



HAL
open science

Three essays on unconventional monetary policies

Valentin Jouvanceau

► **To cite this version:**

Valentin Jouvanceau. Three essays on unconventional monetary policies. Economics and Finance. Université de Lyon, 2019. English. NNT : 2019LYSE2061 . tel-02468833

HAL Id: tel-02468833

<https://theses.hal.science/tel-02468833>

Submitted on 6 Feb 2020

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.



N° d'ordre NNT : 2019LYSE2061

THESE de DOCTORAT DE L'UNIVERSITÉ DE LYON

Opérée au sein de

L'UNIVERSITÉ LUMIÈRE LYON 2

École Doctorale : ED 486 Sciences Économique et de Gestion

Discipline : Sciences économiques

Soutenue publiquement le 10 juillet 2019, par :

Valentin JOUVANCEAU

Three essays on unconventional monetary policies.

Devant le jury composé de :

Stéphane AURAY, Professeur, ENSAI Rennes, Président

Céline POILLY, Professeure des universités, Aix- Marseille Université, Examinatrice

Fabrice COLLARD, Directeur de Recherches CNRS, Toulouse School of Economics, Examineur

Evi PAPPA, Professeur d'université, Universidad Carlos III, Examineur

Aurélien EYQUEM, Professeur des universités, Université Lumière Lyon 2, Directeur de thèse

Contrat de diffusion

Ce document est diffusé sous le contrat *Creative Commons* « [Paternité – pas de modification](#) » : vous êtes libre de le reproduire, de le distribuer et de le communiquer au public à condition d'en mentionner le nom de l'auteur et de ne pas le modifier, le transformer ni l'adapter.

UNIVERSITÉ DE LYON
ÉCOLE DOCTORALE SCIENCES ÉCONOMIQUES ET GESTION

UNIVERSITÉ LUMIÈRE LYON 2
GROUPE D'ANALYSE ET DE THÉORIE ÉCONOMIQUE

THREE ESSAYS ON UNCONVENTIONAL MONETARY POLICIES

Thèse de Doctorat de Sciences Économiques

VALENTIN JOUVANCEAU

en vue de l'obtention du grade de docteur de l'Université de Lyon délivrée par
l'Université Lumière Lyon 2

Jury :

STÉPHANE AURAY	Professeur	ENSAI, <i>Président</i>
FABRICE COLLARD	DR CNRS	Toulouse School of Economics, <i>Rapporteur</i>
AURÉLIEN EYQUEM	Professeur	Université Lumière Lyon 2, <i>Directeur de thèse</i>
STÉPHANE MOYEN	Chercheur	Deutsche Bundesbank, <i>Examineur</i>
EVI PAPPA	Professeur	Université Carlos III de Madrid, <i>Examinatrice</i>
CÉLINE POILLY	Professeur	Aix-Marseille School of Economics, <i>Rapporteuse</i>

Remerciements

Ce travail résume mes cinq années de thèse de doctorat. Plaisir, connaissances et bonheur sont les maîtres mots. Riche en émotions, cette thèse aura affiné mon caractère. Ces lignes sont des remerciements sincères pour les personnes et composantes ayant été essentielles à sa réalisation.

Mes premiers remerciements sont adressés à Aurélien Eyquem, mon directeur de thèse. Sans lui, cette thèse n'existerait pas. Etudiant perdu dans des études de finance, notre rencontre aura été primordiale. Rêveur et englué dans ces futilités, je ne connaissais que très peu la sphère de la recherche académique. Il serait donc mensonger d'affirmer que la macroéconomie était une passion. Néanmoins, le désir d'apprendre m'a toujours accompagné. Cette opportunité a été le déclencheur d'un travail consciencieux. L'intégrité, la passion, le bienveillance et l'esprit d'Aurélien en ont fait un directeur de thèse exemplaire. Aurélien m'aura laissé une grande liberté pour mes recherches, tout en étant présent à chaque moment le nécessitant. Ma gratitude envers lui est immense.

J'exprime également de profonds remerciements à l'égard de Stéphane Moyen, économiste au sein du pôle de recherche de la Banque Centrale d'Allemagne. Stéphane m'a permis de venir effectuer un stage à la Buba lors

de ma dernière année de doctorat. Ces neuf mois ont bouleversé mon travail de recherche, ainsi que ma vie. Après neuf ans à Lyon, c'est à reculons que je suis allé à Francfort. Malgré un mois ou deux d'un ennui immense, les présences de Stéphane et de mes nouveaux amis ont rendu cette expérience extraordinaire. Les discussions avec Stéphane ont été très prolifiques et instructives. Merci pour tout.

Des remerciements : à Sciences Po Aix et Céline Gimet pour m'avoir accueilli lors de mon année d'ATER à Aix-en Provence. À la Banque Centrale d'Allemagne et son centre de recherche pour ces mois de stage. Évidemment, au laboratoire de recherche GATE à Ecully. Des remerciements particuliers à : Bruno, Quentin, Nelly, Tai, Aude, Frédéric, Béatrice et Yamina. Aux professeurs, chercheurs et affiliés : Mateus, Julien, Rym, Lise, Camille, Mathilde, Walid, Frédéric, Sonia, Béatrice, Stéphane, Alain.

Par ailleurs, le GATE m'aura fait découvrir des amis formidables. Sans ordre de préférences : Yohann, Vincent, Liza, Tatiana, Morgan, Claire, Julien, Thomas, Rémi, Marius, Maxime, Chloé, Clément, Tidiane, Jocelyn, Charlotte, Sylvain, Benjamin et Wilfried. Que de rires, de moments incroyables. Cette thèse aurait été si fade sans vous.

Merci à mes amis : Pierre-Louis, Pierre-Antoine, Pierre, Guillaume, Romain, Thomas, Morgann et Jonathan. Merci également à tous mes autres amis, connaissances et collègues que je ne cite pas ici.

Merci à Angélique qui m'a accompagné pendant ces années. Ta contribution est bien plus importante que tu ne pourrais le penser.

Merci à mes amis de Francfort : Irakli, Elias, Gaia et Renata. Et quelle rencontre avec Ieva. Ces mois ont été si intenses, tant d'émotions.

Je remercie ma famille, mais aussi, Cécile, Patrick, Christophe, Delphine et

leurs enfants.

Je termine ces remerciements par mes parents et mon frère. Mon parcours n'est pas dû au hasard. Vous êtes la cause de ma réussite. Je pense avoir reçu une éducation idéale de la part de parents aimants, respectueux et dévoués. Maman et Papa, vos forces de caractère ont été des repères à chaque moment de doute. Tant de sacrifices. Ce travail, sera, je l'espère, le retour que vous espériez. Mon frère, tu es et resteras toujours mon exemple, chimérique et brillant.

Ce travail est dédié en la mémoire de mon regretté grand-père, Rémi.

Contents

Introduction Française	13
Introduction	28
1 Quantitative Easing and Excess Reserves	41
1.1 Introduction	42
1.2 The Baseline Model	47
1.2.1 Commercial banks	47
1.2.2 Shadow banks	52
1.2.3 Aggregation	56
1.2.4 Households and Production	57
1.2.5 Market clearing and Equilibrium conditions	62
1.3 Experiments	62
1.3.1 Calibration	65
1.3.2 Financial crisis experiments	68
1.3.3 Quantitative Easing experiments	70
1.4 Conclusion	82
Appendix A	84

<i>CONTENTS</i>	5
A.1 Additional Figures	84
A.2 Model Equations	93
A.3 Commercial and Shadow Banks' Problems	97
A.4 Data	101
2 New Evidence on the Effects of Quantitative Easing	102
2.1 Introduction	103
2.2 Framework, Estimation and Selection methods	107
2.2.1 Framework	107
2.2.2 Estimation and Selection methods	109
2.3 Data and Forecasting Exercise	112
2.3.1 Data	112
2.3.2 Preliminary discussion	113
2.3.3 Forecasting exercise	115
2.4 Policy analysis	118
2.4.1 Identification of structural shocks	118
2.4.2 Impulse responses analysis	121
2.5 Conclusion	124
Appendix B	126
B.1 Figures	126
B.2 Data	152
B.3 Bayesian linear Gaussian state-space models	153
B.3.1 Bayesian filtering	153
B.3.2 Reminder: Multivariate Gaussian properties	155
B.3.3 Kalman Filter	160

<i>CONTENTS</i>	6
B.3.4 Kalman Smoother	167
B.4 Kalman filter and smoother for multiple TVP-FAVAR	171
B.5 Identification of structural shocks	177
3 State-contingent Forward Guidance	192
3.1 Introduction	193
3.2 The Model	196
3.2.1 The Labor Market	197
3.2.2 The Sequence of Events	197
3.2.3 The Representative Household	198
3.2.4 Firms	199
3.2.5 The Monetary and Fiscal Authorities	202
3.2.6 Market Clearing	202
3.2.7 State-contingent Forward Guidance	203
3.3 Quantitative Evaluations	205
3.3.1 Calibration	205
3.3.2 State-contingent Forward Guidance Experiments	207
3.3.3 Sensitivity analysis	216
3.4 Conclusion	217
Appendix C	218
C.1 Additional Figures	219
C.2 Technical Appendix	226
C.2.1 Equations	226
C.2.2 Solution method	232
Bibliography	249

List of Figures

1.1	Capital quality shock	69
1.2	Securization crisis	71
1.3	Quantitative Easing: public intermediation	74
1.4	Quantitative Easing: public intermediation, cont.	75
1.5	Quantitative Easing: interventions on different markets	76
1.6	Quantitative Easing through Excess Reserves	79
1.7	Quantitative Easing through Excess Reserves, cont.	80
A.1	Assets of the Federal Reserve	84
A.2	Liabilities of the Federal Reserve	85
A.3	Share of excess reserves in total assets, for all Federal Reserve banks	85
A.4	Main interest rates: monetary policy in the US	86
A.5	Cyclical components of business loans and ABS	86
A.6	US detrended data from 2007Q4 to 2017Q4	87
A.7	Capital quality shock, cont.	87
A.8	Securitization crisis, cont.	88
A.9	Securitization crisis, cont.	88

A.10 Capital quality shock, changes in Ξ	89
A.11 Public intermediation (GK) vs Excess Reserves	89
A.12 Public intermediation (GK) vs Excess Reserves, cont.	90
A.13 First scenario (fixed rate), changes in the value of Θ	90
A.14 Excess Reserves, different scenarios	91
A.15 Excess Reserves, different scenarios, cont.	91
A.16 Second scenario (market rate), changes in the value of Θ	92
B.1 Decaying importance in the estimation of covariances	126
B.2 Time-varying probabilities $\pi_{i,t t}$	127
B.3 TOP 1: CPI to QE shocks: Cholesky & BMA	128
B.4 TOP 1: GDP to QE shocks: Cholesky & BMA	128
B.5 TOP 1: QE to QE shocks: Cholesky & BMA	129
B.6 TOP 1: CPI to QE shocks: Cholesky & BMS	129
B.7 TOP 1: GDP to QE shocks: Cholesky & BMS	130
B.8 TOP 1: QE to QE shocks: Cholesky & BMS	130
B.9 TOP 1: CPI to QE shocks: Zero, Sign restrictions & BMA	131
B.10 TOP 1: GDP to QE shocks: Zero, Sign restrictions & BMA	131
B.11 TOP 1: QE to QE shocks: Zero, Sign restrictions & BMA	132
B.12 TOP 1: CPI to QE shocks: Zero, Sign restrictions & BMS	132
B.13 TOP 1: GDP to QE shocks: Zero, Sign restrictions & BMS	133
B.14 TOP 1: QE to QE shocks: Zero, Sign restrictions & BMS	133
B.15 TOP 1: Average IRF of QE structural shocks during QE1 with BMA	134
B.16 TOP 1: Average IRF of QE structural shocks during QE1 with BMS	134
B.17 TOP 1: Average IRF of QE structural shocks during QE2 with BMA	135
B.18 TOP 1: Average IRF of QE structural shocks during QE2 with BMS	135

B.19	TOP 1: Average IRF of QE structural shocks during QE3 with BMA	136
B.20	TOP 1: Average IRF of QE structural shocks during QE3 with BMS	136
B.21	TOP 1: Accumulated IRF of QE structural shocks during QE1 with BMA	137
B.22	TOP 1: Accumulated IRF of QE structural shocks during QE1 with BMS	137
B.23	TOP 1: Accumulated IRF of QE structural shocks during QE2 with BMA	138
B.24	TOP 1: Accumulated IRF of QE structural shocks during QE2 with BMS	138
B.25	TOP 1: Accumulated IRF of QE structural shocks during QE3 with BMA	139
B.26	TOP 1: Accumulated IRF of QE structural shocks during QE3 with BMS	139
B.27	VAR: CPI to QE shocks: Cholesky & BMA	140
B.28	VAR: GDP to QE shocks: Cholesky & BMA	140
B.29	VAR: QE to QE shocks: Cholesky & BMA	141
B.30	VAR: CPI to QE shocks: Cholesky & BMS	141
B.31	VAR: GDP to QE shocks: Cholesky & BMS	142
B.32	VAR: QE to QE shocks: Cholesky & BMS	142
B.33	VAR: CPI to QE shocks: Zero, Sign restrictions & BMA	143
B.34	VAR: GDP to QE shocks: Zero, Sign restrictions & BMA	143
B.35	VAR: QE to QE shocks: Zero, Sign restrictions & BMA	144
B.36	VAR: CPI to QE shocks: Zero, Sign restrictions & BMS	144
B.37	VAR: GDP to QE shocks: Zero, Sign restrictions & BMS	145
B.38	VAR: QE to QE shocks: Zero, Sign restrictions & BMS	145

B.39	VAR: Average IRF of QE structural shocks during QE1 with BMA	146
B.40	VAR: Average IRF of QE structural shocks during QE1 with BMS	146
B.41	VAR: Average IRF of QE structural shocks during QE2 with BMA	147
B.42	VAR: Average IRF of QE structural shocks during QE2 with BMS	147
B.43	VAR: Average IRF of QE structural shocks during QE3 with BMA	148
B.44	VAR: Average IRF of QE structural shocks during QE3 with BMS	148
B.45	VAR: Accumulated IRF of QE structural shocks during QE1 with BMA	149
B.46	VAR: Accumulated IRF of QE structural shocks during QE1 with BMS	149
B.47	VAR: Accumulated IRF of QE structural shocks during QE2 with BMA	150
B.48	VAR: Accumulated IRF of QE structural shocks during QE2 with BMS	150
B.49	VAR: Accumulated IRF of QE structural shocks during QE3 with BMA	151
B.50	VAR: Accumulated IRF of QE structural shocks during QE3 with BMS	151
3.1	Impacts of SCFG and ZLB	210
3.2	Effects of SCFG on output	212
3.3	Extracted discount factor shocks	214
3.4	Counterfactual analysis	215
C.1	Alternative exit thresholds	219
C.2	Alternative entry threshold	220
C.3	Effects of SCFG on output, atlternative entry threshold	221

C.4	Alternative calibration, C2	222
C.5	Alternative calibration, C3	223
C.6	Alternative calibration, C4	224
C.7	Effects of SCFG on output, atlernative calibrations	225

List of Tables

1.1	Parameters values	67
A.1	Data Chapter 1	101
2.1	Top 10 of model specifications	118
2.2	Identification scheme: lower triangular zero restrictions	120
2.3	Identification scheme: mixture of zero and sign restrictions	120
B.1	Data Chapter 2	152
3.1	Parameters of the benchmark model	206
3.2	Alternative calibration	216

Introduction Française

Au cours de la dernière décennie, la Réserve Fédérale Américaine (FED) a déployé de nouveaux instruments monétaires pour faire face à la Grande Récession (GR) et à la “Zero Lower Bound” (ZLB). Un épisode de ZLB est caractérisé par une situation où le niveau du taux nominal de court terme atteint sa borne à zéro. Parmi les nouveaux instruments, l’assouplissement quantitatif (“Quantitative Easing”, QE) a fait l’objet de nombreuses discussions dans la littérature. Le QE peut être défini, au sens large, comme un programme d’achat d’actifs à grande échelle mené par une banque centrale.

En outre, le Federal Open Market Committee (FOMC) a eu recours à des déclarations publiques sur l’orientation future de la politique monétaire, un outil de communication appelé “forward guidance” (FG). Le FG est défini comme *qualitatif/quantitatif et dépendant du temps/dépendant de l’état*.¹ La littérature fait la distinction entre, les déclarations *Delphic* ou *Odyssean*. Une communication Delphic se produit généralement en période d’essor. Elle repose sur un cadre monétaire explicite, comme une règle de Taylor traditionnelle. En d’autres termes, la banque centrale communique la situation fu-

¹Une liste détaillée des communications pendant la GR se trouve à <https://www.federalreserve.gov/monetarypolicy/timeline-forward-guidance-about-the-federal-funds-rate.htm>.

ture de la politique monétaire et ajuste l'instrument de politique monétaire en fonction de l'activité économique. Le guidage Delphic est perçu comme une mauvaise nouvelle en cas de ZLB. Inversement, une orientation Odyssean est généralement donnée en période de crise et est interprétée comme un engagement de la banque centrale à s'écarter de *ou* modifier la règle de politique monétaire en vigueur.

Par exemples, la communication du FOMC de mars 2009 : "(...) *economic conditions are likely to warrant exceptionally low levels of the federal funds rate for an extended period.*" est un FG Delphic, dépendant du temps. Celle de décembre 2012 : "(...) *this exceptionally low range for the federal funds rate will be appropriate at least as long as the unemployment rate remains above 6-1/2 percent.*" est un FG Odyssean, dépendant de l'état de l'économie.

Comme les politiques conventionnelles, ces politiques non conventionnelles sont susceptibles de soutenir l'activité économique et assurer l'ancrage de l'inflation à une cible. De plus, elles peuvent potentiellement aplatir la courbe des taux. Par conséquent,

**DANS QUELLES MESURES LES POLITIQUES MONÉTAIRES NON
CONVENTIONNELLES SOUTIENNENT-ELLES L'ACTIVITÉ ÉCONOMIQUE ?
QUELS EN SONT LES CANAUX DE TRANSMISSION ?**

Cette thèse revient sur les tenants et aboutissants de ce changement de régime monétaire américain. Aux États-Unis, une politique monétaire conventionnelle consiste à cibler le taux des fonds fédéraux (Federal Funds rate, FFR). La cible du taux FFR est normalement comprise entre le taux d'escompte et le taux de dépôt, un système considéré comme un *corridor*. L'objectif d'une politique monétaire conventionnelle est de faire correspondre le taux FFR effectif sur

le marché des prêts interbancaires à la cible du FOMC. Le FFR effectif est le coût moyen des opérations quotidiennes de prêt des institutions de dépôt entre elles. Les prêts du jour au lendemain sont des actifs à court terme et liquides détenus par les institutions de dépôt pour couvrir les réserves obligatoires demandées par la banque centrale, mais aussi les opérations et les besoins commerciaux. Au-delà des réserves obligatoires, la demande de réserves diminue avec le coût d'opportunité des réserves excédentaires. En d'autres termes, les institutions de dépôt réalisent toujours au moins un profit nul en se prêtant des fonds sur le marché interbancaire. La limite supérieure du corridor est le taux primaire (d'escompte) auquel les banques peuvent obtenir des liquidités auprès de la FED. Les banques ne sont évidemment pas incitées à prêter du jour au lendemain à un coût plus élevé que le taux d'escompte. La limite inférieure du corridor est le taux de dépôt sur les réserves excédentaires. Ce dernier était nul aux États-Unis avant la crise financière de 2007. Ainsi, le taux FFR effectif se situe dans une région de demande tout à fait inélastique (sur le plan prix-quantité) ; sur une courbe de demande des réserves ayant une pente négative. De plus, la rareté des réserves est un levier pour la FED pour influencer le taux FFR effectif. Ce levier est déclenché par des opérations d'open market (OMO) sur le marché des prêts interbancaires. Par exemple, une politique monétaire expansionniste, procède comme suit : Tout d'abord, la FED fixe une cible visant à réduire le taux FFR effectif, donc inférieure au taux effectif. Par la suite, la FED augmente les liquidités sur le marché interbancaire par : des opérations de pension ("repurchase agreements"), où l'achat de titres par l'intermédiaire du System Open Market Account (SOMA). Par conséquent, ces opérations exercent une pression à la baisse sur le taux FFR effectif afin qu'il s'aligne à la cible FOMC. Toutefois, au moins trois conditions doivent être remplies pour que ces

politiques soient efficaces : la rareté des réserves, la stabilité des conditions sur le marché interbancaire et l'absence de contraintes sur le niveau du taux FFR effectif (pas de ZLB).

Un krach financier modifie inévitablement les mécanismes de transmission d'une politique monétaire classique. Au début de la crise financière de 2007, le marché interbancaire s'est *resserré* en raison de la faillite d'un important "primary dealer", Lehman Brothers.² En décembre 2008, la FED a fixé l'objectif du FFR à un niveau proche de la ZLB, entre 0 et 25 points de base. Pendant ce temps, elle s'est engagée dans des programmes de *sauvetage*. Entre septembre 2008 et janvier 2010, la FED a fourni un grand nombre de facilités de liquidité à la "discount window", mais aussi directement sur les marchés financiers.³

En décembre 2008, la FED a annoncé le premier de ses programmes de QE (le dernier cycle (QE3) ayant pris fin en octobre 2014). Par ailleurs, le système monétaire est passé d'un système *corridor* à un système *floor*. En octobre 2008, la FED a commencé à verser des intérêts sur les réserves obligatoires et excédentaires pour reprendre les rênes des taux à court terme.⁴ De janvier 2009 à décembre 2015, les intérêts sur les réserves obligatoires (IORR) et les intérêts sur les réserves excédentaires (IOER) ont été fixés à la limite supérieure

²Un "primary dealer" est une contrepartie commerciale de la FED de New York pour la conduite des OMO. Les "primary dealers" sont obligés de participer aux OMO, aux enchères de titres américains et de transmettre des informations sur le marché.

³Discount window programs: Term Auction Facility (TAF), Primary Dealer Credit Facility (PDCF), and Term Securities Lending Facility (TSLF). Market liquidity facilities: Commercial Paper Funding Facility (CPFF), Asset-Backed Commercial Paper Money Market Mutual Fund Liquidity Facility (AMLF), Money Market Investor Funding Facility (MMIFF), and the Term Asset-Backed Securities Loan Facility (TALF). Voir plus d'informations à <https://www.federalreserve.gov/monetarypolicy/bst.htm>

⁴Le Congrès et la FED ont initialement voté l'autorisation de verser des intérêts sur les réserves en vertu de la Financial Service Regulatory Relief Act 2006. La mise en œuvre était normalement prévue pour le 1er octobre 2011. Réf : <https://www.congress.gov/109/plaws/pub1351/PLAW-109pub1351.pdf>.

de l'objectif FFR (voir Figure A.4 de l'Annexe A). Dans ces circonstances, les institutions de dépôt étaient clairement incitées à épargner des réserves excédentaires, tandis que d'autres teneurs de marché, comme les entreprises parrainées par l'État et les courtiers en valeurs mobilières, non admissibles aux paiements de réserves, avaient des possibilités d'arbitrage. En particulier, ils ont pu acheter des réserves à des coûts dérisoires et les prêter sur des marchés secondaires. Dans un tel environnement, le taux sur les réserves est la limite inférieure (*floor*) au taux FFR effectif; alors que l'offre de réserves s'ajuste librement à la demande de marché. Par conséquent, le prix d'équilibre des réserves se situe maintenant dans une région de demande élastique, sur une courbe de demande quasiment plate. Dans ce système monétaire, les réserves sont abondantes, et les OMO conventionnelles plutôt inefficaces.

Ainsi, le QE est apparu comme une solution intéressante. Premièrement, le QE est qualitatif en éliminant les actifs toxiques des marchés financiers. Deuxièmement, le QE était principalement soutenu par des réserves excédentaires, ce qui entraîne une pression à la baisse sur les taux à court terme. Troisièmement, le QE réduit les rendements à long terme, donc aplatit la courbe des taux. Enfin, le QE offre une marge de manœuvre supplémentaire aux institutions de dépôt pour la réaffectation des portefeuilles, le refinancement ou le renflouement des dettes et le recours à la recapitalisation. Néanmoins, le QE et les programmes de sauvetage des liquidités ont eu un effet dramatique sur la taille du bilan de la FED (cf. figures A.2, A.2 et A.3 de l'Annexe A).

Dans la littérature, les contributions empiriques soulignent que les programmes de QE ont des effets considérables sur les conditions financières et l'économie réelle. En particulier, les effets financiers du QE sont identifiés, principalement, par trois approches : les études d'événements, les régressions

sur séries chronologiques et les modèles dynamiques de structure par terme des taux d'intérêts.

Les études d'événements mesurent les réactions des prix/taux des actifs autour de l'annonce du QE, sur des données à très haute fréquence. Avec cette méthode, Krishnamurthy & Vissing-Jorgensen (2011), Gagnon et al. (2011) ou Swanson (2011) constatent une baisse moyenne des rendements à long terme de 100 points de base due au QE1 (décembre 2008 à juin 2010) et 15 points de base pour le QE2 (octobre 2010 à décembre 2012), quel que soit le type d'actifs acheté. Dans les analyses de régression temporelle de Hancock & Passmore (2011), Hamilton & Wu (2012) ou Joyce et al. (2012), les impacts sont deux fois plus faibles, environ 50 points de base en moyenne pour le QE1. Enfin, les modèles dynamiques de structure par terme des taux d'intérêts de Christensen & Rudebusch (2012) et de Bauer & Rudebusch (2013) indiquent une diminution moyenne d'environ 90 points de base sur les primes de risque des bons du Trésor à 10 ans pendant le QE1.

La littérature empirique sur les effets financiers du QE est abondante, alors que peu de contributions examinent les effets macroéconomiques. Les modèles vectoriels autorégressifs (VAR) dans Gambacorta et al. (2014) et Weale & Wieladek (2016) ou le modèle VAR à paramètres variables de Baumeister & Benati (2013) fournissent quelques résultats. Ces études indiquent que le QE1 a contribué à la croissance de l'IPC entre 0,90% et 1,50% et à la croissance du PIB entre 1,08% et 1,40%. Notre chapitre 2 met en évidence les problèmes de surestimation des effets macroéconomiques du QE.

Les effets importants du QE contrastent fortement avec la neutralité relative suggérée par la littérature théorique. L'affirmation d'une neutralité est reprise par les points de vue de Curdia & Woodford (2011) et Williamson (2012, 2016).

Selon Williamson (2018), les programmes de QE consistent en une transformation d'échéances, des actifs liquides de court terme (réserves) en actifs illiquides de long terme (principalement des obligations d'État), un processus habituellement censé être la fonction principale des banques privées. En tant que tel, il soutient que le QE devrait être neutre, selon le théorème de Modigliani (1977). Curdia & Woodford (2011) soutiennent, également, une quasi neutralité en déclarant que le QE est une politique *non pertinente*.⁵ Leurs contributions montrent que la politique monétaire est optimale lorsque le taux d'intérêt sur les réserves est égal au taux cible de la banque centrale, conformément à une règle monétaire à la Friedman & Schwartz (1963). Ils ajoutent que l'offre de réserves doit ensuite être ajustée pour que cette proposition s'applique en tout temps. Par conséquent, ils formulent deux conditions pour lesquelles les programmes de QE ne sont pas pertinents : (i) si la banque centrale échange des réserves contre des titres du Trésor, et (ii) si le QE ne modifie pas les attentes des agents privés concernant les mesures monétaires futures. Néanmoins, ils admettent que ces conditions peuvent être violées en présence de marchés *suffisamment perturbés*. Dans de telles circonstances, les programmes de QE sont susceptibles d'être efficaces, si et seulement s'ils ciblent les actifs privés.

