state that gates out non-relevant information. This mechanism is reinforced in “stateful” LSTMs, a subset of RNNs.¹² The flip side of this feature is the requirement of long series for training, so much so that the roughly 250 observations present in our quarterly series are barely sufficient.¹³ In informing this section we mainly refer to Almosova and Andresen (2019), who first applied stateless RNN-LSTM to inflation forecasting. They show that these models can outperform most traditional forecasting tools and thus are interesting devices to study dynamic properties.

Indeed, they find that these models outperform common forecasting tools at most horizons, prevailing decidedly after the two years horizon. In their investigation, they use monthly raw data to let the model pick up spontaneously any non-linearities in the data – such as seasonality.

Our approach for this application consists of two steps of increasing granularity. First off, we simply feed the whole sample to the LSTM, let it learn freely and then produce a sufficient number of forecasts to compute the usual statistics on inflation persistence. These forecasts will depend on whatever the LSTM learned from the sequence and will allow for a synthetic extension of the

¹²A more detailed and formal introduction to stateful RNN-LSTM is presented in the Appendix.

¹³Properly estimated statistical learning models require about 10^7 data points to train on. With macroeconomic time series, we hardly work with series longer than 300 quarters. Quarterly series are more common in macro applications than weekly and monthly data. The latter are often not available or highly seasonal.

sample size. The output of such trained networks provides insights on likely paths for future inflation and its inertia.

However, to assess the *dynamic change* of persistence we need to train the model on sub-samples of the data. Two options lend themselves to the task: we first split by decades the time series and repeat the analysis just outlined; secondly, we let the LSTM train on a rolling window. This latter will output predictions that can be used to compute persistence and its change over time, much in the spirit of our previous exercises. In the same vein as the Bayesian method, we train the network on a fraction of the data and simulate model-consistent future paths for inflation. These provide additional data points to measure variations in persistence at any given point in time over our sample.

4.4.1 LSTM forecasting

To produce forecasts for our analyses it is important to decide whether to use a direct or indirect approach to forecasting. The latter consists in feeding the model with data up to time t and subsequently with its own previous forecasts, therefore iterating on data and forecast values. Direct forecasting, on the other hand, use specifically designed models to produce forecasts at a given horizon. Marcellino, Stock, and Watson (2006) compares these two approaches to time series forecasting and find that for linear specifications iterated forecasts perform better than direct ones, and improve with longer forecasting horizons. The case of LSTM differs from the framework of Marcellino, Stock, and Watson (2006) as these models are not strictly linear.

To convey this idea, consider the following simplification. LSTM network links past information to present observation through an arbitrary function F :

$$\pi_t = F(\pi_{t-1}, \dots, \pi_{t-p}; W) + \varepsilon_t \quad (4.10)$$

with p being the lags, W collecting network's parameters, and ε representing an arbitrary error, not necessarily Gaussian nor iid. Then, when the model is trained and \hat{W} is optimised, the model boils down to a possibly non-linear function \hat{F} , which can be used to produce forecasts. Naturally, the one period ahead forecast, conditional on time t , reads

$$\hat{\pi}_{t+1|t} = E_t [\hat{F}(\pi_t; \hat{W})] \quad (4.11)$$

Iterating forward, then, equates to

$$\begin{aligned} \hat{\pi}_{t+2|t} &= E_t [\hat{F}(\pi_{t+1}; \hat{W})] \\ &= E_t [\hat{F}(\hat{F}(\pi_t; \hat{W}); \hat{W})] \\ &= E_t [\hat{F}(\hat{F}(F(\pi_{t-1}, \dots, \pi_{t-p}; W) + \varepsilon_t; \hat{W}); \hat{W})]. \end{aligned} \quad (4.12)$$

Potential non-linearities in F and \hat{F} prevents from taking out ε from the expectations operator directly, but rather calls for computationally intensive integration. Furthermore, no assumption is cast upon the distribution of errors, which in turn accrue over the iterations. A more appropriate approach would consist in fitting one model for each forecast horizon, $\hat{F}^{(t+1)}, \dots, \hat{F}^{(t+h)}$, based only on information at time t . Such solution is more in the spirit of direct forecasts. While more appropriate, this avenue is computationally demanding, thus we simply assume ε to have mean zero and iterate on previous forecasts, as in Almosova and Andresen (2019).

4.4.2 LSTM setup

When setting up a LSTM for training, the researcher needs to define its structure, the nodes, and a loss function to evaluate the fit. We study models with one and two layers, and varying numbers of nodes per layer. Satisfactory results can be obtained by a one-layer LSTM with about 75 nodes. Almosova and Andresen (2019) find that the best performance in terms of forecast RSME is produced with 100 nodes.¹⁴ Our preferred loss function is the mean squared error (mse) loss, computed comparing at each step the discrepancies between true and predicted values generated by the network and then used to guide further adjustments in the network's parameters. The choice of such loss function is useful to make direct contact with standard econometric tools, but similar results can be achieved with other compatible loss functions, like mean absolute error (mae).

Weights and biases of the network are optimised to minimize such loss function via the ADAM optimizer (Kingma and Ba, 2014), which is now standard in the field of machine learning (Schmidt, Schneider, and Hennig, 2021). LSTMs feature large numbers of parameters to optimise, usually in the order of thousands if not tens of thousands, and are updated at every iteration. The optimising algorithm explores such highly dimensional parametric space following the gradient of the loss function for as many epochs as the researcher decides to train.¹⁵ This implies that a neural network can be presented several times with the same batch of data and incur in overfitting on the training set with poor out-of-sample performances. We tackle the risk using the early-stopping criterion to govern the adaptive stopping of the optimization. The network is thus shown a 90% subset of the training sample, it is fit with such subsample only. The iterations stop when the loss does not decrease for a given number of iterations on the 10% that was left out for validation. This criterion ensures the generalization of the resulting network. To further improve the generalisation of the results, we impose $L2$ regularisation on the network parameters, nudging weights towards zero in the vein of a Ridge regression.

¹⁴Although it is not clearly stated, one can infer from the parameters count in footnote 8 that such model features a single layer.

¹⁵Other tuning parameters depend largely on the algorithm of choice: with ADAM, we rely on the default values for the perturbation of the first two moments of the stochastic gradient, namely $\beta_1 = .9$, $\beta_2 = .999$. We also impose $L2$ (Ridge) regularisation on the deeper parameters of the network.

4.4.3 Full sample: forecasting inflation and its persistence

We present here the results of a model with one and two layers, 1000 nodes per layer, trained on the full sample of each series. These networks are then used to forecast the inflation rates for the following ten years. Each model is trained on a variety of different periods, with varying volatility, trend, cyclicity, and monetary policy regimes. This exercise is interesting since LSTMs are geared to capture at the same time short period swings and dependencies that unfold on longer horizons. Extending the sample with predictions from these trained LSTMs allows for a first assessment of the feature learnt and also to propose possible future realisations for inflation inertia going forward.

This first set of results builds on the extension of the sample via forecasts produced by networks trained on the full sample. Comparing models with one or two layers, both with 1000 nodes per layer, the latter does not seem to take advantage of the deeper structure, although forecasts are qualitatively closer to past realisations and smoother overall. It is reasonably due to a shortage of data points: deeper networks, despite the parameters regularization imposed on them, navigate a much more highly dimensional parameter space and thus need more variation in the data as well as more observations to devise a minimum in the loss function.

The next sections present the results of our study employing synthetic data generated by a set of LSTMs on diverse subsamples of the data. From a technical point of view, the LSTMs seem to attain a steady-state-like level when used to iterate forward: despite the absence of clearly defined equilibrium linkages and structural shocks, forecasts converge to the sample mean when iterated for sufficiently long horizons.¹⁶

4.4.4 Regressions on LSTM

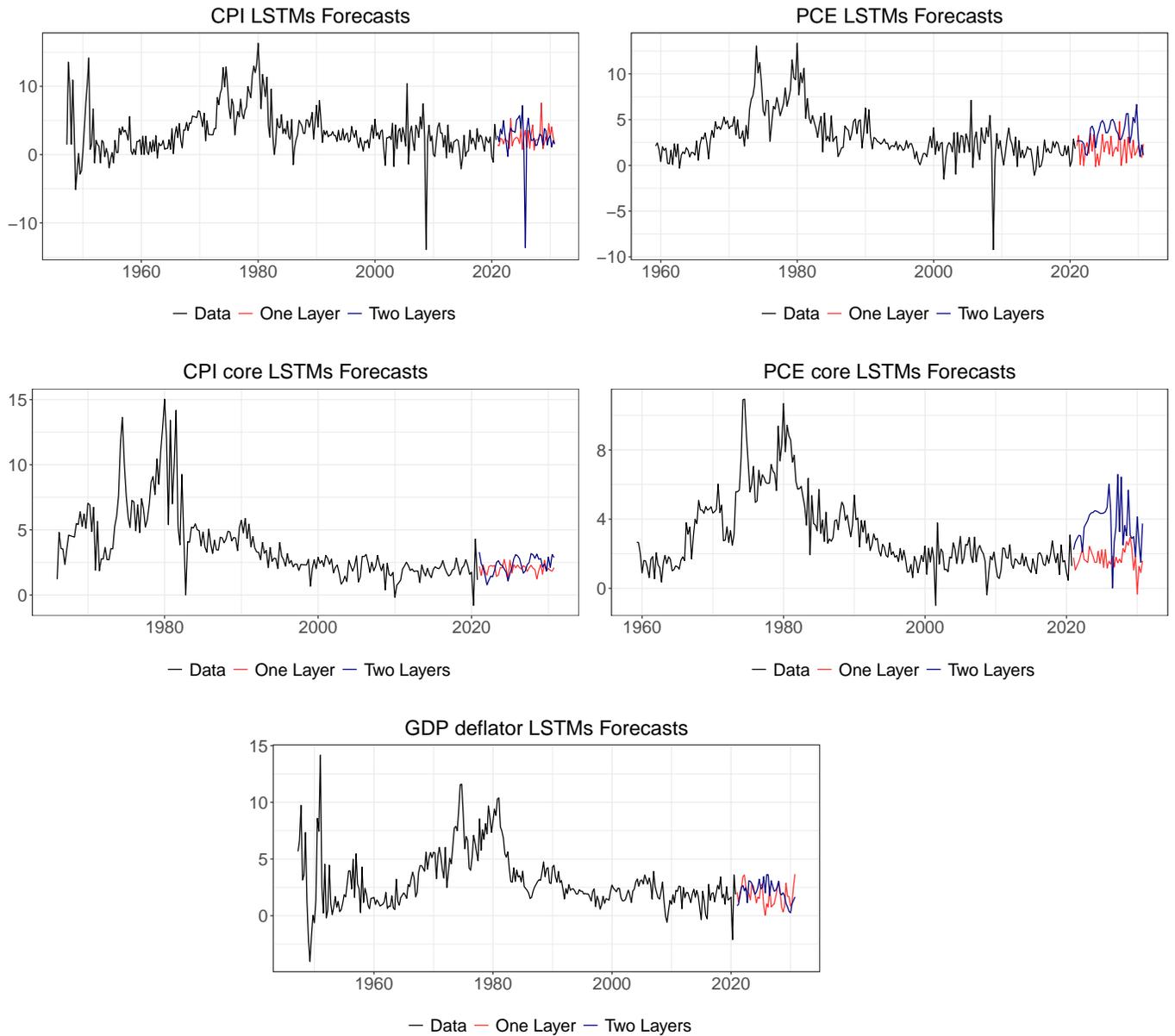
The first approach to measure variations in inflation inertia consists in splitting the series into 10-year non-overlapping subsamples (plus optimal lags). Each subsample is used to train a LSTM network that subsequently forecasts on a 40-quarter (ten years) horizon. Then, we fit an autoregressive to elicit the persistence dynamics incorporated into the LSTMs from inflation observations. Panes in the left column of Fig.(5.8) plot the values from these regressions for the β_1 , alongside with confidence intervals. To complete the analysis, right panes in Fig.(5.8) present the values for the sum of β s for AR(3) models.

Such lag length choice mirrors the optimal lags selected in Section 5.2 via the BIC minimisation. Broadly, the downward trend for inflation inertia is confirmed, with a sharp drop in all series except the GDP deflator series, which flattens slightly and takes a hump-shaped profile. This holds for both the AR(1) and the AR(3) analyses.¹⁷

¹⁶See the presentation of LSTM in the Appendix to clarify the role of the sample mean.

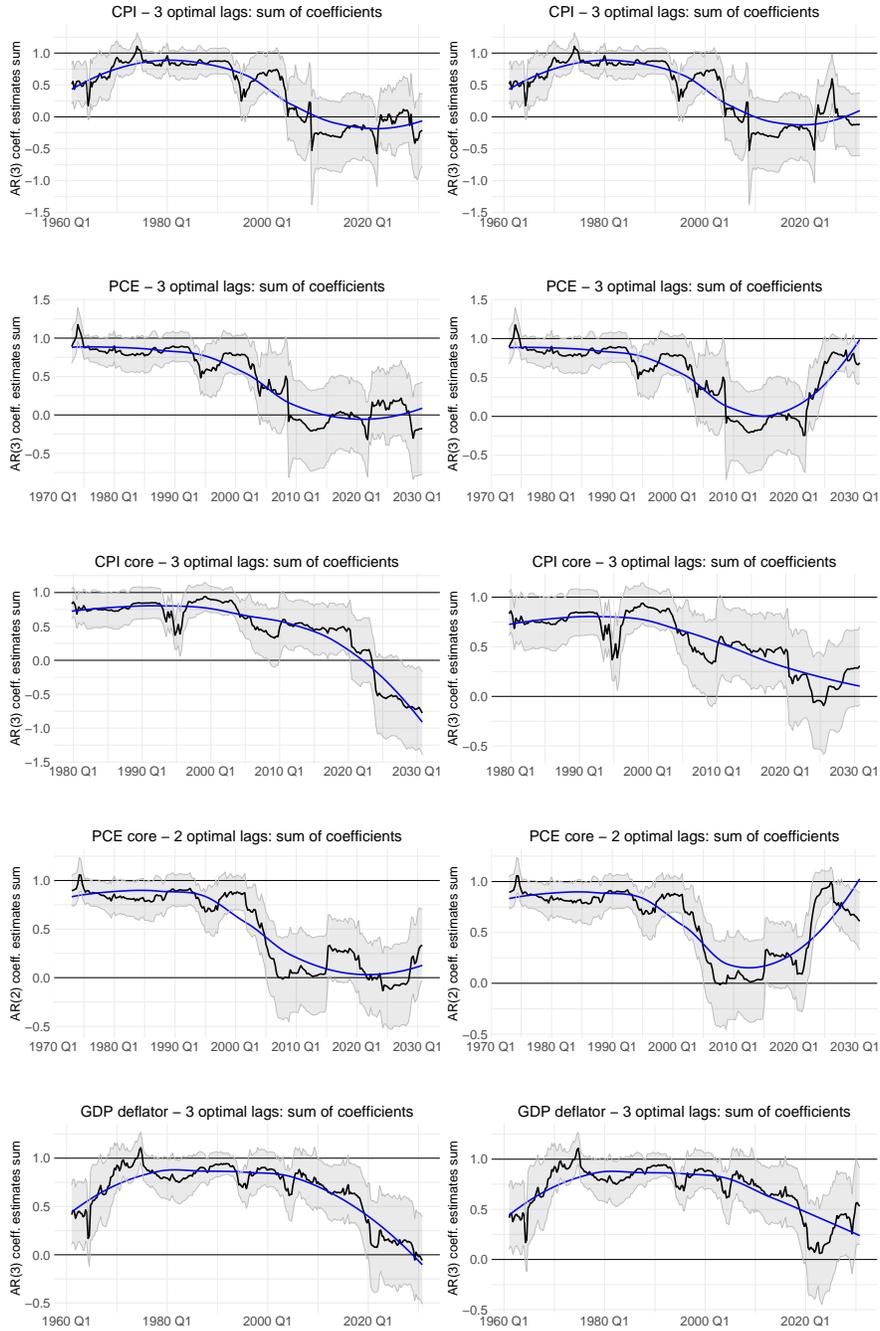
¹⁷Appendix (5.G.1) presents detailed OLS regressions results behind such bar plots.

Figure 4.6 – Full Sample Forecasts



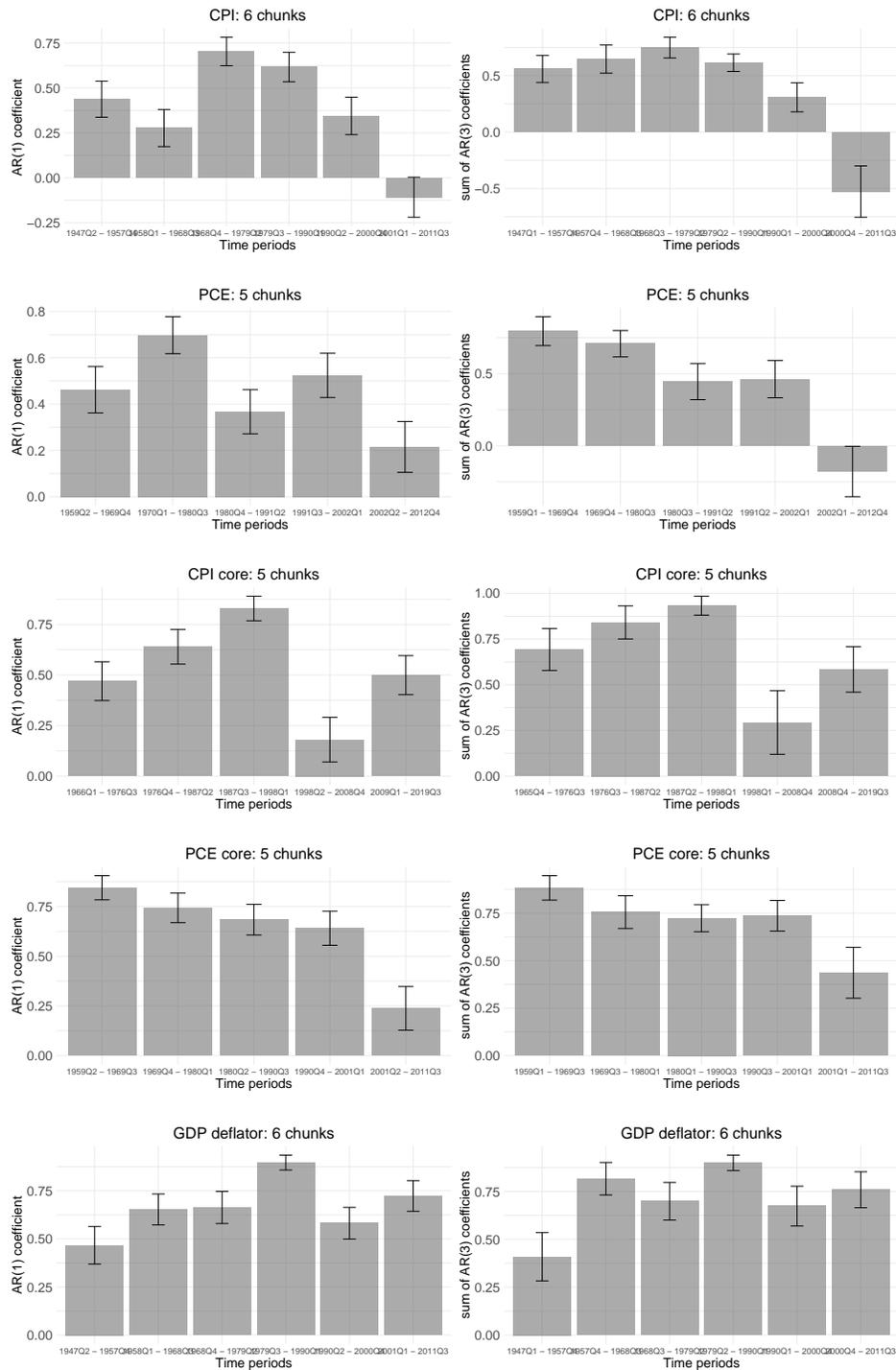
Forecasts from one- and two-layer LSTMs trained on the full sample. Iterative forecasts on 40 quarters.