Compte tenu de la contraction observée des marchés financiers, nous pouvons affirmer que les programmes de QE ont probablement été impactants pendant la GR. Pour étayer ce point de vue, les chercheurs s'appuient sur les

⁵Bernanke (2012) définit le QE *pure* comme : l'accent mis sur le niveau des réserves sans porter d'intérêt à la composition des actifs détenus par la banque centrale. Inversement, il affirme que le QE aux États-Unis était plutôt une politique d'assouplissement qualitatif en raison de sa concentration sur la composition des actifs achetés. Ref : <https://www.federalreserve.gov/newsevents/speech/bernanke20090113a.htm>. Dans cette thèse, nous nous en tenons à une définition large du QE en tant que programme d'achat d'actifs à long terme mené par une banque centrale.

mécanismes théoriques des frictions financières. Dans l'hypothèse de l'habitat préféré (Modigliani & Sutch (1966), Vayanos & Vila (2009)) et selon la théorie de rebalancements de portefeuille de Tobin (1969), le QE est efficace. Dans le premier cas, les marchés sont segmentés en fonction des préférences des investisseurs quant à l'échéance des actifs. Pour que les investisseurs soient disposés à investir dans d'autres segments, le risque encouru doit au moins être compensé par le rendement attendu de l'actif. Dans ce contexte, le QE favorise la demande de substituts imparfaits parce qu'il modifie l'offre de l'actif visé. Ainsi, le rendement de ces proches substituts se réduit.

Les modèles macroéconomiques fondateurs à frictions financières sont les suivants : Kiyotaki & Moore (1997) (KM) et Bernanke et al. (1999) (BGG). Dans BGG, les frictions émergent d'un problème d'aléa moral entre les banques et les entrepreneurs. Dans ce cadre, les banques font face à un coût de surveillance dans l'esprit de Townsend (1979). Cette vérification coûteuse de l'état induit une prime de financement externe endogène, entre le coût de l'endettement extérieur et le coût d'opportunité du financement interne par fonds propres. Chaque entrepreneur s'engage dans un projet qui nécessite un financement externe auprès de banques. Toutefois, les remboursements des prêts sont incertains, car soumis à un choc idiosyncratique. En outre, les entrepreneurs observent parfaitement les rendements du capital, alors que les banquiers ne le peuvent pas. Les banquiers peuvent néanmoins payer un coût de surveillance pour éviter les *lemons*.⁶

Les deux parties concluent donc un accord ayant la forme d'un contrat de prêt. Dans un cas, les entrepreneurs font faillite lorsque la valeur du choc idiosyncratique est supérieure à un seuil spécifié dans le contrat de prêt. Dans

⁶Les *Lemons* sont des voitures défectueuses dans la célèbre étude d'Akerlof (1978)

cette situation, les banquiers collectent le reste du capital net des coûts de surveillance. Dans l'autre cas, les emprunteurs réalisent des bénéfices nets des intérêts de la dette payés aux banquiers. En période de crise, les banquiers doivent donc compenser la probabilité croissante de défaillance par une hausse du coût des crédits. Par conséquent, la situation nette (net worth) des entrepreneurs implique un mécanisme d'accélérateur financier. Imaginons une contraction de l'économie qui détériorerait la situation nette des entrepreneurs. Par la suite, le ralentissement de l'économie est amplifié par l'augmentation de la prime de financement externe, ce qui réduit encore la demande globale. En d'autres termes, les frictions financières induisent des effets de rétroaction.

Christiano et al. (2010, 2014) ajoutent des frictions à la BGG dans un modèle canonique d'équilibre général dynamique monétaire (DSGE), et trouvent que les chocs exogène de risque financiers cause des "mean-preserving interest rate spreads", et sont donc des vecteurs importants des fluctuations macroéconomiques. Dans un tel cadre, le QE est efficace parce qu'il libère une partie de la pression sur les écarts de taux, ce qui facilite les conditions financières des entreprises de diverses façons. D'abord, en augmentant simultanément le prix de l'actif acheté et celui des proches substituts, donc en soutenant leur situation nette. Le renforcement de la situation nette soutient la solvabilité, ce qui permet de contenir la hausse de la prime de financement. Au cours des prochaines périodes, la demande de crédit augmentera donc avec l'effet de rétroaction. Deuxièmement, en soutenant directement l'offre de crédit. En effet, l'offre de réserves excédentaires permet aux banquiers d'émettre des dépôts, ce qui favorise l'offre de nouveaux prêts. Ces effets de richesse et de crédit sont l'essence même du canal du bilan du QE.

KM conçoivent, eux, des frictions financières sous forme de contraintes de

garantie. Dans leur modèle de base, l'économie est peuplée de deux types d'agents, à savoir des agriculteurs et des cueilleurs. L'économie comprend deux biens, les biens durables (terres) et les biens non durables (fruits). La terre et la main-d'œuvre sont les intrants pour la production de fruits. KM imposent aux agriculteurs d'être impatients et d'emprunter des fonds à des cueilleurs patients. Une deuxième hypothèse rend la technologie des agriculteurs idiosyncrasique. En d'autres termes, les compétences des agriculteurs sont limitées, de sorte que la production est à la hauteur de l'effort de travail. Par ailleurs, les agriculteurs sont autorisés à renoncer au travail et à répudier leurs dettes. En conséquence, les cueilleurs exigent des agriculteurs qu'ils mettent en gage leurs terres (en tout ou en partie). Pour que les deux parties puissent participer à ce contrat, la contrainte de garantie doit faire en sorte que tout remboursement de dette soit d'une valeur au moins inférieure à la valeur marchande d'un terrain. A l'équilibre, les agriculteurs sont donc fortement endettés.

KM ajoutent alors un capital reproductible (arbres) dans la fonction de production. Les arbres font l'objet d'une dépréciation et ne peuvent être utilisés comme garantie. Ces caractéristiques permettent aux remboursements de dettes d'être strictement inférieurs à la valeur marchande des terres garanties. Dans un tel cadre, un choc négatif de productivité provoque une rétroaction. La situation nette des agriculteurs diminue en fonction de la chute des prix des terres. Par conséquent, les agriculteurs sont limités dans leurs dépenses d'investissement en raison du resserrement des contraintes de garantie. Dans les périodes suivantes, les revenus du capital chutent en raison de la pénurie d'investissements, ce qui réduit encore la situation nette des agriculteurs. Dans une étude connexe, Iacoviello (2005) étudie les implications macroéconomiques de ces contraintes de garantie sur le marché immobilier. Par la suite, Guerrieri

& Iacoviello (2017) montrent que les contraintes de garantie liées au logement induisent des asymétries macroéconomiques pendant les cycles d'expansion et d'effondrement du marché du logement. En période de boom immobilier, les contraintes de garantie sont relâchées, de sorte que la richesse immobilière contribue à peine à la croissance de la consommation. En période d'effondrement, les contraintes se resserrent et la chute de la valeur du patrimoine immobilier renforce le ralentissement macroéconomique par le biais d'une boucle similaire de rétroaction. Ces résultats suggèrent que le QE est plus efficace pendant les périodes d'effondrement de l'économie. A contrario, le relâchement des contraintes réduisent inévitablement l'efficacité de QE.

Dans un modèle QE DSGE, Gertler & Karadi (2011) (GK) inversent un problème d'aléa moral classique. Dans ce cadre, la prime externe ne découle pas de l'incertitude quant à la qualité des entreprises mais de l'avidité de certains banquiers. Les ménages participent au financement des banques via des dépôts, à condition que les banquiers respectent des contraintes de bilan. Cependant, certains banquiers se comportent mal en détournant des actifs parce que leur durée de vie est limitée. Ainsi, une contrainte incitative doit s'imposer pour que les ménages déposent des dépôts. Cette dernière stipule que la valeur espérée des banques doit être au moins supérieure à la valeur des actifs détournés. La contrainte incitative induit un ratio de levier qui fixe une limite endogène à la taille du bilan des banquiers.

Ces frictions limitent l'arbitrage des actifs sur les marchés financiers et toute baisse exogène des prix des actifs renforce les contraintes de bilan. Cette baisse des prix des actifs produit un mécanisme d'accélérateur financier endogène qui, à son tour, gèle l'offre de crédit des intermédiaires privés. GK étudient les effets du QE dans un tel environnement. Ils conçoivent le QE comme une

intermédiation publique de l'offre de crédit : la banque centrale émet des engagements à court terme soumis à des coûts négligeables. GK trouvent qu'il est optimal pour le bien-être des agents de contourner les frictions financières sur les marchés privés par le QE, si et seulement si, les coûts d'efficience de l'intermédiation de la banque centrale ne compensent pas les gains de bien-être induits par ce QE. Notre chapitre 1 propose un modèle d'équilibre général avec des frictions financières couplant Gertler & Karadi (2011, 2013) et Guerrieri & Iacoviello (2017) pour évaluer les impacts d'une abondance de réserves excédentaires dans l'économie.

La politique du FG est un autre outil pour s'attaquer à la situation de la ZLB. Les déclarations publiques de la banque centrale peuvent être perçues de deux façons opposées. D'une part, le FG peut améliorer les anticipations des agents quant à l'orientation future du taux d'intérêt nominal. Les agents s'attendent à d'autres mesures de stimulation, c'est-à-dire à une baisse future des taux d'intérêt à court terme. Ainsi, la courbe des taux s'aplatit, ce qui entraîne une hausse de l'inflation. Dans ces circonstances, le FG stimule l'économie. D'autre part, les agents peuvent considérer que ces déclarations révèlent de mauvaises nouvelles sur l'état futur de l'économie. L'hypothèse principale est que la banque centrale est mieux informée que les participants du marché. Dans ce cas, une communication exerce une pression à la hausse sur les taux longs, ce qui induit des effets néfastes sur l'activité.

Dans la littérature du FG, les chercheurs sont confrontés à des problèmes d'identification. En effet, le FG se transmet principalement par l'anticipation des agents sur l'orientation future de la politique monétaire et/ou les perspectives macroéconomiques. De ce fait, les études empiriques sont rares. En outre, les modèles théoriques tels que les DSGE, souffrent du *forward guidance puzzle* ;

nommé par Del Negro et al. (2012) suite à leur observation de la surestimation potentielle des effets du FG.

L'étude de Woodford & Eggertsson (2003) développe des justifications théoriques sur l'utilisation du FG. Cependant, la question des effets de la politique monétaire anticipée faisait déjà l'objet d'un débat très animé dans les années 70 et 80. En effet, Mishkin (1982) prouve dans une série d'articles empiriques que l'affirmation de neutralité des politiques monétaires anticipées de Modigliani (1977) ne tient pas. Plus récemment, Gertler & Karadi (2015) ont montré l'importance des chocs inattendus de politique monétaire pour l'identification des mécanismes de transmission monétaire. Ils constatent que les surprises monétaires entraînent de légères pressions sur les taux à court terme, mais qu'elles ont des effets papillon sur les conditions financières. En effet, les primes à terme et les écarts de taux ont tendance à réagir fortement aux surprises qui, à leur tour, affectent l'activité réelle. La contribution importante de cette étude est de s'attaquer au problème de simultanéité des chocs de politiques monétaires : lors d'une période t , les chocs de politiques monétaires pourraient être affectés par des variables financières leurs étant fortement corrélées, mais absentes du modèle. Pour résoudre ce problème, ils utilisent dans un modèle VAR, des informations sur données à haute fréquence pour identifier les surprises monétaires, mais aussi de basse fréquence telles que la production ou l'IPC. La simultanéité disparaît sous hypothèse que les nouvelles quotidiennes sur l'état de l'économie n'affectent pas les déclarations actuelles du FOMC.

L'étude de Campbell et al. (2012) est une référence importante dans la compréhension du FG. Ils affirment que le principal défi du FG, en présence de la ZLB, est de convaincre le public en la capacité de la banque centrale à

tenir ses promesses. Premièrement, ils se demandent si les communications du FOMC contiennent des informations susceptibles de modifier les attentes des agents quant à l'orientation future de la politique monétaire. La réponse collégiale dans la littérature avant le GR est, *oui* (e.g. Gürkaynak et al. (2005)). A cet égard, ils identifient les effets des déclarations du FOMC par l'estimation de deux facteurs : Un facteur dit de cible qui est lié aux variations du taux directeur actuel et un facteur dit de chemin (FG) qui est lié aux variations des taux futurs attendus. Par la suite, ils démontrent qu'une augmentation inattendue du facteur de chemin réduit les prévisions du chômage et augmente les anticipations d'inflation. En d'autres termes, les prévisionnistes croient que le FOMC possède des renseignements plus précis sur l'état de l'économie. Ainsi, ils soutiennent que les déclarations du FOMC contiennent des éléments Delphic jusqu'en 2007. Deuxièmement, ils identifient les effets du FG au moyen d'une règle monétaire classique. Dans ce cadre, le FG est contenu dans un terme d'erreur retardé. Ils constatent que ces deux spécifications (deux facteurs et la règle du taux d'intérêt) expliquent une grande partie des fluctuations de certains actifs financiers (bons du trésor, entreprises). L'inclusion d'une telle règle dans leur DSGE estimée montre que le FG Odyssean est en mesure d'assurer un mandat d'inflation stable à long terme.

Néanmoins, Del Negro et al. (2012) affirment et démontrent que les modèles DSGE surestiment les effets du FG. Ils appellent ce phénomène, le *forward guidance puzzle*. A cet égard, Carlstrom et al. (2015) trouvent que le modèle DSGE standard de Smets & Wouters (2007) produirait une inflation explosive si la ZLB était maintenue pendant neuf trimestres. Toutefois, de nombreuses communications du FOMC ont relayé des informations sur la politique à venir sur plus de huit semestres. Del Negro et al. (2012) résolvent ce problème

dans leur DSGE en assumant que les agents font face à une probabilité de mourir. Ainsi, les changements induits par le FG sur les attentes sont beaucoup moins importants que dans le cas où les agents ont une durée de vie infinie. D'autres articles intéressants, comme Andrade et al. (2015) utilisent un mécanisme d'hétérogénéité de l'information et trouvent que dans le cas d'une communication floue, les Delphic ou Odyssean FG sont moins efficaces. Enfin, Caballero & Farhi (2017) démontrent que le stimulus du FG est atténué lorsque l'économie s'est déjà remise d'une crise. Notre chapitre 3 analyse les effets d'un "state-contingent FG" dans le cadre d'un engagement "Odyssean" sur une cible de taux de chômage.

La thèse est organisée en trois articles de recherche. Le premier chapitre porte sur : "Quantitative Easing and Excess Reserves". Le chapitre 2 présente le papier : "New Evidence on the Effects of Quantitative Easing". Le chapitre 3 est un papier intitulé : "State-contingent Forward Guidance".

Introduction

In the last decade, the US Federal Reserve (FED) has deployed new monetary instruments to tackle the Great Recession (GR) and the zero lower bound (ZLB). A ZLB episode is characterized by a situation where the level of the nominal short-term rate is stuck to zero. Among the new instruments, Quantitative Easing (QE) has been the focus of many discussions in the literature. QE can be broadly defined as a large-scale program of asset purchases conducted by a central bank.

In addition, the Federal Open Market Committee (FOMC) has delivered public statements to the future stance of monetary policy; a communication tool called forward guidance (FG). FG is defined as *qualitative/quantitative* and *time-dependent/state-contingent*.⁷ The literature distinguishes between, *Delphic* or *Odyssean* statements. A Delphic communication generally occurs in normal times. It is based on an explicit monetary framework, such as a traditional Taylor rule. In other words, the central bank communicates how the monetary policy will stand in the future and likely adjusts the policy instrument in regards with the economic activity. Delphic guidance is therefore usually perceived as bad news under a ZLB. Conversely, an Odyssean guidance is

⁷A list of the US FOMC FG can be found at <https://www.federalreserve.gov/monetarypolicy/timeline-forward-guidance-about-the-federal-funds-rate.htm>

generally initiated during trouble times and is interpreted as a commitment for the central bank to deviate from *or* to change the actual policy rule.

Factually, the FOMC statement of March 2009: “(...) *economic conditions are likely to warrant exceptionally low levels of the federal funds rate for an extended period.*” is a time-contingent Delphic FG. The communication of December 2012: “(...) *this exceptionally low range for the federal funds rate will be appropriate at least as long as the unemployment rate remains above 6-1/2 percent.*” is an Odyssean state-contingent FG.

Like conventional policies, these unconventional policies are likely to support the economic activity and to anchor inflation to a target. They also potentially flatten the term-structure of interest rates. Hence,

**TO WHAT EXTENT ARE UNCONVENTIONAL MONETARY POLICIES
SUPPORTING THE ECONOMIC ACTIVITY? WHAT ARE THE TRANSMISSION
CHANNELS TO THE ECONOMY?**

This thesis revisits the ins and outs of this change in the US monetary regime. In the US, a conventional monetary policy consists in the FOMC targeting of the Federal Funds rate (FFR). The FFR target is normally comprised between the discount rate and the deposit rate, a system viewed as a *corridor*. The purpose of a conventional monetary policy is to match the effective FFR in the interbank lending market with the FOMC target. The effective FFR is the average cost of the daily lending operations of depository institutions among themselves. Overnight loans are short-term and liquid assets held by depository institutions to cover the central bank's reserve requirements and business needs. Beyond reserve requirements, the demand for reserves is decreasing in the opportunity cost of the holdings of excess reserves. In other words, depository

institutions make at worst zero profit by lending to each other in the overnight market. The upper bound of the corridor is the primary (discount) rate at which banks can obtain liquidities from the FED. Consequently, there is obviously no incentive for banks to lend overnight at a higher cost than the discount rate. The lower bound of the corridor is the deposit rate in excess reserves. The latter was zero in the US before the 2007 financial crisis. Thus, the effective FFR lies in a quite inelastic demand region (in the price-quantity plane); on a downward-sloping demand curve of reserves. In addition, the scarcity of reserves is a lever for the FED to influence the effective FFR. This lever is triggered through open-market operations (OMO) in the interbank lending market. For instance, a typical expansionary monetary policy proceeds as follows: First, the FED sets a target below the current effective FFR. Second, the FED increases liquidities using either repurchase agreements (temporary borrowing) or purchases of short-term securities through the System Open Market Account (SOMA). As a result, these operations exert a downward pressure on the effective FFR. However, there are at least three conditions for these operations to be effective: the scarcity of reserves, steady conditions on the overnight market and the absence of constraints on the level of the effective FFR (no ZLB).

A financial crash inevitably alters the transmission mechanisms of a conventional monetary policy. At the onset of the 2007 financial crisis, the interbank market squeezed because of the failure of an important primary dealer, Lehman Brothers.⁸ Hence, the FED served as a the lender of last resort. In December 2008, the FED set the FFR target close to the ZLB, between 0 to 25

⁸A primary dealer is a trading counterparty of the New York FED for the conduction of OMO. Primary dealers are forced to participate in OMO, in auctions of US securities and transmit market information.

basis points. Meanwhile, it engaged in extraordinary rescue programs. Between September 2008 and January 2010, the FED provided a large amount of exceptional liquidity facilities in the discount window, but also directly in the financial markets.⁹

In December 2008, the FED announced the first of its QE programs (the last round (QE3) ceased in October 2014). In addition, it switched from a *corridor* to a *floor* system. In October 2008, it started paying interest on required and excess reserves, to take up the reins of the short-term rates.¹⁰ From January 2009 to December 2015, the interest on required reserves (IORR) and the interest on excess reserves (IOER) were fixed to the upper limit of the FFR target (see Figure A.4 in the Appendix A). Under these circumstances, depository institutions had clear incentives to put excess reserves aside, while other market makers such as government-sponsored enterprises and security brokers and dealers, not eligible to reserve payments, had arbitrage opportunities. In particular, they were able to purchase reserves at roughly no cost and lend it in secondary markets. In such environment, the rate on reserves imposes a lower (*floor*) bound to the effective FFR, while the supply of reserves freely adjusts to the demand in the market. Hence the equilibrium price for reserves now lies in an elastic demand region; on a nearly flat demand curve for reserves. In this system, reserves are abundant; therefore conventional OMO are pretty ineffective.

⁹Discount window programs: Term Auction Facility (TAF), Primary Dealer Credit Facility (PDCF), and Term Securities Lending Facility (TSLF). Market liquidity facilities: Commercial Paper Funding Facility (CPFF), Asset-Backed Commercial Paper Money Market Mutual Fund Liquidity Facility (AMLF), Money Market Investor Funding Facility (MMIFF), and the Term Asset-Backed Securities Loan Facility (TALF). See more information at <https://www.federalreserve.gov/monetarypolicy/bst.htm>

¹⁰The Congress and FED initially voted the authorization to pay interest on reserves by the 2006 Financial Service Regulatory Relief Act. The implementation was normally scheduled on October 1, 2011. Ref: <https://www.congress.gov/109/plaws/pub1351/PLAW-109pub1351.pdf>.

Thus, QE appeared as an interesting solution, for the following reasons. First, QE is qualitative by removing toxic assets from the financial markets. Second, QE was mainly backed by excess reserves, which causes a downward pressure on the short-term rates. Third, QE reduces long-term yields and then flatten the yield curve. Last, QE offers an additional degree of freedom for depository institutions to reallocate portfolios, roll over or bail-out debts and resort to recapitalization. Nevertheless, QE and the liquidity rescue programs had a dramatic effect on the size of the FED's balance sheet (see Figures A.1, A.2 and A.3 in the Appendix A).

In the literature, empirical contributions highlight that QE programs have sizable effects on financial conditions and the real economy. In particular, the financial effects of QE are identified, mainly, by three approaches: event studies, time series regressions and dynamic term structure models.

Event studies measure the market reactions in the asset prices/yields around the announcement of QE. With this method, Krishnamurthy & Vissing-Jorgensen (2011), Gagnon et al. (2011) or Swanson (2011) find, regardless of the type of purchased assets, an average decline in long-term yields of 100 basis points during QE1 (December 2008 to June 2010) and 15 basis points in QE2 (October 2010 to December 2012). In the time regressions analyses of Hancock & Passmore (2011), Hamilton & Wu (2012) or Joyce et al. (2012), the impacts are two times smaller, around 50 basis points on average in QE1. Last, the dynamic term structure models of Christensen & Rudebusch (2012) and Bauer & Rudebusch (2013) find an average decrease of roughly 90 basis points of risk premiums of 10-year Treasury bonds in QE1.

The empirical literature on the financial effects of QE is abundant while few contributions inspect the macroeconomic impacts. The use of vector au-

toregression (VAR) models in Gambacorta et al. (2014) and Weale & Wieladek (2016) or a time-varying parameters VAR model of Baumeister & Benati (2013) provide some insights. These studies inform that the QE1 contributed to the growth of CPI in between 0.90% and 1.50%, and to the growth of GDP between 1.08% and 1.40%. Our Chapter 2 highlights overestimation concerns in the literature.

The sizable effects of QE are in stark contrast with the relative neutrality suggested by the theoretical literature. The neutrality assertion is embraced by the views of Curdia & Woodford (2011) and Williamson (2012, 2016). According to Williamson (2018), QE programs consist of a maturity transformation, from short-term liquid assets (reserves) to long-term illiquid assets (mainly government bonds), a process usually supposed to be the primary function of private banks. As such, he argues that QE should be neutral, according to the Modigliani (1977) theorem. Curdia & Woodford (2011) also support this view by stating that a *pure* QE is an *irrelevant proposition*.¹¹ Their contributions find that monetary policy is optimal when the interest rate on reserves equates the central bank's target rate, in line with a Friedman & Schwartz (1963) rule's view. They add that the supply of reserves must then be adjusted for this proposition to hold at all times. As a result, they formulate two conditions under which QE programs are irrelevant: (i) if the central bank swaps reserves for Treasury securities, and (ii) if QE does not change the expectations of private agents about future monetary actions. Nevertheless, they admit that these conditions can be

¹¹Bernanke (2012) defines a *pure* QE like: The focus on the level of reserves without having interest on the composition of assets held by the central bank. Conversely, he asserts that QE in the US was rather a *credit easing* policy due to its focus on the composition of the purchased assets. Ref: <https://www.federalreserve.gov/newsevents/speech/bernanke20090113a.htm>. In this thesis, we stick to a broad definition of QE as programs of long-term asset purchases conducted by a central bank.

violated in the presence of *sufficiently impaired markets*. In such circumstances, QE programs are likely to be effective, if and only if targeted to private assets.

Given the squeeze in financial markets, we can argue that QE programs have somehow been effective during the GR. To support this view, scholars rely on the theoretical mechanisms of financial frictions. Under the preferred-habitat assumption (Modigliani & Sutch (1966), Vayanos & Vila (2009)) and/or according to Tobin (1969)'s portfolio balance theory of the term structure of interest rates, QE is effective/relevant. In the former, markets are segmented by the preferences of investors over the maturity of assets. For investors to be willing to invest in other segments, the incurred risk must at least be offset by the expected return in the asset. In this context, QE fosters demand for imperfect substitutes because it changes the supply of the targeted asset; therefore, reduces the yields of these close substitutes.

The seminal macroeconomic models of financial frictions are Kiyotaki & Moore (1997) (KM) and Bernanke et al. (1999) (BGG). In BGG, frictions emerge from a moral hazard problem between conflicting banks and entrepreneurs. In this framework, banks face a monitoring cost in the spirit of Townsend (1979). This costly state verification induces an endogenous external finance premium, between the cost of raising external debt and the opportunity cost of internal equity financing. More precisely, each entrepreneur enters a project that requires external funding. Entrepreneurs borrow funds from bankers. However, loan repayments are uncertain because the projects' returns are subject to an idiosyncratic shock. In addition, entrepreneurs perfectly observe returns, while bankers cannot. Bankers can nevertheless pay a monitoring cost to avoid for *lemons*.¹²

¹²*Lemons* are defective cars in the famous study of Akerlof (1978)

Both parties reach an agreement in the form of a loan contract. In one case, entrepreneurs go bankrupt when the value of the idiosyncratic shock is greater than an optimal threshold specified by the loan contract. In this situation, bankers collect the remainder of capital net of monitoring cost. In the other case, borrowers earn profits net debt interest paid to the bankers. In a bust period, bankers must compensate the increasing likelihood of default probability by a rise in the cost of loans. Hence, the procyclicality of the net worth of entrepreneurs implies a financial accelerator mechanism. For instance, a contraction in the economy would usually deteriorate the net worth position of entrepreneurs. Subsequently, the downturn is amplified by the increase in the external finance premium, which further reduces aggregate demand. In other words, financial frictions allow for feedback effects.

Christiano et al. (2010, 2014) include frictions à la BGG in a monetary dynamic general equilibrium model, and find that risk shocks induce mean-preserving interest rate spreads; therefore, are important drivers of macroeconomic fluctuations. In such framework, QE is effective because it releases some of the pressure on spreads, which eases financial conditions for firms in various ways. First, by contemporaneously rising the price of the purchased asset and those of close substitutes, hence by supporting their net worth. In turn, a stronger net worth sustains the credit-worthiness of firms, containing the rise in the external premium. In the next periods, the demand for credit increases with the feedback effect. Second, by directly supporting the supply of credit. Indeed, the supply of excess reserves allows bankers to issue deposits which boosts the supply of new loans. These wealth and credit effects are the essence of the balance sheet channel of QE.

KM design financial frictions in the form of collateral constraints. In their

basic model, the economy is populated by two types of agents, namely farmers and gatherers. The economy comprises two goods, durable assets (lands) and nondurable commodities (fruits). Land and labor are the inputs for the production of fruits. KM impose that the impatient farmers must borrow funds from patient gatherers. A second assumption makes farmers' technology idiosyncratic. In other words, farmers' skills are bounded, so that production is up to the labor effort. In addition, farmers are allowed to give up working and repudiate their debts. As a result, gatherers demand that farmers pledge their land against lending. For both sides to find an agreement, the collateral constraint must insure that any debt repayment is worth at least lower than the market value of a land. At the equilibrium, farmers are thus greatly leveraged.

KM then add reproducible capital (trees) into the farmer's production function. Trees are subject to depreciation and cannot be used as collateral. These features allow for debt repayments to be strictly lower than the market value of collateralized land. In such framework, an unanticipated negative productivity shock causes a feedback loop. Farmers' net worth diminish because of the drop in land prices. Hence, farmers are limited in their investment spending due the tightening in collateral constraints. In the next periods, the revenues from capital fall because of the shortage in investment, which further reduces farmers' net worth. In a related work, Iacoviello (2005) studies the macroeconomic implications of such collateral constraints on the housing market. Later, Guerrieri & Iacoviello (2017) show that housing collateral constraints induce macroeconomic asymmetries during housing boom and bust cycles. In a housing boom, collateral constraints are slack, therefore the housing wealth barely contributes to the consumption growth. During a bust, constraints tighten so that the collapse in the housing wealth enhances the macroeconomic downturn

through the feedback loop. These results suggest that QE is more effective during busts. Conversely, in a situation of slack constraints, the effectiveness of QE programs is inevitably reduced.

In a QE DSGE model, Gertler & Karadi (2011) (GK) reverse a standard moral hazard problem, in a way that the premium does not emerge from the uncertainty in the quality of firms but from some bankers' greed. Households participate in the funding of banks via deposits, under the condition that bankers comply with balance sheet constraints. Nevertheless, some bankers misbehave by diverting assets because their expected lifetime is limited. Thus, an incentive constraint must bind for households to deposit. The latter states that the expected franchise value of the banks must be at least greater than the value of the diverted assets. The incentive constraint induces a leverage ratio that imposes an endogenous limit to the size of the balance sheet of bankers.

These frictions limit the arbitrage in financial markets among assets, and any exogenous drop in asset prices tightens the balance sheet constraints. This drop in asset prices produces an endogenous financial accelerator mechanism, that in turn freezes the supply of credit by private intermediaries. GK study the effects of QE in such environment. They design QE as a *public intermediation* of the supply of credit: the central bank issues short-term liabilities subject to negligible efficiency costs. GK find that it is welfare optimal to bypass financial frictions in the private markets if and only if the efficiency costs of the central bank intermediation do not offset the welfare gains induced by QE. Our Chapter 1 proposes a general equilibrium model with financial frictions bridging Gertler & Karadi (2011, 2013) and Guerrieri & Iacoviello (2017) to evaluate the impacts of a flush of excess reserves in the economy.

The FG policy is another tool to tackle a situation of ZLB. The public state-

ments of the central bank can be perceived in two opposite ways. On the one hand, the FG can correct the expectations of agents about the future stance of the nominal interest rate. Agents anticipate further stimulus; that is, future pressure on the short-term rates. As such, the yield curve flattens, which causes a rise in inflation. In these circumstances, the FG stimulates the economy. On the other hand, agents can consider that these statements reveal bad news about the future state of the economy. The main assumption is that the central bank is more informed than market participants. In this case, a statement puts an upward pressure in the long-term rates, which induces detrimental effects in the economy.

In the literature of FG, researchers are confronted to identification problems. Indeed, the FG mainly transmits through the anticipation of agents regarding the future stance of the monetary policy and/or the macroeconomic outlook. Because of that, empiric studies are scarce. In addition, theoretical models such as DSGE, suffer from the so-called *forward guidance puzzle*; named by Del Negro et al. (2012) following their observation of the overestimation of the effects of FG.

Seminal theoretical justifications of the use of FG, are developed in the study of Woodford & Eggertsson (2003). However, the question about the impacts of the anticipated monetary policy was already a *hot* debate in the seventies and eighties. Indeed, Mishkin (1982) proves in a series of empirical papers to that the short-run neutrality assertion of Modigliani (1977) about an anticipated monetary policy, does not hold. More recently, Gertler & Karadi (2015) show evidence on the importance of monetary surprises in the identification of the monetary transmission mechanism. They find that monetary surprises cause small deviations in short-term rates, but have butterfly effects in finan-

cial conditions. Indeed, term premia and credit spreads tend to greatly react in response to surprises, which in turn affect the real activity. The important contribution of this study is to tackle the simultaneity problem of the monetary policy shocks: within a period t , monetary policy shocks might be highly correlated to financial variables, absent of the model. To deal with this issue, they use high frequency data to identify monetary surprises in a VAR model with low frequency data such as the output or CPI. The underlying assumption to tackle simultaneity, is that news about the state of the economy cannot affect the current statements of the FOMC.

The study of Campbell et al. (2012) is an important reference about FG. They assert that the main challenge of the FG in the presence of ZLB, is for the central bank to convince the public about its ability to follow its promises. First, they question if FOMC communications contain information capable of shifting the expectations of agents about the future stance of the monetary policy. Before the GR, the collegial answer in the literature is, *yes* (e.g. Gürkaynak et al. (2005)). In that respect, they identify the effects of FOMC statements by the estimation of two factors: A target factor which is related to the changes in the current policy rate and a path factor (FG) which is related to the changes in the expected future rates. Subsequently, they demonstrate that an unexpected increase in the path factor, reduces forecasts of unemployment and increases the expectations in inflation. Phrased differently, forecasters believe that the FOMC possesses more precise information about the state of the economy. As such, they argue that FOMC statements contain Delphic components until 2007. Second, they identify the effects of FG through a classical interest rate rule. In this framework, FG is modeled in an lagged error term. They find that these two specification (two factors and interest rule) explain a large part of

the fluctuations in asset prices (Treasury and corporate bonds). The inclusion of such rule in their estimated DSGE highlights that Odyssean FG are able to ensure a mandate of a steady inflation in the long-run.

Nevertheless, Del Negro et al. (2012) argue and demonstrate that DSGE models overestimate the effects of FG. They call this phenomenon, the *forward guidance puzzle*. In that respect, Carlstrom et al. (2015) find that a standard Smets & Wouters (2007) DSGE model would produce explosive output and inflation paths if the ZLB is maintained during nine quarters. However, many FOMC statements contained information regarding a change/stability of the monetary policy in a span of at least eight quarters. To solve for that puzzle, agents face a probability of dying in their DSGE. As such, the shifts in expectations induced by FG, are much smaller than in a case with infinitely lived agents. Other interesting papers, such as Andrade et al. (2015) use information heterogeneity and find that in a case of a blurry communication, either Delphic or Odyssean are less effective. Last, Caballero & Farhi (2017) demonstrate that the stimulus of FG is muted when the economy already recovered from a crisis. Our Chapter 3 analyzes the effects of a state-contingent FG under an Odyssean commitment to a target in the unemployment rate.

The thesis is organized in three research papers. The paper of the first chapter is: "Quantitative Easing and Excess Reserves". The chapter 2 offers a paper named: "New Evidence on the Effects of Quantitative Easing". The chapter 3 is a paper called: "State-contingent Forward Guidance".

Chapter 1

Quantitative Easing and Excess Reserves

Abstract

What are the impacts of a flush of interest-bearing excess reserves to the real economy? Surprisingly, the theoretical literature remains silent about this question. We address this issue in a new Keynesian model with various financial frictions and reserve requirements in the balance sheet of bankers. Modeling QE by the supply of excess reserves allows for endogenous changes in the relative supply of financial assets. We find that this mechanism is crucial to identify and disentangle between the portfolio balance, the credit and the asset prices channels of QE. Further, we demonstrate that the macroeconomic effects of QE are rather weak and mainly transmitted through the asset prices channel.

GATE Working Paper 1910 – March 2019.

1.1 Introduction

A lot remains unknown about the effects of Quantitative Easing (QE), from the size of the effects that it had, to the transmission channels through which programs worked. At the onset of the Great Recession (GR thereafter), the FED lowered dramatically the effective fund rate until it hit the zero lower bound. While engaging in large rescue programs providing exceptional liquidity facilities in the discount window and in the interbank market¹, subsequently the FED announced the first of its QE asset purchase programs.² QE appeared as a potentially good solution for several reasons. To start with, QE could be qualitative and remove toxic assets from the capital markets. Further, QE could reduce long-term nominal yields and then flatten the yield curve. Finally, QE could offer an additional degree of freedom for banks to reallocate portfolios, roll over or bail-out debts and resort to recapitalization. Nevertheless, QE and the liquidity rescue programs had a dramatic effect on the size of the FED's balance sheet (see Figures A.1 and A.2 in Appendix A). In particular, the expansion of the liabilities has been covered by the issuance of deposits; that is, interest-bearing excess reserves (see Figures A.3 and A.4 in Appendix A).³ Hence,

¹Discount window programs: Term Auction Facility (TAF), Primary Dealer Credit Facility (PDCF), and Term Securities Lending Facility (TSLF). Market liquidity facilities: Commercial Paper Funding Facility (CPFF), Asset-Backed Commercial Paper Money Market Mutual Fund Liquidity Facility (AMLF), Money Market Investor Funding Facility (MMIFF), and the Term Asset-Backed Securities Loan Facility (TALF). See more information in <https://www.federalreserve.gov/monetarypolicy/bst.htm>

²Forward guidance was also a great part of the unconventional monetary instruments. Often defined as *Delphic* and *Odyssean* forward guidance, these explicit commitments to the future stance of monetary policy consistently paced the monetary announcements during the ZLB period.

³The interest rate on required reserves (IORR) and the interest rate on excess reserves (IOER) were settled by the Emergency Economic Stabilization Act of 2008

By what avenues excess reserves transmit to the real economy? In other words: To what extents such maturity transformation (swaps of excess reserves for assets) affects macroeconomic variables?

So far in the theoretical literature, the absence of excess reserves limits/overlooks the identification of the transmission channels of QE, as will be clear later. We address this issue by using a general equilibrium model that incorporates the two most popular types of financial frictions in the theoretical literature, namely banking frictions and collateral constraints. Our model is equipped to quantify the relative importance of three theoretical transmissions channels of QE programs, the portfolio balance, bank lending (credit) and asset prices channels. Our contributions are twofold.

First, we provide a more realistic way of modeling the implementation of QE programs. In their seminal papers, Gertler & Karadi (2011, 2013) (GK) design QE as a *public intermediation* in the supply of credit: the central bank elastically issues short-term liabilities subject to negligible efficiency costs to supply private capital at a lower interest rate. We offer an alternative, and include reserve requirements in the balance sheet of bankers. In addition, we impose limits to arbitrage to cause an excess return in reserves. In Curdia & Woodford (2011), the nominal interest rate is always greater or equal to the interest rate on reserves. Here, we assume the opposite, in line with the practice of the FED. Then, we design QE programs as the exogenous supply of excess reserves. This way of modeling QE programs is much closer to the actual QE programs than any alternative in the literature.

Second, we disentangle the different transmission channels of QE, not only with the introduction of excess reserves, but also by having various financial

frictions. The model includes banking frictions in the shadow banking system, banking frictions on commercial banks à la GK, and a collateral constraint in the housing market in the spirit of Guerrieri & Iacoviello (2017). The latter basically sets an upper borrowing limit for impatient households. In their framework, impatient agents borrow directly from patient households. In our model, we impose a private financial intermediation in the funding process. Doing so not only directly constrains the housing consumption of impatient agents by the market value of their collateral, but also indirectly constrains it through the moral hazard problem that bankers face.

For the sake of reality towards the episode of GR, we take into account a shadow banking system à la Meeks et al. (2017) that provides asset backed securities (ABS) after the securitization of loan bundles. Shadow banks support the supply of credit in the economy by providing high-quality collateral to bankers, and the mechanism is in line with the observed countercyclicality of the size of the shadow banking system (see Figure A.5 in Appendix A). In many contributions that include financial frictions, the GR is modeled as a negative capital quality shock. We adopt a more realistic mechanism and consider a fall in the quality of ABS. In this greater risk environment, bankers proceed to portfolio reallocations that lead them to reduce the supply of loans, and cut the demand for ABS. As a consequence, shadow banks engage in fire sales of securitized assets.

The latter is important in terms of results. Indeed, we find that our model successfully replicates key features of the GR when driven by a negative shock on the quality of ABS. In this greater risk environment, we are able to capture an endogenous mechanism within the financial system that causes ABS to be procyclical, reconciling theory with the data. We demonstrate that in

such environment, a public intermediation à la GK limits the understanding of the transmission channels of QE programs because it bypasses the endogenous portfolio reallocations, and neglects the endogenous response of the financial system to the program. In addition, it greatly overstates – roughly by a factor of 1 to 10 – the macroeconomic effects of QE programs. The main reason is that QE à la GK overlooks the relative changes in the balance sheets. Finally, we find that most macroeconomic effects generated by QE programs are channeled through the wealth effect – asset prices.

In this paper, we find that the portfolio balance channel of QE programs is always present, working as the first stage of a multi-stage rocket. But on top of the portfolio balance channel, the model allows to gauge the relative contribution of the credit channel and the asset prices channel. The introduction of excess reserves is crucial because it allows to trigger changes in the relative supply of assets; an essential transmission mechanism of QE within the financial system.

While the model performs reasonably well in replicating the GR, it is not its main focus. Indeed, we abstract from important New-Keynesian features such as wage rigidities, important non-linearities like the ZLB or the occasionally binding constraints. However, our results do not qualitatively suffer from these lacks, since the addition of such features would rather amplify the magnitude of the effects of QE programs, making our evaluation here a lower-bound estimate.

In the literature, the transmission channels of QE programs are still being discussed. Indeed, the economic literature mostly focused on the quantitative effects of QE programs empirically (see Baumeister & Benati (2013) or Weale

& Wieladek (2016) among others).⁴ In most cases, empirical contributions find that QE programs have very large effects on financial conditions and the real economy. However, these results are in stark contrast with the relative neutrality of QE suggested by the theoretical literature. The neutrality assertion is embraced by the monetarist's views of Curdia & Woodford (2011) and Williamson (2012, 2016), unless financial markets are segmented or if there are financial frictions. For instance, under the preferred-habitat assumption (Modigliani & Sutch (1966), Vayanos & Vila (2009)) and/or according to Tobin (1969)'s portfolio balance theory of the term structure of interest rates, QE should be effective/relevant. In the former, markets are segmented by the preferences of investors over asset maturities and yields. For investors to be willing to invest in other segments, the incurred risk must at least be offset by the expected return in the asset. In this context, QE fosters demand for imperfect substitutes because of the change in the supply of the targeted asset, and QE programs reduce the yields of close substitutes. In the latter, a financial crisis raises private and public spreads over riskless assets because of financial frictions, and QE programs help alleviate those. QE was also shown to have large effects in models of financial frictions: in GK a premium on risky assets emerges as the result of bankers diverting a fraction of the total assets intermediated. The friction limits the arbitrage in financial markets among assets, and any exogenous drop in asset prices tightens the balance sheet constraints. This drop produces an endogenous financial accelerator mechanism, that in turn freezes the supply of credit by private intermediaries.

The paper is organized as follows. Section 2 presents the baseline model. Section 3 investigates the effects of securitization in the transmission of a GK

⁴For further empirical literature on QE see Borio & Zabai (2018).

negative capital quality shock, and investigates the effects of a fall in the quality of ABS. It also quantifies the effects of QE programs, modeled either as in GK (public credit intermediation), or as a rise in excess reserves. Section 4 concludes.

1.2 The Baseline Model

Our baseline model is a generalization of the seminal paper of GK. Extensions capture key features of the housing market and the shadow banking system to allow for a better understanding of the origins of the GR – lining-up with usual narratives – and a better understanding of the effects of QE programs.

In the model, patient households hold financial assets and acquire housing. They fund impatient households and firms in their investment projects. However, patient workers are unable to perfectly enforce payments from borrowers, and resort to commercial banks. Commercial banks are financial intermediaries specialized in lending activities. Due to financial frictions, bankers are limited in their actions. Bankers must hold sufficient amounts of high-quality collateral to meet their balance sheet constraints. Shadow banks trade loan bundles for high-quality ABS. This section offers details about the respective behavior of commercial and shadow banks, households, producers, and the central bank.

1.2.1 Commercial banks

Commercial banks offer multiple services to three types of agents. First, they grant loans to intermediate good producers. Second, they facilitate government expenditures by holding long-term government bonds. Last, they provide

mortgage loans to impatient households. Each bank is owned and managed by a fraction of patient households, namely, bankers. Their assets are backed by the issuance of deposits and net worth (equity). However, financial frictions limits the amount of funds patient households are willing to deposit. Accordingly, a rational banker would circumvent frictions by entirely relying on equity. We then bound the expected lifetime of bankers with an i.i.d probability σ to create a rotation within the workforce. The financial frictions arise from a typical moral hazard problem à la GK in which bankers can be shady and divert a fixed fraction of their assets. Depositors demand that bankers pledge a fraction of each asset, and the composition of the bankers balance sheet shapes the relationship of trust between both parties. Bankers hold a variety of assets, that may differ in their degrees of pledgeability. In particular, primary loans induce strong suspicion in the eyes of patient households. Bankers then must hold high-quality collateral such as ABS.

The asset side of the individual balance sheet of a banker (superscript c) is made of the remainder of primary loans s_t^c , ABS a_t^c , long-term government bonds b_t^c and mortgage loans to impatient households m_t^c . Liabilities consist of deposits d_t^c and net worth n_t^c . The balance sheet identity writes:

$$Q_t s_t^c + q_t^a a_t^c + q_t^b b_t^c + q_t^h m_t^c = d_t^c + n_t^c \quad (1.1)$$

in which lower case symbols represent individual quantities. Let be Q_t , q_t^a , q_t^b and q_t^h the respective market prices of assets. The net worth evolves as:

$$\begin{aligned} n_t^c &= R_t^s Q_{t-1} s_{t-1}^c + R_t^a q_{t-1}^a a_{t-1}^c + R_t^b q_{t-1}^b b_{t-1}^c + R_t q_{t-1}^h m_{t-1}^c - R_t d_{t-1}^c \\ &= (R_t^s - R_t) Q_{t-1} s_{t-1}^c + (R_t^a - R_t) q_{t-1}^a a_{t-1}^c + (R_t^b - R_t) q_{t-1}^b b_{t-1}^c + R_t n_{t-1}^c \end{aligned} \quad (1.2)$$

where R_t^s is the real stochastic return on primary loans, R_t^a on ABS and R_t^b on long-term government bonds. The return on deposits is the real interest rate R_t . In the interest of parsimony we do not design financial frictions in the housing market. In other words, mortgage loans are riskless, hence their cost is the real interest rate.

Going deeper into details, the timeline of a banker's decisions proceeds as follows: In the dawn of t , a banker issues deposits, grants loans to firms and impatient households, funds the government and receives ABS. She then makes a choice between two options: Be honest or shady. An honest banker accumulates net worth until all stakeholders receive their payoffs in $t + 1$. Conversely, a dubious banker sells a fraction of her assets to outside investors. For the sake of discretion, the amount of fraud cannot exceed an upper bound θ^c . Her home household enjoys the booty in period $t + 1$. Unfortunately, this larceny leads the bank to a fragile position. Indeed, rational depositors stop funding the banks in this context. As a consequence, bankers are in danger of bankruptcy. Both parties benefit from reaching an agreement, that takes the form of an incentive compatibility constraint, that states that the discounted franchise value of the bank must be at least as large as a fraction of diverted assets:

$$V_t^c \geq \theta^c (Q_t s_t^c + \Xi q_t^a a_t^c + \Delta q_t^b b_t^c) \quad (1.3)$$

in which Ξ and Δ are weights related to the assets' pledgeability. By fixing $\theta^c > \theta^c \Delta > \theta^c \Xi \in [0,1]$, we assume that ABS have the highest collateral quality. In line with Meeks et al. (2017), it is quite natural to think that securitization mitigates risk through diversification. In addition, ABS increase the liquidity of the underlying assets. It is also reasonable to define a higher collateral value

for long-term government bonds than private loans, public issuers being less likely to default than private issuers. The franchise value of a bank V_t^c depends on the sum of accumulated net worth conditional on the death and survival probabilities:

$$V_{t-1}^c = \mathbb{E}_{t-1} \Lambda_{t-1,t} ((1 - \sigma)n_t^c + \sigma V_t^c) \quad (1.4)$$

which is a standard Bellman equation. Variable $\Lambda_{t-1,t}$ is the stochastic discount factor derived from the patient households program. The banker's objective is then to maximize the franchise value (1.4) subject to the incentive constraint (1.3), the balance sheet identity (1.1) and the law of motion of net worth (1.2). Using the method of undetermined coefficients, we guess and later verify that the bank's franchise value V_t^c is linear in the time-varying coefficients μ_t^c , μ_t^a , μ_t^b , η_t^c . Accordingly, we conjecture that the value function is in period t :

$$V_t^c = \mu_t^c Q_t s_t^c + \mu_t^a q_t^a a_t^c + \mu_t^b q_t^b b_t^c + \eta_t^c n_t^c \quad (1.5)$$

Hereafter, insert the particular solution (1.5) into the recurrence relation (1.4) and differentiate using the method of Lagrange multipliers. Let λ_t^c be the associated Lagrangian multiplier of the incentive constraint (1.3). The differentiations yield the following first-order necessary conditions, respectively for s_t^c , a_t^c , b_t^c , λ_t^c :

$$(1 + \lambda_t^c) \mu_t^c = \lambda_t^c \theta^c \quad (1.6)$$

$$\mu_t^a = \mu_t^c \Xi \quad (1.7)$$

$$\mu_t^b = \mu_t^c \Delta \quad (1.8)$$

$$Q_t s_t^c + \Xi q_t^a a_t^c + \Delta q_t^b b_t^c = \phi_t^c n_t^c \quad (1.9)$$

where,

$$\phi_t^c = \frac{\eta_t^c}{\theta^c - \mu_t^c} \quad (1.10)$$

The LHS of equation (1.6) shows the marginal benefit from expanding what we call primary loans. The RHS indicates the marginal cost of fraud. Due to the fact $1 > \Delta > \Xi$, the holdings of ABS and government bonds are marginally less attractive than loans. In this respect, as long as $\mu_t^c < \theta^c$, it is always profitable for bankers to divert assets. The latter is also a necessary condition for the incentive constraint to bind, otherwise, bankers would indefinitely supply loans to offset frictions. In other words, the incentive constraint imposes that the holding of assets is curbed by a maximum equity multiplier ϕ_t^c (see equation (1.10)). Hence, as long as the incentive constraint binds, the size of a banker balance sheet is endogenously constrained by fluctuations in its net worth. The latter also induces an equality between the gains from larceny $\theta^c \phi_t^c$ and the costs of losing franchise value $\mu_t^c \phi_t^c + \eta_t^c$.

At the end of period $t + 1$, a shady banker entirely consumes her net worth with probability $(1 - \sigma)$. In this case the marginal gain of net worth is then one. Conversely, an honest banker ensures the continuity of her services. Hence, either the marginal gain from larceny or the marginal cost of losing franchise value apply. Thereby, the weighted shadow value of a banker net worth is defined by:

$$\Omega_{t+1}^c = 1 + \sigma(\theta^c \phi_{t+1}^c - 1) \quad (1.11)$$

Finally, the time-varying coefficients linear in $s_t^c, a_t^c, b_t^c, n_t^c$ satisfy the Bellman

equation with:⁵

$$\mu_t^c = \mathbb{E}_t \Lambda_{t,t+1} \Omega_{t+1}^c (R_{t+1}^s - R_{t+1}) \quad (1.12)$$

$$\mu_t^a = \mathbb{E}_t \Lambda_{t,t+1} \Omega_{t+1}^c (R_{t+1}^a - R_{t+1}) \quad (1.13)$$

$$\mu_t^b = \mathbb{E}_t \Lambda_{t,t+1} \Omega_{t+1}^c (R_{t+1}^b - R_{t+1}) \quad (1.14)$$

$$\eta_t^c = \mathbb{E}_t \Lambda_{t,t+1} \Omega_{t+1}^c R_{t+1} \quad (1.15)$$

which implies that financial frictions induce discounted excess returns. To give the reader a preview of the model mechanics, imagine an exogenous financial crash that lowers the price of capital. The market value of securities is reduced, hence the bankers' net worth falls. Assets tighten in proportion to the equity multiplier. In addition, equation (1.6) indicates that a tightening in the constraint substantially increases the marginal gain from expanding primary loans; that is, the marginal cost of fraud. The latter then raises the excess return in loans and the cost of capital. In a nutshell, an exogenous crash in asset prices is in turn amplified by the endogenous tightening in the size of the balance sheet. This feedback effect further reduces the lending activities of the commercial banks.

1.2.2 Shadow banks

The shadow banking sector is composed by competitive shadow banks in the exclusive charge of the securitization of primary loans. A shadow bank purchases loan bundles in a secondary market to create high-quality ABS. Each shadow bank is owned and managed by a fraction of patient households, namely, brokers. The existence of shadow banks is motivated by the obligation

⁵Appendix B provide details for commercial and shadow banks problems

of banks to hold high-quality collateral. Hence, the entire stock of ABS is sold to bankers. In this way, the flow of ABS is confined within the financial system. However, financial frictions limit trade between bankers and brokers. A broker survives with probability σ and exits her occupation with probability $(1 - \sigma)$. Some brokers are tempted to divert a fraction of their assets. Therefore, bankers demand that brokers produce a sufficient amount of ABS.

The individual balance sheet of a broker (superscript b) comprises securitized assets s_t^b , ABS a_t^b and net worth (equity) n_t^b with:

$$Q_t s_t^b = q_t^a a_t^b + n_t^b \quad (1.16)$$

in which the accumulated net worth follows:

$$\begin{aligned} n_t^b &= R_t^s Q_{t-1} s_{t-1}^b - R_t^a q_{t-1}^a a_{t-1}^b \\ &= (R_t^s - R_t^a) Q_{t-1} s_{t-1}^b + R_t^a n_{t-1}^b \end{aligned} \quad (1.17)$$

For the sake of parsimony, the real returns and market prices in the same class of assets are equivalent across markets.⁶

The timeline of a broker's lifetime is as follows. At the beginning of t , a broker purchases loan bundles and transforms those into ABS. She then chooses to be honest or shady. An honest broker pursues securitization until the final payoffs in $t + 1$. A dubious broker diverts a fraction of her securitized assets. She fixes her greediness to an upper bound θ^b to avoid drawing attention. The related booty goes to her home household at the end of $t + 1$. Unfortunately, the future of the shadow bank is at risk due to this embezzlement. Indeed, bankers are perfectly familiar with this ploy being faced with such a dilemma them-

⁶It is conceivable to relax this assumption by adding an intermediation between bankers and brokers such as a special purpose vehicle or impose transaction costs.

selves. In this situation, bankers decide to cut offering loan bundles. However, bankers are in need for high-quality collateral to obtain funds from depositors. Thus, they agree to the condition that the discounted franchise value of the shadow bank must be at least as large as the amount of fraud. The resulting incentive constraint writes:

$$V_t^b \geq \theta^b Q_t s_t^b \quad (1.18)$$

in which the franchise value V_t^b is the sum of accumulated net worth conditional on the exiting and survival probabilities:

$$V_{t-1}^b = \mathbb{E}_{t-1} \Lambda_{t-1,t} \left((1 - \sigma) n_t^b + \sigma V_t^b \right) \quad (1.19)$$

The broker's objective is to maximize the franchise value (1.19) subject to the incentive constraint (1.18), the balance sheet identity (1.16) and the law of motion of net worth (1.17). Using the method of undetermined coefficients, the particular solution is linear in the time-varying coefficients ν_t^b, η_t^b . Accordingly, we conjecture that the value function satisfies:

$$V_t^b = \nu_t^b Q_t s_t^b + \eta_t^b n_t^b \quad (1.20)$$

Then substitute the particular solution (1.20) into the Bellman equation (1.19). Let λ_t^b be the associated Lagrangian multiplier of the incentive constraint (1.18). By differentiating, the first-order necessary conditions for, s_t^b and λ_t^b are:

$$(1 + \lambda_t^b) \nu_t^b = \lambda_t^b \theta^b \quad (1.21)$$

$$Q_t s_t^b = \phi_t^b n_t^b \quad (1.22)$$

where,

$$\phi_t^b = \frac{\eta_t^b}{\theta^b - \nu_t^b} \quad (1.23)$$

The LHS of equation (1.21) is the marginal gain from increasing the securitization activity. The RHS indicates the marginal cost of embezzlement. As long as $\nu_t^b < \theta^b$, it is always profitable for brokers to divert assets. This is a necessary condition for the incentive constraint to bind, otherwise, brokers would extend securitization to infinite amounts to offset frictions. Hence, an isomorphic equity multiplier ϕ_t^b applies to the brokers balance sheets (see equation (1.23)). The binding constraint induces an equality between the gains from diverting securitized assets $\theta^b \phi_t^b$ and the costs of losing the franchise value $\mu_t^b \phi_t^b + \eta_t^b$.