Figure 4.7 – Full Sample LSTM Predicted Persistence



Left column: AR(k) on data and one-layer LSTM forecast (2020 onward). Right column: AR(k) on data and two-layer LSTM forecast (2020 onward).

Figure 4.8 – LSTM on Decades – Persistence On Data and Forecasts



Left column: $AR(1)$ estimates for β_1 . Right column: $AR(3)$ estimates for sum of β_i . Subsamples are non overlapping and encompass 10 years of data, plus appropriate lags. 'Time periods' report the start-end dates for the subsample of the actual data. Forecasts start from the end date of the sample and run on iteratively for the following decade. Each LSTM network has one layer, 500 neurons, MSE loss function, early-stopping, and is trained for 2000 epochs.

4.4.5 Rolling LSTMs

This Section presents results from a set of networks trained on a rolling subset of data. Each window spans ten years and is used to train a LSTM net; then, indirect forecasts for the next 40 quarters are produced and used to compute persistence statistics. Once the process is over, the window moves one quarter ahead, drops the oldest observations and a new network is trained. In the same vein of the rolling window in Section (5.2), using subsets of data and brand new networks allows the assessment of dynamic changes in the underlying data generating process. The idea is to track closely different features of inflation that LSTMs detect in the data and replicate in the forecasts. The advantage over distinct samples is that it allows for a visual detection of such changes – smooth or abrupt ones – at the cost of a precise timing of structural breaks. This procedure is done for AR(1) and AR(3) models.

In general, reported persistence is significantly spikier, with large swings in both higher and lower levels of inertia. Partly, this comes from the relatively small set of data points used to train the networks, which in turn pick up and possibly over-represent local features.¹⁸ Nevertheless, these analyses provide interesting insights on inflation persistence itself, and on different properties of the series employed.

At a first pass, left panes in Fig.(5.9) display a higher persistence in the initial decades of the covered period, while the scenario is more mixed for recent observations. CPI and PCE (headline and core) show a broad decreasing trend, although the estimates for the first lag coefficient β_{1t} seem to rebound slightly upwards, around 2010. However, the rebound typically starts from negative estimates of the first autocorrelation coefficient. Interestingly, estimates for CPI and PCE display stark U-shaped dynamics in the years from the early 90s to early 00s, as is particularly clear for core PCE. The GDP deflator stands out from the rest of the series, since it posts a decidedly downward trajectory for the last available decade. Consistently with our previous analyses, the estimates for the deflator suggest that some other factor came into play prior to international trade pressures or commodities fluctuations.

A concurrent explanation for such dynamics hinges on policy interventions. With a 10-year rolling window, though, estimates formed in the early 90s are based on a subsample starting in the early 80s, when Volcker impulses a steep turn in the inflation processes. Although appealing, this can explain only part of the dynamics in such decade. Volcker intervention brought down inflation to a moderate level, thus zeroing its structural inertia until its level was under 5%. This policy can well explain the slump and subsequent rebound, while it gives no hints on the second and steady fall in persistence toward the end of the 90s. As previously shown, commodities' volatility plays a minor role, as such dynamics is observed in both core PCE and CPI. Similarly, international trade factors might come at the right timing and act as catalysts to processes already

¹⁸The same exercise can be carried out doubling the window width to 80 quarters – or decreasing significantly the nodes so to avoid over-parametrisation.

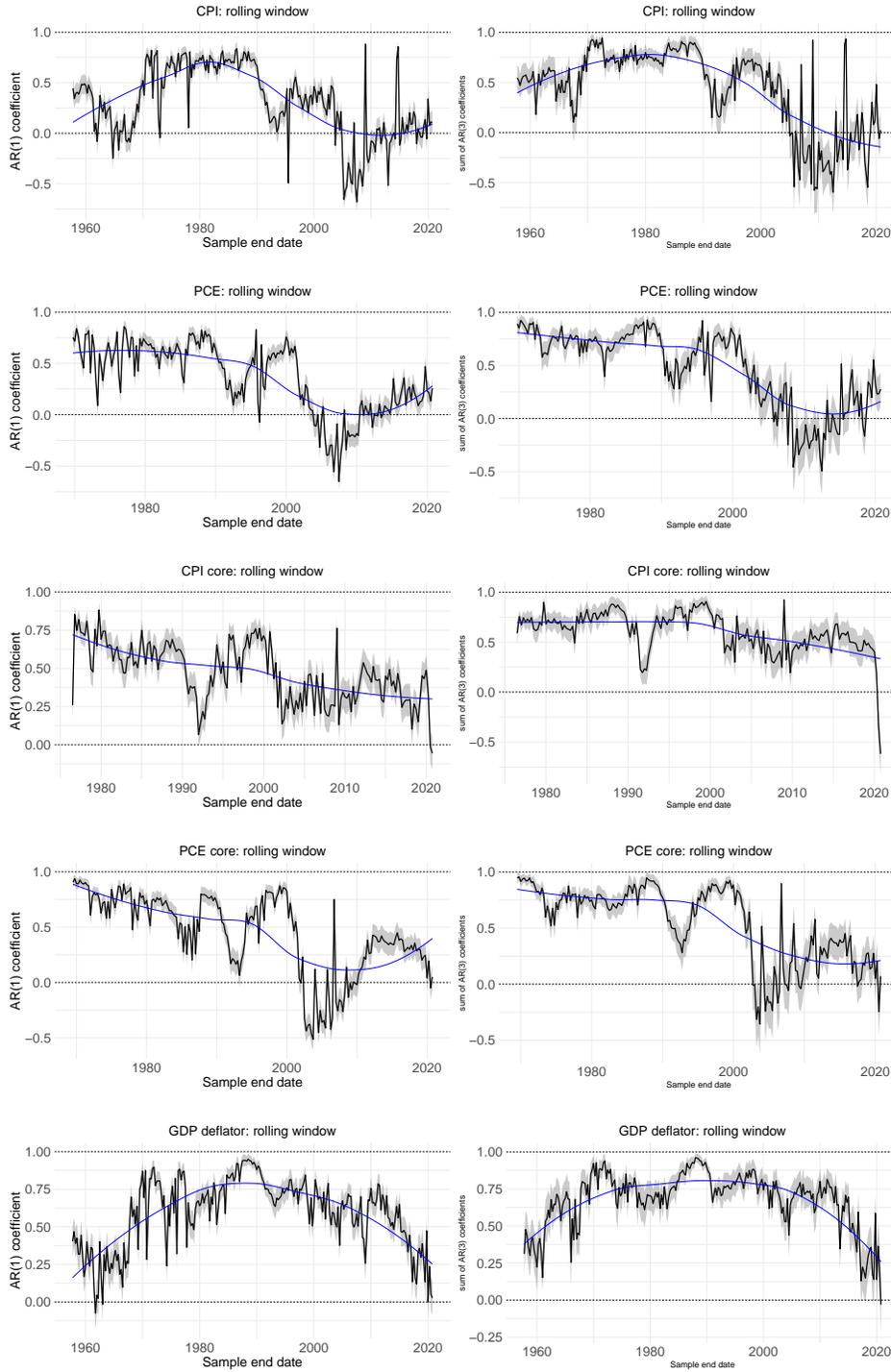
in place. International competition, especially from China, affect tradable sectors, which overlap substantially with manufacturing, thus accelerating a process of sectoral reallocation.

The overall trend, though, can be eyeballed with the help of a polynomial smoother, which points in all plots to a generalised decrease in inflation persistence. It must be noted that during the last decade there seems to be a slow, gradual uptake of inertia, compatible with the frequentist analyses presented in Sec.(5.2). A notable exception is the series for the GDP deflator, which tracks more closely the *composition* of the US economy: it shows all along the sample a concave trend in persistence, and thus substantiates the claim of structural transformation.

Looking at the right panes in Fig.(5.9), which depict the aforementioned procedure for the sum of the coefficients of an AR(3) process, we can further validate our result of decreasing inflation persistence.

A common pattern is pervasive in all estimates: since early 2000 all estimates present higher volatility in the point estimates and higher uncertainty around these. The end of the Great Moderation period, with fairly stable inflation, appears to set the inflation process on a less predictable ground, with generally large swings in its persistence and basically a quasi-white noise process at times. The root causes for such behaviour are still unclear, but the lively debate on the Phillips Curve, consumers and financiers diverging perception of inflation, and trends at the firm and macro levels do point to some candidate explanations.

Figure 4.9 – Rolling LSTM – Persistence on Data and Forecasts



Left column: plots for AR(1) autocorrelation on a 10-year rolling window augmented with optimal lags. Right column: plots for coefficients sum from an AR(3). Within each window a small LSTMs is trained (2000 epochs, 1 layer, 500 nodes, MSE loss) and then used to iteratively forecast the next 40 quarters (10 years) since last actual observation. Then an autoregressive model is estimated on this extended window, the autocorrelation is stored and plotted as black solid line in correspondence of the last actual data point date, with shaded areas reporting 95% confidence intervals around the point estimate. Blue solid line represent a LOESS polynomial fit to highlight long term trends.

4.5 Conclusion

Inflation behaviour has been widely investigated in recent years, within Phillips curves and statistical frameworks, yet no conclusive consensus has emerged. Regarding inflation dynamics, it is unclear whether inflation persistence has stabilised (Fuhrer, 2011; Pivetta and Reis, 2007) or declined (Stock and Watson, 2007). Even less established is the debate around the determinants of inflation dynamics (Mavroeidis, Plegborg-Moller, and Stock, 2014). Persistence, or equivalently inertia, is a fundamental property to consider when fiscal or monetary policies are devised and evaluated, as it encapsulates how responsive prices are to interventions.

In this work we revise and extend previous analyses of inflation persistence for the US macroeconomy. We extend the set of inflation measures to include GDP deflator, which tracks closely the US economy's structure, Consumer Price Index, and Personal Consumption Expenditure index – core and headline. This extension allows the isolation of a number of potential confounding factors: international trade effects from imported goods and services, volatility effects from energy and food items, evolving structure of the US industrial composition. After using autoregressive and Bayesian tools, we extend the methodological toolkit drawing from the deep neural networks field. We adapt Long-Short Term Memory (LSTM) recursive neural networks to leverage their predictive performance and flexible management of nonlinearities – time-varying lag structure, seasonality, short-lived cyclical fluctuations, and long term trends. This class of deep, recursive neural networks already outperforms classic forecasting tools for time series (Almosova and Andresen, 2019; Verstyuk, 2020). We train several of these nets over the full, split, and rolling samples and leverage their flexibility to extend observations and thus study US inflation persistence since WWII.

We show that inflation persistence substantially decreased since the mid-'90s. This pattern holds irrespective of the revised measure of inflation we use. The timing suggests that it is not fully explained by international trade or commodities prices: Persistence peaks around the second half of the '90s, before China's WTO accession and before the increase in energy and food volatility. This is confirmed when we look at the data in a more flexible way. We find evidence that inflation series currently behave similarly to a white noise process, showing a decreasing connection with past values. We also report on the increased statistical uncertainty associated with headline series: volatility in commodities prices further decrease the predictability of overall inflation. Conversely, the GDP deflator displays a smooth, hump-shaped decrease in persistence, suggestive of longer trends in the economy.

In light of the encouraging performance of LSTM models applied to time series econometrics, a promising avenue of research is the further extension of such tools to a wider array of methods. Furthermore, we leave for future research the investigation on the root causes for decreasing inertia.

4.A A primer on the RNN-LSTM framework

In the last section we introduce RNN-LSTM models and use them to gain a deeper understanding of inflation dynamics. This section gives a succinct presentation of these deep-learning tools, presents the optimisation algorithm employed and shortly discusses the hyperparameters tuning assumptions. These paragraphs draw from Almosova and Andresen (2019), Greff et al. (2015), Jozefowicz, Zaremba, and Sutskever (2015), Karpathy (2015), and Verstyuk (2020). For a thorough, formal presentation of Neural Networks within the statistical learning framework, refer to Hastie, Tibshirani, and Friedman (2009).

From statistical learning to long-short term memory models Artificial Neural Networks (ANNs) are systems combining nodes, layers, relationships, biases, and activation functions. By design, they mimic in their structure the human brain with interconnected neurons (nodes) and connections thereof (layers). Recurrent Neural Networks (RNN) are a subclass of ANN, especially devised to deal with *sequences*. Within the RNN class are the Long-Short Term Memory models, which address more complex sequential structures involving varying time dependency and indexed observations.

In a nutshell, each layer is populated with nodes (or neurons) that form a linear combination of the layer's input. Therefore a sequence of layers boils down to a sequence of linear combinations of the original input. Such combination is flexible enough to approximate any nonlinear function with an arbitrary degree of precision.

The general structure of a neural network with M layers and N_m , $m \in 1, \dots, M$, nodes per layer is the following, unrolling the hidden layers

$$\begin{aligned}
 \mathbf{h}_1 &= \mathbf{g}_1 (\mathbf{b}_1 + \mathbf{W}_1 \mathbf{x}) \\
 \mathbf{h}_2 &= \mathbf{g}_2 (\mathbf{b}_2 + \mathbf{W}_2 \mathbf{h}_1) \\
 \mathbf{h}_3 &= \mathbf{g}_3 (\mathbf{b}_3 + \mathbf{W}_3 \mathbf{h}_2) \\
 &\dots \\
 \mathbf{h}_M &= \mathbf{g}_M (\mathbf{b}_M + \mathbf{W}_M \mathbf{h}_{M-1}) \\
 \hat{\mathbf{y}} &= \mathbf{g}_{M+1} (\mathbf{b}_{M+1} + \mathbf{W}_{M+1} \mathbf{h}_M)
 \end{aligned} \tag{4.13}$$

where $\mathbf{x} \in \mathbb{R}^K$ is a K -dimensional vector of inputs rescaled to have $E\mathbf{x} = \mathbf{0}$ and $V\mathbf{x} = \mathbf{1}$, while $\mathbf{g}_m : \mathbb{R}^{N_m} \mapsto \mathbb{R}^{N_m}$ are activation function mapping layers output from one layer to the following downstream.¹⁹ $\mathbf{b}_m \in \mathbb{R}^{N_m}$ and $\mathbf{W}_m \in \mathbb{R}^{N_m \times (N_m - 1)}$ are the biases and weights of the m -th layer. These last objects will be the target for the optimisation and will be tuned as to minimise a loss

¹⁹For $m \neq 1$, in such case $\mathbf{g}_1 : \mathbb{R}^K \mapsto \mathbb{R}^{N_2}$.

function. Finally, $\hat{\mathbf{y}}$ is the predicted vector of the network, to be compared and evaluated against the observed one, \mathbf{y} .

We can roll up the network in a more succinct way by function composition

$$\hat{\mathbf{y}} = \mathbf{g}_M \circ \dots \circ \mathbf{g}_1(\mathbf{x}) \quad (4.14)$$

The general idea is that the input \mathbf{x} is passed sequentially through the layers as a conveyor belt and it is transformed – possibly in nonlinear ways – by the mediation of weights, biases and activation functions that are encapsulated in each layer \mathbf{g}_m to finally predict a value $\hat{\mathbf{y}}$. Tuning \mathbf{b}_m and \mathbf{W}_m will eventually improve the predictions of the network.

While the above structure describes a generic RNN, the inner workings of each \mathbf{g}_m make LSTMs apt to dealing with complex time series. In particular, each LSTM layer is composed of gates, states, memory, and output cells. In short, the m -th layer, inherits state and output from the previous one and contains

$$\begin{aligned} \mathbf{i}_t &= \sigma(\mathbf{b}_i + \mathbf{W}_i \mathbf{h}_{m-1,t} + \mathbf{U}_i \mathbf{h}_{t-1}) \\ \mathbf{f}_t &= \sigma(\mathbf{b}_f + \mathbf{W}_f \mathbf{h}_{m-1,t} + \mathbf{U}_f \mathbf{h}_{t-1}) \\ \mathbf{o}_t &= \sigma(\mathbf{b}_o + \mathbf{W}_o \mathbf{h}_{m-1,t} + \mathbf{U}_o \mathbf{h}_{t-1}) \\ \mathbf{c}_t &= \mathbf{f}_t \odot \mathbf{c}_{t-1} + \mathbf{i}_t \odot \tanh(\mathbf{b}_c + \mathbf{W}_c \mathbf{h}_{m-1,t} + \mathbf{U}_c \mathbf{h}_{t-1}) \\ \mathbf{h}_t &= \mathbf{o}_t \odot \tanh(\mathbf{c}_t) \end{aligned} \quad (4.15)$$

where \mathbf{i} is the input cell fed with the output of the previous layer $\mathbf{h}_{m-1,t}$ and past observations \mathbf{h}_{t-1} . $\sigma(\cdot)$ is a sigmoid function that squashes its inputs into the $[-1, 1]$ interval to avoid exploding behaviour. Then, \mathbf{f}_t is the forget gate, deciding what part of information retain from the past. \mathbf{o}_t is the output gate, which decides upon the final output of the cell. While these three cells produce activation vectors that signal what to retain, to forget, and to pass on, the last two cells are more involved. \mathbf{c}_t is the cell hidden state, it is updated upon the previous cell state and the new, retained information: this results from simultaneously forgetting something from the previous state: $\mathbf{f}_t \odot \mathbf{c}_{t-1}$; and updating from the current input: $\mathbf{i}_t \odot \tanh(\cdot)$. Finally, the last cell combines all of the above in the final output \mathbf{h}_t . In this formulation \odot stands for the element-wise multiplication, while, similarly to $\sigma(\cdot)$, \tanh is used to regularise values in a given space. In all of the above, $\mathbf{U}_s, \mathbf{W}_s, \mathbf{b}_s$ are weights matrices and biases relative to that particular cell.

The advantage of LSTMs over other infrastructures is to be found in the additive (instead of multiplicative) update of the hidden state \mathbf{c}_t , which prevents the issue of vanishing gradient when performing backpropagation.

Lastly, the training of these networks is performed deciding a loss function L that will be minimised by adjusting the chained weights and biases \mathbf{W}, \mathbf{b} so to obtain predictions $\hat{\mathbf{y}}$ as close as possible to the observed values \mathbf{y} . Typical loss functions are Mean Squared Error (mse) or Mean Absolute Error (mae), which are also helpful to make direct contact to the econometrics field. In the mse case, thus, the objective is

$$\min_{\mathbf{W}, \mathbf{b}} L(\mathbf{W}, \mathbf{b}; \hat{\mathbf{y}}, \mathbf{y}) = \frac{1}{T} \sum_{t=1}^T [\hat{\mathbf{y}}_t(\mathbf{W}, \mathbf{b}) - \mathbf{y}_t]^2 \quad (4.16)$$

To minimise such loss function, the full chained gradient is computed and weights and biases are adjusted accordingly, while the networks is evaluated on a variable subset of data (ie, batches of contiguous data points). The adjustment via the chained gradient from the final predicted y back to the earliest layers of the network is precisely backpropagation.