As for bankers, a shady broker entirely consumes her net worth at the end of period $t + 1$ with probability $(1 - \sigma)$. In this situation, the marginal gain of net worth is one. An honest broker builds on equity with probability σ until $t + 1$. In this case, either the marginal gain from embezzlement or the marginal cost of squeezing the franchise value apply. The weighted shadow value of a broker net worth is thus:

$$\Omega_{t+1}^b = 1 + \sigma(\theta^b \phi_t^b - 1) \quad (1.24)$$

The time-varying coefficients in s_t^b and n_t^b satisfy the Bellman equation according to:

$$\mu_t^b = \mathbb{E}_t \Lambda_{t,t+1} \Omega_{t+1}^b (R_{t+1}^s - R_{t+1}^a) \quad (1.25)$$

$$\eta_t^b = \mathbb{E}_t \Lambda_{t,t+1} \Omega_{t+1}^b R_{t+1}^a \quad (1.26)$$

The above system of equations provides a general description of this financial system. In the preceding example of the exogenous crash in the price of capital,

we highlighted the feedback effect in the absence of shadow banks. Henceforth, both net worth of bankers and brokers are seriously impaired by the fall in the price of capital. Indeed, primary securitized loans are similarly market priced. The incentive constraints tighten and feed the downturn.

1.2.3 Aggregation

By definition, the components of the equity multipliers ϕ_t^c and ϕ_t^b are identical across commercial and shadow banks. Thus, the sum of the individual balance sheet constraints yield the following aggregate identities:

$$Q_t S_t^c + \Xi q_t^a A_t^c + \Delta q_t^b B_t^c = \phi_t^c N_t^c \quad (1.27)$$

$$Q_t S_t^b = \phi_t^b N_t^b \quad (1.28)$$

The same principle applies to the net worth of bankers and brokers. The aggregate net worth consist of the retained earnings of survivors and the initial provision for newcomers:⁷

$$\begin{aligned} N_t^c &= \sigma \left((R_t^s - R_t) Q_{t-1} S_{t-1}^c + (R_t^a - R_t) q_{t-1}^a A_{t-1}^c \right) \\ &\quad + \sigma \left((R_t^b - R_t) q_{t-1}^b B_{t-1}^c + R_t N_{t-1}^c \right) + \chi^c \\ N_t^b &= \sigma \left((R_t^s - R_t^a) Q_{t-1} S_{t-1}^b + R_t^a N_{t-1}^b \right) + \chi^b \end{aligned} \quad (1.29)$$

Let Z_t be the gross profits per unit of capital as will be clear later. The parameter δ represents the depreciation rate of capital. Accordingly, the real return on

⁷ $\chi^c = \omega^c \left(\bar{Q}\bar{S} + \bar{q}^a \bar{A}^c + \bar{q}^b \bar{B}^c \right)$ and $\chi^b = \omega^b \left(\bar{Q}\bar{S}^b \right)$, in which the bars indicate the steady-state values.

primary loans is:

$$R_{t+1}^c = \frac{Z_{t+1} + (1 - \delta)Q_{t+1}}{Q_t} \zeta_{t+1} \quad (1.30)$$

where ζ_{t+1} is an exogenous capital quality shock à la GK. Accordingly, the excessive payoff of capital corresponds to the sum of gross profits and the market value of capital stock net depreciation. The ABS are derivative products of primary loans, hence their real return is:

$$R_{t+1}^a = \frac{Z_{t+1} + (1 - \delta)q_{t+1}^a}{q_t^a} \zeta_{t+1} \quad (1.31)$$

Finally, the long-term government bonds return a periodic unit of currency. Let P_t be the aggregate price level of the economy. Thus, the real return on long-term government bonds is:

$$R_{t+1}^b = \frac{\frac{1}{P_t} + q_{t+1}^b}{q_t^b} \quad (1.32)$$

1.2.4 Households and Production

There is a unit mass of households in two different groups, namely, patient and impatient households. In the group of patient households, people switch occupations each period t . A fixed proportion f is a worker while the remaining part is either a banker or a broker. Conversely, impatient households cannot work for financial intermediaries. The utility flows for each type of household

are then given by:

$$u_t = \mathbb{E}_t \sum_{i=0}^{\infty} \beta^i \left(\Gamma_c \log(C_{t+i} - \varepsilon^c C_{t+i-1}) + \bar{j} \bar{\Gamma}_h \log(H_{t+i} - \varepsilon^h H_{t+i-1}) - \frac{1}{1+\eta} L_t^{1+\eta} \right) \quad (1.33)$$

$$u'_t = \mathbb{E}_t \sum_{i=0}^{\infty} (\beta')^i \left(\Gamma'_c \log(C'_{t+i} - \varepsilon^c C'_{t+i-1}) + \bar{j} \bar{\Gamma}'_h \log(H'_{t+i} - \varepsilon^h H'_{t+i-1}) - \frac{1}{1+\eta} L_t'^{1+\eta} \right) \quad (1.34)$$

in which the prime symbols denote the impatient households' variables. Variable C_t stands for the consumption of nondurable goods, H_t accounts for housing consumption and L_t for hours worked. Utility flows are discounted by the usual $\beta > \beta'$ factors. Parameter j reflects (constant) housing preferences and $\varepsilon^{c,h}$ captures habit formation.⁸ Patient households face a budget constraint of the form:

$$\begin{aligned} C_t + D_t^h + q_t^h (H_t - H_{t-1}) + Q_t \left(S_t^h + \frac{1}{2} \kappa (S_t^h - \bar{S}^h)^2 \right) + q_t^b \left(B_t^h + \frac{1}{2} \kappa (B_t^h - \bar{B}^h)^2 \right) \\ = W_t L_t + T_t + R_t D_{t-1}^h + R_t^s Q_{t-1} S_{t-1}^h + R_t^b q_{t-1}^b B_{t-1}^h + \Pi_t \end{aligned} \quad (1.35)$$

where W_t stands for the real wage, T_t denotes lump-sum taxes and Π_t defines dividends from the ownership of commercial banks, shadow banks and capital producing firms. Patient households accumulate deposits D_t^c subject to the moral hazard problem exposed earlier. They also purchase corporate debt and long-term government bond, subject to potential transaction costs. They operate frictionless arbitrage when the convex transaction cost κ is zero. Conversely, they hold assets to their steady-state levels (\bar{S}^h, \bar{B}^h) when the cost tends towards

⁸The factors $\Gamma_c, \Gamma'_c, \Gamma_h, \Gamma'_h$ scale consumption and housing so that steady-state values of consumption and housing do not depend on habit factors. Hence, $\Gamma_c = (1 - \varepsilon^c)/(1 - \beta \varepsilon^c)$, $\Gamma'_c = (1 - \varepsilon^c)/(1 - \beta' \varepsilon^c)$, $\Gamma_h = (1 - \varepsilon^h)/(1 - \beta \varepsilon^h)$, $\Gamma'_h = (1 - \varepsilon^h)/(1 - \beta' \varepsilon^h)$

infinity. The impatient households budget constraint is:

$$C'_t + q_t^h(H'_t - H'_{t-1}) + R_t M_{t-1}^h = W'_t L'_t + M_t^h \quad (1.36)$$

meaning that impatient agents must borrow amounts M_t^h from commercial banks to secure housing consumption H'_t . Following Guerrieri & Iacoviello (2017), we introduce a collateral constraint on housing demand:

$$M_t^h \leq \gamma_m M_{t-1}^h + (1 - \gamma_m) \chi q_t^h H'_t \quad (1.37)$$

where $\gamma_m > 0$ governs the inertia of the borrowing upper bound χ . This constraint states that some $(1 - \gamma_m)$ impatient agents can roll over debts.⁹ For simplicity, housing is in a fixed supply. In this context, housing is a durable good that serves as a collateral for consumption. Hence, in trouble times, the fall in housing prices is further amplified by the constraint and largely contributes to the drop in consumption.¹⁰

The production sector comprises different entities that produce the final output of the economy. Competitive intermediate good producers are one of those. Intermediate producers sell goods to retailers by using raw capital and labor in the two groups of households. A standard Cobb-Douglas function captures the production of intermediate goods:

$$Y_t^m = A_t (\zeta_t K_t)^\alpha L_t^{(1-\alpha)(1-\varphi)} \bar{L}_t^{(1-\alpha)\varphi} \quad (1.38)$$

⁹This specification is based on findings in Guerrieri & Iacoviello (2017), and implies that housing prices lead changes in the level of debt.

¹⁰In our framework, the collateral constraint is always binding for practical reasons. The latter means that we do not capture the asymmetric effects of housing booms and busts à la Guerrieri & Iacoviello (2017). In the absence of slackness of the constraint, housing booms will then overfeed consumption. However, Guerrieri & Iacoviello (2017) find that exogenous shocks that are not directly related to housing prices and/or consumption cause limited asymmetries. Accordingly, our qualitative results are relatively immune.

where φ is the relative wage share of impatient households.¹¹ Variable A_t is an exogenous productivity measure. After production, firms need to invest in new capital for the next period. Hence, they borrow by issuing state-contingent claims in the capital market. Securities are backed by primary and securitized loans or are in the form of corporate debts in the bond market. Accordingly, the equilibrium condition on the capital market writes:

$$Q_t K_{t+1} = Q_t (S_t^c + S_t^b + S_t^h) \quad (1.39)$$

Under perfect competition, firms earn zero profits state by state, hence the marginal gain of capital is:

$$Z_t = P_t^m \alpha \frac{Y_t^m}{K_t} \quad (1.40)$$

in which P_t^m is the relative intermediate output price. The accumulation process of capital consists of the leftover capital stock net depreciation and the investment in new units of capital, as follows:

$$K_{t+1} = \zeta_{t+1} (1 - \delta) K_t + I_t \quad (1.41)$$

Note that firms are not directly exposed to financial frictions. However, they indirectly face borrowing constraint due to the relationship between commercial banks and patient households. Therefore, the level of the cost of capital is higher than in a standard new-Keynesian model.

Capital producers are in charge of the creation of the new unit of capital. After production, they sell it to the intermediate good producers at the market price Q_t . The production of capital is subject to adjustment costs. Thus, the

¹¹When $\varphi = 0$, the weight of impatient households in production goes to zero.

profit maximization program is given by:

$$\max \mathbb{E}_t \sum_{i=1}^{\infty} \Lambda_{t,t+i} \left\{ Q_i I_i - \left[1 + f \left(\frac{I_i}{I_{i-1}} \right) \right] I_i \right\} \quad (1.42)$$

which by optimization yields to a standard Tobin's Q ratio between the assets' market value and the replacement cost:

$$Q_t = 1 + f \left(\frac{I_t}{I_{t-1}} \right) + \left(\frac{I_t}{I_{t-1}} \right) f' \left(\frac{I_t}{I_{t-1}} \right) - \mathbb{E}_t \Lambda_{t,t+1} \left(\frac{I_{t+1}}{I_t} \right)^2 f' \left(\frac{I_{t+1}}{I_t} \right) \quad (1.43)$$

A continuum of j different retailers of a unit mass closes the production sector. The following CES aggregator function yields to the final output:

$$Y_t = \left[\int_0^1 Y_{jt}^{\frac{\varepsilon-1}{\varepsilon}} dj \right]^{\frac{\varepsilon}{\varepsilon-1}} \quad (1.44)$$

Following sticky prices principle, we introduce nominal rigidities à la Calvo (1983). Accordingly, retailers change the price of the final output with probability $(1 - \gamma)$. Otherwise, retailers index prices to the inflation rate π_t . The optimal reset price P_t^* is then conditional on these probabilities with:

$$\max \mathbb{E}_t \sum_{i=0}^{\infty} \gamma^i \beta^i \Lambda_{t,t+i} \left[\frac{P_t^*}{P_{t+i}} \prod_{k=1}^i (1 + \pi_{t+k-1})^{\gamma^p} - P_{t+i}^m \right] Y_{jt+i} \quad (1.45)$$

The first-order necessary condition writes:

$$\sum_{i=0}^{\infty} \gamma^i \beta^i \Lambda_{t,t+i} \left[\frac{P_t^*}{P_{t+i}} \prod_{k=1}^i (1 + \pi_{t+k-1})^{\gamma^p} - \frac{1}{1 - 1/\varepsilon} P_{t+i}^m \right] Y_{jt+i} = 0 \quad (1.46)$$

for which we derive the aggregate price level dynamic by using the law of large numbers:

$$P_t = \left[(1 - \gamma) (P_t^*)^{1-\varepsilon} + \gamma \left(\pi_{t-1}^{\gamma^p} P_{t-1} \right)^{1-\varepsilon} \right]^{\frac{1}{1-\varepsilon}} \quad (1.47)$$

1.2.5 Market clearing and Equilibrium conditions

The market clearing conditions for primary loans, ABS, long-term government bonds, housing, mortgage loans and deposits are:

$$\begin{aligned} K_{t+1} &= S_t^c + S_t^b + S_t^h & A_t^c &= A_t^b & \bar{B} &= B_t^c + B_t^h \\ H_t + H_t' &= 1 & M_t^h &= M_t^c & D_t^h &= D_t^c \end{aligned} \quad (1.48)$$

The aggregate resource constraint is:

$$Y_t = C_t + C_t' + \left[1 + f \left(\frac{I_t}{I_{t-1}} \right) \right] I_t + \bar{G} \quad (1.49)$$

The government budget constraint writes:

$$\bar{G} + R_t^b \bar{B} = T_t + \bar{B} \quad (1.50)$$

Last, the central bank conducts standard monetary policy following a Taylor rule of the form:

$$i_t = \bar{i} + \kappa_\pi \pi_t + \kappa_y (\log Y_t - \log Y_t^*) + \epsilon_t \quad (1.51)$$

where \bar{i} is the steady-state level of the nominal interest rate, Y_t^* is the flexible-price output and ϵ_t is an exogenous monetary shock.

1.3 Experiments

In this section we explore the ability of the model to mimic features of the GR and QE episodes. As stated before, our main goal is not quantitative but qualitative. Indeed, for the sake of parsimony, we ignore standard new-Keynesian features such as wage rigidities, and abstract from any securitization process

in the housing market. In addition, we abstract from occasionally binding constraints, that would certainly refine the results. However, we do believe that the addition of these features would strengthen rather than reduce our key findings, and leave these refinements for further research.

The first two experiments consist in highlighting the comovements between primary and securitized loans. In the literature, much of the empirical and theoretical evidences support a countercyclical nature of the shadow banking system (see empirical evidences in Nelson et al. (2018) and Meeks et al. (2017) among others). The Figure A.5 in Appendix A depicts the cyclical relationship between the sum of business non-financial corporate and non-corporate business loans by depository institutions and securitized other loans and advances by issuers of ABS.¹² Securitized loans are mostly countercyclical and highly volatile until 2004. A clear break appears after 2004 (see Figure A.6 in Appendix A for further details). This change coincides with the contractionary US monetary policy of the summer 2004 in response to the concerns about housing prices. Usually, securitization helps support the supply of credit by the banking sector. Issuers of ABS are mainly highly leveraged bankruptcy-remote companies that hold pools of assets. Pools cover a large range of assets such as consumer credits, Treasury securities and mortgages, and business loans. Their liabilities essentially consist of the issuance of pass-through securities, namely ABS. In other words, the interests, the principal payments and the amortization pass to the investors through the securities. In the end, these entities are inclined to cover more risk because they are only intermediaries between bor-

¹²Other loans and advances by issuers of ABS are non-financial business loans securitized by depository institutions and finance companies and syndicated loans to non-financial corporate businesses. Ref: https://www.federalreserve.gov/apps/fof/Guide/z1_tables_description.pdf

rowers and investors. After securitization, the assets are more liquid and more marketable. Nevertheless, ABS are associated to higher risk. Indeed, if interest rates increase, borrowers are more likely to default which induces a great drop in the market value of the ABS. On the contrary, if interest rates decrease, borrowers are more likely to refinance debts which causes a fall in the interests of the ABS. In our view, the former case is a plausible explanation of the procyclical spiral in the credit market for firms between 2005Q1 and 2009Q1. Indeed, the traditional sector has incentives to expand securitization in such a period of asset bubbles and deflating interest rates. Thus, they benefit from the decrease in the cost of capital and the soaring in asset prices by clearing the risk out of their balance sheets. But there is a point where the explosion of the assets bubble and defaults inevitably cause the disruption of this mechanism. The latter is the main interest of our simulated experiments. In this respect, Figure A.6 in Appendix A displays the cyclical components in both business and securitized loans in deviation from their levels in 2007Q4. Their joint rise in the beginning of the recession are certainly related to the strong inertia in the credit market for firms. In other words, the downturn in the real economy leads the crash in the market. To convince the reader, we expose the model to an exogenous degradation of the collateral value of ABS and find an intrinsic procyclical spiral.

As exposed earlier, the remainder of the exercises serves to dissociate over the transmission channels of QE. To do so, we refine the standard specification of QE by introducing banking reserve requirements and the supply of excess reserves to the model.

1.3.1 Calibration

The parameters of the baseline model are calibrated along the values of GK, Meeks et al. (2017) and the estimates of Guerrieri & Iacoviello (2017).¹³ These values are adjusted to a quarterly frequency. They mostly reflect economic conditions before the GR in the US.

The estimates of Guerrieri & Iacoviello (2017) are: the discount factors for patient β and impatient households β' , the capital share α , the capital depreciation rate δ , the habit parameters in consumption ε^c and housing ε^h , the weight for housing preferences \bar{j} , the wage share between patient and impatient workers φ , the maximum loan-to-value ratio χ , the labor disutility η and the factor of inertia γ^m . The inverse elasticity of investment η_i , the price rigidity parameters γ and γ^p and the Taylor rule coefficients $\kappa_\pi, \kappa_X, \rho$, are taken from GK.

The parameters related to the financial system are set to hit the following data targets: the survival probability σ matches an expected lifetime for bankers and brokers of roughly ten years. The parameters $\theta^c, \zeta^c, \Delta$ and Ξ are specific to the commercial banks and are adjusted to target data on the excess returns $(\bar{R}^s - \bar{R}), (\bar{R}^a - \bar{R}), (\bar{R}^b - \bar{R})$ and the equity multiplier $\bar{\phi}^c$. The excess return on primary loans $(\bar{R}^s - \bar{R})$ corresponds to the average Moody's AAA corporate bond yield relative to yield on 10-year Treasury constant maturity between 1990Q1 and 2007Q4; that is, 100 basis points. According to estimates from Meeks et al. (2017), we fix the excess return on ABS $(\bar{R}^a - \bar{R})$ to 25 basis points.¹⁴ The excess return on long-term government bonds $(\bar{R}^b - \bar{R})$ is of 50

¹³Readers are reminded that the main contribution of this paper is not quantitative, hence the estimation of parameters is not of high interest. In addition, most of the parameter values are adjusted to match data targets.

¹⁴This value reflects an average pre-crisis ABS spread over swap rates for high-quality securitized assets such as credit card, equipment or mortgages, see Meeks et al. (2017) for further

basis points, in line with the average estimates of the term premium on a 10-year zero coupon bond by Adrian et al. (2013) between 2000Q1 and 2007Q1.

The average equity capital to total assets for commercial banks is 9.5 between 2000Q1 and 2007Q1.¹⁵ However, in the model, the moral hazard problem overstates risk within the financial system, hence fluctuations in asset prices. In that respect, GK argue that values between 4 and 6 compensate the absence of the risk-sharing components between borrowers and commercial banks. We follow their lead and fix the steady-state value for the bankers' equity multiplier $\bar{\phi}^c$ to 4. The two parameters θ^b , ζ^b dedicated to the shadow banks are adjusted to match a broker's equity multiplier of 8. For the same above reasons, this value reflects half of the average of investment banks' leverage ratio reported in GK and Gertler and Kiyotaki (2015).

The portfolio adjustment cost κ ensures that steady-state returns are equivalent along the capital markets; in other words, credit and bond markets. In doing so, borrowers will most likely turn to the bond market to bypass financial frictions during bust periods. In addition, the bond market represents a large fraction of the capital market according to the steady-state values \bar{S}^h/\bar{K} and \bar{B}^h/\bar{B} . Those are in line with the calibration of GK as well as the shares of long-term government bonds and public expenditure to output (\bar{B}/\bar{Y} , \bar{G}/\bar{Y}). Parameter values are summarized in Table 1.1 below.

details.

¹⁵This figure is obtained using time series for the Total Equity Capital for Commercial Banks in United States and Total Assets, All Commercial Banks in the F.R.E.D database.

Table 1.1: Parameters values

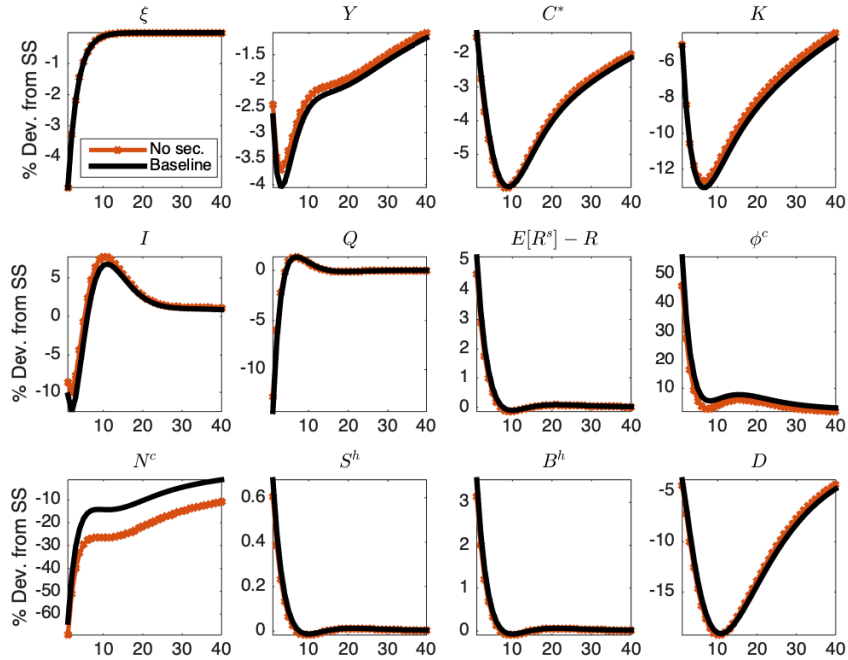
Parameter values	
Capital share in the production function, α	0.300
Discount rate patient households, β	0.995
Discount rate impatient households, β'	0.992
Capital depreciation rate, δ	0.025
Habit in consumption, ε^c	0.684
Habit in housing, ε^h	0.879
Wage share patient/impatient φ	0.501
Labor desutility, η	1.000
Inertia in the borrowing constraint, γ^m	0.694
Maximum loan-to-value ratio, χ	0.900
Inverse elasticity of investment, η_i	1.728
Elasticity of substitution between goods, ε	4.167
Probability of price rigidity, γ	0.779
Measure of price indexation, γ^p	0.241
Taylor Rule inflation coefficient, κ_π	1.500
Taylor Rule markup coefficient, κ_χ	-0.125
Survival probability, σ	0.972
Divertibility of primary loans, θ_c	0.382
Divertibility of securitized assets, θ_b	0.271
Relative divertibility of government bonds, Δ	0.500
Relative divertibility of ABS, Ξ	0.250
Transfers to new bankers, ξ_c	0.002
Transfers to new brokers, ξ_b	0.001
Bond market, transaction costs, κ	1.000
Weight on housing preferences, \bar{j}	0.040
Steady-state values	
Proportion of securities held by patient households, \bar{S}^h / \bar{K}	0.500
Proportion of gov. bond held by patient households, \bar{B}^h / \bar{B}	0.750
Government bonds to output, \bar{B} / \bar{Y}	0.450
Government expenditures to output, \bar{G} / \bar{Y}	0.200
Share of securitized assets on capital, \bar{S}^b / \bar{K}	0.150

1.3.2 Financial crisis experiments

Figures 1.1 and A.7 in Appendix A, report the impulse response functions (IRF) to an exogenous capital quality shock à la GK. This shock either captures a huge depreciation of the price of capital or a fall in the quality of capital. In this context, we compute the IRF of the model with and without shadow banks to quantify the marginal contribution of the securitization. The results indicate that the securitization weakly amplifies the macroeconomic fluctuations. This result can be understood along the following lines. The decline in the market price of primary and securitized loans deteriorates the net worth of bankers and brokers. Brokers' balance sheet constraint explodes due to their high degree of leverage. As a consequence, brokers proceed to fire sales of assets to dampen losses, and cease supplying ABS, which raises their market price. The leverage constraint also leads bankers to experience losses. The marginal gains from diverting primary loans shoots up. Hence, bankers roll over debts by purchasing cheaper primary securities, which further diminish the demand for deposits. The shock causes a sharp reduction in the supply of ABS, which serves as a buffer mechanism inside the financial system. The latter is particularly visible in the response of bankers' net worth. However, this buffer mechanism cannot entirely offset the contraction in investment and output.

While this GK exercise is widely used in the literature to mimic the effects of the GR, it cannot generate the procyclical spiral in the credit market for firms that actually characterized the GR. In the next experiment, we look at the effects of a direct disruption in the securitization mechanism, that actually captures this salient feature of the GR.

Figure 1.1: Capital quality shock



IRF of the baseline model to a 5% unanticipated capital quality shock with and without the securitization mechanism. The black solid lines show the responses in the baseline model with securitization. The red solid lines with crosses are responses related to the model without securitization. C^* is the sum of consumption by patient and impatient households. Each response is expressed in percent deviation from steady-state.

The next experiment lowers the collateral value of ABS by exogenously increasing the weight Ξ in the bankers balance sheet constraints (see Equation (1.3)). In the interest of gauging the quantitative properties of the baseline model, the shock targets the 2009Q2 detrended drop of real GDP (see Figure A.6 in Appendix A).¹⁶ Figure 1.2 shows the matching of the IRF.¹⁷ The exogenous change in the weight parameter Ξ distorts the optimal condition

¹⁶The persistence parameter of the shock is 0.96 in order to match a persistent decline in real GDP. We loop over the variance parameter of the shock to minimize the distance between the data target and the deviation in output in the model.

¹⁷Figures A.8 and A.9 in Appendix A provide other key IRF.

in Equation (1.7). The marginal benefit from diverting ABS suddenly jumps to a huge level, and bankers swiftly cut their demand for ABS. In response, the market price of ABS falls and triggers a big contraction in the net worth of bankers. Bankers then sell large amounts of assets in return for short-term liabilities, reflecting a greater discipline. As a consequence, the net worth of brokers falls.