Optimiser The choice of the optimising algorithm is paramount in such framework. In light of the wealth of parameters to fine-tune in order to minimise L , the dimensionality of the parameters space easily scales up to orders of millions of dimensions. Therefore, the optimisation must efficiently explore such space and avoid local minima when possible. Nowadays, the ADAM algorithm has proven to be reliable and efficient in these terms (see Ruder (2016) for a thorough overview of several optimisation algorithms). In a nutshell, it is an adaptation of the stochastic gradient descent algorithm, where on top of the gradient directions there are stochastic perturbations. From Kingma and Ba (2014), it boils down to

$$\begin{aligned} m_t &= \beta_1 m_{t-1} + (1 - \beta_1) g_t \\ v_t &= \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \\ \hat{m}_t &= \frac{m_t}{1 - \beta_1^t} \\ \hat{v}_t &= \frac{v_t}{1 - \beta_2^t} \\ [\mathbf{W}, \mathbf{b}]_{t+1} &= [\mathbf{W}, \mathbf{b}]_t - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} \end{aligned} \quad (4.17)$$

Where g and g^2 are the mean and uncentered variance of the gradient at the current step, while m and v are their moving averages, and \hat{m}, \hat{v} are unbiased estimates of these two moments. Finally, the parameters θ are updated as in the last equation. The parameters $\beta_1, \beta_2, \epsilon, \eta$ are the decay rates, a smoothing term, and the learning rate, which essentially governs the change in the parameters.

Implementing this algorithm involves selecting a number of iterations (in ML jargon, *epochs*) and let the algorithm update the parameters for long enough to explore the minima.

Batches, early stopping, validation A handful of choices completes the setup of our exercise with LSTMs: batch size, early stopping on validation, and regularisation.

Batch size governs the subsamples that are fed to the model *at once*. While during each epoch the model is shown the whole dataset, the researcher can choose to pass smaller chunks of data in order to let the model pick up relevant patterns that are common across batches. To grasp the idea, consider seasonality in monthly data: if every May presents a spike and we only train the model on individual months, it will take a longer time to catch such seasonality than if we train it on batches 12 or 24 months at once. The tradeoff in choosing the batch size, thus, is between convergence speed and learning: if the batch size equals the number of observations, the training is faster but less refined and eventually the model is more exposed to overfitting. If the batch size is very small the training is slow and the model might miss key patterns, but out-of-sample performance might benefit.

A key factor of batches is that they are drawn at random from the whole dataset, much in the spirit of bootstrapping. This stochastic subsampling introduces a source of randomness that helps with the out-of-sample generalisation of the trained model.

Once the batch size is defined,²⁰ the researcher can either set a number of epochs and let the model train, or set a stopping rule and set an upper bound for the iterations. In this work we adopt the Early Stopping rule on validation data. When the model starts training, a subsample of the training data is kept apart and not used for learning. At the end of each epoch, the model performance is evaluated on such validation set via the loss function. The training therefore stops when the loss stops decreasing on the validation data or the number of epochs is reached. This procedure ensures a higher level of generalisation of the network and might save some computational time during training.

Lastly, at the end of each epoch, during the parameters update, we impose $L2$ regularisation on the parameters. LSTMs and RNNs in general present several thousands of parameters, and overfitting is often a real risk. To minimise such threat, we add a penalisation to parameters, in the spirit of ridge regressions. In short, the loss function L is augmented to nudge the optimisation to retain only relevant parameters:

$$\min_{\mathbf{W}, \mathbf{b}} L(\mathbf{W}, \mathbf{b}; \hat{\mathbf{y}}, \mathbf{y}) + \lambda \sum_{i=1}^P \|\mathbf{W}_i, \mathbf{b}_i\|^2 \quad (4.18)$$

where we take the square of the $l2$ norm, P is the total number of parameters, and λ governs the penalty relevance. In particular, λ can be interpreted as the penalty given to model complexity. It is typically set between 0 and .1: higher values nudge model's weights to be close (but not exactly equal) to 0. λ helps to balance the trade-off between generalisation to new observations and overfitting the training data.

²⁰It must be noted that, for technical reasons, the batch size must evenly divide training and test samples.

Data transformation To properly train the LSTMs, it is necessary to prepare the dataset with some transformations. After splitting the full sample into two parts, training and test, data must be rescaled to match zero mean and unitary standard deviation. Importantly, this is done *separately* for training and test subsamples. Once the model is trained and produces forecasts, these are still scaled to be of null mean and unitary standard deviation, hence they must be reconverted to the original data magnitude. Interestingly, LSTMs seem to produce forecasts similar to traditional IRFs, in such there is a mean-reversing force when the forecast horizon is long enough.

Taking stock Finally, after reviewing these components of the LSTMs setup, we can sum up the values for the hyperparameters used in our exercise in Tab.(5.2).

Table 4.2 – Hyperparameters

Instance	Hyperparameter	Value
Full sample, 1 layer	nodes	1000
	epochs	5000
	batch size	highest prime factor of sample size
	lags	15
	early stopping	yes
	trainable parameters	~ 4mln
Full sample, 2 layers	nodes per layer	750
	epochs	5000
	batch size	highest prime factor of sample size
	lags	15
	early stopping	yes
	trainable parameters	~ 7mln
10y subsamples, 1 layer	nodes	500
	epochs	2000
	batch size	highest prime factor of sample size
	lags	10
	early stopping	no
	trainable parameters	~ 1mln
10y rolling window, 1 layer	starting sample	optimal lags + 10 years
	nodes	500
	epochs	2000
	batch size	highest prime factor of sample size
	lags	10
	early stopping	no
	trainable parameters	~ 1mln
Common across setups		
ADAM optimiser	β_1	.9 (def)
	β_2	.999 (def)
	η	.001 (def)
	ϵ	$1e - 7$ (def)
early stopping	validation share	last 10% of batch
	tolerance	$1e - 5$
	patience	20% of epochs
forecast horizon	quarters ahead	40

4.B Year-on-year series

When computing the inflation rate, one can either choose to compute the annualised rate of change between two contiguous quarters (quarter-on-quarter, *qoq*), or compute the change from the corresponding quarter of the previous year (year-on-year, *yoy*). While in principle these methods yield broadly the same inflation rates, *yoy* series display a rather different, higher level of persistence, as found when computing the optimal lags via the BIC minimisation. Formally, the two rates result from

$$\pi_t^{qoq} = 400 \times \ln(P_t/P_{t-1}) \qquad \pi_t^{yoy} = 100 \times \ln(P_t/P_{t-4})$$

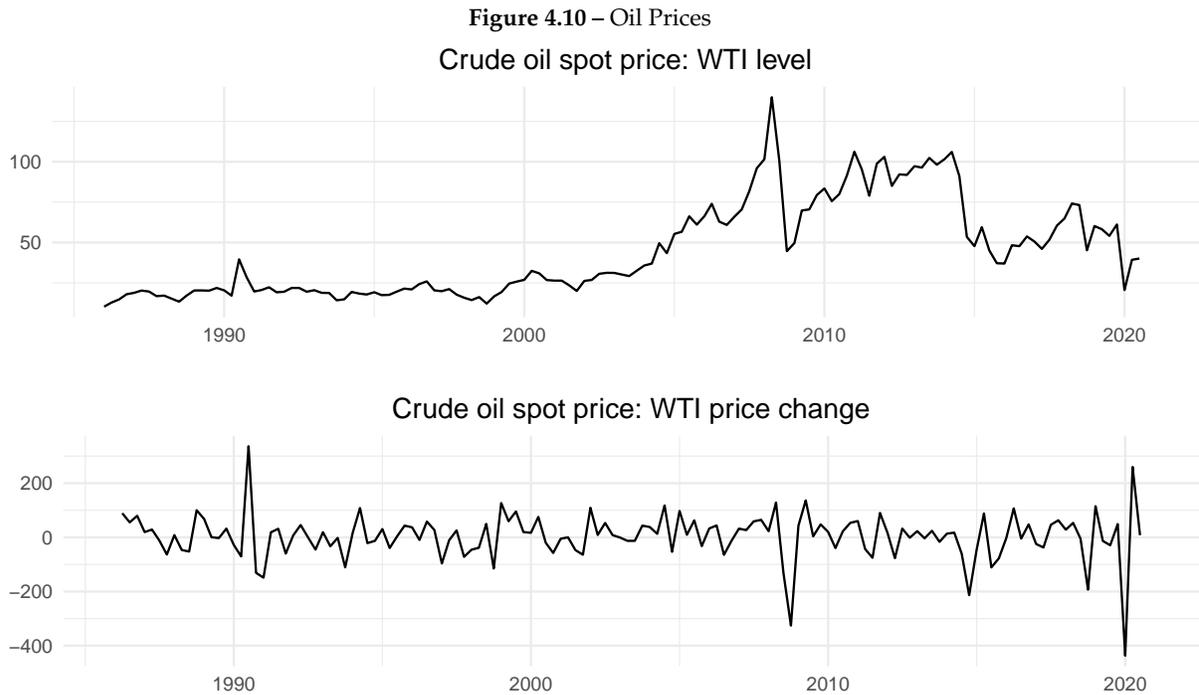
To compare the discrepancies in persistence, Tab.(5.3) reports the values for optimal lags in the two sets of series.

Table 4.3 – Optimal lags from BIC minimisation

	GDP Defl.	CPI headline	CPI core	PCE headline	PCE core
k_{qoq}^*	3	3	3	3	2
k_{yoy}^*	9	9	18	9	18

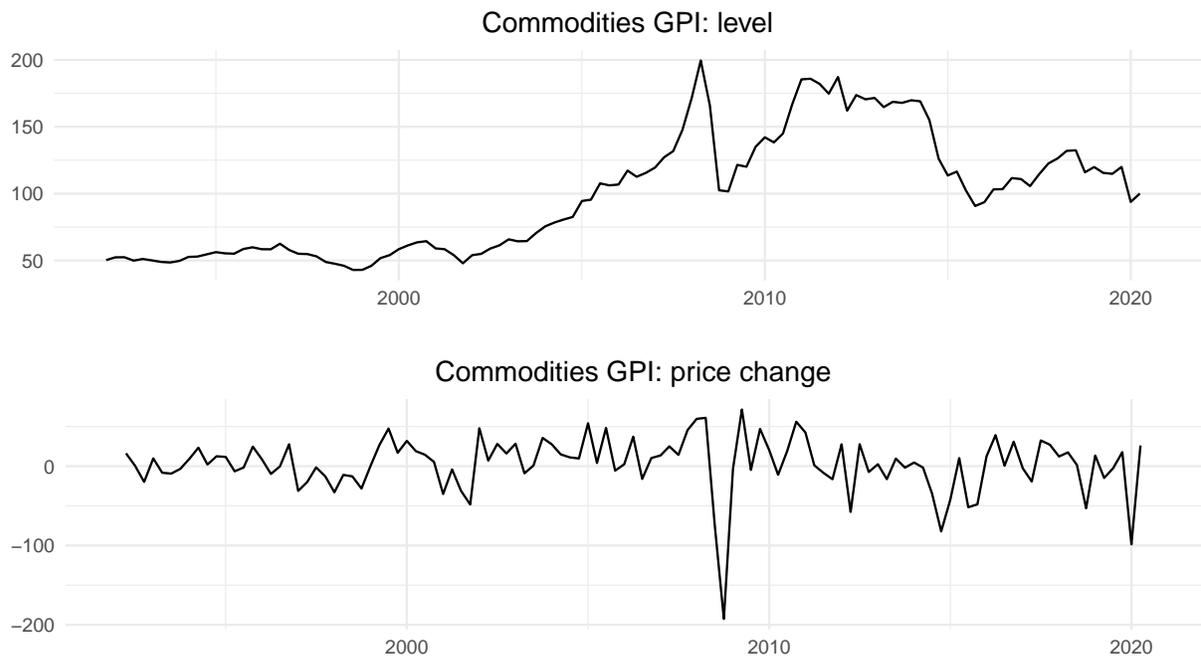
4.C Isolating oil and commodities' inflation from headline

Figs.(5.10,5.11) present data on oil and commodities prices, along with their variations. This allows the appreciation of the main differences between headline and core series for CPI and PCE, with core series excluding the items included in these plots.



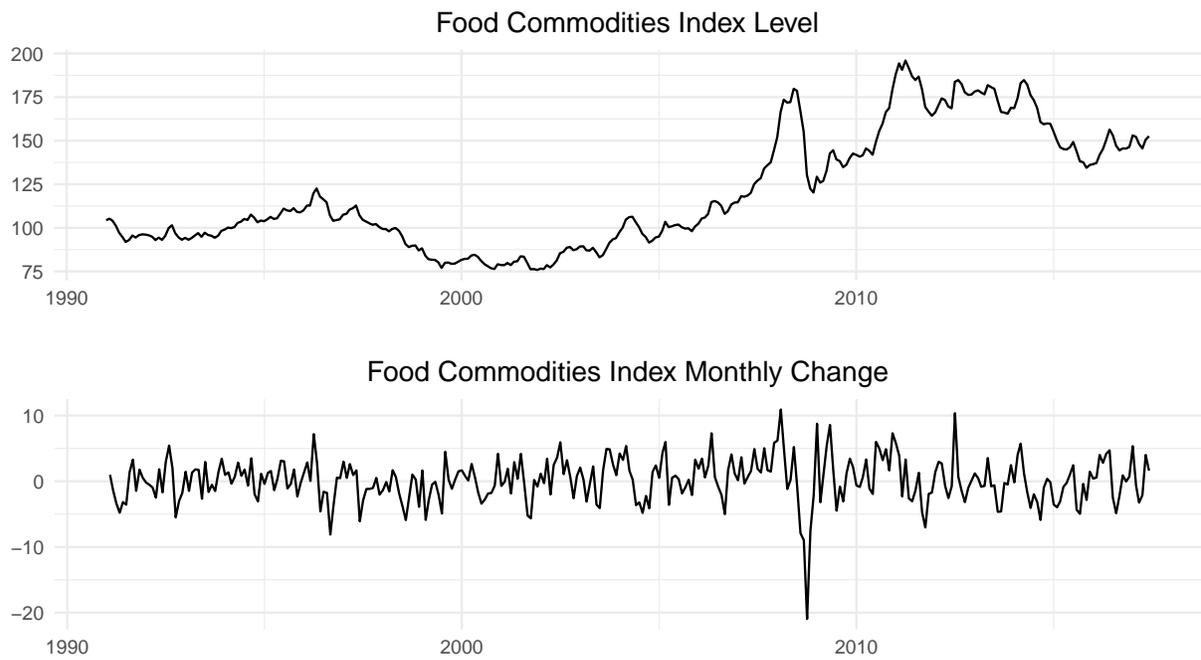
West Texas Intermediate spot price, level (top), and qoq annualized percent change (bottom). Source: FRED, St. Louis Fed.

Figure 4.11 – Commodities Prices



Global Price Index for commodities, level (top), and qoq annualized percent change (bottom). This series includes prices for oil, gas, metals, grains, among others. Source: FRED, St. Louis Fed.

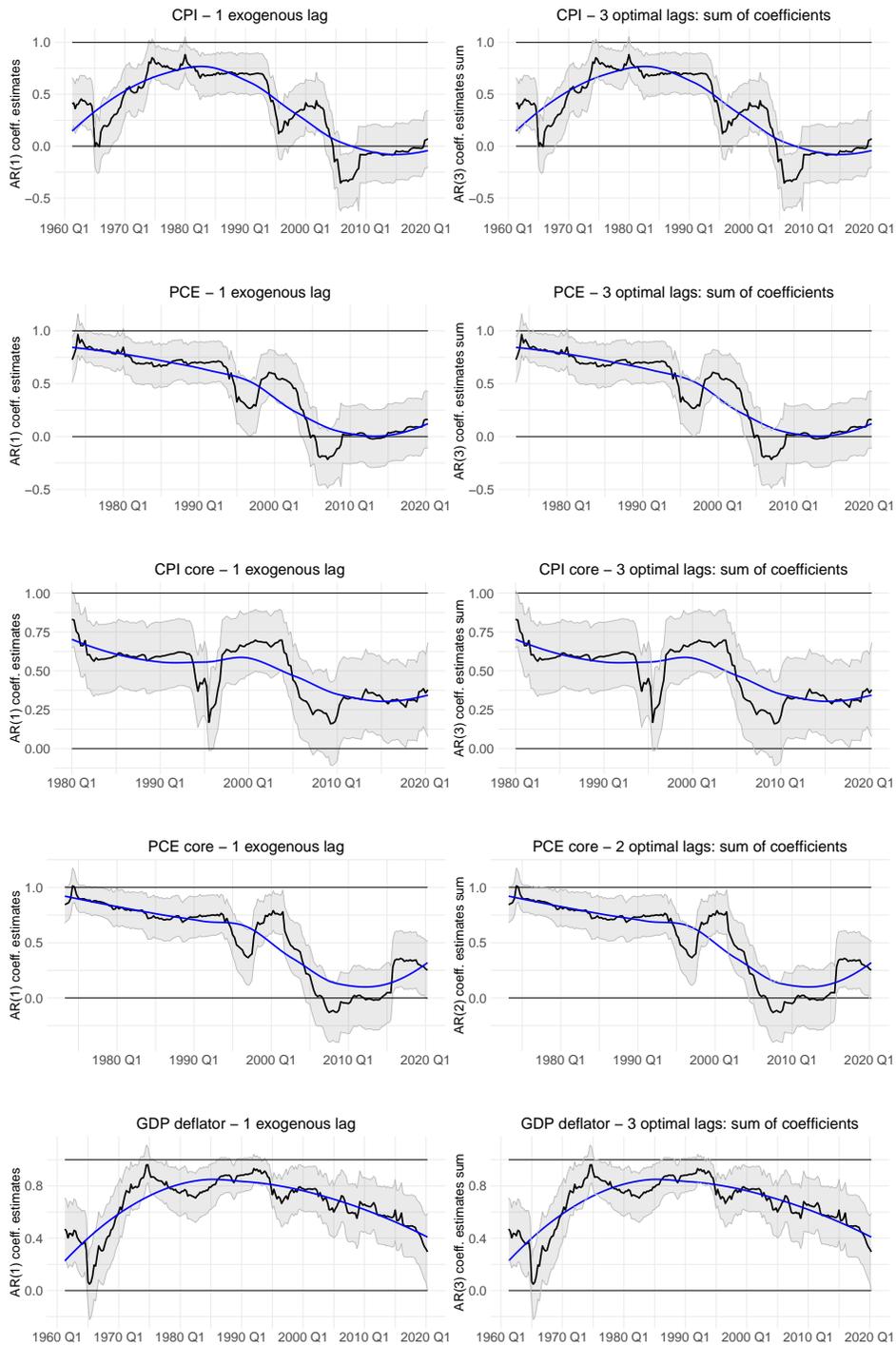
Figure 4.12 – Food Commodities Prices



World Food Commodities Index for industrial raw materials, produce, beverage; level (top), and month on month annualized percent change (bottom). Source: IMF.

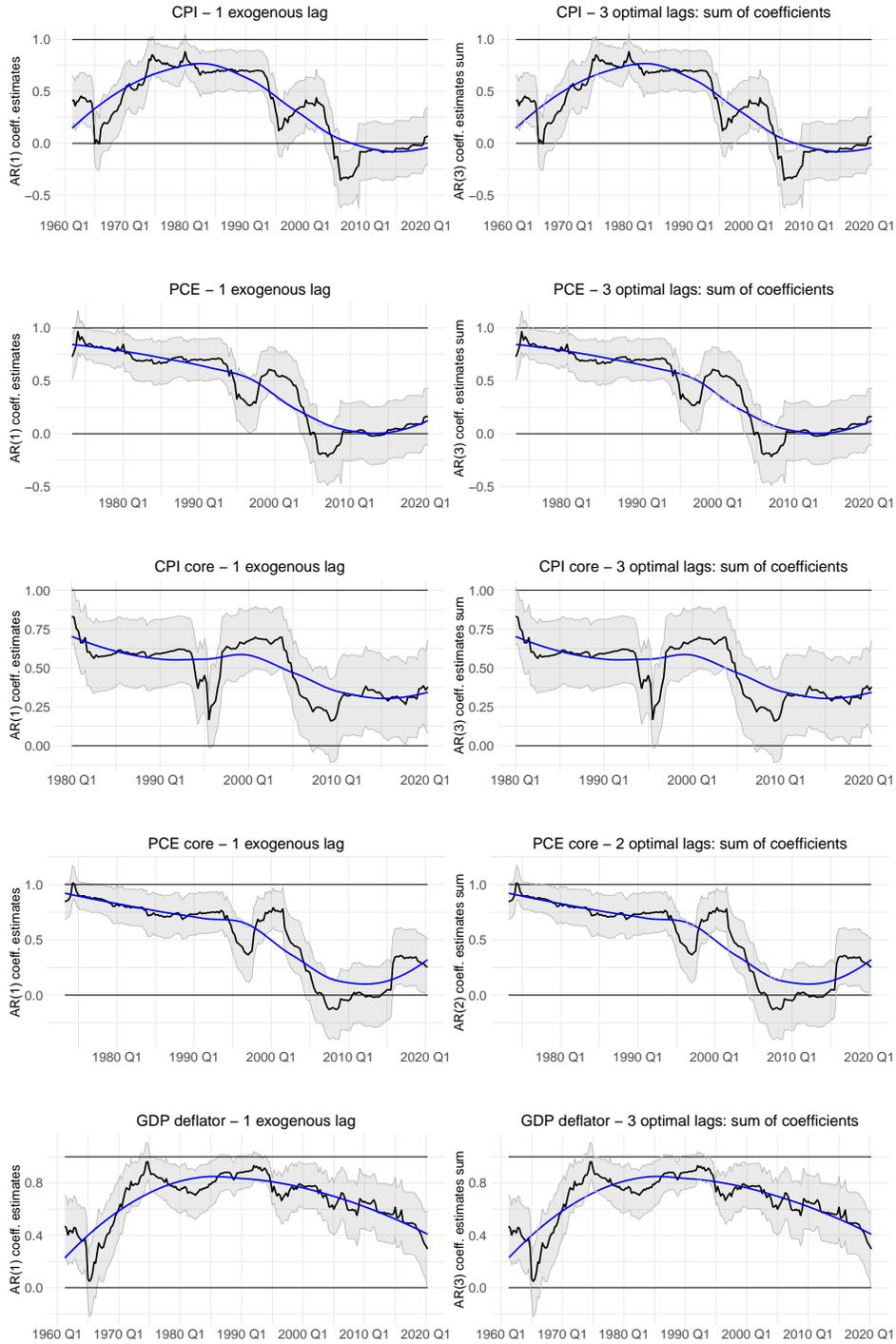
4.D Robustness with varying window width

Figure 4.13 – Frequentist Persistence – 20 Quarters Window



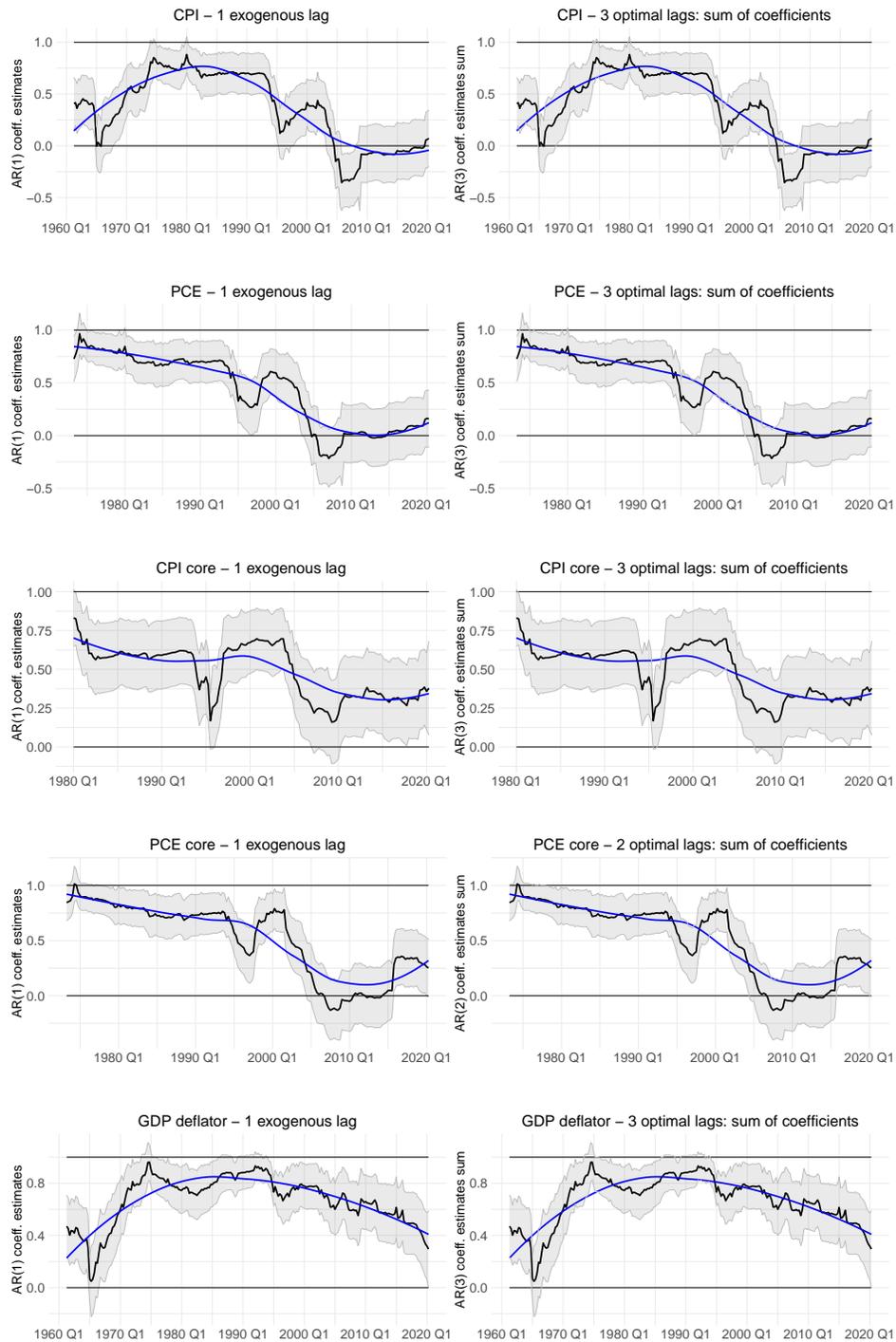
$AR(1)$ (left column) and $AR(k^*)$ (right column) estimates for CPI, PCE, GDP deflator series, rolling window of 20 quarters (5 years). Black line plots $\beta_{1,t}$ point estimates over time, red lines are 5% confidence intervals, blue line is LOESS fit to (polynomially-)smooth out point estimates, grey bands are its 5% confidence intervals.

Figure 4.14 – Frequentist Persistence – 40 Quarters Window



$AR(1)$ (left column) and $AR(k^*)$ (right column) estimates for CPI, PCE, GDP deflator series, rolling window of 40 quarters (10 years). Black line plots $\beta_{1,t}$ point estimates over time, red lines are 5% confidence intervals, blue line is LOESS fit to (polynomially-)smooth out point estimates, grey bands are its 5% confidence intervals.

Figure 4.15 – Frequentist Persistence – 80 Quarters Window

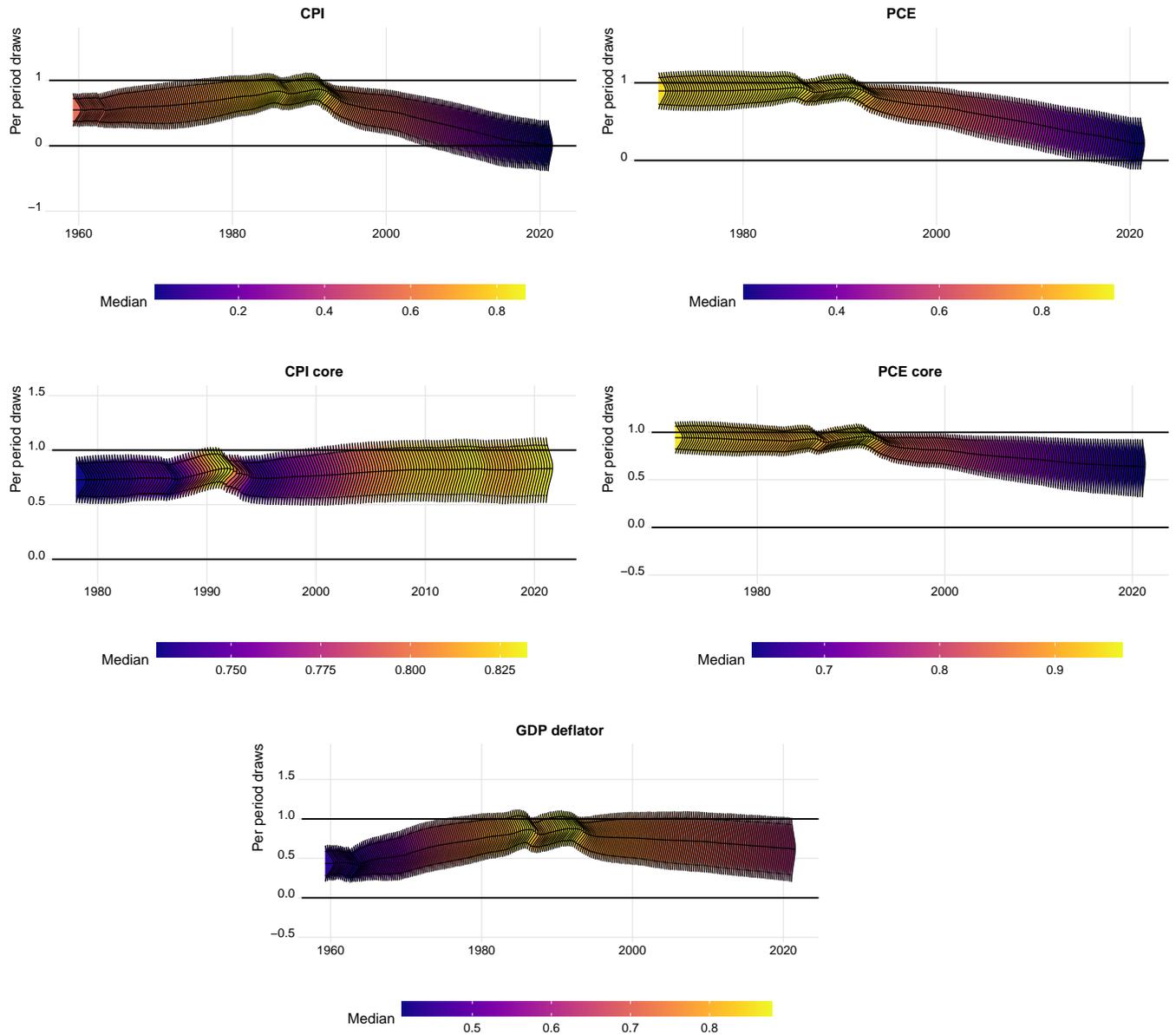


$AR(1)$ (left column) and $AR(k^*)$ (right column) estimates for CPI, PCE, GDP deflator series, rolling window of 80 quarters (20 years). Black line plots $\beta_{1,t}$ point estimates over time, red lines are 5% confidence intervals, blue line is LOESS fit to (polynomially-)smooth out point estimates, grey bands are its 5% confidence intervals.

4.E Draw distributions

These plots provide further insights on the densities of persistence produced by the Bayesian analysis presented in Sec.(5.3).

Figure 4.16 – Bayesian Draws – Per Period Full Densities



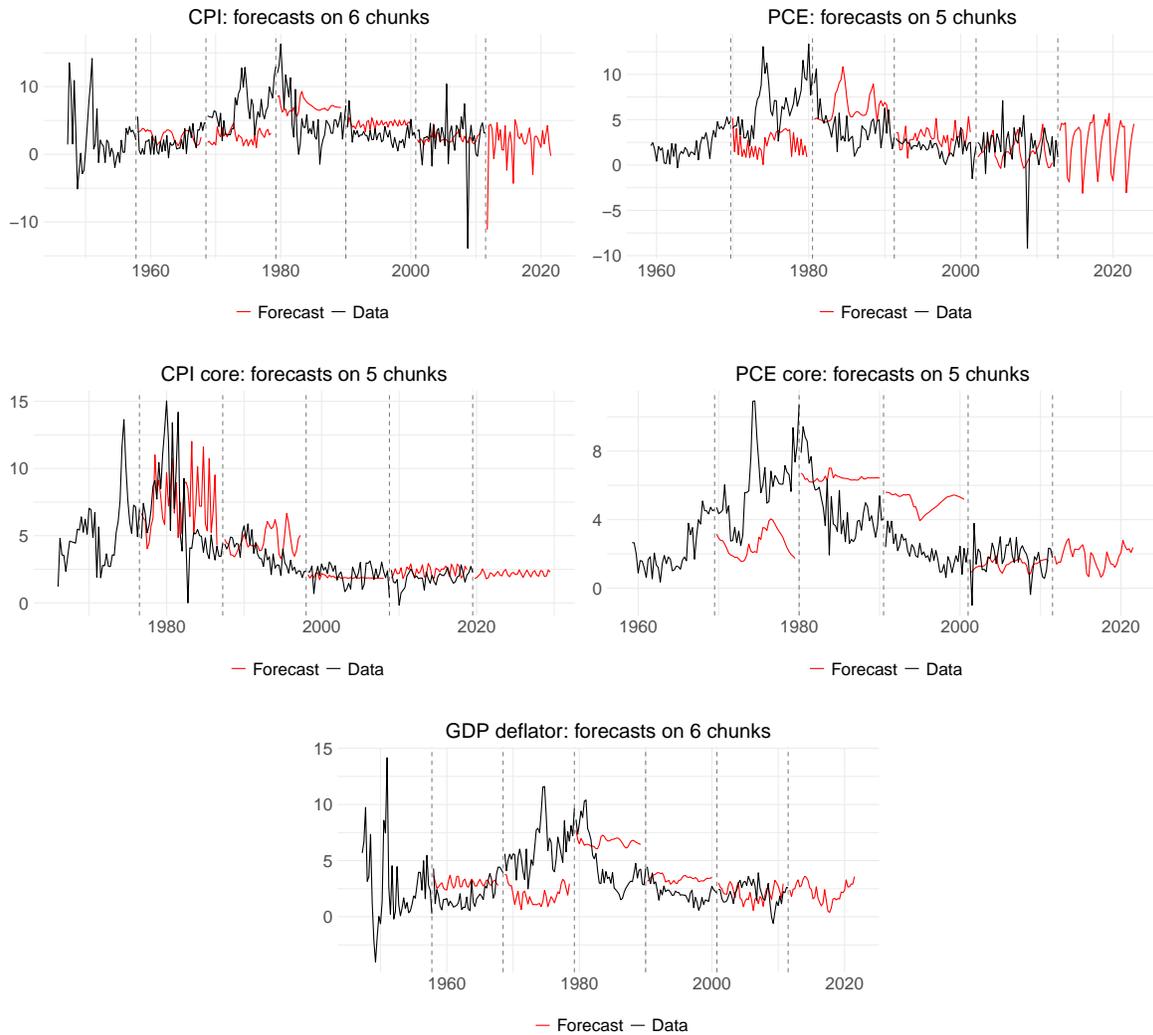
Per-period estimated densities: sum of $AR(3)$ draws distributions per quarter. Persistence computed on simulated forward trajectories based on data up to t . 300k total iterations, 150k burn-in, resulting in 150k conserved draws per period. Each density also reports 5%, 50%, and 95% percentiles. Colour depends on median value for each time t density.

4.F LSTM data and forecasts

This section collects plots for the forecasts of all LSTMs mentioned in the main body of the paper.

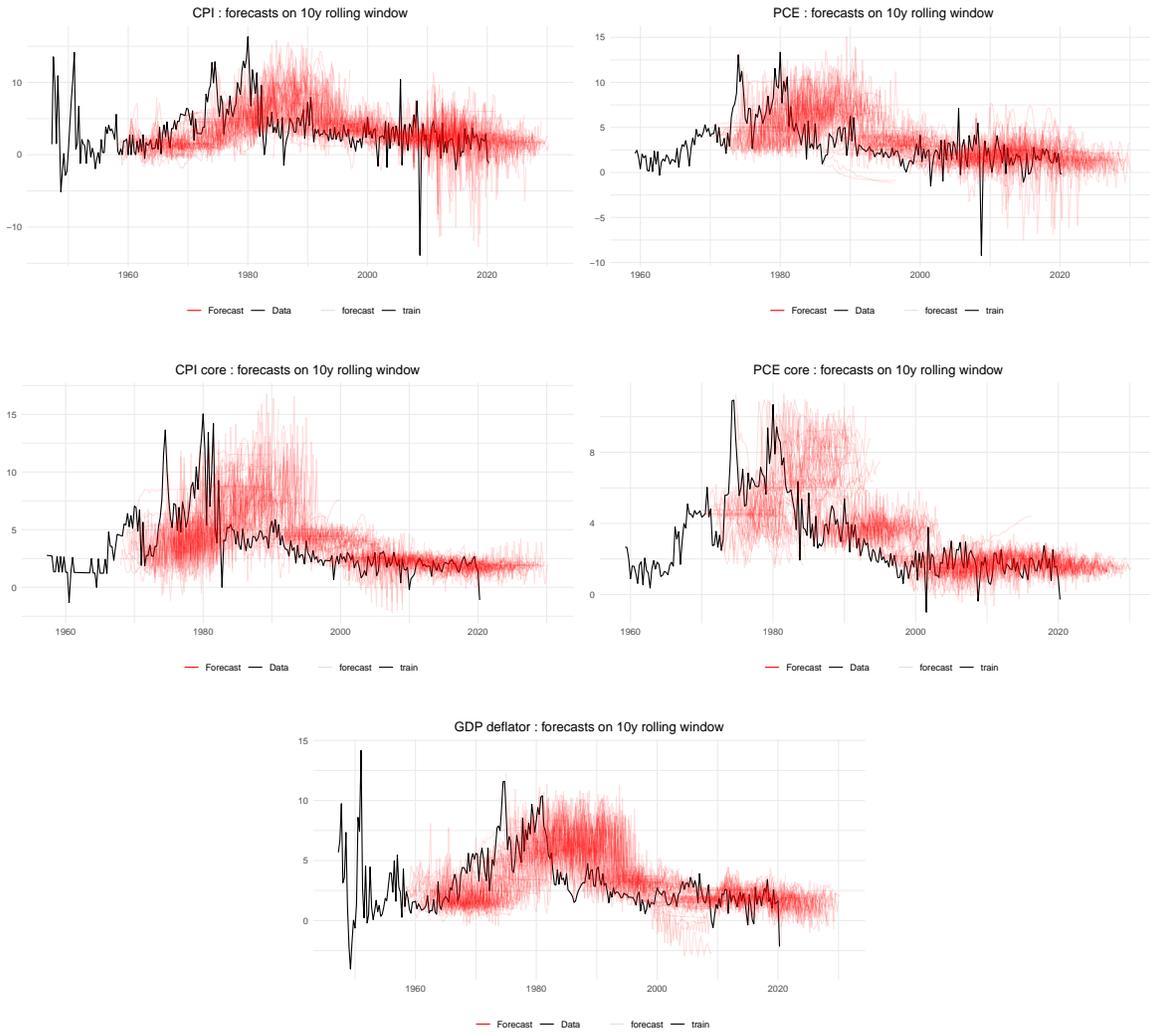
4.F.1 LSTM predictions on non-overlapping subsamples

Figure 4.17 – LSTM Forecasts on Decade Subsamples



Indirect forecasts from an LSTM trained on 10 years of data, plus appropriate lags defined by BIC minimisation. One layer, 500 nodes, early stopping criterion with 2000 epochs upperbound. Forecast horizon is $h = 40$: first prediction is produced from last available data point and then iterated forward. Dashed vertical lines mark data subsamples' end date.