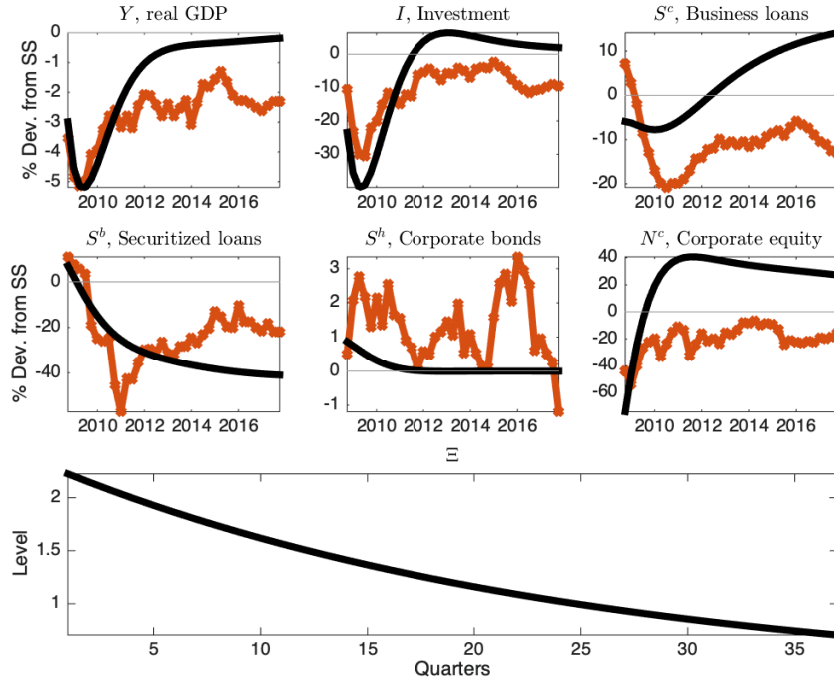
The procyclicality spiral applies when bankers are intrinsically subject to a greater risk exposure; that is, when the collateral quality of ABS shrinks. Along the lines of Gertler et al. (2012), bankers adopt more discipline in the presence of a riskier environment. Indeed, the Figure A.10 in Appendix A displays evidence that macroeconomic fluctuations are more pronounced in a safer environment. In our view, the baseline model generates convincing qualitative results and adequate quantitative properties. Thus, we believe that the model is a suitable environment to study for the theoretical transmission channels of QE programs.

1.3.3 Quantitative Easing experiments

In order to perform QE experiments, we impose some modification to the baseline model, and first introduce QE shocks à la GK. A more realistic QE experiment is conducted after this first-pass analysis. To this end, a balance sheet of the central bank is introduced to the model. The monetary authority finances its operations by issuing costless short-term debts D_t^s , bought by patient households. The balance sheet of the central bank then writes:

$$D_t^s = Q_t S_t^s + q_t^a A_t^s + q_t^b B_t^s \quad (1.52)$$

Figure 1.2: Securization crisis



The black solid lines represent IRF to an unanticipated securitization shock, characterized by an increase in Ξ . The red solid line with crosses are detrended US data given in percent change from 2007Q4 level. Each response related to the DSGE model is expressed in percent deviation from steady-state, except for Ξ . Details about the data are given in Appendix D.

which is simply zero in the absence of asset purchases. The equilibrium conditions in the financial markets must now read:

$$K_{t+1} = S_t^c + S_t^b + S_t^h + S_t^g \quad (1.53)$$

$$A_t^c + A_t^g = A_t^b \quad (1.54)$$

$$\bar{B} = B_t^c + B_t^h + B_t^g \quad (1.55)$$

The consolidated government budget constraint is now modified:

$$\bar{G} + R_t^b \bar{B} = T_t + \bar{B} + (R_t^s - R_t) Q_{t-1} S_{t-1}^g + (R_t^a - R_t) q_{t-1} A_{t-1}^g + (R_t^b - R_t) q_{t-1} B_{t-1}^g \quad (1.56)$$

In a baseline scenario, we only allow for the central bank to purchases assets in the primary market; that is, in the balance sheet of bankers.¹⁸ These asset purchases are driven by exogenous first-order disturbances, as follows:

$$\begin{aligned} S_t^g &= \psi_t^c (S_t^c (+S_t^b + S_t^h)), & A_t^g &= \psi_t^a (A_t^c (+A_t^b)), & B_t^g &= \psi_t^b (\bar{B}^c (+B_t^h)), \\ \psi_t^i &= \rho^i \psi_{t-1}^i + \varepsilon_t^i, & i &= \{c, a, b\}. \end{aligned} \quad (1.57)$$

in which blue colored variables refer to the interventions of the central bank into the balance sheet of brokers and in the bond market. By relaxing the degrees of interventions, we can easily disentangle how each type of purchase, in each market, is effective in supporting the real economy. In their studies, GK and Meeks et al. (2017) find that the effectiveness of asset purchases is an increasing function of the value of the asset weights in the incentive constraints. Figures 1.3 and 1.4 display similar results in our model.¹⁹ Unsurprisingly, a purchase of primary loans is marginally the most effective policy because these assets have, by definition the greatest leverage effect. Hence, QE à la GK is almost transmitted one for one through the rise in credit supply (see the equilibrium relations (1.53), (1.54) and (1.55)), abstracting from the relative changes in the supply of assets. This mechanism artificially overestimates the impacts

¹⁸Superscript *g* indicates central bank related variables. Unlike GK, we ignore efficiency costs per unit of asset purchases, as these costs have negligible implications for the overall dynamics and therefore are of interest only for welfare analysis.

¹⁹In the baseline scenario, we match the increase in the total assets of Federal Reserve banks as a percentage of GDP, *i.e.* 7.59% between 2008Q3 and 2008Q4 to investigate the macroeconomic implications of QE

of QE, as will be clear later.

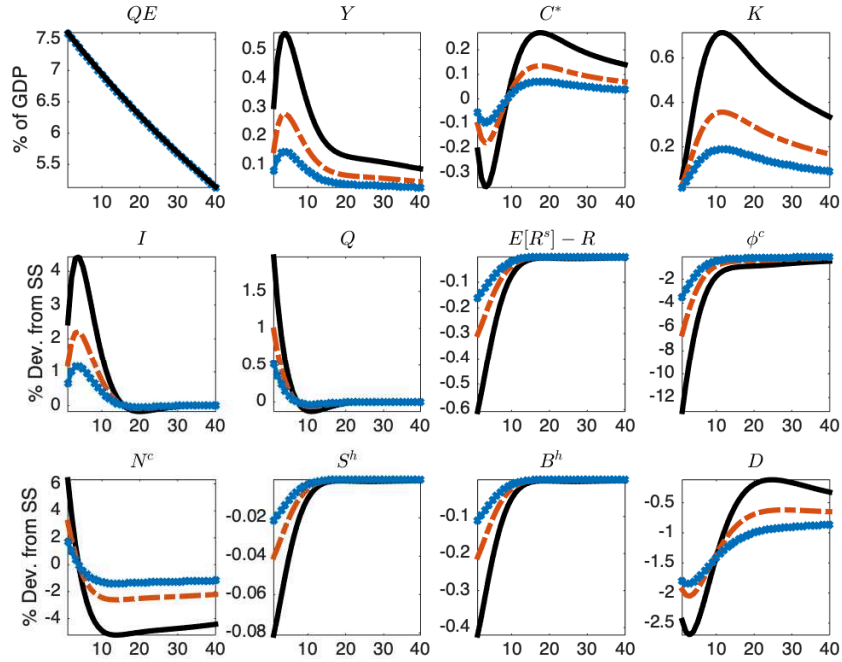
Figure 1.5 allows for variations in market interventions, and shows that the effects of QE programs depending on the type of asset are similar, regardless of the market on which the intervention takes place. For instance, the impact of primary securities purchases S^c in the primary market are equivalent to the effects of a purchases of households securities S^h in the bond market.²⁰ As such, this way of designing QE intervention clearly shuts down some of the theoretical transmission channels of QE. In particular, it remains silent on the implications of QE programs within the financial system, as it does not trigger changes in the relative supply of assets. These shortcomings are overturned by refining the design of QE intervention, introducing banking reserves in the model and using the supply of excess reserves to implement QE programs. This way of designing QE programs in a DSGE model with financial frictions is both more realistic with respect to the actual implementation of QE programs, and allows for more refined understanding of the different channels through which QE affects the financial system, and *in fine* the real economy.

In addition, the introduction of the supply of excess reserves allows us to monitor the portfolio reallocations within the balance sheets of financial intermediaries. Indeed, it will make it possible to assess the implications of an excess return on both types of reserves, something that the GK mechanism overlooks. The addition of reserve requirements changes the optimization program of a representative banker. The balance sheet identity and the law of motion of net worth become:

$$Q_t s_t^c + q_t^a a_t^c + q_t^b b_t^c + q_t^h m_t^c + q_t^{rr} rr_t^c = d_t^c + n_t^c \quad (1.58)$$

²⁰This result applies even with changes in the transaction costs κ .

Figure 1.3: Quantitative Easing: public intermediation

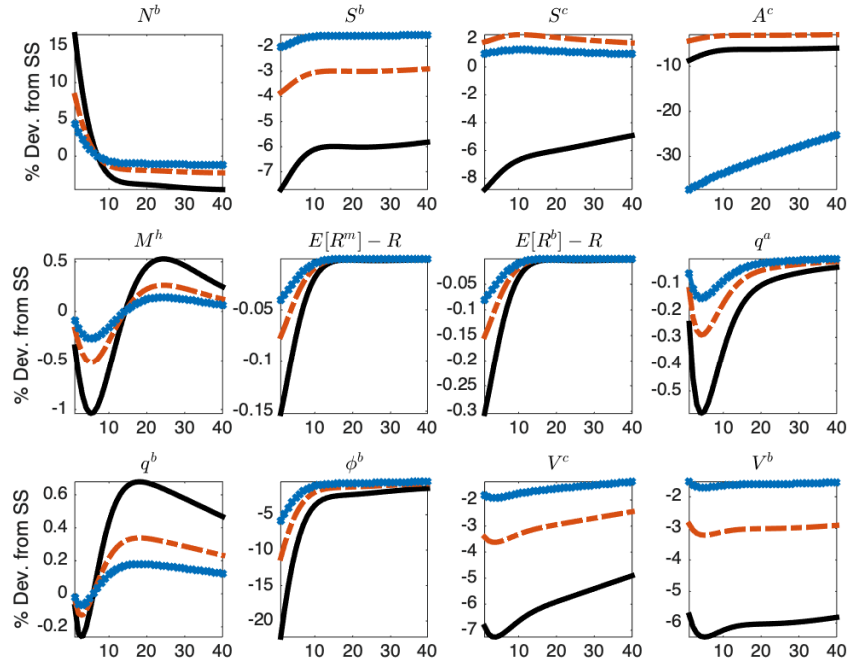


IRF to unanticipated QE shocks with a mechanism à la GK. The shocks are calibrated to match the observed deviation of total assets of Federal Reserve to GDP (7.59% of GDP between 2008Q3 and 2008Q4). The black solid lines are the responses to primary loan purchases. The red dashed line are purchases of long-term government bonds. The blue solid lines with crosses are ABS purchases. Each response is expressed in percent deviation from steady-state, except those of QE, that are given in percentage points of annual GDP.

$$\begin{aligned}
 n_t^c &= R_t^s Q_{t-1} s_{t-1}^c + R_t^a q_{t-1}^a a_{t-1}^c + R_t^b q_{t-1}^b b_{t-1}^c + R_t q_{t-1}^h m_{t-1}^c \\
 &+ R_t^{rr} q_{t-1}^{rr} r r_{t-1}^c - R_t d_{t-1}^c \\
 &= (R_t^s - R_t) Q_{t-1} s_{t-1}^c + (R_t^a - R_t) q_{t-1}^a a_{t-1}^c + (R_t^b - R_t) q_{t-1}^b b_{t-1}^c \\
 &+ (R_t^{rr} - R_t) q_{t-1}^{rr} r r_{t-1}^c + R_t n_{t-1}^c
 \end{aligned} \tag{1.59}$$

where q_t^{rr} is a market price for banking reserves, determined by the general equilibrium. The market price for banking reserves will help to dissociate the different transmission channels of QE, as will be clearer later. Realistically, the

Figure 1.4: Quantitative Easing: public intermediation, cont.



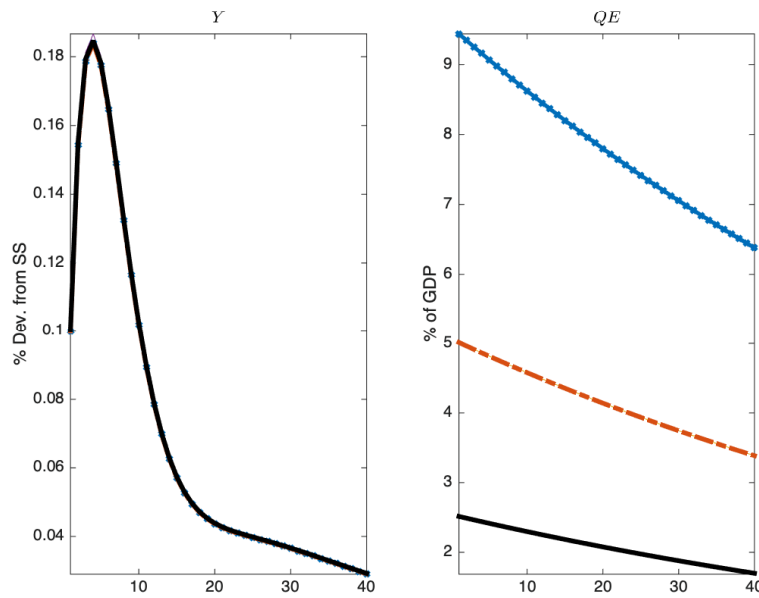
baseline (first) scenario defines the excess return on both, requirements and excess reserves, as an upper target of the nominal interest rate. In this scenario, the excess return on reserves is given by:

$$\bar{R}_t^{rr} = v \times i_t \quad (1.60)$$

where v sets the upper limit. In the second scenario (market rate), banking reserve requirements pay a fixed coupon as follows:

$$R_{t+1}^{rr} = \frac{\zeta + q_{t+1}^{rr}}{q_t^{rr}} \quad (1.61)$$

Figure 1.5: Quantitative Easing: interventions on different markets



IRF to unanticipated QE shocks with a mechanism à la GK. The shocks are calibrated to match a 0.10% increase in output in the first period. The black solid lines depict QE purchases of securities in the primary, secondary and bond markets. The red dashed lines are purchases of ABS in the primary and the secondary markets. The blue solid lines with crosses show long-term government bond purchases in the primary and bond markets. For each class of assets, there is only one visible line, because the responses are combined despite changes of market interventions.

Banking reserves are defined as a fixed fraction of households deposits:²¹

$$rr_t^c = \varphi^{rr} d_t^c \quad (1.62)$$

where the requirement ratio φ^{rr} is up to 10%.²² We also modify the incentive constraint to allow bankers to divert a fraction of reserve requirements, which is a way for us to model a limits to arbitrage.²³ We adjust the parameter Θ to

²¹ ζ is set so that the steady-state level of the market price of reserves is one.

²²This reserve requirement ratio corresponds to the ratio imposed to the highest tranche of depository institutions. Ref: <https://www.federalreserve.gov/monetarypolicy/reservereq.htm>.

²³Some empirical evidences of such frictions are documented in Martin et al. (2015).

match the average interest spread of 10 basis points between the IORR and the effective FFR (see Figure A.4 in Appendix A). The banker's incentive constraint and the linear guessed solutions of underdetermined coefficients are modified and become:

$$V_t^c \geq \theta^c (Q_t s_t^c + \Xi q_t^a a_t^c + \Delta q_t^b b_t^c + \Theta q_t^{rr} rr_t^c) \quad (1.63)$$

$$V_t^c = \mu_t^c Q_t s_t^c + \mu_t^a q_t^a a_t^c + \mu_t^b q_t^b b_t^c + \mu_t^r q_t^{rr} rr_t^c + \eta_t^c n_t^c \quad (1.64)$$

which yields an additional first-order condition, and a slight change in the incentive constraint. The time-varying coefficient μ_t^r ; that is, the marginal cost of diverting reserves is:

$$\mu_t^r = \mu_t^c \Theta \quad (1.65)$$

$$Q_t s_t^c + \Xi q_t^a a_t^c + \Delta q_t^b b_t^c + \Theta q_t^{rr} rr_t^c = \phi_t^c n_t^c \quad (1.66)$$

$$\mu_t^r = \mathbb{E}_t \Lambda_{t,t+1} \Omega_{t+1}^c (R_{t+1}^{rr} - R_{t+1}) \quad (1.67)$$

where the second equation (1.66) is the new balance sheet equation.²⁴ Aggregating among bankers, the law of motion of net worth is now:

$$N_t^c = \sigma \left((R_t^s - R_t) Q_{t-1} S_{t-1}^c + (R_t^a - R_t) q_{t-1}^a A_{t-1}^c + (R_t^b - R_t) q_{t-1}^b B_{t-1}^c \right) \quad (1.68)$$

$$+ (R_t^{rr} - R_t) q_{t-1}^{rr} RR_{t-1}^{rr} + R_t N_{t-1}^c \Big) + \chi^c \quad (1.69)$$

We close the model by the QE program, designed as an excess supply of re-

²⁴ $\chi^c = \omega^c (\bar{Q} \bar{S}^c + \bar{q}^a \bar{A}^c + \bar{q}^b \bar{B}^c + \bar{q}^{rr} \bar{R} \bar{R}^{rr})$

serves driven by first-order disturbances:²⁵

$$ER_t^g = v_t^c Q_t S_t^g + v_t^a q_t^a A_t^g + v_t^b q_t^b B_t^g \quad (1.70)$$

$$v_t^i = \rho^i v_{t-1}^i + \varepsilon_t^i, \quad i = \{c, a, b\}. \quad (1.71)$$

The consolidated government budget constraint and the equilibrium condition for total reserves are:

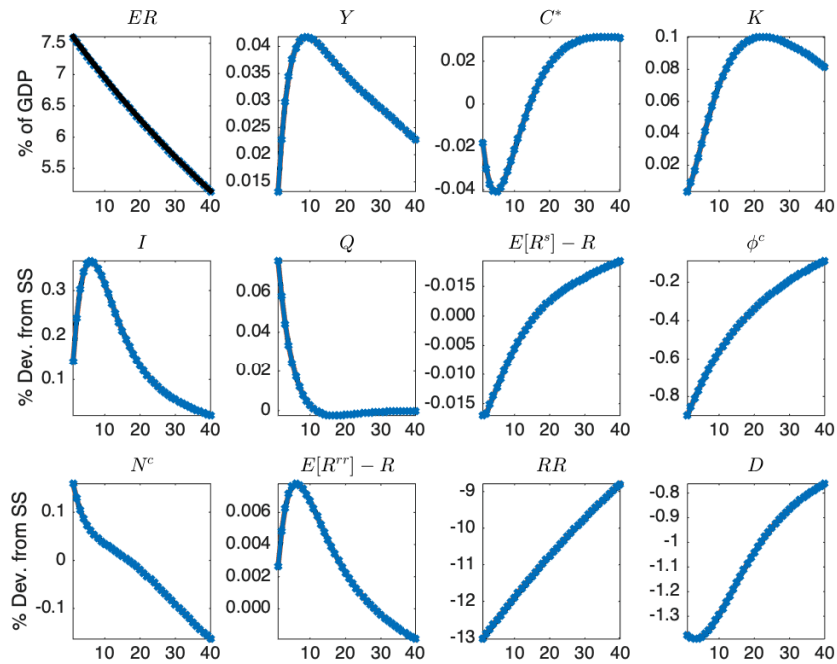
$$\bar{G} + R_t^b \bar{B} + R_t^{rr} TR_t^g = T_t + (R_t^{rr} - R_t) ER_t^g + \bar{B} + TR_t^g \quad (1.72)$$

$$TR_t^g = RR_t^c + ER_t^g \quad (1.73)$$

Before we compare the results obtained with this design of QE program through excess reserves to the GK specification of QE programs, we consider the effects of an excess reserves QE experiment on each type of asset purchases in the baseline framework (first scenario). Figures 1.6 and 1.7 reveal that QE programs have similar effects regardless of the type of asset purchases. The mechanism proceeds as follows: The liabilities of the central bank is exogenously expanded with excess reserves. Reserve requirements are mechanically diminished due the equilibrium relation in equation (1.73). Consequently, in any case, QE here, exerts a relaxing leverage effect proportional to the weight Θ in the balance sheet of bankers (see equation (1.66)). We believe that this decrease in reserve requirements is analogous to a purchase of assets in the balance sheet of bankers. Here the loosening of constraints is thus indirect. The latter entails a leveraging process. As a result, bankers take more risks and supply more primary loans. These reallocations dampen the demand of

²⁵The persistence parameters $\rho^i, i = \{c, a, b\}$ have similar values as in the GK-like QE experiments; that is 0.96. The market price of reserves q_t^{rr} has to be removed in each of these equations to obtain the first scenario; that is, the upper target of the nominal interest rate.

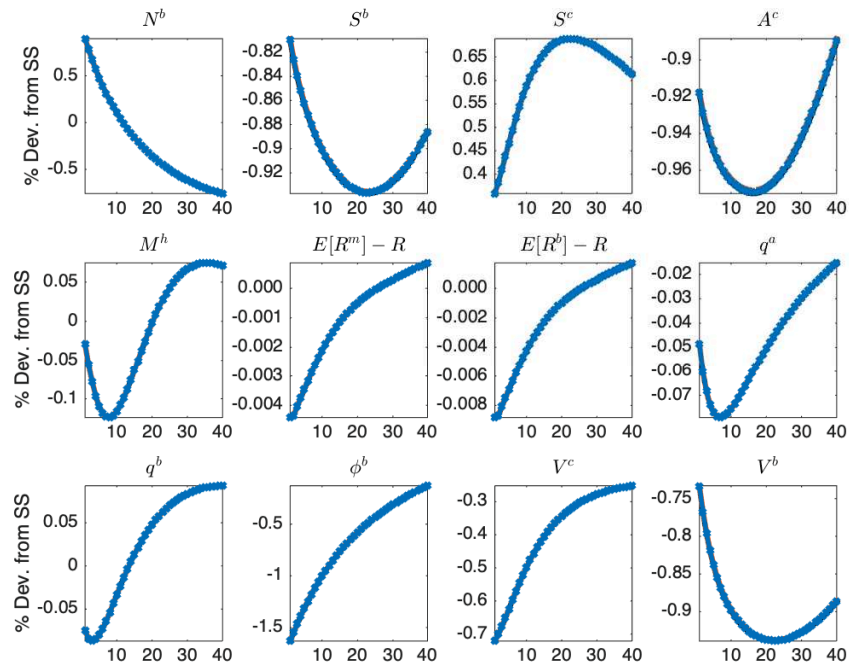
Figure 1.6: Quantitative Easing through Excess Reserves



IRF to unanticipated QE shocks through excess reserves. In this scenario, the interest rate in reserves is fixed and proportional to the nominal interest rate. The purchases are similar for any type of asset purchase. For each panel, responses overlap and form a single blue solid line with crosses. Each response is expressed in percent deviation from steady-state, except for excess reserves that are given in percent of GDP.

bankers for ABS, and brokers suffer from this shortage. However, their net worth is partially protected by the soaring in the market price of capital. From a macroeconomic perspective, the increase in asset prices causes wealth effects for households and firms. In addition, both also benefit from the sharp fall in the cost of capital, and from the increase in the collateral value of housing. In brief, QE is clearly effective in fostering financial conditions in the traditional sector. Nevertheless, the macroeconomic effects of QE are extremely small quantitatively in comparison of the huge increase in the supply of excess reserves.

Figure 1.7: Quantitative Easing through Excess Reserves, cont.



Given the small effects of a QE program modeled as an excess of reserves, we compare our results to an equivalent QE program with a mechanism à la GK. Figures A.11 and A.12 in Appendix A displays a program of ABS purchases under both specifications.²⁶ We remind the reader that the ABS purchase is the less effective purchase in a GK specification. The results show that the fluctuations are significantly greater with a GK mechanism than when QE is modeled as an excess of reserves. The way QE programs are modeled is thus crucial to determine the overall qualitative and quantitative effects of QE programs, and mechanisms à la GK can seriously overstate their importance. In addition,

²⁶We report ABS purchases because responses are closer in magnitude for both specifications. Indeed, other asset purchases with a mechanism à la GK, such as primary loans and long-term government bonds, are by far more effective, as already shown in Figures 1.3 and 1.4.

we compute gradual changes in the weight Θ that as a reminder, symbolizes the steady-state level in the excess return of reserves. Our goal is to find a value for Θ that matches the deviation in output under both specifications. Figure A.13 in Appendix A displays the results. A simple back of the envelope calculation indicates that Θ must be roughly 5 for output to get close to the target – the effect in output obtained with a GK mechanism. The latter would imply a steady-state excess return on reserves of 500 basis points, very far from empirically realistic values.²⁷ These elements indicate that the macroeconomic effects of QE programs might be significantly overestimated in the literature, that mostly relies on GK mechanisms.²⁸

Let us now focus on a potential drawback of the modeling of QE as an excess of reserves in the first scenario, namely the fact that it is rather tricky to identify the relative importance of the various transmission channels of QE. The portfolio balance channel – changes in the composition of portfolios of financial intermediaries – unquestionably triggers fluctuations within the financial system in the above experiment. In addition, the reduction in the cost of capital and the rise in primary securities can be thought of as a credit channel. Finally, the rise in asset prices generates strong wealth effects. However, the first (baseline) scenario of QE modeled as an excess of reserves does not help distinguish the different channels at play. Our second scenario addresses this issue.

In the second scenario, we separate the transmission channels of a QE program modeled as an excess of reserves. To do so, we switch to the market rate scenario. The idea is to write a model in which asset prices exhibit more dynamics and affect the financial system more significantly. On the one hand,

²⁷Such a large steady-state spread also leads to equilibrium indeterminacy issues.

²⁸see Jouvanceau (2019) for an empirical discussion on this issue.

larger macroeconomic fluctuations would be supportive evidence of the asset prices effects of QE programs (wealth effect). On the other hand, smaller fluctuations would indicate that the credit channel dominates. The results in Figures A.14 and A.15 in Appendix A support the view that wealth effects are more important. Indeed, the portfolio reallocations are qualitatively similar to the first scenario. However, the fluctuations in asset quantities are dramatically reduced, while deviations in output and investment are of close magnitude. In addition, the baseline and market scenario also exhibit differences in the persistence of the effects. The latter are mostly explained by the sluggish rises in capital, that alters the bank lending transmission channel – this is certainly at play but is dominated in magnitude by the wealth effects. Finally, as in the GK vs. excess reserves comparison exercise, we play with the value of Θ . Figure A.16 in Appendix A emphasizes that macroeconomic fluctuations are clearly increasing in the fluctuations in asset prices.

All of these experiments support the view that the relative changes in asset prices clearly alter bankers' decisions. On the one hand, the portfolio rebalancing effects of QE alleviate financial frictions and foster financial conditions in the traditional sector. On the other hand, aggregate demand is slightly stimulated by wealth effects, that in turn boost households' consumption and private investment.

1.4 Conclusion

In this paper, we show that financial frictions are useful to understand the GR and the effects of QE programs. We extend the seminal paper of GK and refine the financial sector to include securitization (in the form of ABS and a

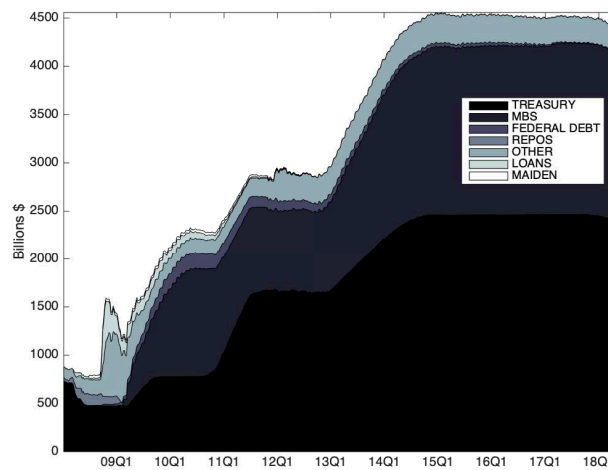
shadow banking system) and housing, with a collateral constraint that affects households consumption. These elements help explain the procyclical credit market spiral in the wake of the GR. In this context, we demonstrate that a greater exposure to risk triggers discipline in bankers' decisions of portfolio re-allocations. Shadow and commercial bankers accumulate losses which further and endogenously magnify the size of the subsequent economic downturn.