Figure 4.18 – LSTM Rolling Window Forecasts



Indirect forecasts from a several LSTMs trained on a 10-year rolling window, plus appropriate lags defined by BIC minimisation. One layer, 500 nodes, early stopping criterion with 2000 epochs upperbound. Forecast horizon is $h = 40$: first prediction is produced from last available data point and then iterated forward.

4.G LSTM analyses

4.G.1 OLS regressions on decades

The following tables present detailed information on the OLS regressions that produce the results plotted in subsection 5.4.4. Each table's column represents a subsample of actual data points augmented with the forecast produced by the trained LSTM. Hence, the number of observations results from $40 + k$ data points and 40 forecasts, where k is the optimal number of lags for each series. The left-most column is the oldest subsample, right-most is the closest in time.

Table 4.4 – CPI decades regressions with LSTM forecasts

	<i>Dependent variable:</i>					
	CPI					
	1947Q2 1957Q4	1958Q1 1968Q3	1968Q4 1979Q2	1979Q3 1990Q1	1990Q2 2000Q4	2001Q1 2011Q3
1 st lag	.438*** (.100)	.277*** (.103)	.703*** (.079)	.617*** (.082)	.345*** (.104)	-.108 (.111)
Constant	1.452*** (.403)	1.604*** (.272)	2.050*** (.568)	1.802*** (.465)	1.705*** (.293)	2.313*** (.423)
Observations	82	82	82	82	82	82
R ²	.192	.083	.496	.414	.120	.012
Adjusted R ²	.182	.071	.490	.407	.109	-.001
F Statistic (df = 1; 80)	19.017***	7.215***	78.882***	56.537***	10.914***	.932

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 4.5 – PCE decades regressions with LSTM forecasts

	<i>Dependent variable:</i>				
	PCE				
	1959Q2 1969Q4	1970Q1 1980Q3	1980Q4 1991Q2	1991Q3 2002Q1	2002Q2 2012Q4
1 st lag	.777*** (.071)	.541*** (.093)	.410*** (.093)	.300*** (.106)	.0004 (.112)
Constant	.526*** (.194)	3.267*** (.685)	2.196*** (.382)	1.174*** (.205)	2.118*** (.321)
Observations	82	82	82	82	82
R ²	.602	.299	.197	.091	0.00000
Adjusted R ²	.597	.290	.187	.080	-.012
F Statistic (df = 1; 80)	121.230***	34.052***	19.637***	8.005***	.00001

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 4.6 – GDP Deflator decades regressions with LSTM forecasts

	<i>Dependent variable:</i>					
	GDP deflator					
	1947Q2 1957Q4	1958Q1 1968Q3	1968Q4 1979Q2	1979Q3 1990Q1	1990Q2 2000Q4	2001Q1 2011Q3
1 st lag	.467*** (.098)	.653*** (.081)	.663*** (.083)	.897*** (.039)	.582*** (.082)	.723*** (.080)
Constant	1.466*** (.364)	.659*** (.177)	2.171*** (.550)	.352** (.171)	.854*** (.182)	.605*** (.184)
Observations	82	82	82	82	82	82
R ²	.222	.451	.442	.870	.383	.507
Adjusted R ²	.212	.444	.435	.869	.376	.501
F Statistic (df = 1; 80)	22.777***	65.679***	63.292***	537.595***	49.747***	82.417***

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 4.7 – Core CPI decades regressions with LSTM forecasts

	<i>Dependent variable:</i>				
	CPI core				
	1966Q1 1976Q3	1976Q4 1987Q2	1987Q3 1998Q1	1998Q2 2008Q4	2009Q1 2019Q3
1 st lag	.816*** (.064)	.656*** (.084)	.828*** (.061)	.200* (.110)	.501*** (.097)
Constant	1.463** (.635)	1.867*** (.513)	.442** (.181)	1.773*** (.252)	.981*** (.196)
Observations	82	82	82	82	82
R ²	.672	.430	.698	.040	.249
Adjusted R ²	.668	.423	.694	.028	.239
F Statistic (df = 1; 80)	163.902***	60.265***	184.684***	3.330*	26.504***

Note: *p<0.1; **p<0.05; ***p<0.01

Table 4.8 – Core PCE decades regressions with LSTM forecasts

	<i>Dependent variable:</i>				
	PCE core				
	1959Q2 1969Q3	1969Q4 1980Q1	1980Q2 1990Q3	1990Q4 2001Q1	2001Q2 2011Q3
1 st lag	.876*** (.055)	.733*** (.076)	.684*** (.076)	.377*** (.104)	.368*** (.105)
Constant	.368** (.169)	1.741*** (.500)	1.323*** (.345)	.976*** (.188)	1.185*** (.214)
Observations	81	81	81	81	81
R ²	.765	.544	.509	.143	.133
Adjusted R ²	.762	.538	.503	.133	.122
F Statistic (df = 1; 79)	257.018***	94.107***	81.810***	13.226***	12.163***

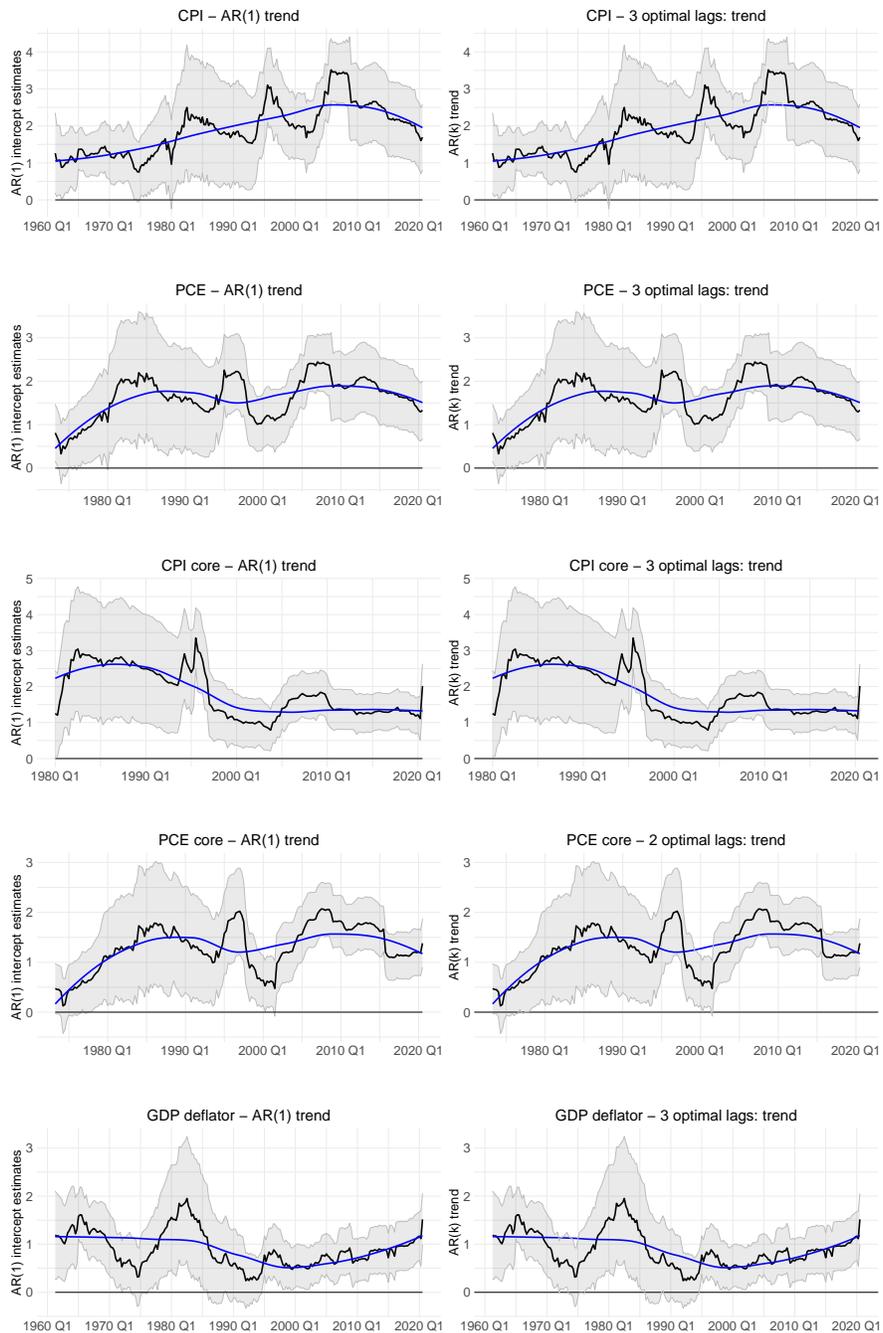
Note: *p<0.1; **p<0.05; ***p<0.01

4.H Trend estimates

This section collects estimates on trend inflation that results from two approaches explored in the body of the chapter. For both the frequentist and the LSTM applications, a side product of estimating autoregressive models of varying order is the intercept.

4.H.1 Frequentist application

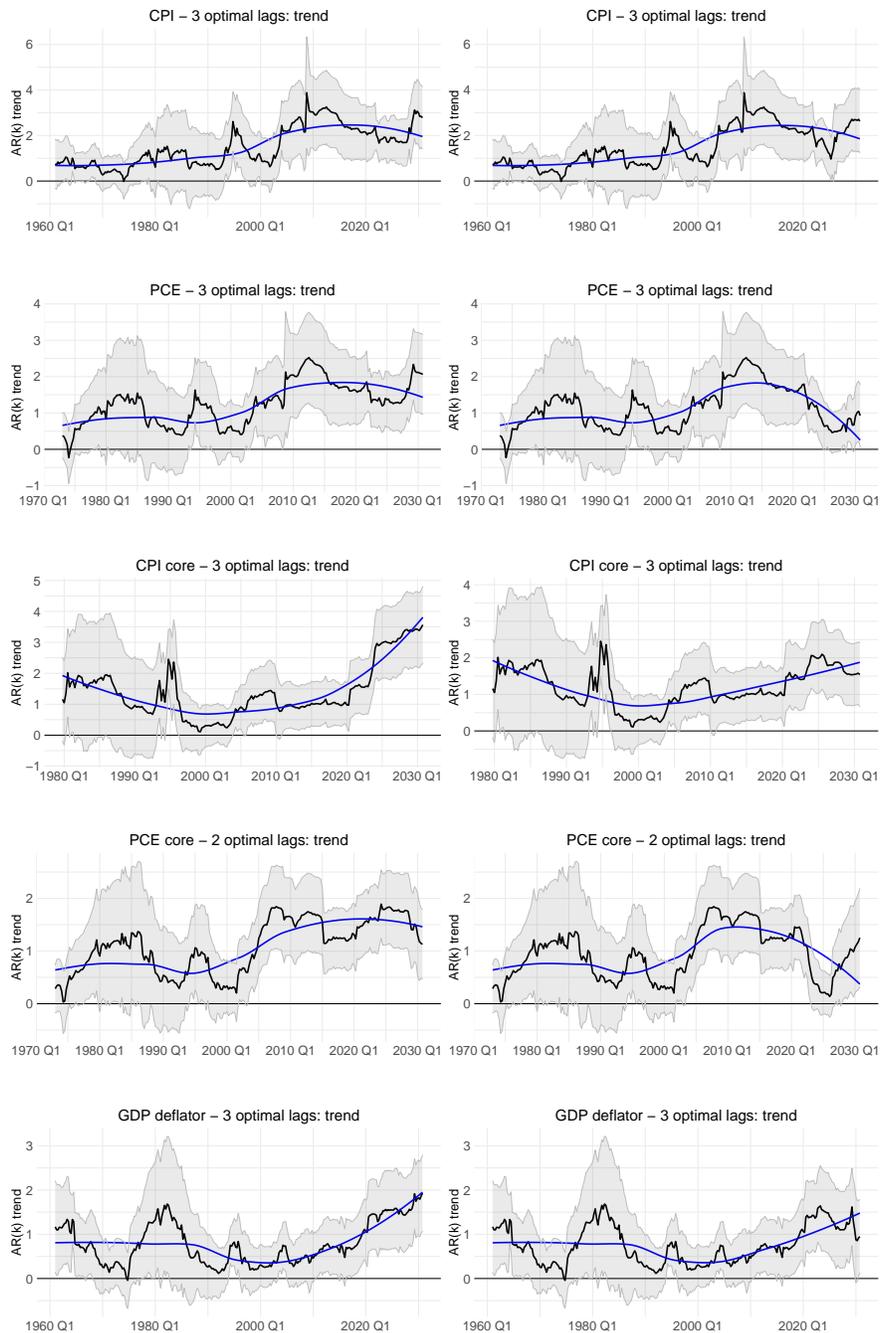
Figure 4.19 – Trend Inflation – Frequentist Measure



Frequentist estimates of trend inflation, computed as intercept of an autoregressive process. Left column: AR(1). Right column: $AR(k^*)$ with lags selected by BIC minimisation.

4.H.2 LSTM output: full sample

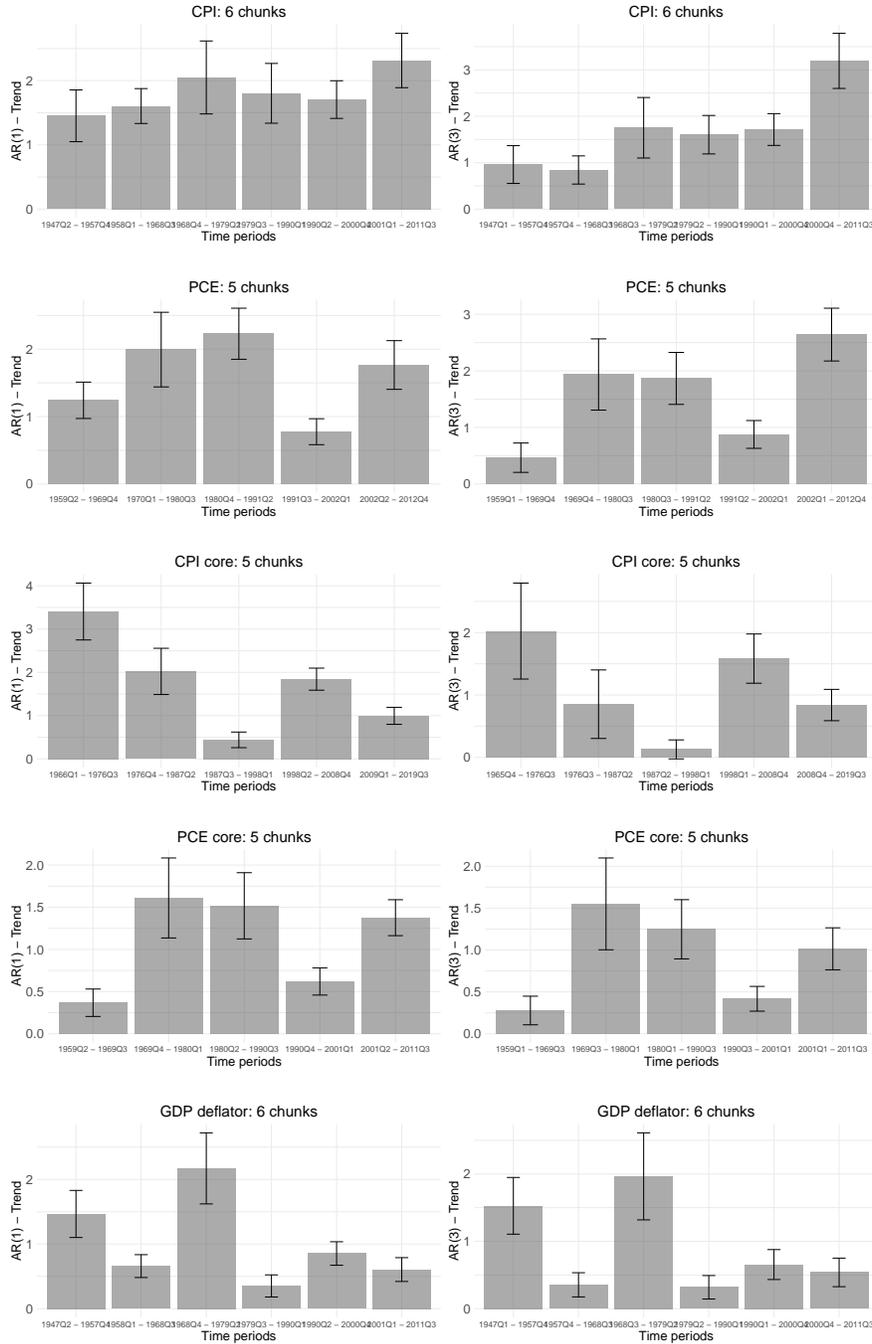
Figure 4.20 – Trend Inflation – LSTM on Full Sample



LSTM trained on the full sample of data and then iterated forward to forecast 40 data points. Left column: one-layer net. Right column: two-layer net. All estimates are from an $AR(k^*)$ with lags minimising BIC.

4.H.3 LSTM output: subsample analysis

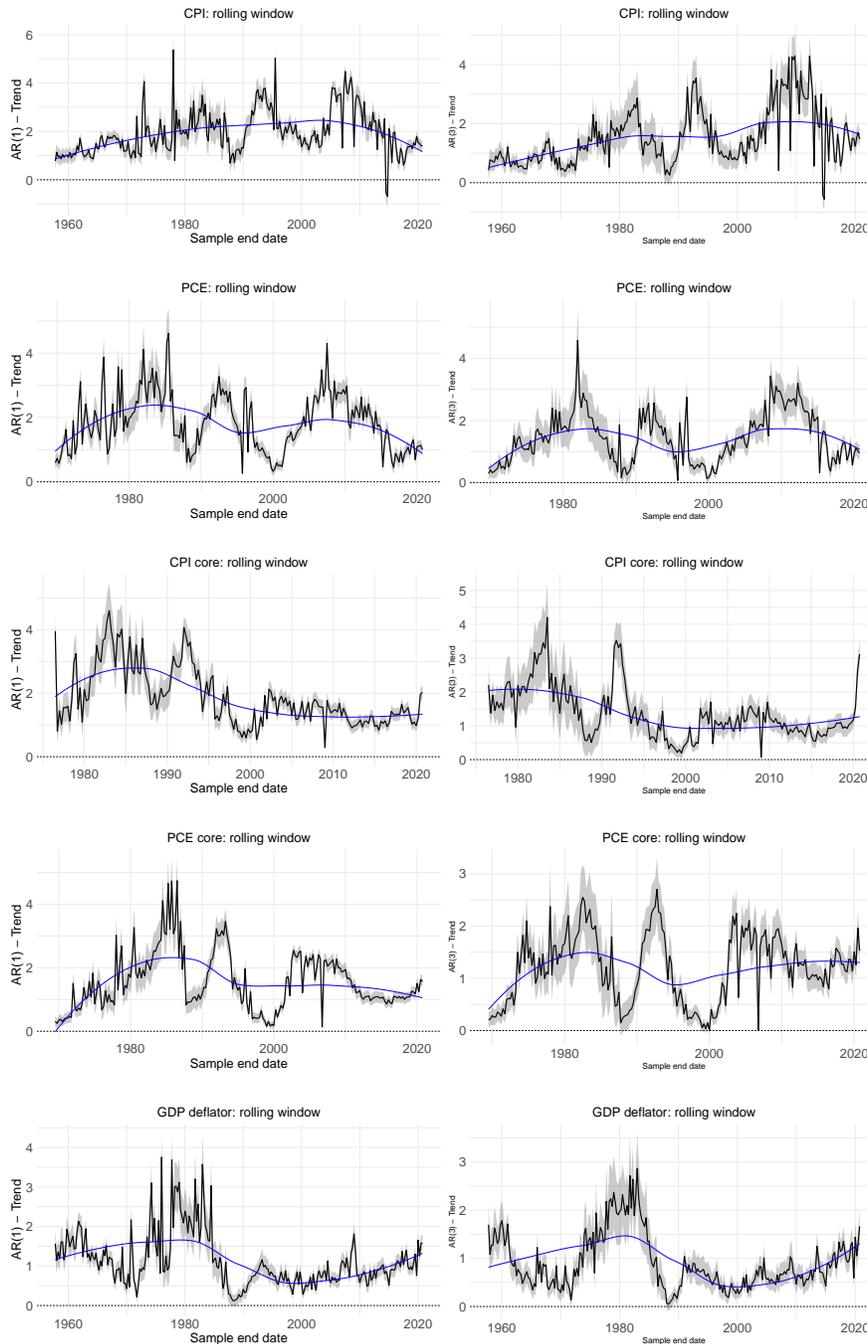
Figure 4.21 – Trend Inflation – LSTM on Decade Subsamples



LSTMs trained on ten-year subsamples, plus appropriate lags, then iterated forward to produce 40 data points. Left column: $AR(1)$. Right column: $AR(k^*)$ with lags minimising BIC.