We then conduct QE experiments. Our first contribution is to model QE programs as being implemented through an excess of reserves, closer to the actual implementation schemes. We show that implementation matters for the effects of QE programs: the financial and macroeconomic effects are clearly dampened when QE is implemented through an excess of reserves – despite huge amounts of simulated excess reserves – compared to a simplified implementation scheme à la GK. Our second contribution is to allow for multiple transmission channels of QE programs, and gauge their relative importance. Our assumptions allow for a portfolio balance channel, a credit channel and an asset prices (wealth) channel of QE. Our results suggest that QE programs are mainly effective through wealth effects; that is, through the asset prices channel. Finally, our qualitative results are robust to quantitative refinements.

Appendix A

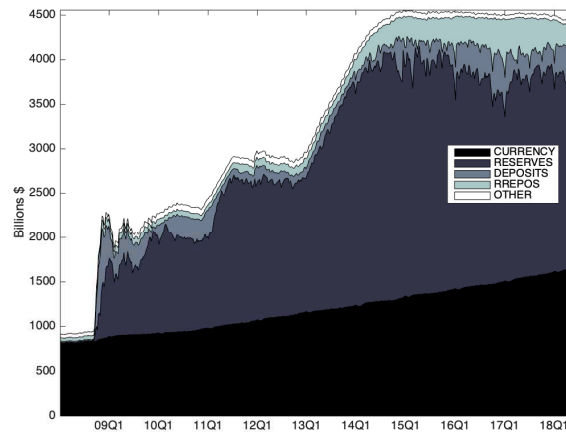
A.1 Additional Figures

Figure A.1: Assets of the Federal Reserve



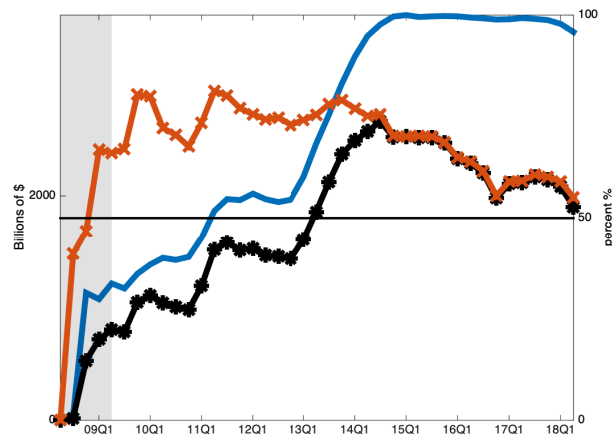
TREASURY: US Treasury securities, MBS: Mortgage-backed securities, FEDERAL DEBT: Federal agency debt securities, REPOS: Repurchase agreements, MAIDEN: Net portfolio holdings of Maiden Lane LLC, OTHER = Float + Central bank liquidity swaps + Other Federal Reserve assets + Foreign currency denominated assets + Gold stock + Special drawing rights certificate account + Treasury currency outstanding + Unamortized premia and discounts on securities held outright.

Figure A.2: Liabilities of the Federal Reserve



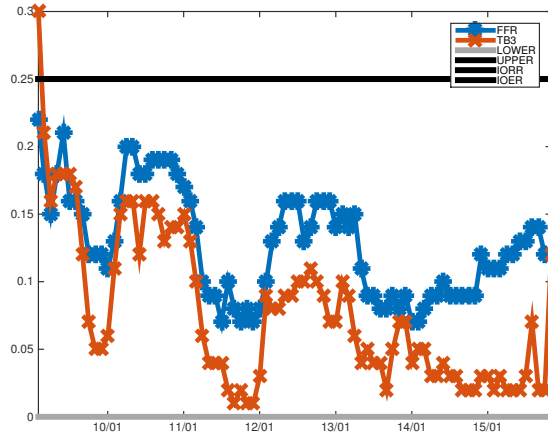
DEPOSITS: Deposits with F.R. Banks, other than reserve balances, RREPOS: Reverse repurchase agreements, OTHER = Treasury cash holdings + Other liabilities and capital

Figure A.3: Share of excess reserves in total assets, for all Federal Reserve banks



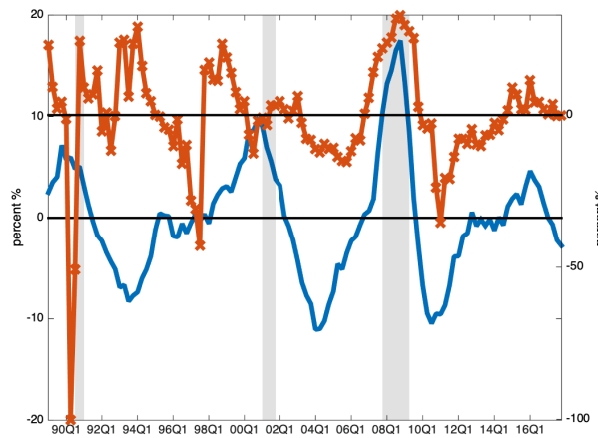
The left-hand scale and blue solid line represents excess reserves of depository institutions. The left-hand scale and black line with stars shows the total assets of all Federal Reserves banks. The right-hand scale in red solid line with crosses is the share of excess reserves in total assets. The gray shaded area is the Great Recession period. Further details on data are given in Appendix D.

Figure A.4: Main interest rates: monetary policy in the US



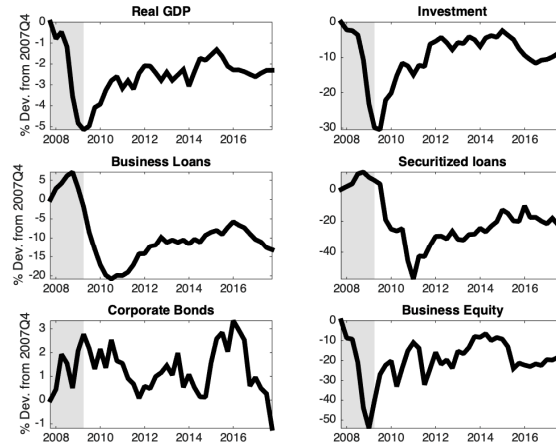
The upper limit of the federal funds target range and interest rates on both required and excess reserves are combined in the black line. The blue line is the effective federal funds rate. The red line shows the 3-month Treasury bill rate.

Figure A.5: Cyclical components of business loans and ABS



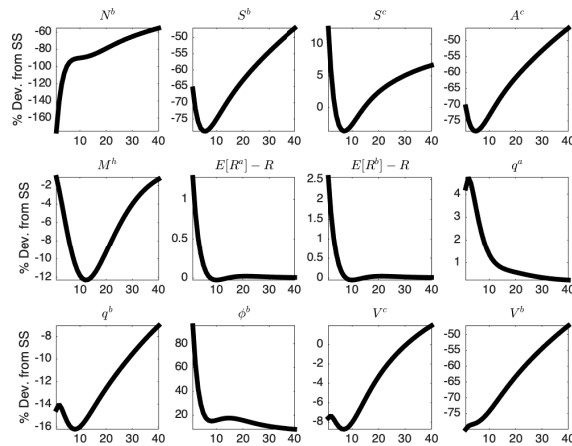
The left-hand scale and blue solid line represent the sum of business non-financial corporate or non-corporate business loans by depository institutions. The right-hand scale and red solid line with crosses shows securitized other loans and advances by issuers of asset-backed securities. Both series are deflated with the implicit price deflator computed by the US Bureau of Economic Analysis. Cyclical components are filtered using a HP-filter with a smoothing parameter of 1600. Further details on the data are given in Appendix D.

Figure A.6: US detrended data from 2007Q4 to 2017Q4



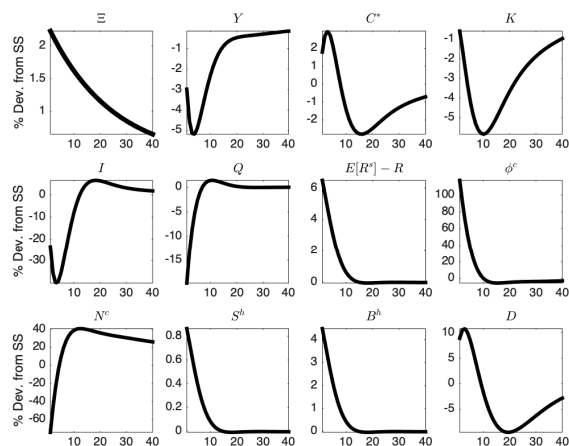
Data are first deflated with the implicit price deflator and then detrended using a HP-filter with a smoothing parameter of 1600. The gray shaded area shows the Great Recession period. All variables are expressed in percent deviation from their 2007Q4 level. Further details on the data are given in Appendix D.

Figure A.7: Capital quality shock, cont.



This figure displays additional IRF related to the baseline model with securitization only. Variables are expressed in percentage deviation from the steady-state.

Figure A.8: Securitization crisis, cont.



This figure depicts key IRF after an unanticipated securitization crisis. Variables are expressed in percentage deviation from the steady-state.

Figure A.9: Securitization crisis, cont.

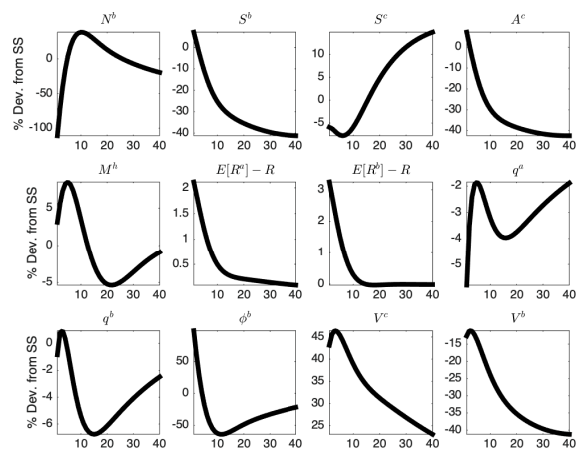
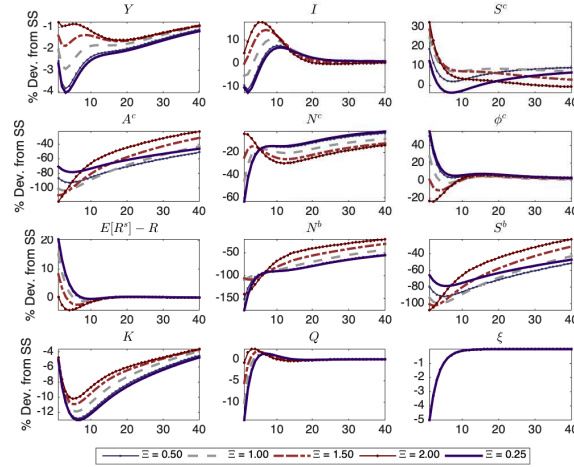
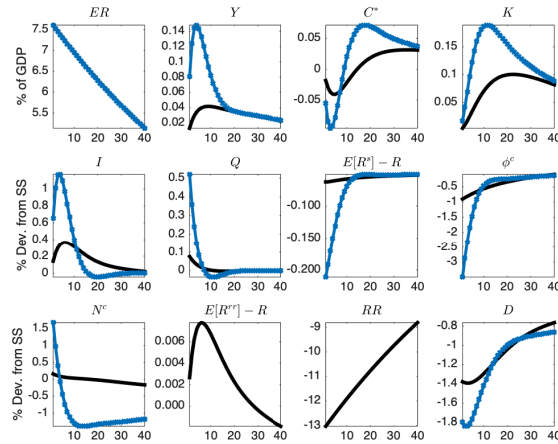


Figure A.10: Capital quality shock, changes in Ξ



IRF to unanticipated capital quality shock with gradual changes in the values of Ξ . A low value of Ξ capture an environment with low risk. Each line corresponds to a different value given in the legend. Variables are expressed in percentage deviation from the steady-state.

Figure A.11: Public intermediation (GK) vs Excess Reserves



IRF to unanticipated QE shocks with a GK implementation and with excess reserves. The blue solid lines with crosses show ABS purchases with a GK implementation. The gray dashed lines represent purchases of any type of assets in our first (baseline) excess reserves scenario. Variables are expressed in percentage deviation from the steady-state, except for QE that is given in percent of GDP.

Figure A.12: Public intermediation (GK) vs Excess Reserves, cont.

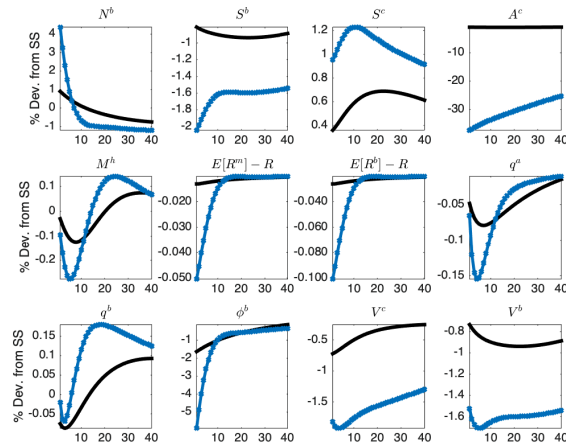
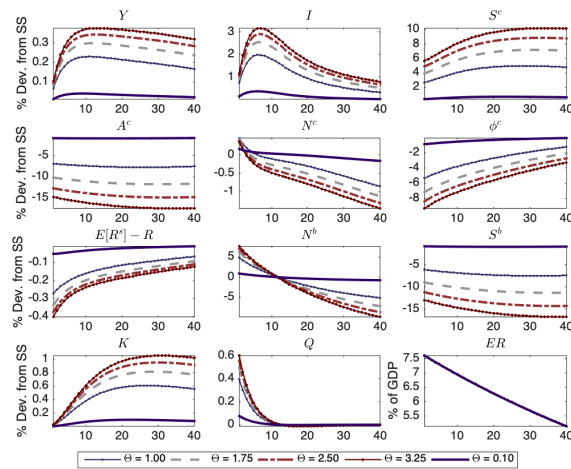
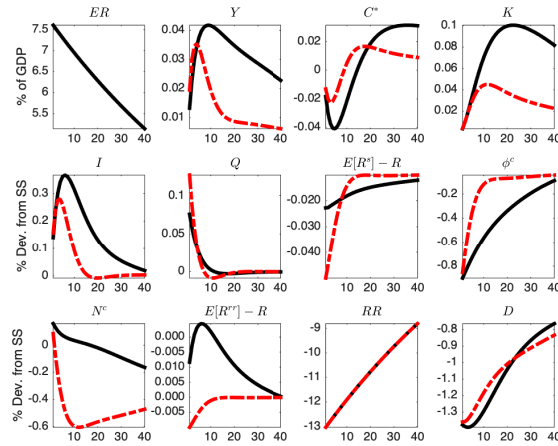


Figure A.13: First scenario (fixed rate), changes in the value of Θ



IRF to unanticipated QE shocks implemented with excess reserves in the first (baseline) scenario with multiple values of Θ . Each line corresponds to a different value given in the legend. Variables are expressed in percentage deviation from the steady-state, except for excess reserves that is given in percent of GDP.

Figure A.14: Excess Reserves, different scenarios



IRF comparing the two excess reserves scenarios. The black solid lines show responses in the first scenario, *i.e.* with a fixed interest rate. The red solid lines with crosses represent the second scenario, in which the return on reserves is stochastic.

Figure A.15: Excess Reserves, different scenarios, cont.

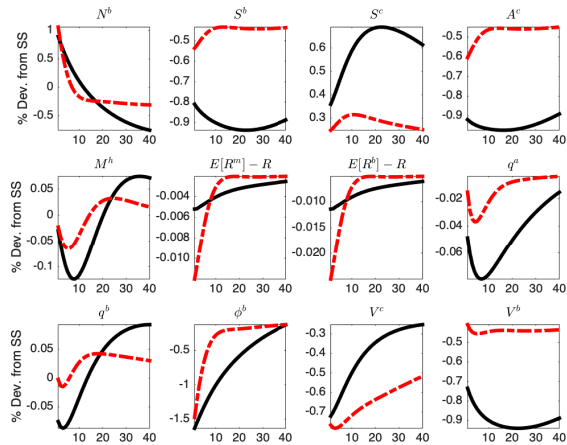
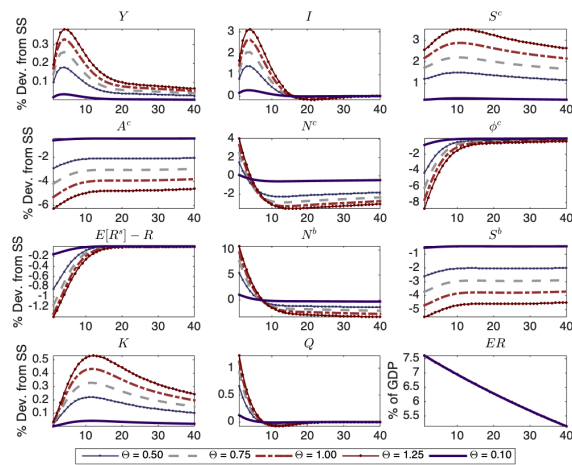


Figure A.16: Second scenario (market rate), changes in the value of Θ



The values of Θ are less flexible in this scenario, to satisfy Blanchard-Kahn conditions. However, the macroeconomic effects are much larger. Accordingly, the QE program seems to be effective mostly through the portfolio balance channel and wealth effects for private agents.

A.2 Model Equations

The set of equations for the baseline model without excess reserves is:

- Patient households:

$$\begin{aligned} C_t + D_t^h + q_t^h(H_t - H_{t-1}) + Q_t \left(S_t^h + \frac{1}{2}\kappa(S_t^h - \bar{S}^h)^2 \right) + q_t^b \left(B_t^h + \frac{1}{2}\kappa(B_t^h - \bar{B}^h)^2 \right) \\ = W_t L_t + T_t + R_t D_{t-1}^h + R_t^s Q_{t-1} S_{t-1}^h + R_t^b q_{t-1}^b B_{t-1}^h + \Pi_t \end{aligned} \quad (\text{A.1})$$

$$W_t u_{c,t} = u_{L,t} \quad (\text{A.2})$$

$$\mathbb{E}_t \Lambda_{t,t+1} R_{t+1} = 1 \quad (\text{A.3})$$

$$\Lambda_{t,t+1} = \beta \frac{u_{c,t+1}}{u_{c,t}} \quad (\text{A.4})$$

$$S_t^h = \bar{S}^h + \Lambda_{t,t+1} \frac{(R_{t+1}^s - R_{t+1})}{\kappa} \quad (\text{A.5})$$

$$B_t^h = \bar{B}^h + \Lambda_{t,t+1} \frac{(R_{t+1}^b - R_{t+1})}{\kappa} \quad (\text{A.6})$$

$$q_t^h u_{c,t} = u_{h,t} + \beta \mathbb{E}_t q_{t+1}^h u_{c,t+1} \quad (\text{A.7})$$

- Impatient households:

$$C_t' + q_t^h(H_t' - H_{t-1}') + R_t M_{t-1}^h = W_t' L_t' + M_t^h \quad (\text{A.8})$$

$$M_t^h = \gamma_m M_{t-1}^h + (1 - \gamma_m) \chi q_t^h H_t' \quad (\text{A.9})$$

$$W_t' u_{c',t} = u_{L',t} \quad (\text{A.10})$$

$$q_t^h u_{c',t} = u_{h',t} + \beta' \mathbb{E}_t q_{t+1}^h u_{c',t+1} + u_{c',t} \lambda_t^h (1 - \gamma_m) \chi q_t^h \quad (\text{A.11})$$

$$(1 - \lambda_t^h) u_{c',t} = \beta' \mathbb{E}_t \left(R_{t+1} - \gamma_m \lambda_{t+1}^h u_{c',t+1} \right) \quad (\text{A.12})$$

where λ_t^h is the multiplier of the borrowing constraint (1.37) normalized by the marginal utility of consumption of impatient households $u_{c',t}$.

- Intermediate goods producers:

$$Y_t^m = A_t (\xi_t K_t)^\alpha L_t^{(1-\alpha)(1-\varphi)} L_t'^{(1-\alpha)\varphi} \quad (\text{A.13})$$

$$Z_t = P_t^m \alpha \frac{Y_t^m}{K_t} \quad (\text{A.14})$$

$$W_t = P_t^m (1-\alpha)(1-\varphi) \frac{Y_t^m}{L_t} \quad (\text{A.15})$$

$$W_t' = P_t^m (1-\alpha)\varphi \frac{Y_t^m}{L_t'} \quad (\text{A.16})$$

$$R_{t+1}^c = \frac{Z_{t+1} + (1-\delta)Q_{t+1}}{Q_t} \xi_{t+1} \quad (\text{A.17})$$

$$R_{t+1}^a = \frac{Z_{t+1} + (1-\delta)q_{t+1}^a}{q_t^a} \xi_{t+1} \quad (\text{A.18})$$

- Capital producers:

$$K_{t+1} = \xi_{t+1}(1-\delta)K_t + I_t \quad (\text{A.19})$$

$$Q_t = 1 + f\left(\frac{I_t}{I_{t-1}}\right) + \left(\frac{I_t}{I_{t-1}}\right) f'\left(\frac{I_t}{I_{t-1}}\right) - \mathbb{E}_t \Lambda_{t,t+1} \left(\frac{I_{t+1}}{I_t}\right)^2 f'\left(\frac{I_{t+1}}{I_t}\right) \quad (\text{A.20})$$

- Final goods producers:

$$Y_t^m = D_t Y_t \quad (\text{A.21})$$

$$D_t = \gamma D_{t-1} \pi_{t-1}^{-\gamma^p \varepsilon} \pi_t^\varepsilon + (1-\gamma) \left(\frac{1 - \gamma \pi_{t-1}^{\gamma^p (1-\gamma)} \pi_t^{\gamma-1}}{1-\gamma} \right)^{-\frac{\varepsilon}{1-\gamma}} \quad (\text{A.22})$$

$$F_t = Y_t P_t^m + \mathbb{E}_t \left[\beta \gamma \Lambda_{t,t+1} \frac{\pi_t^{-\gamma^p \varepsilon}}{\pi_{t+1}^{-\varepsilon}} F_{t+1} \right] \quad (\text{A.23})$$

$$J_t = Y_t + \mathbb{E}_t \left[\beta \gamma \Lambda_{t,t+1} \frac{\pi_t^{\gamma^p (1-\varepsilon)}}{\pi_{t+1}^{(1-\varepsilon)}} J_{t+1} \right] \quad (\text{A.24})$$

$$\pi_t^* = \frac{\varepsilon}{\varepsilon-1} \frac{F_t}{J_t} \pi_t \quad (\text{A.25})$$

$$\pi_t^{1-\varepsilon} = \gamma \pi_{t-1}^{\gamma p(1-\varepsilon)} + (1-\gamma)(\pi_t^*)^{1-\varepsilon} \quad (\text{A.26})$$

$$i_t = R_{t+1} \pi_{t+1} \quad (\text{A.27})$$

- Commercial banks:

$$\mu_t^c = \mathbb{E}_t \Lambda_{t,t+1} \Omega_{t+1}^c (R_{t+1}^s - R_{t+1}) \quad (\text{A.28})$$

$$\mu_t^a = \mathbb{E}_t \Lambda_{t,t+1} \Omega_{t+1}^c (R_{t+1}^a - R_{t+1}) \quad (\text{A.29})$$

$$\mu_t^b = \mathbb{E}_t \Lambda_{t,t+1} \Omega_{t+1}^c (R_{t+1}^b - R_{t+1}) \quad (\text{A.30})$$

$$\eta_t^c = \mathbb{E}_t \Lambda_{t,t+1} \Omega_{t+1}^c R_{t+1} \quad (\text{A.31})$$

$$\phi_t^c = \frac{\eta_t^c}{\theta^c - \mu_t^c} \quad (\text{A.32})$$

$$Q_t S_t^c + \Xi q_t^a A_t^c + \Delta q_t^b B_t^c = \phi_t^c N_t^c \quad (\text{A.33})$$

$$\begin{aligned} N_t^c &= \sigma \left((R_t^s - R_t) Q_{t-1} S_{t-1}^c + (R_t^a - R_t) q_{t-1}^a A_{t-1}^c + (R_t^b - R_t) q_{t-1}^b B_{t-1}^c + R_t N_{t-1}^c \right) \\ &+ \omega^c \left(\bar{Q} \bar{S} + \bar{q}^a \bar{A}^c + \bar{q}^b \bar{B}^c \right) \end{aligned} \quad (\text{A.34})$$

$$Q_t S_t^c + q_t^a A_t^c + q_t^b B_t^c + q_t^h M_t^c = D_t^c + N_t^c \quad (\text{A.35})$$

$$\Omega_{t+1}^c = (1-\sigma) + \sigma(\mu_t^c \phi_t^c + \eta_t^c) \quad (\text{A.36})$$

- Shadow banks:

$$\nu_t^b = \mathbb{E}_t \Lambda_{t,t+1} \Omega_{t+1}^b (R_{t+1}^s - R_{t+1}^a) \quad (\text{A.37})$$

$$\eta_t^b = \mathbb{E}_t \Lambda_{t,t+1} \Omega_{t+1}^b R_{t+1}^a \quad (\text{A.38})$$

$$\phi_t^b = \frac{\eta_t^b}{\theta^b - \nu_t^b} \quad (\text{A.39})$$

$$Q_t S_t^b = \phi_t^b N_t^b \quad (\text{A.40})$$

$$N_t^b = \sigma \left((R_t^s - R_t^a) Q_{t-1} S_{t-1}^b + R_t^a N_{t-1}^b \right) + \omega^b \left(\bar{Q} \bar{S}^b \right) \quad (\text{A.41})$$

$$Q_t S_t^b = q_t^a A_t^b + N_t^b \quad (\text{A.42})$$

$$\Omega_{t+1}^b = (1 - \sigma) + \sigma(\nu_t^b \phi_t^b + \eta_t^b) \quad (\text{A.43})$$

- Market clearing and miscellaneous:

$$K_{t+1} = S_t^c + S_t^b + S_t^h + S_t^g \quad (\text{A.44})$$

$$A_t^c + A_t^g = A_t^b \quad (\text{A.45})$$

$$\bar{B} = B_t^c + B_t^h + B_t^g \quad (\text{A.46})$$

$$H_t + H_t' = 1 \quad (\text{A.47})$$

$$M_t^h = M_t^c \quad (\text{A.48})$$

$$D_t^h = D_t^c \quad (\text{A.49})$$

$$\bar{G} + (R_t^b - 1)\bar{B} = T_t + (R_t^s - R_t)Q_{t-1}S_{t-1}^g + (R_t^a - R_t)q_{t-1}^a A_{t-1}^g + (R_t^b - R_t)q_{t-1}^b B_{t-1}^g \quad (\text{A.50})$$

$$R_{t+1}^b = \frac{\frac{1}{P_t} + q_{t+1}^b}{q_t^b} \quad (\text{A.51})$$

$$Y_t = C_t + C_t' + \left[1 + f\left(\frac{I_t}{I_{t-1}}\right) \right] I_t \quad (\text{A.52})$$

$$i_t = \bar{i} + \kappa_\pi \pi_t + \kappa_y (\log Y_t - \log Y_t^*) + \epsilon_t \quad (\text{A.53})$$

A.3 Commercial and Shadow Banks' Problems

COMMERCIAL BANKS

Given the balance sheet identity (1.1) and the evolution of net worth (1.2), the banker's optimization problem comes down to a Bellman equation. Using the method of undetermined coefficients, we guess that the solution is linear in the time-varying coefficients:

$$V_t^c = \mu_t^c Q_t s_t^c + \mu_t^a q_t^a a_t^c + \mu_t^b q_t^b b_t^c + \eta_t^c n_t^c \quad (\text{A.54})$$