4.H.4 LSTM output: rolling window

Figure 4.22 – Trend Inflation – LSTM on Rolling Windows



LSTMs trained on 56-quarter rolling windows, plus appropriate lags, then iterated forward to produce 40 data points. Left column: $AR(1)$. Right column: $AR(k^*)$ with lags minimising BIC.

Chapter 5

Conclusion

This dissertation is a snapshot of an ongoing research work, taken at the time of the writing. Research work is chiefly *in fieri*, with multiple spin-offs, new ideas and projects coming up, old ones that become appealing once again. Thus, this chapter is intended to summarise the key points of the dissertation in relation to future works.

5.1 Policy Rules and Liquidity

Chapter 3 combines aggregate, quarterly real-time data to study how the US central bank pursued its mandate of price stability and full employment. Departing from traditional specifications of the interest rate rules, we verify whether the Federal Reserve Bank is concerned with liquidity in the financial dimension of the macroeconomy. The robustness of a standard, model-derived Taylor rule is weaker than previously found. Estimating a two-state Markov switching model of the baseline policy rule we find multiple policy regimes: the timing, though, contrasts with the established narrative of an accommodating stance that triggered high inflation during the 1970s. When assuming a standard rule, it turns out that the Fed complied with the Taylor principle, at least from the 1960s onwards, but with varying intensity. This finding challenges the hypothesis of policy-induced sunspot equilibria for the high inflation episodes, as well as the stringency of the Taylor principle.

Yet, augmenting the interest rate rule with financial liquidity proxies provides a different scenario. To account for risk and maturity factors we approximate liquidity with spreads of the average yields of BAA bonds and quarterly S&P500 returns, over US Treasury Bills of comparable duration. Liquidity proxies turn out to be strong and significant predictors of movements in the policy interest rate: as liquidity in the economy decreases, the Fed swiftly eases its policy rate. This relation holds consistently across the whole sample, exogenous subsamples, and endogenising the change in regimes via a Markov switching model. The role of liquidity, moreover, interacts with the weight of inflation expectations, decreasing their coefficient sizeably.

All in all, the Taylor principle turns out significantly undermined, while liquidity plays a relevant role in the policy stance of the Fed. The analysis, though, lends itself to several improvements. First off, to clean the results from residual endogeneity and complement the discrete regime switches approach, it is possible to apply the methods proposed by Boivin (2006) and Kim and Nelson (2006). Time-varying parameters estimates would not provide precise timing but are informative of long tendencies. Accounting for potential heteroskedasticity also sheds light on the varying nature of shocks that hit the economy: this allows for a more informative comparison between periods and a thorough evaluation of policy effectiveness.

On the data side, the next improvement sits in more granular information on the financial market. Namely, bond-level data are essential to further remove risk and term premia from the measures of liquidity. Ideally, transaction-level data would allow for an almost-perfect assessment of liquidity variations over time, but such data are more often than not expensive and cover short periods of time. At-issuance data, on the other hand, represent a fair middle ground. Bloomberg offers such information: it informs on the aggregate and idiosyncratic perceived risk with cross-sectional data for amount, rating, expected yield, among other factors. Exploiting these cross-sections over time is a promising exercise, which will shed a brighter light on the effects of liquidity for monetary policy.

5.2 Optimal Liquidity Management

Chapter 4 extends the baseline NKDSGE to account for liquid assets and sketches a central bank that, potentially, controls the handle of aggregate nominal liquidity. It also sets interest rates according to expected inflation only, similarly to the ECB's mandate of price stability. The interest rate rule that summarises monetary policy targets the yield on liquid assets. The monetary policy stance is relaxed and does not need to comply with the Taylor Principle to ensure the determinacy of the model. While it matches the dynamics of the baseline NKDSGE model under active monetary policy, the model sheds light on the effects of a passive policy stance. When the central bank departs from the Taylor Principle and reacts less than one-to-one to expected inflation there is virtually no evidence of degenerate dynamics, nor sunspot equilibria. Indeed, aggregate dynamics, especially inflation, display increased persistence for weak stances, alongside increased volatility. Recessions are more severe but short-lived under active regimes.

When simulating a sudden real liquidity crisis similar to that of 2008, though, the policy stance entails different speeds of recovery. Under an active stance, the impact on output is limited and recovery is relatively speedy. The policy rate stands below its equilibrium level for a longer period, similar to real liquidity, and inflation responds mildly. With a passive central bank, the impact on output and inflation is an order of magnitude greater, although all aggregates converge back to their steady-state levels faster than in the active regime.

The model with liquidity builds on the kernel of NKDSGE but enriches it with a wider set of assets. The reduced-form approach, though, calls for sound microfoundation, akin to Geromichalos and Herrenbrueck (2021). Moreover, liquid bonds are assumed to be in zero net supply, and freely tradable with cash. Besides the microfoundation, the model with liquidity naturally lends itself to several extensions, in close contact with recent developments in actual policymaking. Liquid bonds, whose yield is targeted by the central bank, are natural candidates to add a fiscal authority, for instance. Similarly, in a separate and complementary way, they would represent a full-fledged banking sector, which would amplify the role of liquidity in such a setting.

The most relevant extension, though, sits in relaxing the assumption of a central bank that fixes once and for all the total nominal liquidity in the economy. Since 2008 major central banks actively expanded by orders of magnitude their balance sheets, *de facto* flooding the economy with central bank reserves to secure trading and revive commercial credit supply. It is therefore natural to think of a central bank that operates with two instruments: the interest rate rule and the total amount of nominal liquidity, as it happens with unconventional policy tooling. Then optimal rules for both instruments can be derived in usual frameworks and analysed against other rules, measuring the welfare effects of each.

By virtue of its NKDSGE core, the model with liquidity can include existing blocks from established extensions: capital accumulation, sticky wages, financial accelerator are but a few that would deepen the analysis of the role of liquidity and its consequences for policy.

5.3 Inflation Persistence

Chapter 5 investigates the dynamics of US inflation over time. It applies traditional and novel methods to assess variations behind the inflation data generating process. From a purely statistical, univariate point of view, it finds that inflation persistence was relatively high until about 1995 and then declined steadily until the present, when inflation behaves closely to a white noise process. The result emerges across CPI, PCE (both headline and core), and GDP deflator measures of inflation; it also holds across frequentist, Bayesian, and deep-learning econometric methods.

The use of several and complementary measures of inflation allows for the identification of suggestive sources for such downward trend. First off, comparing headline and core inflation indexes is informative on the influence of commodities, energy, and food prices. These items are particularly volatile, especially since the 2000s, which could influence the underlying measure of persistence. Both headline and core series do display downward trends for inflation persistence, which allow excluding commodities as the source for the decreased inertia. Secondly, CPI and PCE indexes cover different baskets of goods to capture price level changes. Namely, the CPI basket emerges from surveys on consumers' purchases, whilst the PCE basket surveys business sales. Moreover, they also differ in the weighting and updating scheme assigned to each basket

item. These differences are even more relevant since the CPI is the base for updating a number of social security benefits, while the PCE is the policy reference for monetary policy. Thus, despite often moving in locksteps, they inform separately fiscal and monetary policy: this is useful to separate the effects of changes in monetary policy in particular, although it eventually affects all prices. Finally, the GDP deflator is computed on the output of the US economy, by local agents. It is, therefore, an index devoid of imported inflation and less prone to globalisation forces, at least for final goods and services. Input prices and their pass-through, though, are a source of potential bias.

Several improvements can enrich and reinforce the findings of the chapter, for example adopting a non-parametric measure of persistence such as the half-life of a shock to its process. Moreover, the chapter is intimately connected to the ongoing debate on the Phillips curve: taking a statistical stand, it leaves many causal questions unsatisfied.

From this work spring two lines of research, outlined separately.

5.3.1 Macroeconometrics with Deep-Learning

The first research avenue covers the machine learning toolkit that powers the last part of chapter 5. The chapter provides a first example of DL application to more genuinely macroeconomic tasks, besides the typical forecasting ones.

Currently, the blackbox nature of machine learning and artificial intelligence methods has hampered the adoption in the economic profession: it is hard to interpret a trained neural network in terms of causal inference, despite its remarkable performances in classification and forecasting.

Notwithstanding, over the past decade artificial intelligence has become a debated topic both in academic circles and in the general public. Its application is increasing and the effectiveness is promising, growing hand in hand with data availability, computational power, and algorithmic improvement. Recent years have seen an increased penetration of machine learning in economic research, with interesting results in forecasting (Jung, Patnam, and Ter-Martirosyan, 2018; Kock and Teräsvirta, 2016; Makridakis, Spiliotis, and Assimakopoulos, 2018; McAdam and McNelis, 2005; Medeiros et al., 2019; Nakamura, 2005), macroeconomics (Almosova and Andresen, 2019; Bajari et al., 2015; Chakraborty and Joseph, 2017; Giannone, Lenza, and Primiceri, 2018; Goulet Coulombe et al., 2019; Korobilis, 2018; Verstyuk, 2020), econometrics (Athey, 2018; Athey and Imbens, 2015, 2019; Mullainathan and Spiess, 2017; Varian, 2014), asset pricing (Gu, Kelly, and Xiu, 2020), and computational economics (Fernandez-Villaverde and Guerron-Quintana, 2020; Maliar, Maliar, and Winant, 2019).

Specifically, the project aims at showing that Long-Short Term Memory models can be cast as a generalisation of Vector Autoregressive models in analysing macroeconomic time series. LSTMs are currently used mainly for text analysis and natural language processing and were specifically devised to flexibly handle time-varying dependencies and highly non-linear structures, typical

of text and speech. For example, they power auto-correction and completion in smartphones, language translation, digital assistants, and image labelling.

Recent developments in VAR literature extended to non-linear structures such as regime-switching, thresholds, sign and inequality restrictions, time-varying parameters. These improvements, though, require the imposition of restrictions on the data or the model and careful a priori judgement of the researcher, limiting somehow the applicability and interoperability of such methods. The advantage of LSTMs in such framework is that this class of models can easily and naturally accommodate nonlinearities (eg seasonality, smooth and sudden regime shifts, thresholds, unit roots, cointegration, etc) without direct intervention or restrictions from the researcher. LSTMs represent an efficient and effective complement to existing methods, thanks to their flexible joint-handling of nonlinearities in the data. This flexibility offers an edge over traditional methods in that LSTMs let data speak freely: this result is achieved at once, without imparting any particular structures to the data.

A straightforward example of application, making contact with VARs, is impulse responses. As much as estimated VARs, a trained network can be fed with a transitory or persistent shock to one or multiple time series to study how such perturbation propagates through the economy. Critically, while VARs rest on the concept of steady-state, LSTMs do not need such restriction and thus can inform on the trajectory of macroaggregates out of equilibrium.

Similarly, LSTMs can be employed as laboratories for policy experiments and data-based counterfactuals, in close competition with Synthetic Control Methods. Training on data prior to the UK Brexit vote, a LSTM can produce informed forecasts on GDP and provide a comparison with actual economic performance, so to assess the cost of such a decision.

5.3.2 Inflation and Structural Change

Within the Phillips curve framework, current inflation is determined by swings in economic activity and expected inflation. Dynamics are therefore passed on essentially from output shocks and trends. chapter 5 finds evidence of decreased inflation persistence but only offer suggestive channels. The smooth, hump-shaped dynamics of the GDP deflator is suggestive of the transformations that the US economy is undergoing. Structural change is one of these long-run processes that is changing the fabric of many advanced economies.

In a nutshell, it is the slow and smooth reallocation of production factors and value-added across the three main sectors of the economy along a balanced growth path. At the early stage, as income increases and the aggregate economy starts developing, the factor share allocated to agriculture decreases. These factors relocate towards the manufacturing and services sectors, with the former absorbing most factors initially. Then, manufacturing typically reaches a tipping point and starts a decreasing path, when the economy as a whole fully matured into an industrial economy. From that tipping point, services absorb increasingly more factors and the sector grows its share.

This dynamic is observed across a number of advanced economies, at a global scale, although some developing countries appear to be transitioning directly from agriculture- to services-oriented economies (Rodrik, 2015).

Although structural is well studied in growth and development, both empirically and theoretically (see Herrendorf, Rogerson, and Valentinyi (2014) for a comprehensive review). Usually, the main drivers behind structural change are found in technology or preferences. For the latter, departing from homothetic preferences allows income increases on the balanced growth path to affect consumption shares for the three sectors. This in turn affects production and factor allocation, accordingly.

For the technological channel, differences in sectoral productivities are the modelling device of choice (Ngai and Pissarides, 2007). In such a setting and for given consumption elasticities, relatively more productive sectors will progressively need fewer inputs, which are absorbed from less productive ones. The technological mechanism is akin to Baumol (1967) disease: eventually, the least productive sector prevails in the economy and raises the aggregate price level to that of the sector. The sectoral dynamics are evident, as Figure 6.1 shows: services prices rise faster than manufacturing and labour shares steadily relocate towards service-providing sectors.

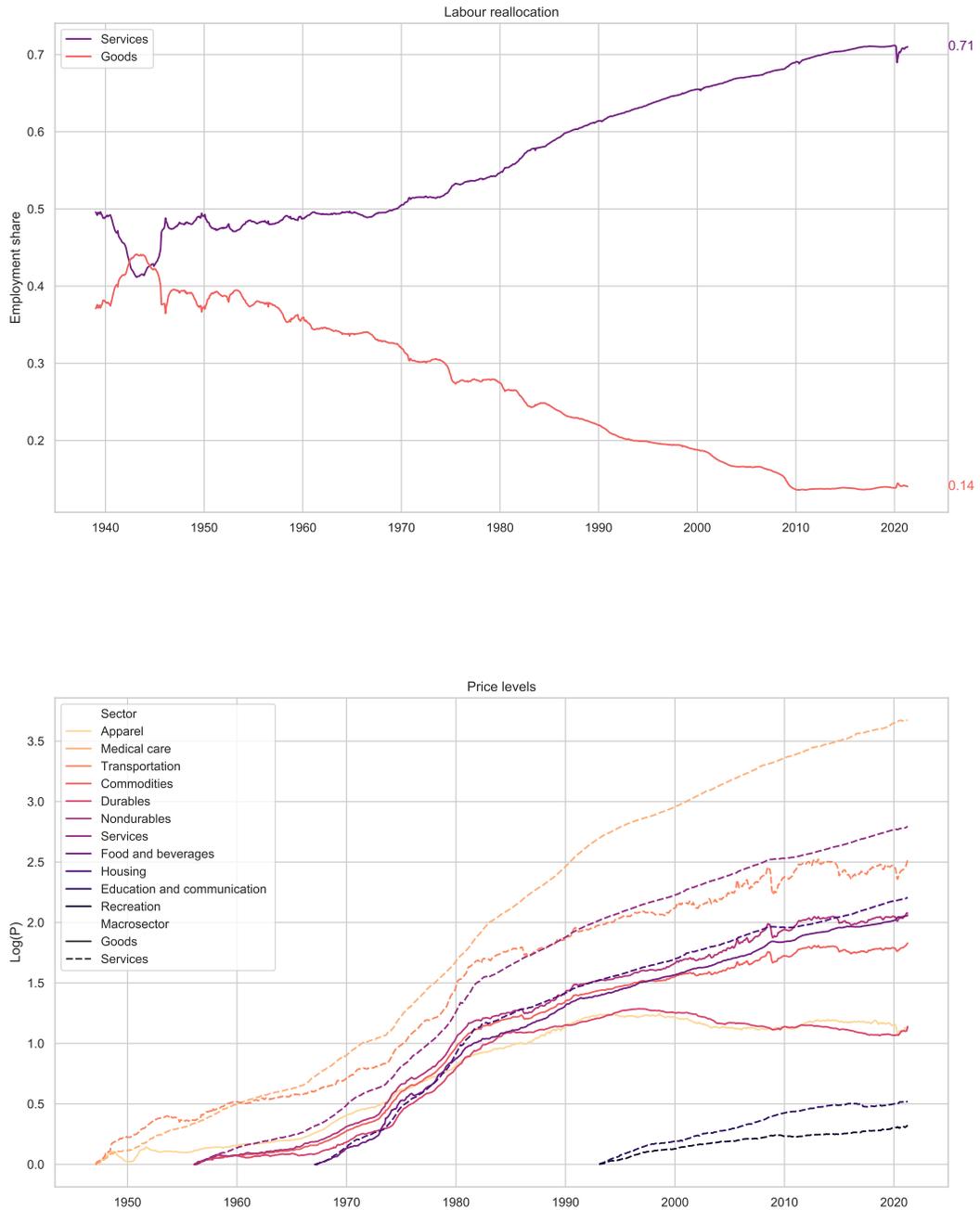
As developed economies transition towards greater services shares, hence, the process behind the output gap mirrors such transition. Consequently, differences in the structure of services and manufacturing are passed on to the inflation process.

Structural change pertains to growth phenomena, thus its analysis in monetary macroeconomics is underdeveloped, with few notable exceptions that mainly focus on the effects of different sectors on monetary policy and transmission (Bouakez, Cardia, and Ruge-Murcia, 2009; Carvalho, Lee, and Park, 2021; Galesi and Rachedi, 2019; Nakamura and Steinsson, 2010).

These works explore how a multisectoral setting affect shock transmission, often with the addition of input-output linkages. Monetary policy is influenced as long as sectors differ in price rigidity: aggregate inflation responds more slowly to aggregate policy shocks when sectoral linkages and heterogeneity are considered.

Thus, policy-makers ought to consider sectoral heterogeneity when designing interventions, since it significantly alters the reaction to policy shocks. Along the same lines, though, sectoral heterogeneity, input-output linkages, and reallocation significantly alter the inflation process. The issue is often analysed through comparative *statics* lenses as in Galesi and Rachedi (2019), but it is even more relevant in its *dynamics*. Indeed, the idea is to study an economy at the juncture of endogenous structural change and monetary policy, to analyse how a growing services share – in inputs or output shares – affects the dynamics of observed, aggregate inflation. The challenge is to conciliate sectoral heterogeneity in productivity (trend and variability of the technological process), in pricing functions (sectors with stickier prices), and factor elasticities (labour or capital intensive sectors) with mechanisms typical of long-run growth models. Studying how the inflation

Figure 5.1 – US Factor Reallocation and Sectoral Prices



Top panel: employment share of macro sectors in the US economy, goods producing macrosector vs services providing. Bottom panel: log of sectoral price index, rebased to 1 for the first observation. Source: Bureau of Labour Statistics.

process changes dynamically in response to the underlying sectoral recomposition is a compelling avenue for research, since it would shed light on why inflation responds less to the output gap and displays decreased persistence and predictability. The more so as structural change could significantly affect the dynamics of trend inflation, which in turn affect much of the observed behaviour of inflation. Studies like Ascari and Sbordone (2014) and Stock and Watson (2007) do indeed point to changes in level and volatility of trend inflation as main drivers of decreased inflation predictability, but remain silent on the root origin of such decline.

Studying how sectoral inflation dynamics and structural change affect aggregate inflation reconciles long term trends – as the reallocation of output and employment from goods-producing sectors to services – and short term fluctuations due to productivity shocks and nominal frictions. It can help explain why inflation has become less persistent in recent decades. A theoretical model focusing on heterogeneous sectoral characteristics as productivity growth and volatility, price rigidity degree, and factors intensity is warranted to study how the transmission mechanism changes when the economy endogenously pivots towards services. The main question to answer is how should monetary policy take into account these sectoral unbalances to maintain or improve its effectiveness.

In light of the effects that sectoral heterogeneity plays in determining aggregate outcomes, it is increasingly interesting to deepen the analysis of heterogeneity on the firms' side of the economy. While much work has been done on moving beyond the representative consumer framework in macroeconomic research, comparatively less effort has been devoted to understanding how firm heterogeneity affects the economy. Especially in policy analyses, it is paramount to understand how firms with different characteristics react to fiscal incentives or monetary shocks. Credit access, productivity, monopoly power, skill intensity, network relations are just some dimensions of heterogeneity worth investigating in relation to policy.

In close relation with chapter 3 and chapter 4, credit and liquidity positions are fundamental factors for firms' survival. Small, labour-intensive enterprises constitute the vast majority of firms in developed countries. These firms have thus little capital to pledge as collateral for credit, and often operate with thin liquidity buffers. Hence, understanding heterogeneity is crucial to devise effective policy mixes to uphold healthy firms while easing factor reallocation towards these latter.

Bibliography

- Aiyagari, Rao and Neil Wallace (Oct. 1992). "Fiat Money in the Kiyotaki-Wright Model". In: *Economic Theory* 2.4, pp. 447–464.
- Aiyagari, S. Rao, Neil Wallace, and Randall Wright (1996). "Coexistence of money and interest-bearing securities". In: *Journal of Monetary Economics* 37.3, pp. 397–419.
- Almosova, Anna and Niek Andresen (2019). "Nonlinear inflation forecasting with recurrent neural networks". In: *WP*.
- AMLF (2016). URL: <https://www.federalreserve.gov/regreform/reform-amlf.htm>.
- Andrle, Michal (May 2013). *What Is in Your Output Gap? Unified Framework & Decomposition into Observables*. IMF Working Papers 2013/105. International Monetary Fund.
- Ascari, Guido and Argia M. Sbordone (Sept. 2014). "The Macroeconomics of Trend Inflation". In: *Journal of Economic Literature* 52.3, pp. 679–739.
- Athey, Susan (2018). "The Impact of Machine Learning on Economics". In: *The Economics of Artificial Intelligence: An Agenda*. NBER Chapters. National Bureau of Economic Research, Inc, pp. 507–547.
- Athey, Susan and Guido W. Imbens (2015). *Machine Learning for Estimating Heterogeneous Causal Effects*. Research Papers 3350. Stanford University, Graduate School of Business.
- (2019). "Machine Learning Methods That Economists Should Know About". In: *Annual Review of Economics* 11.1, pp. 685–725.
- Auer, Raphael, Claudio Borio, and Andrew Filardo (2017). *The globalisation of inflation: the growing importance of global value chains*. BIS Working Papers 602. Bank for International Settlements.
- Autor, David H., David Dorn, and Gordon H. Hanson (2016). "The China Shock: Learning from Labor-Market Adjustment to Large Changes in Trade". In: *Annual Review of Economics* 8.1, pp. 205–240.
- Bajari, Patrick et al. (2015). "Machine Learning Methods for Demand Estimation". In: *American Economic Review* 105.5, pp. 481–85.
- Baumol, William (1967). "Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis". In: *American Economic Review*.
- Baxa, Jaromír, Roman Horváth, and Bořek Vašíček (2013). "Time-varying monetary-policy rules and financial stress: Does financial instability matter for monetary policy?" In: *Journal of Financial Stability* 9.1, pp. 117–138.
- Benati, Luca and Paolo Surico (2008). "Evolving U.S. Monetary Policy and The Decline of Inflation Predictability". In: *Journal of the European Economic Association* 6.2-3, pp. 634–646.
- Benhabib, Jess, Stephanie Schmitt-Grohe, and Martin Uribe (2001). "The perils of Taylor rules". In: *Journal of Economic Theory*.

- Benhabib, Jess, Stephanie Schmitt-Grohé, and Martín Uribe (May 2002). "Chaotic Interest-Rate Rules". In: *American Economic Review* 92.2, pp. 72–78.
- Bernanke, Ben S. (2002). "Asset-price "bubbles" and monetary policy". In: *Remarks before the New York Chapter of the National Association for Business Economics*.
- Bernanke, Ben and Mark Gertler (Mar. 1989). "Agency Costs, Net Worth, and Business Fluctuations". In: *American Economic Review* 79.1, pp. 14–31.
- Bernanke, Ben, Mark Gertler, and Simon Gilchrist (1996). "The financial accelerator and the flight to quality". In: *The review of economics and statistics*, pp. 1–15.
- Bianchi, Francesco (2013). "Regime switches, agents' beliefs, and post-WWII US macroeconomic dynamics". In: *Review of Economic Studies*.
- Bianchi, Francesco and Andrea Civelli (2015). "Globalization and Inflation: Evidence from a Time Varying VAR". In: *Review of Economic Dynamics* 18.2, pp. 406–433.
- Bilbiie, Florin O., Fabio Ghironi, and Marc J. Melitz (Oct. 2019). "Monopoly Power and Endogenous Product Variety: Distortions and Remedies". In: *American Economic Journal: Macroeconomics* 11.4, pp. 140–74.
- Bilbiie, Florin O. and Roland Straub (2013). "Asset market participation, monetary policy rules, and the Great Inflation". In: *The Review of Economics and Statistics*.
- Blanchard, Olivier J. and Charles M. Kahn (1980). "The solution of linear difference models under rational expectations". In: *Econometrica*.
- Boivin, Jean (2006). "Has US monetary policy changed? Evidence from drifting coefficients and real-time data". In: *Journal of Money, Credit and Banking*.
- Bouakez, Hafedh, Emanuela Cardia, and Francisco J. Ruge-Murcia (Nov. 2009). "The Transmission Of Monetary Policy In A Multisector Economy". In: *International Economic Review* 50.4, pp. 1243–1266.
- Caballero, Ricardo J., Emmanuel Farhi, and Pierre-Olivier Gourinchas (2016). "Safe Asset Scarcity and Aggregate Demand". In: *American Economic Review*.
- (2017). "The safe asset shortage conundrum". In: *Journal of Economic Perspectives*.
- Calvo, Guillermo A. (1983). "Staggered prices in a utility-maximizing framework". In: *Journal of Monetary Economics*.
- (2016). *Macroeconomics in times fo liquidity crises: searching for economic essentials*.
- Calvo, Guillermo A. and Carlos A. Végh (1996). "Disinflation and interest-bearing money". In: *The Economic Journal*.
- Calvo, Guillermo and Carlos Végh (1990a). "Interest Rate Policy in a Small Open Economy: The Predetermined Exchange Rates Case". In: *Staff Papers (International Monetary Fund)* 37.4, pp. 753–776.
- (Dec. 1990b). "Money Supply and Interest Rate Policy in a New-Keynesian Framework". In: 1990/119.
- Canova, Fabio and Luca Sala (2009). "Back to square one: identification issues in DSGE models". In: *Journal of Monetary Economics*.
- Canzoneri, Matthew B. and Behzad T. Diba (2005). "Interest rate rules and price determinacy: the role of transactions services of bonds". In: *Journal of Monetary Economics*.
- Canzoneri, Matthew et al. (Dec. 2008a). "Monetary Aggregates and Liquidity in a Neo-Wicksellian Framework". In: *Journal of Money, Credit and Banking* 40.8, pp. 1667–1698.
- (2008b). "Monetary and fiscal policy coordination when bonds provide transactions services". In: *CEPR WP 6814*.

- Canzoneri, Matthew et al. (2011). “The role of liquid government bonds in the great transformation of American monetary policy”. In: *Journal of Economic Dynamics and Control* 35.3, pp. 282–294.
- Carvalho, Carlos, Jae Won Lee, and Woong Yong Park (Jan. 2021). “Sectoral Price Facts in a Sticky-Price Model”. In: *American Economic Journal: Macroeconomics* 13.1, pp. 216–256.
- Carvalho, Carlos, Fernanda Nechio, and Tiago Tristao (2018). “Taylor Rule Estimation by OLS”. In: *Federal Reserve of San Francisco Working Paper Series 2018-11*.
- Chakraborty, Chiranjit and Andreas Joseph (2017). *Machine learning at central banks*. Bank of England working papers 674. Bank of England.
- Chen, Long, David A. Lesmond, and Jason Wei (Feb. 2007). “Corporate Yield Spreads and Bond Liquidity”. In: *Journal of Finance* 62.1, pp. 119–149.
- Chordia, Tarun, Asani Sarkar, and Avanidhar Subrahmanyam (2005). “An Empirical Analysis of Stock and Bond Market Liquidity”. In: *Review of Financial Studies* 18.1, pp. 85–129.
- Chow, Gregory C. (1960). “Tests of equality between sets of coefficients in two linear regressions”. In: *Econometrica*.
- Ciccarelli, Matteo and Chiara Osbat (2017). “Low inflation in the euro area: Causes and consequences”. In: *ECB Occasional Paper 181*.
- Clarida, Richard, Jordi Gali, and Mark Gertler (2000). “Monetary policy rules and macroeconomic stability: evidence and some theory”. In: *Quarterly Journal of Economic*.
- Clower, Robert (1967). “A Reconsideration Of The Microfoundations Of Monetary Theory”. In: *Economic Inquiry* 6.1, pp. 1–8.
- Cochrane, John H. (2011). “Determinacy and identification with Taylor rules”. In: *Journal of Political Economy*.
- (2020). *The Fiscal Theory of the Price Level*.
- Cogley, Timothy, Giorgio E. Primiceri, and Thomas J. Sargent (2008). “Inflation-gap persistence in the US”. In: *American Economic Journal: Macroeconomics*.
- Cogley, Timothy and Thomas J. Sargent (2002). “Evolving post-World War II US inflation dynamics”. In: *NBER Macroeconomics Annual 2001*. Ed. by Mark Gertler and Kennet Rogoff.
- (2005). “Drift and volatilities: monetary policies ad outcomes in the post World War II US”. In: *Review of Economic Dynamics*.
- Cogley, Timothy and Argia Sbordone (2009). “Trend inflation, indexation, and inflation persistence in New Keynesian Phillips curve”. In: *American Economic Review*.
- Coibion, Olivier and Yuriy Gorodnichenko (2015). “Is the Phillips curve alive and well after all? Inflation expectations and the missing disinflation”. In: *American Economic Journal – Macroeconomics*.
- Coibion, Olivier, Yuriy Gorodnichenko, and Rupal Kamdar (Dec. 2018). “The Formation of Expectations, Inflation, and the Phillips Curve”. In: *Journal of Economic Literature* 56.4, pp. 1447–91.
- Coibion, Olivier, Yuriy Gorodnichenko, and Tiziano Ropele (2019). “Inflation expectations and firm decisions: new causal evidence”. In: *WP*.
- Cooley, Thomas F and Gary D Hansen (Sept. 1989). “The Inflation Tax in a Real Business Cycle Model”. In: *American Economic Review* 79.4, pp. 733–748.
- (Aug. 1991). “The Welfare Costs of Moderate Inflation”. In: *Journal of Money, Credit and Banking* 23.3, pp. 483–503.
- CPIFF (2016). URL: <https://www.federalreserve.gov/regreform/reform-cpff.htm>.

- D'Agostino, Antonello and Paolo Surico (Mar. 2009). "Does Global Liquidity Help to Forecast U.S. Inflation?" In: *Journal of Money, Credit and Banking* 41.2-3, pp. 479–489.
- Davig, Troy and Eric Leeper (2006). "Fluctuating macro policies and the fiscal theory". In: *NBER Macroeconomics Annual*.
- (2011). "Monetary-fiscal policy interactions and fiscal stimulus". In: *European Economic Review*.
- Del Negro, Marco et al. (2017). "Safety, liquidity, and the natural rate of interest". In: *Brookings Papers on Economic Activity*.
- Diba, Behzad and Olivier Loisel (2021). "Pegging the interest rate on bank reserves: A resolution of New Keynesian puzzles and paradoxes". In: *Journal of Monetary Economics* 118.C, pp. 230–244.
- Dick-Nielsen, Jens, Peter Feldhütter, and David Lando (2012). "Corporate bond liquidity before and after the onset of the subprime crisis". In: *Journal of Financial Economics* 103.3, pp. 471–492.
- Eggertsson, Gauti B. and Paul Krugman (2012). "Debt, Deleveraging, and the Liquidity Trap: A Fisher-Minsky-Koo Approach". In: *The Quarterly Journal of Economics* 127.3, pp. 1469–1513.
- Fernandez-Villaverde, Jesus and Pablo Guerron-Quintana (2020). "Estimating DSGE models: recent advances and future challenges". In: *NBER WP w27715*.
- Fernández-Villaverde, Jesús, Pablo Guerrón-Quintana, and Juan F. Rubio-Ramírez (Apr. 2010). *Fortune or Virtue: Time-Variant Volatilities Versus Parameter Drifting in U.S. Data*. NBER Working Papers 15928. National Bureau of Economic Research, Inc.
- (2015). "Estimating dynamic equilibrium models with stochastic volatility". In: *Journal of Econometrics* 185.1, pp. 216–229.
- Fernández-Villaverde, Jesús, Samuel Hurtado, and Galo Nuño (2020). *Financial Frictions and the Wealth Distribution*. CESifo Working Paper Series 8482. CESifo.
- Fuhrer, Jeffrey (2011). "Inflation Persistence". In: *Handbook of Monetary Economics*.
- Gabaix, Xavier (2016). "A behavioral New Keynesian model". In: *Mimeo, New York University*.
- Galesi, Alessandro and Omar Rachedi (2019). "Services Deepening and the Transmission of Monetary Policy". In: *Journal of the European Economic Association* 17.4, pp. 1261–1293.
- Gali, Jordi (2015). *Monetary Policy, Inflation, and the Business Cycle*.
- Gali, Jordi and Mark Gertler (1999). "Inflation dynamics: a structural econometric analysis". In: *Journal of Monetary Economics*.
- Gali, Jordi, Mark Gertler, and J. David Lopez-Salido (2001). "European inflation dynamics". In: *European Economic Review*.
- Garcia-Schmidt, Mariana and Micheal Woodford (2019). "Are low interest rates deflationary? A paradox of perfect-foresight analysis". In: *American Economic Review*.
- Geromichalos, Athanasios and Lucas Herrenbrueck (2021). "The liquidity-augmented model of macroeconomic aggregates: A new monetarist DSGE approach". In: *Review of Economic Dynamics*.
- Gertler, Mark and Peter Karadi (Jan. 2011). "A model of unconventional monetary policy". In: *Journal of Monetary Economics* 58.1, pp. 17–34.
- Gertler, Pavel, Roman Horváth, and Júlia Jonášová (Apr. 2020). "Central Bank Communication and Financial Market Comovements in the Euro Area". In: *Open Economies Review* 31.2, pp. 257–272.
- Giannone, Domenico, Michele Lenza, and Giorgio E. Primiceri (2018). *Economic predictions with big data: the illusion of sparsity*. Staff Reports 847. Federal Reserve Bank of New York.
- Goulet Coulombe, Philippe et al. (2019). *How is Machine Learning Useful for Macroeconomic Forecasting?* CIRANO Working Papers 2019s-22. CIRANO.

- Grandmont, Jean-Michel and Yves Younes (1972). "On the Role of Money and the Existence of a Monetary Equilibrium". In: *Review of Economic Studies* 39.3, pp. 355–372.
- Greenspan, Alan (2002). "Economic volatility". In: *Remarks at symposium sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole*.
- Greff, Klaus et al. (2015). "LSTM: A Search Space Odyssey". In: *CoRR abs/1503.04069*.
- Gu, Shihao, Bryan Kelly, and Dacheng Xiu (Feb. 2020). "Empirical Asset Pricing via Machine Learning". In: *The Review of Financial Studies* 33.5, pp. 2223–2273.
- Hall, Robert and Ricardo Reis (2016). "Achieving price stability by manipulating the Central Bank's payment on reserves". In: *NBER Working Paper*.
- Hamilton, James D. (1989). "A new approach to the economic analysis of nonstationary time series and the business cycle". In: *Econometrica*.
- (1994). *Time series analysis*.
- Hastie, Trevor, Robert Tibshirani, and Jerome Friedman (2009). *The Elements of Statistical Learning*.
- Havranek, Tomas et al. (2015). "Cross-country heterogeneity in intertemporal substitution". In: *Journal of International Economics* 96.1, pp. 100–118.
- Herrendorf, Berthold, Richard Rogerson, and Akos Valentinyi (2014). "Growth and Structural Transformation". In: *Handbook of Economic Growth*.
- Holmström, Bengt and Jean Tirole (Sept. 2011). *Inside and Outside Liquidity*. Vol. 1. MIT Press Books 0262015781. The MIT Press.
- Houweling, Patrick, Albert Mentink, and Ton Vorst (June 2005). "Comparing possible proxies of corporate bond liquidity". In: *Journal of Banking & Finance* 29.6, pp. 1331–1358.
- Iacopetta, Maurizio and Raoul Minetti (Jan. 2019). "Asset Dynamics, Liquidity, And Inequality In Decentralized Markets". In: *Economic Inquiry* 57.1, pp. 537–551.
- Iacoviello, Matteo and Raoul Minetti (Jan. 2003). "Financial Liberalization And The Sensitivity Of House Prices To Monetary Policy: Theory And Evidence". In: *Manchester School* 71.1, pp. 20–34.
- (Mar. 2008). "The credit channel of monetary policy: Evidence from the housing market". In: *Journal of Macroeconomics* 30.1, pp. 69–96.
- Istrefi, Klodiana, Florens Odendahl, and Giulia Sestieri (2020). *Fed Communication on Financial Stability Concerns and Monetary Policy Decisions: Revelations from Speeches*. Working papers 779. Banque de France.
- Jarociński, Marek and Elena Bobeica (2017). *Missing disinflation and missing inflation: the puzzles that aren't*. Working Paper Series 2000. European Central Bank.
- Jozefowicz, Rafal, Wojciech Zaremba, and Ilya Sutskever (2015). "An Empirical Exploration of Recurrent Network Architectures". In: ed. by Francis Bach and David Blei. Vol. 37, pp. 2342–2350.
- Jung, Jin-Kyu, Manasa Patnam, and Anna Ter-Martirosyan (2018). *An Algorithmic Crystal Ball: Forecasts-based on Machine Learning*. IMF Working Papers. International Monetary Fund.
- Karpathy, Andrej (2015). *The unreasonable effectiveness of recurrent neural networks*. URL: <http://karpathy.github.io/2015/05/21/rnn-effectiveness/>
- Kidger, Patrick and Terry J. Lyons (2019). "Universal Approximation with Deep Narrow Networks". In: *CoRR abs/1905.08539*.
- Kim, Chang-Jin and Charles R. Nelson (2006). "Estimation of a forward-looking monetary policy rule: A time-varying parameter model using ex post data". In: *Journal of Monetary Economics* 53.8, pp. 1949–1966.
- Kingma, Diederik P. and Jimmy Ba (2014). *Adam: A Method for Stochastic Optimization*.