Now insert the guessed solution into the Bellman equation then maximize the objective function subject to the incentive constraint (1.3). Using a Lagrangian multiplier λ_t^c , the problem writes:

$$\mathcal{L} = (1 + \lambda_t^c) \left(\mu_t^c Q_t s_t^c + \mu_t^a q_t^a a_t^c + \mu_t^b q_t^b b_t^c + \eta_t^c n_t^c \right) - \lambda_t^c \theta^c (Q_t s_t^c + \Xi q_t^a a_t^c + \Delta q_t^b b_t^c) \quad (\text{A.55})$$

The first-order conditions associated with s_t^c , a_t^c , b_t^c , λ_t^c are:

$$(1 + \lambda_t^c) \mu_t^c = \lambda_t^c \theta^c \quad (\text{A.56})$$

$$\mu_t^a = \mu_t^c \Xi \quad (\text{A.57})$$

$$\mu_t^b = \mu_t^c \Delta \quad (\text{A.58})$$

$$Q_t s_t^c + \Xi q_t^a a_t^c + \Delta q_t^b b_t^c = \phi_t^c n_t^c, \quad \text{if } \lambda_t^c > 0 \quad (\text{A.59})$$

where,

$$\phi_t^c = \frac{\eta_t^c}{\theta^c - \mu_t^c} \quad (\text{A.60})$$

Using the FOCs (A.57), (A.58) and (A.59), step by step, one can rewrite the guessed solution as:

$$\begin{aligned}
V_t^c &= \mu_t^c Q_t s_t^c + \mu_t^c \Xi q_t^a a_t^c + \mu_t^c \Delta q_t^b b_t^c + \eta_t^c n_t^c \\
V_t^c &= \mu_t^c (Q_t s_t^c + \Xi q_t^a a_t^c + \Delta q_t^b b_t^c) + \eta_t^c n_t^c \\
V_t^c &= \mu_t^c \phi_t^c n_t^c + \eta_t^c n_t^c
\end{aligned} \tag{A.61}$$

which in turn can be plugged into the Bellman equation as follows:

$$\mu_t^c Q_t s_t^c + \mu_t^a q_t^a a_t^c + \mu_t^b q_t^b b_t^c + \eta_t^c n_t^c = \mathbb{E}_t \Lambda_{t,t+1} \left((1 - \sigma) + \sigma (\mu_{t+1}^c \phi_{t+1}^c + \eta_{t+1}^c) \right) n_{t+1}^c \tag{A.62}$$

Define the shadow price of a banker net worth as:

$$\Omega_{t+1}^c = (1 - \sigma) + \sigma (\mu_{t+1}^c \phi_{t+1}^c + \eta_{t+1}^c) \tag{A.63}$$

Then, using the law of motion of a banker net worth (1.2):

$$\begin{aligned}
\mu_t^c Q_t s_t^c + \mu_t^a q_t^a a_t^c + \mu_t^b q_t^b b_t^c + \eta_t^c n_t^c &= \mathbb{E}_t \Lambda_{t,t+1} \Omega_{t+1}^c n_{t+1}^c \\
\mu_t^c Q_t s_t^c + \mu_t^a q_t^a a_t^c + \mu_t^b q_t^b b_t^c + \eta_t^c n_t^c &= \mathbb{E}_t \Lambda_{t,t+1} \Omega_{t+1}^c \left((R_{t+1}^s - R_{t+1}) Q_t s_t^c + (R_{t+1}^a - R_{t+1}) q_t^a a_t^c \right. \\
&\quad \left. + (R_{t+1}^b - R_{t+1}) q_t^b b_t^c + R_{t+1} n_t^c \right)
\end{aligned} \tag{A.64}$$

which gives the solutions:

$$\mu_t^c = \mathbb{E}_t \Lambda_{t,t+1} \Omega_{t+1}^c (R_{t+1}^s - R_{t+1}) \tag{A.65}$$

$$\mu_t^a = \mathbb{E}_t \Lambda_{t,t+1} \Omega_{t+1}^c (R_{t+1}^a - R_{t+1}) \tag{A.66}$$

$$\mu_t^b = \mathbb{E}_t \Lambda_{t,t+1} \Omega_{t+1}^c (R_{t+1}^b - R_{t+1}) \tag{A.67}$$

$$\eta_t^c = \mathbb{E}_t \Lambda_{t,t+1} \Omega_{t+1}^c R_{t+1} \tag{A.68}$$

SHADOW BANKS

Given the balance sheet identity (1.16) and the evolution of net worth (1.17), the broker's optimization problem comes down to a Bellman equation. Using the method of undetermined coefficients, we guess that the solution is linear in the time-varying coefficients:

$$V_t^b = v_t^b Q_t s_t^b + \eta_t^b n_t^b \quad (\text{A.69})$$

Now insert the guessed solution into the Bellman equation then maximize the objective function subject to the incentive constraint (1.18). Using a Lagrangian multiplier λ_t^b , the problem writes:

$$\mathcal{L} = (1 + \lambda_t^b) (v_t^b Q_t s_t^b + \eta_t^b n_t^b) - \lambda_t^b \theta^b Q_t s_t^b \quad (\text{A.70})$$

The first-order conditions associated with s_t^b, λ_t^b are:

$$(1 + \lambda_t^b) v_t^b = \lambda_t^b \theta^b \quad (\text{A.71})$$

$$Q_t s_t^b = \phi_t^b n_t^b, \quad \text{if } \lambda_t^b > 0 \quad (\text{A.72})$$

where,

$$\phi_t^b = \frac{\eta_t^b}{\theta^b - v_t^b} \quad (\text{A.73})$$

Using the FOC (A.71), one can rewrite the guessed solution as:

$$V_t^b = v_t^b \phi_t^b n_t^b + \eta_t^b n_t^b \quad (\text{A.74})$$

which in turn can be plugged into the Bellman equation as follows:

$$v_t^b Q_t s_t^b + \eta_t^b n_t^b = \mathbb{E}_t \Lambda_{t,t+1} \left((1 - \sigma) + \sigma (v_{t+1}^b \phi_{t+1}^b n_{t+1}^b + \eta_{t+1}^b) \right) n_{t+1}^b \quad (\text{A.75})$$

Define the shadow price of a broker net worth as:

$$\Omega_{t+1}^b = (1 - \sigma) + \sigma(v_{t+1}^b \phi_{t+1}^b + \eta_{t+1}^b) \quad (\text{A.76})$$

Then, using the law of motion of a broker net worth (1.17):

$$\begin{aligned} v_t^b Q_t s_t^b + \eta_t^b n_t^b &= \mathbb{E}_t \Lambda_{t,t+1} \Omega_{t+1}^b n_{t+1}^b \\ v_t^c Q_t s_t^b + \eta_t^b n_t^b &= \mathbb{E}_t \Lambda_{t,t+1} \Omega_{t+1}^b ((R_{t+1}^s - R_{t+1}^a) Q_t s_t^b + R_{t+1}^a n_t^b) \end{aligned} \quad (\text{A.77})$$

which gives the solutions:

$$v_t^b = \mathbb{E}_t \Lambda_{t,t+1} \Omega_{t+1}^b (R_{t+1}^s - R_{t+1}^a) \quad (\text{A.78})$$

$$\eta_t^b = \mathbb{E}_t \Lambda_{t,t+1} \Omega_{t+1}^b R_{t+1}^a \quad (\text{A.79})$$

A.4 Data

Table A.1: Data Chapter 1

Description	Units	Sources
Gross Domestic Product	Billions	US B.E.A.
Gross Private Domestic Investment	Billions	US B.E.A.
Gross Domestic Product: Implicit Price Deflator	Index 2012	US B.E.A.
Nonfinancial Corporate Business; Depository Institutions loans; Liability	Millions	US Z.1 F.A.
Nonfinancial Noncorporate Business; Depository Institutions loans; Liability	Millions	US Z.1 F.A.
Nonfinancial Corporate Business; Debt Securities; Liability	Millions	US Z.1 F.A.
Nonfinancial Corporate Business; Corporate Equities; Liability	Millions	US Z.1 F.A.
Other loans and advances by issuers of ABS; asset	Millions	US Z.1 F.A.
All Federal Reserve Banks: Total Assets	Millions	H.4.1 F.A.R.B.
Excess Reserves of Depository Institutions	Millions	H.3 A.R.D.I.M.B.
US Treasury securities	Millions	H.4.1 F.A.R.B.
Federal agency debt securities	Millions	H.4.1 F.A.R.B.
Mortgage-backed securities	Millions	H.4.1 F.A.R.B.
Unamortized premiums on securities held outright	Millions	H.4.1 F.A.R.B.
Unamortized discounts on securities held outright	Millions	H.4.1 F.A.R.B.
Repurchase agreements	Millions	H.4.1 F.A.R.B.
Loans	Millions	H.4.1 F.A.R.B.
Net portfolio holdings of Maiden Lane LLC	Millions	H.4.1 F.A.R.B.
Float	Millions	H.4.1 F.A.R.B.
Central bank liquidity swaps	Millions	H.4.1 F.A.R.B.
Other Federal Reserve assets	Millions	H.4.1 F.A.R.B.
Foreign currency denominated assets	Millions	H.4.1 F.A.R.B.
Gold stock	Millions	H.4.1 F.A.R.B.
Special drawing rights certificate account	Millions	H.4.1 F.A.R.B.
Treasury currency outstanding	Millions	H.4.1 F.A.R.B.
Currency in circulation	Millions	H.4.1 F.A.R.B.
Reverse repurchase agreements	Millions	H.4.1 F.A.R.B.
Treasury cash holdings	Millions	H.4.1 F.A.R.B.
Deposits with F.R. Banks, other than reserve balances	Millions	H.4.1 F.A.R.B.
Effective Federal Funds Rate	Percent	H.15. B.G.F.R.S.
3-Month Treasury Bill: Secondary Market Rate	Percent	H.15. B.G.F.R.S.
Federal Funds Target Range - Upper Limit	Percent	B.G.F.R.S.
Federal Funds Target Range - Lower Limit	Percent	B.G.F.R.S.
Interest Rate on Required Reserves	Percent	B.G.F.R.S.
Interest Rate on Excess Reserves	Percent	B.G.F.R.S.

Initialisms/Acronyms: ABS = Asset-Backed-Securities, B.E.A = Bureau of Economic Analysis, F.A. = Financial Accounts, F.A.R.B. = Factors Affecting Reserve Balances, A.R.D.I.M.B. = Aggregate Reserves of Depository Institutions and the Monetary Base, B.G.F.R.S = Board of Governors of the Federal Reserve System.

Chapter 2

New Evidence on the Effects of Quantitative Easing

Abstract

Have the macroeconomic effects of QE programs been empirically overestimated? Using a large set of model specifications that differ in the degree of time-variation in parameters, the answer is yes. Our forecasting exercise suggests that it is crucial to allow for time-variation in parameters, but not for stochastic volatility. In an analysis of structural QE shocks, we find that QE1 had larger macroeconomic effects than QE2 and QE3, but much smaller than usually found in the literature.

GATE Working Paper 1912 – March 2019

2.1 Introduction

The Great Recession and the subsequent zero lower bound episode led the Federal Reserve to engage in QE programs. Those consisted in buying several types of assets to restore the functioning of credit markets and foster the real economy. It is recognized in the literature that QE pressured long-term yields and successfully improved financial conditions for banks and firms.¹ In addition, in the below-mentioned studies, the macroeconomic effects of QE are found to be large. However, these results emerge from vector autoregression (VAR) with constant parameters models. These dynamic systems are, in our view, subject to various potential issues, that lead us to raise the following question:

To what extent are these estimates reliable?

The VAR specification is routinely used rather than justified upon the underlying generating process of data. As such, the approach is susceptible to lead to important misspecification errors, especially when looking at the way in which QE programs were conducted in details. Indeed, the Federal Reserve had quite different objectives for each of the QE programs implemented.

In a nutshell, the QE1 program was an emergency plan to revitalize frozen private credit markets. The QE2 program, along with the Operation Twist, was employed to foster economic conditions and combat deflationary pressures. The QE2 program differed from its predecessor, as it only focused on purchasing long-term Treasury bonds rather than toxic assets. The QE3 program was

¹Ref: Gagnon et al. (2011), Baumeister & Benati (2013) and Christensen & Rudebusch (2012), among others.

state-contingent: intended to continue indefinitely until the economic fundamentals improve.

Consequently one can only expect differences in the macroeconomic effects, and in the underlying transmission channels (TC) of QE programs. Hence, we argue that models with constant coefficients might not be well suited to capture what we consider as *structural changes* in the macroeconomic impacts of QE programs. Along the same line, it is also worth testing for stochastic volatility.

VAR models allow for many degrees of freedom thanks to a relatively low number of variables. However, the use of sparse information might induce large overestimation of the identified effects. Furthermore, standard estimation procedures usually ignore model uncertainty. Beyond the fact that many scholars arbitrarily select a set of variables, the habit is to add or switch several variables as robustness checks instead of discriminating over all variable combinations. These approaches are questionable in light of the results of Hoeting et al. (1999), who demonstrate that considering model uncertainty greatly improves out-of-sample forecasting.

Therefore, there are good reasons to think that many empirical studies on QE programs are potentially affected by one or all of these problems. In this paper, we evaluate these issues using a large set of model specifications with different degrees of time-variation in parameters. Models are also estimated with a QE factor, capturing the different TC of QE. The term *specification* refers to the modeling assumptions of the systems. The set comprises systems that capture structural changes and stochastic volatilities. As such, the usual VAR specification with constant parameters, and homoskedasticity, is a particular case where additional restrictions are imposed. Overall, our evaluation comprises three parts.

First, we use a grid-search algorithm to select the *optimal* specification across all the possible variants in the set of model specifications. The ability to forecast macroeconomic and QE variables determines this *optimal* specification.²

Second, for each of the model specification, we account for model uncertainty in the time-varying relationship between QE, the real economy, and the associated TC of QE. More precisely, the unknown TC factor is estimated using all the possible combinations of observables related to the TC of QE. Each of these variables receives a time-varying weight in function of its forecasting power. Furthermore, the selection over variables depends on dynamical selection procedures.

Third, we discuss the implications of QE: (i) We gauge the time-varying forecasting importance of the portfolio balance, signaling, balance sheet and risk-taking channels of QE and (ii) We quantify the macroeconomic effects of QE programs based on the identification of structural shocks. The identification is achieved by means of Cholesky decompositions and a mix of zero and sign restrictions. Our findings are multiple.

The forecasting exercise suggests that the time-variation in parameters is of great importance. However, a stochastic volatility does not help improve the fit with the data. We find that the portfolio balance and signaling channels gain forecasting power over time. Conversely, the balance sheet channel has a steady implication, while the risk-taking channel is marginal. Further, despite the increasing forecasting importance of these two channels, the structural

²Given the large set of models to estimate, we resort to a very flexible estimation technique using the algorithm of Koop & Korobilis (2014), the task being computationally overwhelming otherwise. This method allows estimations to take a couple of days instead of several months, and consists of a two-step dual linear Kalman filter and smoother. The algorithm is detailed in the online appendix.

impulse responses show that the macroeconomic effects of QE are not linearly and monotonically increasing with the magnitude of QE shocks, and over time. Indeed, we find that the QE1 program has relatively more effects on real variables than QE2 and QE3 programs, suggesting that QE programs might have asymmetric macroeconomic effects.

The transmission channels of QE are relatively well identified in the theoretical literature, but there are discussions about their importance and in the size of the associated macroeconomic effects.³ On the one hand, Curdia & Woodford (2011) claim that QE should be neutral and have no macroeconomic effects. On the other hand, studies using a mechanism à la Gertler & Karadi (2011) find that QE programs should have large effects. From an empirical perspective, QE programs are found to have strong effects on the real allocations; as in Gambacorta et al. (2014) and Weale & Wieladek (2016) (WW). We believe that our results – QE programs had significant macroeconomic effects but decreasing over time – may reconcile theoretical with empirical results. Indeed, QE had significant effects during the period of economic distress during which QE1 was implemented. At that time, financial frictions were exacerbated, easing the transmission of the QE1 program. Conversely, QE has smaller macroeconomic effects during the recovery period, when QE2 and QE3 programs were implemented, because financial constraints were slack (or slacker) at that time.

The paper proceeds as follows. Section 2.2 presents the general framework, the estimation and selection methods. Section 2.3 describes the data, and reports the results of the forecasting exercise. Section 2.4 offers a discussion of the policy analysis. Section 2.5 concludes.

³Woodford et al. (2012) provides a long discussion to feed the debate.

2.2 Framework, Estimation and Selection methods

2.2.1 Framework

In the literature, researchers have frequently neglected to check for the adequacy between model specification and the generating processes of data, using mostly low-dimensional VAR systems. However, the scarceness of information is generally little advised in econometrics. The short sample of QE programs that we situate between 2008M12 and 2014M11 is clearly characterized by different asset targets but also a mixture of unconventional monetary tools such as “forward guidance” and rescue plans. One can therefore consider an evolution in the transmission channels of QE, as well as the presence of heteroskedasticity. Our goal is gauge the marginal contributions of the addition of: (i) Structural changes through time-varying parameters, (ii) Heteroskedasticity through the stochastic volatility in covariance matrices.

A model misspecification can lead to substantial biases; therefore, misleading inference about forecasting and the magnitude of impulse responses. Thus, we first generalize the standard VAR process to include time-varying parameters. Second, we augment each model specification with a time-varying unobserved factor; that is, variables entering the estimate can switch over time.

Hereafter, we perform a model *specification* and a model *selection*. The model specification consists of testing different modeling assumptions across a set of variants. The changes of specifications will depend on the value of a bunch of hyperparameters, as later explained.

The model selection periodically picks out a combination, or averages on all the possible combinations of variables that compose the factor. Pre-

cisely, we perform Bayesian/Dynamical Model Averaging (BMA, DMA) and Bayesian/Dynamical Model Selection (BMS, DMS) using the algorithm of Koop & Korobilis (2014) (KK). The discriminating criterion is the forecasting power of each of the variable entering the estimation of the TC factor, as detailed later.

Over the variants, the general model specification is an TVP-FAVAR with stochastic volatility. For $t = 1, \dots, T$, let y_t be an $s \times 1$ vector of macroeconomic and QE variables. Let f_t be the unobserved TC factor. The observables used for the estimation of f_t are comprised in x_t , an $n \times 1$ vector. Hence, the model selection is performed across the $2^n - 1$ possible combinations of x_t .⁴ Consequently, the M_i for $i = 1, \dots, I$ general p -lags models, are of the form of:

$$x_{i,t} = \lambda_{i,t}^f f_{i,t} + \lambda_{i,t}^y y_t + \mu_{i,t} \quad (2.1)$$

where $\lambda_{i,t}^f$ are $n \times 1$ vectors of factor loadings and $\lambda_{i,t}^y$ are $n \times s$ matrices of regression coefficients. The Gaussian errors $\mu_{i,t} \sim \mathcal{N}(0, V_{i,t})$ are time-varying. Equations (2.1) form linear space equations, the systems are completed by i state equations:

$$\begin{pmatrix} f_{i,t} \\ y_t \end{pmatrix} = B_{i,t,1} \begin{pmatrix} f_{i,t-1} \\ y_{t-1} \end{pmatrix} + \dots + B_{i,t,p} \begin{pmatrix} f_{i,t-p} \\ y_{t-p} \end{pmatrix} + \varepsilon_{i,t} \quad (2.2)$$

where $(B_{i,t,1}, \dots, B_{i,t,p})$ are time-varying VAR coefficients and $\varepsilon_{i,t} \sim \mathcal{N}(0, Q_{i,t})$ are Gaussian errors. The VAR coefficients and factor loadings evolve according

⁴We obviously preclude the empty set.

to random walk equations:

$$\lambda_{i,t} = \lambda_{i,t-1} + v_{i,t} \quad (2.3)$$

$$\beta_{i,t} = \beta_{i,t-1} + \eta_{i,t} \quad (2.4)$$

where $\lambda_{i,t} = ((\lambda_{i,t}^f)^T, (\lambda_{i,t}^y)^T)^T$, $\beta_{i,t} = (\text{vec}(B_{i,t,1})^T, \dots, \text{vec}(B_{i,t,p})^T)^T$, and $v_{i,t} \sim \mathcal{N}(0, W_{i,t})$, $\eta_{i,t} \sim \mathcal{N}(0, R_{i,t})$ are also Gaussian distributed. All of the above errors are by assumption, uncorrelated over time.

2.2.2 Estimation and Selection methods

We estimate a *great* number of different model specifications. For each of the model specification, the procedure also involves estimating all the i combinations of x_t . Thus, the entire estimation procedure is computationally overburdening, even inconceivable with simulation methods such as MCMC algorithms. Consequently, the simulation-free algorithm of KK, which consists of a dual linear Kalman filter and smoother (KFS), is a doable procedure among the alternatives. The algorithm proceeds as follows:⁵ For each i , a first KFS updates the parameters $\theta_t = (\lambda_t, \beta_t)$ given a principal component estimate of the factor \hat{f}_t . Afterwards, a second KFS updates f_t given the estimates of θ_t .

The error covariance matrices (V_t, Q_t, W_t, R_t) are also estimated using simulation-free methods. The covariances V_t and Q_t are computed with an exponentially weighted moving average method. In this estimator, two decay hyperparameters, κ_1 and κ_2 , manage the degree of stochastic volatility in, respectively, the space (2.1) and state (2.2) equations. In turn, W_t and Q_t are estimated using a variance discounting method à la Koop & Korobilis (2013),

⁵A complete description of the algorithm is provided in the online appendix.

in which two forgetting hyperparameters, κ_3 and κ_4 , control the time-variation of, respectively, λ_t and β_t . The special case where, $\kappa_1 = \kappa_2 = 1$, imposes V_t and Q_t to be time-invariant. In that situation, the model specifications are homoskedastic. Similarly, λ_t and β_t are constant when $\kappa_3 = \kappa_4 = 1$. In that case, the model specification is a FAVAR. The *optimal* values of the κ_i are determined by a grid-search algorithm, as explained below.

This Bayesian filtering and smoothing algorithm requires prior conditions. We choose uninformative priors for two reasons: the recommendations of KK and the stationarity of data. In addition, a training sample would be misleading because QE was exclusively implemented during the span of the sample. The initial conditions are given by:

$$\begin{aligned} f_0 &\sim \mathcal{N}(0, 4), & \beta_0 &\sim \mathcal{N}(0, R_\beta), & R_\beta &= \frac{\gamma}{r^2}, \\ \lambda_0 &\sim \mathcal{N}\left(0, 4 \times \begin{matrix} I \\ (s+1) \times (s+1) \end{matrix}\right), & V_0 &\equiv \begin{matrix} I \\ n \times n \end{matrix}, & Q_0 &\equiv \begin{matrix} I \\ (s+1) \times (s+1) \end{matrix}. \end{aligned} \quad (2.5)$$

where $r = 1, \dots, p$. R_β is a diagonal covariance matrix in the flavor of Minnesota priors, penalizing distant lags. In addition, V_t is diagonal so that $\mu_{i,t}$ errors are idiosyncratic; therefore, the estimation of the factor absorbs information only related to TC variables.

To sum up, the algorithm is common to all of the variants of the model specifications. Each specification differs in the values of the κ_i . For instance, a model is an TVP-FAVAR when either κ_3 or κ_4 is lower than 1, or a specification is heteroskedastic if either κ_1 or κ_2 is lower than 1. In the grid-search algorithm, we restrict the κ_i to three values $\{0.96, 0.98, 1.00\}$, in line with the recommendations of KK, and the three implied distinct cases: The parameters are static for $\kappa_i = 1$, slowly varying for $\kappa_i = 0.98$ and fairly moving $\kappa_i = 0.96$.

As seen in Figure B.1 in the appendix, a $\kappa_1 = 0.96$, induces that the data point from a year ago weighs 60% as much as the $t - 1$ point in the estimation of the covariance V_t ; causing a high stochastic volatility in the equation (2.1).

We use Bayesian/Dynamical Selection/Averaging techniques (BMA, BMS, DMA, DMS) to shed light on the time-varying forecasting importance of each of the TC variables. A DMS dynamically selects the combination of x_t with the highest forecasting performance (for each t). Conversely, DMA is a time-varying weighted average of all the combinations, where the weights evolve in function of the forecasting power of each combination. To do so, let be $\pi_{t|t-1,i}$ the predicted and $\pi_{i,t|t}$ the updated weights determined by the following forms:⁶

$$\pi_{i,t|t-1} = \frac{\pi_{i,t-1|t-1}^\alpha}{\sum_{j=1}^I \pi_{j,t-1|t-1}^\alpha} \quad (2.6)$$

$$\pi_{i,t|t} = \frac{\pi_{i,t|t-1} p_i(y_t | Data_{1:t-1})}{\sum_{j=1}^I \pi_{j,t|t-1} p_j(y_t | Data_{1:t-1})} \quad (2.7)$$

where $\alpha \in [0,1]$ is a forgetting factor that controls the degree of switching between the combinations.⁷ The updated probabilities $\pi_{i,t|t}$ are time-varying in function of the predictive likelihood $p_i(y_t | Data_{1:t-1})$. As for the κ_i , a low value of α leads to a high degree of time-variation. The case where $\alpha = 1$ induces recursive BMA or BMS. BMA/BMS are recursive windowing methods, while DMA/DMS are dynamical updating methods in which weights are exponentially decaying over time ($0 < \alpha < 1$).

⁶The method to predict, and update the time-varying weight is derived from Raftery et al. (2010).

⁷The weights are equally initialized such that $\pi_{i,0|0} = \frac{1}{I}$.

2.3 Data and Forecasting Exercise

2.3.1 Data

We collect US monthly data from the Federal Reserve Economic Database.⁸ The period of the estimation spans the QE programs of the Federal Reserve, from 2008M12 to 2014M11. Three observables are comprised in y_t : The Consumer Price Index (CPI), the real Gross Domestic Product (GDP), and the total assets of all Federal Reserve banks (used as a proxy for QE).⁹ These series are transformed in first log-difference.¹⁰

The TC factor is extracted using eight time series related to four TC of QE; the portfolio balance channel, the signaling channel, the balance sheet channel and the risk-taking channel. The portfolio balance channel is captured by the ten-year Treasury term premium estimates of Adrian et al. (2013), and the Moody's BAA corporate bond yield relative to the ten-year Treasury constant maturity rate. The signaling channel is proxied with the estimates of the ten-year expected average short-term rates by Adrian et al. (2013) and the ten-year breakeven inflation rate. The balance sheet channel consists of households and nonfinancial corporate business net worth series, both taken in first log-difference and interpolated from quarterly frequencies. The risk-taking channel is captured by the growth rates of St. Louis stress and the CBOE volatility indexes.

All model variants have two lags. The total number of different estimated

⁸Ref: <https://fred.stlouisfed.org/>, more details about data are given in the appendix C.

⁹The monthly frequency of the real GDP is approximated using a piecewise cubic hermite interpolator. We believe that such interpolation is more reliable than any use of monthly proxies such as industrial or activity indexes.

¹⁰The series are also standardized in the forecasting exercise.

models is $3^5 \times (2^8 - 1) = 61965$, given the five hyperparameters and eight x_t variables.

2.3.2 Preliminary discussion

The macroeconomic implications of QE programs have been studied at length in the literature. However, the most intriguing is that empirical studies find strong and robust macroeconomic effects of QE (Baumeister & Benati (2013), WW), while some theoretical studies support weak or neutral impacts of QE (Curdia & Woodford (2011), Jouvanceau (2019)). In that respect, Curdia & Woodford (2011) argue that QE is *irrelevant* absent of financial frictions, but potentially effective through a signaling channel. In other words, QE is effective if it is able to change expectations about the future stance of monetary policy. Empirically, Krishnamurthy & Vissing-Jorgensen (2011) or Christensen & Rudebusch (2012) find supportive evidence about these signaling components. We test for the relevance of this channel by including measures of market expectations such as the ten-year expected average short-term rates and the ten-year breakeven inflation rate.