- Kiyotaki, Nobuhiro and John Moore (Apr. 1997). "Credit Cycles". In: *Journal of Political Economy* 105.2, pp. 211–248.
- Kiyotaki, Nobuhiro and Randall Wright (1991). "A contribution to the pure theory of money". In: *Journal of Economic Theory* 53.2, pp. 215–235.
- Kocherlakota, Naranayana (2016). "Fragility of purely real macroeconomic models". In: *NBER Working Paper*.
- Kock, Anders Bredahl and Timo Teräsvirta (2016). "Forecasting macroeconomic variables using neural network models and three automated model selection techniques". In: *Econometric Reviews* 35.8-10, pp. 1753–1779.
- Korobilis, Dimitris (2018). *Machine Learning Macroeconometrics: A Primer*. Working Paper series 18-30. Rimini Centre for Economic Analysis.
- Krippner, Leo (2015). *Zero lower bound term structure modeling: A practitioner's guide*.
- Kurozumi, Takushi and Willem Van Zandweghe (Aug. 2019). *A Theory of Intrinsic Inflation Persistence*. Working Papers 201916. Federal Reserve Bank of Cleveland.
- Kydland, Finn E. and Edward C. Prescott (1982). "Time to Build and Aggregate Fluctuations". In: *Econometrica* 50.6, pp. 1345–1370.
- Lagos, Ricardo, Guillaume Rocheteau, and Randall Wright (June 2017). "Liquidity: A New Monetarist Perspective". In: *Journal of Economic Literature* 55.2, pp. 371–440.
- Leeper, Eric M. (Feb. 1991). "Equilibria under 'active' and 'passive' monetary and fiscal policies". In: *Journal of Monetary Economics* 27.1, pp. 129–147.
- Leshno, Moshe et al. (1993). "Multilayer feedforward networks with a nonpolynomial activation function can approximate any function". In: *Neural Networks* 6.6, pp. 861–867.
- Levieuge, Grégory (2002). "Banques centrales et prix d'actifs : une étude empirique". In: *Revue française d'économie*, volume 16, n. 4, 2002.
- Lhuissier, Stephane (2018). "The regime-switching volatility of Euro Area business cycles". In: *Macroeconomic Dynamics*.
- Lhuissier, Stephane and Fabien Tripier (2019). "Regime-dependent effects of Uncertainty shocks: a structural interpretation". In: *Banque de France WP 714*.
- Lin, Hai, Junbo Wang, and Chunchi Wu (Mar. 2011). "Liquidity risk and expected corporate bond returns". In: *Journal of Financial Economics* 99.3, pp. 628–650.
- Long, John B. and Charles I. Plosser (1983). "Real Business Cycles". In: *Journal of Political Economy* 91.1, pp. 39–69.
- Lucas Robert E, Jr and Nancy L Stokey (May 1987). "Money and Interest in a Cash-in-Advance Economy". In: *Econometrica* 55.3, pp. 491–513.
- Lucas, Robert (Apr. 1972). "Expectations and the neutrality of money". In: *Journal of Economic Theory* 4.2, pp. 103–124.
- (Dec. 1975). "An Equilibrium Model of the Business Cycle". In: *Journal of Political Economy* 83.6, pp. 1113–1144.
- (Jan. 1977). "Understanding business cycles". In: *Carnegie-Rochester Conference Series on Public Policy* 5.1, pp. 7–29.
- (Apr. 1980). "Equilibrium in a Pure Currency Economy". In: *Economic Inquiry* 18.2, pp. 203–220.
- Lucas, Robert E. (1982). "Interest rates and currency prices in a two-country world". In: *Journal of Monetary Economics* 10.3, pp. 335–359.
- Lucas, Robert Jr (Jan. 1976). "Econometric policy evaluation: A critique". In: *Carnegie-Rochester Conference Series on Public Policy* 1.1, pp. 19–46.

- Makridakis, Spyros, Evangelos Spiliotis, and Vassilios Assimakopoulos (2018). “Statistical and Machine Learning forecasting methods: Concerns and ways forward”. In: *PLOS ONE* 13.3, pp. 1–26.
- Maliar, Lilia, Serguei Maliar, and Pablo Winant (2019). “Will artificial intelligence replace computational economists any time soon?” In: *CEPR discussion paper DP14024*.
- Mankiw, N. Gregory, Ricardo Reis, and Justin Wolfers (2003). *Disagreement about Inflation Expectations*. Working Paper 9796. National Bureau of Economic Research.
- Marcellino, Massimiliano, James H. Stock, and Mark W. Watson (2006). “A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series”. In: *Journal of Econometrics*.
- Mavroeidis, Sophocles, Mikkel Plegborg-Moller, and James H. Stock (2014). “Empirical evidence on inflation expectations in the New Keynesian Phillips Curve”. In: *Journal of Economic Literature*.
- McAdam, Peter and Paul McNelis (2005). “Forecasting inflation with thick models and neural networks”. In: *Economic Modelling* 22.5, pp. 848–867.
- Medeiros, Marcelo C. et al. (2019). “Forecasting Inflation in a Data-Rich Environment: The Benefits of Machine Learning Methods”. In: *Journal of Business & Economic Statistics*, pp. 1–22.
- Michaillat, Pascal and Emmanuel Saez (2015). “An economical business-cycle model”. In: *NBER WP 19777*.
- (May 2021). “Resolving New Keynesian Anomalies with Wealth in the Utility Function”. In: *The Review of Economics and Statistics* 103.2, pp. 197–215.
- Miles, David et al. (2017). *And yet it moves: Inflation and the Great Recession*.
- MMIFF (2016). URL: <https://www.federalreserve.gov/regreform/reform-mmiff.htm>.
- Mullainathan, Sendhil and Jann Spiess (2017). “Machine Learning: An Applied Econometric Approach”. In: *Journal of Economic Perspectives* 31.2, pp. 87–106.
- Murray, Christian J., Alex Nikolsko-Rzhevskyy, and David H. Papell (2015). “Markov switching and the Taylor principle”. In: *Macroeconomic Dynamics*.
- Nakamura, Emi (2005). “Inflation forecasting using a neural network”. In: *Economics Letters* 86.3, pp. 373–378.
- Nakamura, Emi and Jón Steinsson (2010). “Monetary Non-neutrality in a Multisector Menu Cost Model”. In: *The Quarterly Journal of Economics* 125.3, pp. 961–1013.
- Ngai, Rachel and Christopher Pissarides (2007). “Structural change in a multisector model of growth”. In: *American Economic Review*.
- Oet, Mikhail V. and Kalle Lyytinen (2017). “Does financial stability matter for the Fed in setting US monetary policy?” In: *Review of Finance*.
- Orphanides, Athanasios (2001). “Monetary policy rules based on real-time data”. In: *American Economic Review*.
- (2004). “Monetary policy rules, macroeconomic stability, and inflation: a view from the trenches”. In: *Journal of Money, Credit, and Banking*.
- PDCF (2016). URL: <https://www.federalreserve.gov/regreform/reform-pdcf.htm>.
- Peek, Joe, Eric S. Rosengren, and Geoffrey M. B. Tootell (2016). “Does Fed policy reveal a ternary mandate?” In: *FRB of Boston WP 16-11*.
- Phillips, William (1958). “The Relation Between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861–19571”. In: *Economica* 25.100, pp. 283–299.

- Pivetta, Frederic and Ricardo Reis (2007). "The persistence of inflation in the United States". In: *Journal of Economic Dynamics and Control*.
- Primiceri, Giorgio E. (2005). "Time varying structural vector autoregressions and monetary policy". In: *Review of Economic Studies*.
- Rackauckas, Christopher et al. (2020). "Universal Differential Equations for Scientific Machine Learning". In: *arXiv preprint arXiv:2001.04385*.
- Ravenna, Federico and Carl E. Walsh (Mar. 2006). "Optimal monetary policy with the cost channel". In: *Journal of Monetary Economics* 53.2, pp. 199–216.
- Roberts, John M (Nov. 1995). "New Keynesian Economics and the Phillips Curve". In: *Journal of Money, Credit and Banking* 27.4, pp. 975–984.
- Rodrik, Dani (2015). "Premature deindustrialisation". In: *Journal of Economic Growth*.
- Rotemberg, Julio J. (1982). "Sticky prices in the United States". In: *Journal of Political Economy*.
- Ruder, Sebastian (2016). "An overview of gradient descent optimization algorithms". In: *CoRR abs/1609.04747*.
- Schestag, Raphael, Philipp Schuster, and Marliese Uhrig-Homburg (2016). "Measuring Liquidity in Bond Markets". In: *Review of Financial Studies* 29.5, pp. 1170–1219.
- Schmidt, Robin M., Frank Schneider, and Philipp Hennig (2021). *Descending through a Crowded Valley – Benchmarking Deep Learning Optimizers*.
- Schmitt-Grohe, Stephanie, Jess Benhabib, and Martin Uribe (Mar. 2001). "Monetary Policy and Multiple Equilibria". In: *American Economic Review* 91.1, pp. 167–186.
- Schmitt-Grohe, Stephanie and Martin Uribe (Feb. 2000). "Price level determinacy and monetary policy under a balanced-budget requirement". In: *Journal of Monetary Economics* 45.1, pp. 211–246.
- (Sept. 2007). "Optimal simple and implementable monetary and fiscal rules". In: *Journal of Monetary Economics* 54.6, pp. 1702–1725.
- Sidrauski, Miguel (Part II, 1969). "Rational Choice and Patterns of Growth". In: *Journal of Political Economy* 77.4, pp. 575–585.
- Sims, Christopher A. (2010). "Chapter 4 - Rational Inattention and Monetary Economics". In: ed. by Benjamin M. Friedman and Michael Woodford. Vol. 3. *Handbook of Monetary Economics*. Elsevier, pp. 155–181.
- Sims, Christopher A. and Tao Zha (2006). "Were there regime switches in US monetary policy?" In: *American Economic Review*.
- Smets, Frank and Rafael Wouters (2007). "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach". In: *American Economic Review*.
- Solow, Robert M. (1956). "A Contribution to the Theory of Economic Growth". In: *The Quarterly Journal of Economics* 70.1, pp. 65–94.
- Stock, James H. and Mark W. Watson (2007). "Why Has U.S. Inflation Become Harder to Forecast?" In: *Journal of Money, Credit and Banking* 39.s1, pp. 3–33.
- Sullivan, Peter (2016). "Qualitative Taylor rule". In: *WP*.
- Surico, Paolo (Nov. 2008). "The Cost Channel Of Monetary Policy And Indeterminacy". In: *Macroeconomic Dynamics* 12.5, pp. 724–735.
- Svensson, Lars E O (Oct. 1985). "Money and Asset Prices in a Cash-in-Advance Economy". In: *Journal of Political Economy* 93.5, pp. 919–944.
- (2017a). "Cost-benefit analysis of leaning against the wind". In: *Journal of Monetary Economics*.

- Svensson, Lars E. O. (2017b). "Leaning against the wind: costs and benefits, effects on debt, leaning in DSGE models, and a framework for comparison of results". In: *International Journal of Central Banking*.
- Tabuada, Paulo and Bahman Ghahesifard (2020). *Universal Approximation Power of Deep Residual Neural Networks via Nonlinear Control Theory*.
- TAF (2016). URL: <https://www.federalreserve.gov/regreform/reform-taf.htm>.
- Taylor, John B. (1993). "Discretion versus policy rules in practice". In: *Carnegie-Rochester Conference Series on Public Policy*.
- (2014). *After unconventional monetary policy. Testimony before the Joint Economic Committee of Congress at the Hearing on "Unwinding quantitative easing: how the Fed should promote stable prices, economic growth and job creation"*.
- Trehan, Bharat (2015). "Survey Measures of Expected Inflation and the Inflation Process". In: *Journal of Money, Credit and Banking* 47.1, pp. 207–222.
- TSLF (2016). URL: <https://www.federalreserve.gov/regreform/reform-tslf.htm>.
- Varian, Hal R. (2014). "Big Data: New Tricks for Econometrics". In: *Journal of Economic Perspectives* 28.2, pp. 3–28.
- Verstyuk, Sergiy (2020). "Modeling multivariate time series in economics: from auto-regressions to recurrent neural networks". In: *WP Mimeo*.
- Walras, Léon (1874). *Elements of Pure Economics*. Corbaz.
- Walsh, Carl E. (2003). *Monetary Theory and Policy*.
- Wischniewsky, Arina, David-Jan Jansen, and Matthias Neuenkirch (2019). "Financial stability and the Fed: evidence from congressional hearings". In: *CESifo WP 7657*.
- Woodford, Michael (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy*.
- Wu, Jing Cynthia and Fan Dora Xia (2016). "Measuring the macroeconomic impact of monetary policy at the zero lower bound". In: *Journal of Money, Credit and Banking*.

List of Tables

3.1	Taylor Rule estimates	50
3.2	Full sample estimates	52
3.3	Exogenous splits, three samples	54
3.4	Optimal segmentation and break dates	57
3.5	Estimates Markov Switching model	58
3.6	Transition matrices	59
3.7	Data sources details for the US	63
3.8	Correlations	64
3.9	Full Sample GMM	70
4.1	Calibration for model simulations	82
4.2	Estimates on simulated data	88
4.3	Optimal lags	89
5.1	Optimal lags selection via BIC	125
5.2	Hyperparameters	147
5.3	Optimal lags from BIC minimisation	148
5.4	CPI decades regressions with LSTM forecasts	159
5.5	PCE decades regressions with LSTM forecasts	160
5.6	GDP Deflator decades regressions with LSTM forecasts	160
5.7	Core CPI decades regressions with LSTM forecasts	161
5.8	Core PCE decades regressions with LSTM forecasts	161

List of Figures

1.1	Historical inflation rates	3
1.2	US Phillips Curve and Monetary Policy	9
1.3	Fed’s Liquidity Facilities	10
1.4	Inflation Expectations	14
1.5	Inflation Persistence	19
2.1	Taux d’inflation historiques	22
2.2	Courbe de Phillips américaine et politique monétaire	29
2.3	Les mesures de liquidité de la Fed	30
2.4	Attentes d’inflation	35
2.5	Persistence de l’inflation	40
3.1	US Macroeconomy	43
3.2	Liquidity-Augmented Taylor Rule Variables for the US	49
3.3	Structural Breaks	56
3.4	Markov Smoothed Probabilities	60
3.5	Nowcasts and Forecasts	67
3.6	Measures of output gap.	68
3.7	OLS Residuals	69
3.8	CUMSUM plots	70
4.1	TFP Shock – IRFs	83
4.2	Interest Rate Shock – IRFs	84
4.3	TFP Shock IRFs – Both Regimes	85
4.4	Interest Rate Shock IRFs – Both Regimes	86
4.5	$AR(k^*)$ Lags	90
4.6	Liquidity Shock	92
4.7	h with Cobb-Douglas liquidity function	104
4.8	Sensitivity on ψ – IRFs	105
4.9	Sensitivity on ϕ – IRFs	106
4.10	Sensitivity on γ – IRFs	108
4.11	Sensitivity on γ – $AR(k^*)$ on Inflation	109
4.12	Sensitivity on γ – Inflation Persistence and Volatility	110
4.13	Autoregressive process lags	112
4.14	Extreme Policies in baseline NKDSGE	114

5.1	US Inflation Data	120
5.2	Headline and Core Inflation	121
5.3	AR (1) Persistence	123
5.4	AR (k^*) Persistence	126
5.5	Bayesian Estimates	130
5.6	Full Sample Forecasts	135
5.7	Full Sample LSTM Predicted Persistence	136
5.8	LSTM on Decades – Persistence On Data and Forecasts	137
5.9	Rolling LSTM – Persistence on Data and Forecasts	140
5.10	Oil Prices	149
5.11	Commodities Prices	150
5.12	Food Commodities Prices	151
5.13	Frequentist Persistence – 20 Quarters Window	153
5.14	Frequentist Persistence – 40 Quarters Window	154
5.15	Frequentist Persistence – 80 Quarters Window	155
5.16	Bayesian Draws – Per Period Full Densities	156
5.17	LSTM Forecasts on Decade Subsamples	157
5.18	LSTM Rolling Window Forecasts	158
5.19	Trend Inflation – Frequentist Measure	163
5.20	Trend Inflation – LSTM on Full Sample	164
5.21	Trend Inflation – LSTM on Decade Subsamples	165
5.22	Trend Inflation – LSTM on Rolling Windows	166
6.1	US Factor Reallocation and Sectoral Prices	173

Abstract

Summary

Inflation Dynamics: Policies and Determinants.

This dissertation focuses on the interplay of monetary policy, liquidity, and inflation dynamics from an empirical and theoretical point of view.

The first chapter empirically investigates the role of financial liquidity in the monetary policy history of the US Federal Reserve Bank. It combines real-time data and Markov switching models to reproduce the Fed's information set and study regime changes. A liquidity-augmented Taylor rule fits well the US data, but its estimates also contrast the consensus on destabilising passive policy stances: passive regimes coexist with moderate inflation after controlling for liquidity.

The second chapter extends a simple NKDSGE model to account for liquid assets. The central bank sets total nominal liquidity and targets interest rates on liquid assets. Passive reactions to expected inflation do not trigger indeterminacy but entail more persistent and volatile inflation. Moreover, passive policies slow down the recovery from a recession.

The third chapter investigates how inflation persistence varied over time in the US macroeconomy. It adapts deep-learning methods, besides more standard ones, to address this question, and consistently finds that inflation inertia has significantly decreased from its peak around 1995. Inflation currently behaves similarly to a memoryless white noise. Policy changes, international trade, or volatile commodities do not seem to determine such a decrease, which predates all of those.

Keywords: Monetary Policy, Inflation, Liquidity, Markov switching, Bayesian Time Series, Deep Learning, New Keynesian Models, Central Banking.

Résumé

Dynamique de l'Inflation : Politiques et Déterminants.

Cette thèse se concentre sur l'interaction entre la politique monétaire, la liquidité et la dynamique de l'inflation.

Le premier chapitre étudie empiriquement le rôle de la liquidité financière dans l'histoire de la politique monétaire de la Réserve Fédérale Américaine. Il conjugue des données en temps réel et des modèles de changement de régime Markoviens pour reproduire l'ensemble des informations de la Fed et étudier les changements de régime. Une règle de Taylor augmentée de la liquidité s'adapte bien aux données américaines, mais contredit le consensus sur l'effet déstabilisant d'une politique passive : les régimes passifs coexistent avec une inflation modérée, une fois tenu compte de la liquidité financière.

Le deuxième chapitre étend un simple modèle NKDSGE pour tenir compte des actifs liquides. La banque centrale fixe la liquidité nominale totale et cible les taux d'intérêt sur les actifs liquides. Les réactions passives à l'inflation attendue ne déclenchent pas l'indétermination, mais entraînent une inflation plus persistante et plus volatile.

Le troisième chapitre étudie l'évolution de la persistance de l'inflation dans la macroéconomie américaine. Il adapte des méthodes de deep-learning, en plus des méthodes plus classiques, pour répondre à cette question, et constate que l'inertie de l'inflation a considérablement diminué depuis son pic autour de 1995, systématiquement à travers les méthodes. L'inflation se comporte actuellement presque comme un bruit blanc sans mémoire. Les changements de politique, le commerce international ou la volatilité des matières premières ne semblent pas déterminer cette baisse, qui est antérieure à tous ces facteurs.

Mots-clés : Politique Monétaire, Inflation, Liquidité, Modèles de Markov, Séries Temporelles Bayésiennes, Deep Learning, Modèles Néokeynésiens, Banques Centrales.