According to Bernanke (2012), QE transmits through the portfolio balance channel. The latter roots in the presence of financial frictions. For Vayanos & Vila (2009) financial frictions emerge in investors' preferred-habitats. In this theory markets are segmented because of the preferences of investors regarding the maturities of assets. For investors to invest in another segment, the expected return must offset the incurred risk; therefore assets are not perfect substitutes. In such environment, QE shifts the yields of close substitutes due to the change in the supply of a targeted asset. In particular, QE exerts a

downward pressure on the duration risk of assets; hence, reducing long-term rates. WW provide empirical evidence of the portfolio balance of QE. In this work, we assume that the ten-year Treasury term premium and the ten-year corporate bond spread are thus conceivable variables to evaluate the potential contribution of this channel.

Gertler & Karadi (2011) (GK) adopt a financial accelerator mechanism in a new-Keynesian model to capture the squeeze in the credit supply in the aftermath of a financial crisis. In this framework, the supply of credit is a function of the size of bankers' balance sheets; causing an excess premium. GK argue that QE is largely effective by loosening balance sheet constraints, hence pressuring excess premiums. However, Jouvanceau (2019) nuances these findings and asserts that most of the macroeconomic effects of QE are weaker than found by GK, and are mainly transmitted through income effects – asset prices. Along these theories, the fluctuations in the net worth of private agents are crucial indicators of, both, the financial accelerator and the income effects. For these reasons, we use series of households and nonfinancial corporate business net worth to proxy the balance sheet channel of QE; which encompasses these effects.

Another transmission channel of QE programs is the risk-taking channel. In theory, risk-taking behavior is negatively correlated with the level of short-term rates. Dell'Ariccia et al. (2017) provide empirical evidence of this phenomenon. From that, one can infer that a downward pressure on long-term rates has opposite effects on the attitude toward risk; characterized by search for yields. In addition, the excess of liquidity leaves room for the funding of uncreditworthy agents. We thus control for the risk components by including the St. Louis stress and the CBOE volatility indexes.

Knowing the outlines for each of these channels, we gauge for their relative importance in the following forecasting exercise.

2.3.3 Forecasting exercise

The forecasting exercise determines the *optimal* model specification in the set of the possible variants. The hyperparameter optimization is conducted using a grid-search algorithm. The grid comprises 3^5 combinations of the five hyperparameters $(\kappa_1, \kappa_2, \kappa_3, \kappa_4, \alpha)$ for the values $\{0.96, 0.98, 1.00\}$. As a reminder, the hyperparameters define the model specification and the degree of time-variation in the coefficients and between the combinations. In particular, the decay factors κ_1 and κ_2 handle the stochastic volatility in, respectively, the space (2.1) and the state (2.2) equations. The forgetting factors κ_3 and κ_4 control the time-variation of, respectively, the factor loadings and the VAR coefficients in the random walk equations (2.3) and (2.4). The forgetting factor α shapes the rate of switching between the combinations of x_t (*cf.* equation (2.7)).

For instance, the case where $(\kappa_i = \alpha = 0.96)$ defines an TVP-FAVAR model with high stochastic volatility and fast DMA/DMS. The situation where $(\kappa_i = \alpha = 0.98)$ is also a TVP-FAVAR, but with slower time-variation in the coefficients, stochastic volatility and DMA/DMS. Last, the case where $(\kappa_i = \alpha = 1.00)$ induces an homoskedastic FAVAR with BMA/BMS procedures. The pattern to build all the model variants is thus straightforward. The out-of-sample forecasting is applied to the $3^5 \times (2^8 - 1) = 61965$ estimated models.

The two metrics of forecasting evaluation are the mean squared forecast errors (MSFE) and the one-step ahead predictive likelihoods (PL). The forecasting

period spans from 2009M3 to 2014M11- h for $h = 1, 2, 3, 4$ months ahead.¹¹ Our ranking procedure proceeds as follows. In a first step, we compute the arithmetic means of MSFE and PL over the h horizons and the dimension s of y_t .¹² In a second step, the means of the step one are ranked in an ascending/descending order for PL/MSFE and BMA/DMA or BMS/DMS. Finally, we compute the average of the rankings of the previous step. Consequently, the top 1 specification is the one with the lowest average ranking across the metrics of evaluation, over DMA/DMS or BMA/BMS.

Table 2.1 presents the top 10 of model specifications throughout the forecasting performance. The first line shows the averaged metrics of the top 1 specification. Other specifications are normalized by the results of the top 1 specification, so that any entry lower than 1 indicates a worse forecasting performance. The top 1 specification is an TVP-FAVAR with a weak stochastic volatility in the space equation (2.1), homoskedasticity in the state equation (2.2) and a fast time-variation in the VAR and the loading coefficients. In addition, BMA/BMS methods are favored. In particular, BMA should not be misunderstood. BMA does not imply that the x_t combinations are always equally weighted in the estimation of the factor.¹³ BMA is a windowing recursive model averaging method.¹⁴ To give the reader an overview, the worst specification is an heteroskedastic FAVAR, that performs roughly 15% worst than the top 1 specification. We believe that such underperformance is rather considerable given the size of the sample. This worst specification is characterized by a

¹¹Forecasts are iterated for $h > 1$.

¹²The PL are one-step ahead predictions, therefore are not averaged.

¹³This is the case if $\alpha = 0$.

¹⁴Further explanations about BMA/BMS/DMA/DMS are detailed in the paper of Raftery et al. (2010).

fast decaying volatility in the space equation and a weak volatility in the state equation. Moreover, it demands constant VAR and loading coefficients. The corresponding hyperparameters are $\alpha = 1.00$, $\kappa_1 = 0.96$, $\kappa_2 = 0.98$, $\kappa_3 = 1.00$, $\kappa_4 = 1.00$.

Conversely, the differences in the forecasting power between the top 10 specifications are clearly insignificant. These models share a fast time-variation in the β_t coefficients and a fair change in the loadings λ_t . In addition, six out of the ten specifications highlight the benefit of the absence of stochastic volatility, while the remaining four require a weak stochastic volatility, but only in the space equation. From these findings, we assert that QE could have had asymmetric transmission to the real economy.

In addition, the updated time-varying probabilities $\pi_{i,t|t}$ in the Figure B.2 (in the appendix) shed light in the evolution of the transmission of QE.¹⁵ The portfolio balance and the signaling channels have increasing forecasting power over time. Conversely, the risk-taking channel is completely muted while the balance sheet channel has a steady implication.

The importance of the portfolio balance and signaling channels is shared with the above-cited literature. However, the asymmetric transmission of QE, is a new contribution. Overall, this exercise shows the importance of testing for various model specification and uncertainty so that to avoid misleading inference. In that direction, the next section highlights the consequence of misspecification in the quantification of the macroeconomic effects of QE derived from an analysis of structural QE shocks.

¹⁵Each x_t variable enters the estimation of the factor with a probability of $\frac{\sum_{k=1}^n \binom{n-1}{k-1}}{\sum_{k=1}^n \binom{n}{k}} = 50.20\%$, with $n = 8$.

Table 2.1: Top 10 of model specifications

TOP	PL DMA	PL DMS	MSFE DMA	MSFE DMS	SPECIFICATIONS	α	κ_1	κ_2	κ_3	κ_4
1	0.4929	0.4924	0.6416	0.6190	TVP-FAVAR	1.00[†]	0.98	1.00	0.96	0.96
2	0.9997	0.9998	1.0009	1.0029	TVP-FAVAR*	1.00	1.00	1.00	0.96	0.96
3	0.9999	0.9995	0.9978	0.9968	TVP-FAVAR*	1.00	1.00	1.00	0.98	0.96
4	0.9995	1.0014	0.9966	0.9851	TVP-FAVAR	0.98	0.98	1.00	0.96	0.96
5	1.0001	0.9965	0.9969	0.9901	TVP-FAVAR	1.00	0.98	1.00	0.98	0.96
6	0.9993	1.0006	0.9973	0.9898	TVP-FAVAR*	0.98	1.00	1.00	0.96	0.96
7	0.9996	1.0003	0.9951	0.9847	TVP-FAVAR*	0.98	1.00	1.00	0.98	0.96
8	0.9997	1.0009	0.9943	0.9796	TVP-FAVAR	0.98	0.98	1.00	0.98	0.96
9	0.9993	1.0006	0.9960	0.9821	TVP-FAVAR*	0.96	1.00	1.00	0.96	0.96
10	1.0000	0.9985	0.9939	0.9868	TVP-FAVAR*	1.00	1.00	1.00	1.00	0.96

[†] When $\alpha = 1$, BMA and BMS are conducted.

*Are fully homoskedastic model specifications.

MSFE are the mean squared forecast errors. PL are the predictive likelihoods. The first line displays the mean of the MSFE over the 4 h -steps and the one-step ahead PL for the top 1 model specification. The remainder are normalized by the corresponding values of the top 1 specification. Hence, any entry lower than 1 indicate a worse forecasting performance.

2.4 Policy analysis

2.4.1 Identification of structural shocks

The above analysis makes it clear that simple VARs are most likely misspecified, and that TVP-FAVAR models provide a better fit with the data. As such, one can arguably question the results from the literature that stem from simple VARs regarding the macroeconomic effects of QE programs. In this section, we compute the structural QE shocks in the top 1 specification and compare it to the literature. For the sake of parsimony, we limit the analysis to the structural shocks related to the y_t variables.¹⁶ Consequently, the VMA form of the top 1

¹⁶The refinement to the entire set of x_t combinations would imply $(2^8 - 1)$ identification restrictions. The latter is feasible for simple triangular zero restrictions but computationally huge for a mix of zero and sign restrictions. This would require further technical progress that are beyond the scope of this paper.

TVP-VAR, is sufficient to pin down the desired structural shocks:

$$z_t = \begin{pmatrix} f_t \\ y_t \end{pmatrix} = B_t(L)^{-1}\varepsilon_t \quad (2.8)$$

where $B_t(L) = I - \sum_{l=1}^p B_{l,t}L^l$ and L is the lag operator. The errors ε_t are the reduced-form shocks. The uncovered structural shocks can be defined as $\vartheta_t = \mathcal{S}_{0,t}\varepsilon_t$ in which $\mathcal{S}_{0,t}$ are the structural contemporaneous parameters so that the state covariance writes $Q_t = (\mathcal{S}_{0,t}\mathcal{S}_{0,t}^T)^{-1}$.¹⁷ In line with the theoretical concepts of Rubio-Ramirez et al. (2010) (RMZ), we propose two sorts of linear restrictions on the space of structural parameters. The first scheme restrict the matrix of the short-run impulse responses to be lower triangular by a sequence of Cholesky factorization. This system is globally identified regarding necessary and sufficient rank conditions.¹⁸ However, despite its tractability, such recursive identification is at odds with theoretical predictions.

Hence, we consider an additional and agnostic scheme of restrictions. The latter is a mix of zero and sign restrictions. In the presence of sign restrictions, an TVP-VAR is not identified. To tackle this issue, the algorithm of Binning (2013) allows to compute a set of admissible impulse responses that satisfy the zero and sign restrictions. The flexibility of this restriction scheme allows for some shocks to be unrestricted while the remainder are locally identified. That way, we assert that the related outcomes are robust elements of a proper quantification about the macroeconomic effects of QE programs. The restriction schemes are summarized in the Tables 2.2 and 2.3.

¹⁷Assuming that $\mathbb{E}_t[\vartheta_t\vartheta_t^T] = I$.

¹⁸Proofs and algorithms are detailed in the online appendix

Table 2.2: Identification scheme: lower triangular zero restrictions

	AS	AD	QE	TC
CPI	1	0	0	0
GDP	x	1	0	0
QE	x	x	1	0
FAC	x	x	x	1

A "x" indicates that a structural shock is left unrestricted. A "0" is a zero restrictions. AS = aggregate supply, AD = aggregate demand, TC = transmission channels.

Table 2.3: Identification scheme: mixture of zero and sign restrictions

	AS	AD	QE	TC
CPI	+	+	x	x
GDP	-	+	x	x
QE	0	0	+	x
FAC	x	x	x	x

The "+", "-" and "0" are the signs and zero restrictions to the structural shocks.

In the first identification scheme, the restrictions have the following interpretation: The structural shocks of the transmission channels of QE have no contemporaneous effects on inflation, GDP and the total of assets of all Federal Reserve banks. The structural QE shocks have no short-run effects on inflation and GDP. The structural aggregate demand shocks have no contemporaneous impacts on inflation.

In a plain english, the restrictions in the second identifying scheme mean: The structural aggregate supply and demand shocks have no contemporaneous effects on the total assets held by the Federal Reserve banks. In addition, we impose that the structural aggregate supply shocks have a positive impact on inflation and a negative effect on GDP. An aggregate demand shocks increase inflation and GDP. Last, structural QE shocks are characterized by an increase

in the total of assets of all Federal Reserve banks.

2.4.2 Impulse responses analysis

The structural shocks are computed over the period 2009M3-2014M11 by using the smoothed estimates of the parameters $\beta_{i,t}$ and $Q_{i,t}$ of the top 1 model specification. The time-varying structural IRF are first presented in three-dimensional surface plots. Second, we dissociate the responses by averaging the IRF for each of the QE programs.¹⁹ In details, the QE1 was in place between 2009M3 to 2010M6 while the QE2 was spreading over 2010M10 to 2012M12. Finally, the QE3 lasted between 2013M1 and 2014M11.

Figures from B.3 to B.14 in the appendix display the 3D IRF of unitary QE structural shocks. Remember that in the lower triangular scheme, inflation and GDP do not contemporaneously respond to QE shocks while in the mix of zero and sign restrictions, inflation and GDP are left unrestricted. Moreover, the 68% and 90% percentiles of draws are computed as a tool for confidence intervals.²⁰

For the ease of visibility, we only show the median responses in the surface plots, the confidence intervals are displayed in the 2D plots only. A first inspection of the results highlights that the IRF are qualitatively and quantitatively evolving over time. In particular, the effects on inflation are high and persistent in the QE1 period. Conversely, the effects are sharp, weaker and less persistent during QE2 and QE3. The effects on the growth rate of GDP seem qualitatively

¹⁹The corresponding matrices of IRF are huge due to the number of dimensions (sample size, horizons, number of variables, number of shocks, draws, number of combinations). Hence, we restrict the number of draws to 100 in both bootstrap and sign restriction algorithms to avoid memory issues. A higher number of draws marginally refines the smoothness IRF.

²⁰In the Cholesky scheme, the draws are obtained with a bootstrap algorithm. In the mixture of zero and sign restrictions, the draws are given by the algorithm of Binning (2013). A sketch of this algorithm is detailed in the online appendix.

similar to those of inflation, but mostly delayed by a quarter. The magnitudes are lower on average.

Figures from B.15 to B.20 in the appendix show the averaged IRF for each of the QE programs. In QE1, we thus learn that the statistical significance of the responses of inflation occurs after a semester, especially in the mixture of restrictions. The average responses of GDP display a similar pattern. In other words, QE shocks have delayed impacts on inflation and GDP. This pattern is reinforced in the QE2 period; GDP and inflation being entirely in phase. However, in that period, the persistence in the effects is distinctively less important than during QE1. The IRF of QE3 resemble to those of QE2. These findings are robust to the use of BMS.

Looking at these IRF, one notices that the magnitude of the effects of QE shocks on GDP and inflation decline over time. However, one should remain cautious, since the total amounts of purchased assets differed in each of the QE programs. Consequently, we gauge the total effects of QE shocks by computing accumulated IRF.

Figures from B.21 to B.26 in the appendix display the corresponding outputs. In both schemes, the IRF are significant in the width of two standard-deviations around the medians. Hence, we compute the total effects of QE shocks by proceeding to simple back of envelope calculations. First, we average the accumulated IRF across schemes and BMA/BMS. Second, we multiply the long-term steady value of medians of CPI and GDP by the total change in the assets for each of QE programs.²¹

²¹The long-term horizon used for the calculations is 60. The total assets held by the Federal Reserves changed by 17.76% in the QE1 program, by 25.88% in the QE2 program and by 51.33% in the QE3 program.

The calculations then suggest that QE1 roughly lead to a total rise of 1.30% in CPI and 1.65% in GDP. These findings are close to the exposed results of previous studies. For example, Baumeister & Benati (2013) and WW respectively estimate that QE1 increased CPI by 0.90% and 1.50%, and GDP by 1.08% and 1.40%. However, being computed with BVAR models with data from 2009M3 to 2014M5, the estimates of WW (among others) are somehow problematic for further inference.²² In particular, they find that the maximum impacts of a unitary structural QE shock induces an increase of 0.58% in GDP and 0.62% in CPI.²³ Hence, by linearity, in their study, the total effects would be 2.10% and 3.92% for CPI and 1.97% and 3.67% for GDP in the QE2 and QE3 programs, respectively.²⁴ In our top 1 specification the corresponding total effects are up to {1.57%, 2.03%} (CPI, GDP) for the QE2 program, and {2.16%, 2.95%} for the QE3 program, which is overall significantly lower.

We proceed to further comparison by also estimating VAR models augmented with the estimated factors of the top 1 specification.²⁵ In this specification, the total effects are of {1.45%, 2.11%, 3.94%} for CPI and {1.95%, 2.79%, 5.28%} for GDP, in QE1, QE2 and QE3, respectively. It is straightforward but, having constant effects over time; the impact of QE are linearly increasing with the magnitude of the amounts purchased. However, in our top 1 specification with time-varying parameters, the total effects on CPI {1.29%, 1.57%, 2.16%} and GDP {1.65%, 2.03%, 2.95%} are clearly not monotonic.

Overall, the multipliers are {0.072%, 0.060%, 0.042%} for CPI and {0.092%, 0.078%, 0.057%}

²²In addition, they assert that: “asset purchases did not become less effective over time”

²³In the study of WW, QE structural shocks are expressed in percent of GDP (2009Q1).

²⁴These estimates are computed by considering that the total assets of QE purchases was worth 3.40% and 6.32% of 2009Q1 GDP in QE2 and QE3, respectively.

²⁵The corresponding IRF of these VAR models are depicted in the appendix.

for GDP in QE1, QE2, QE3, respectively.²⁶

The lesson from the above analysis is that the effects of QE shocks estimated by VAR specifications are certainly overestimated. In addition, QE had asymmetric effects on the real economy. In particular, QE was more effective when the markets were disrupted as when the QE1 program was implemented; that is, when financial frictions were exacerbated. Conversely, QE had marginally lower macroeconomic effects when the QE2 and QE3 programs were executed; even though the portfolio balance and the signaling channels gained forecasting importance over time.

2.5 Conclusion

In the empirical literature regarding the macroeconomic effects of QE, studies rely on simple VAR specification. In this paper, we expose the potential issues of such modeling assumptions. We find that they can induce misleading insights about the transmission of QE to the real economy, and may overestimate the macroeconomic effects of QE programs.

These results emerge from an *optimal* selection among a large set of model specifications that allows for structural changes in parameters and stochastic volatility. The *optimal* specification is found using a grid-search algorithm over hyperparameters that rule the variants. Specifications are ranked according to their ability to predict macroeconomic and QE variables. All models are augmented with an unobserved factor that represents the different theoretical TC of QE. In addition, we take uncertainty into account by using all possible

²⁶Multipliers represent the median responses of CPI and GDP given an 1% accumulated increase in QE shocks.

combinations of the TC variables which serve in the estimation of the factor. Afterward, dynamic selection and averaging methods inform about the time-varying forecasting importance of each TC.

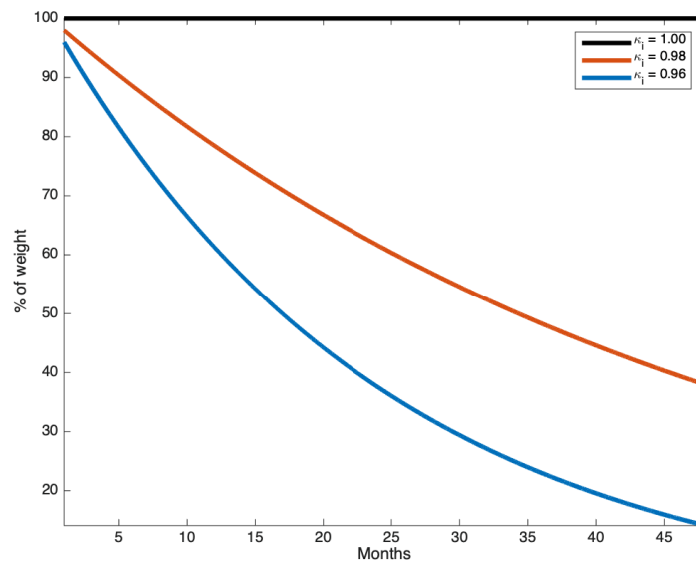
On the one hand, our results suggest that the time-variation in parameters improves the fit of the estimation. On the other hand, stochastic volatility is not critical. Given the importance of time-variation in parameters, we find that QE was asymmetrically transmitted to the real economy across the different QE programs. This assertion is further supported by a policy analysis of structural QE shocks. Indeed, we find that the macroeconomic effects of QE do not increase monotonically in the amount of purchases. In particular, QE was more impacting during QE1, when financial markets were severely impaired, from 2009M3 to 2010M6.

On the contrary, despite the increasing forecasting importance of the portfolio balance and signaling channels between 2010M10 to 2014M11 (QE2 and QE3 periods), the estimated macroeconomic impacts of QE are lower. We argue that the slackness of financial constraints, in boom cycles, potentially accounts for the reduced macroeconomic effects of QE. These asymmetries pave the way for additional and interesting theoretical studies, that could look at the effects of QE depending on the slackness of financial constraints.

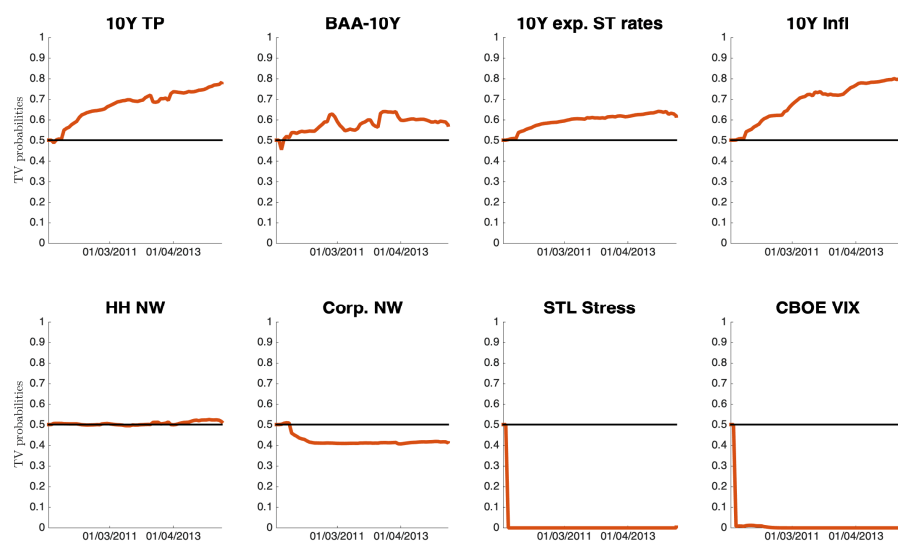
Appendix B

B.1 Figures

Figure B.1: Decaying importance in the estimation of covariances

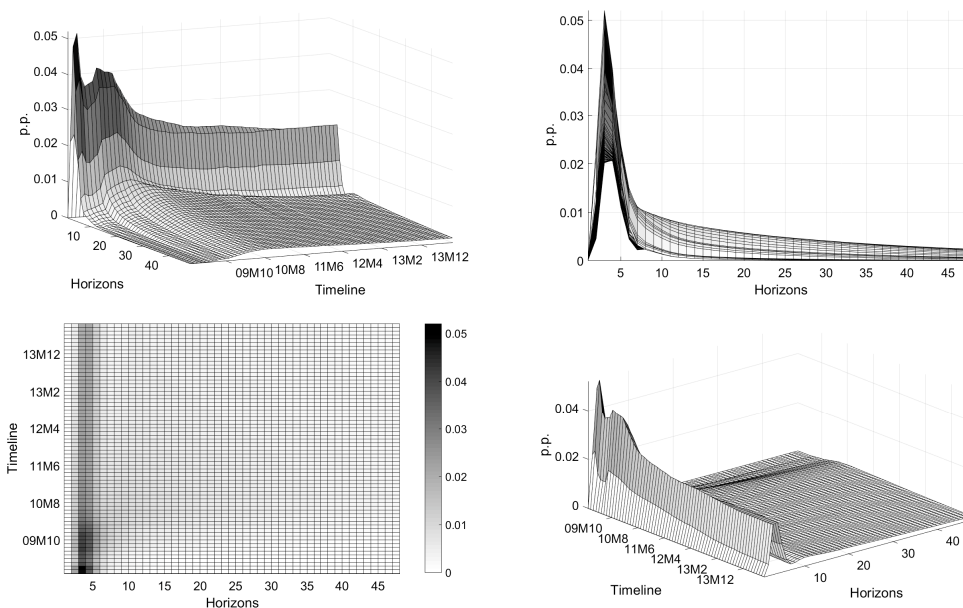


These lines display the decaying importance of data from x months ago in the estimation of each covariance, for each of the κ_i cases.

Figure B.2: Time-varying probabilities $\pi_{i,t|t}$ 

The red solid lines show the time-varying probabilities $\pi_{i,t|t}$ in the top 1 specification. The black solid lines are the theoretical probabilities of equally included x_t variables. Details about data and acronyms are given in the appendix.

Figure B.3: TOP 1: CPI to QE shocks: Cholesky & BMA



IRF in the surface plots are displayed using the median of either bootstrap or sign restrictions draws.

Figure B.4: TOP 1: GDP to QE shocks: Cholesky & BMA

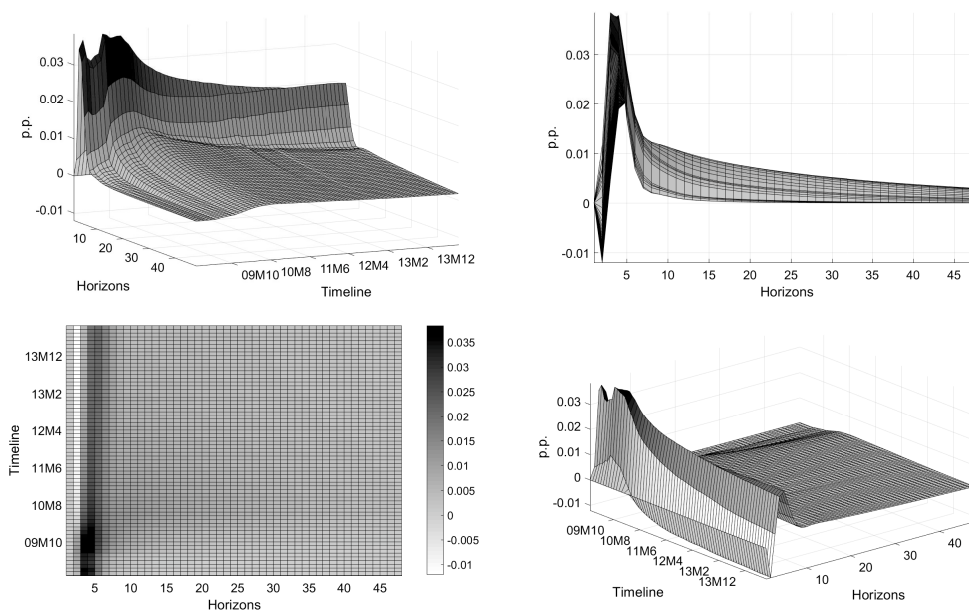


Figure B.5: TOP 1: QE to QE shocks: Cholesky & BMA

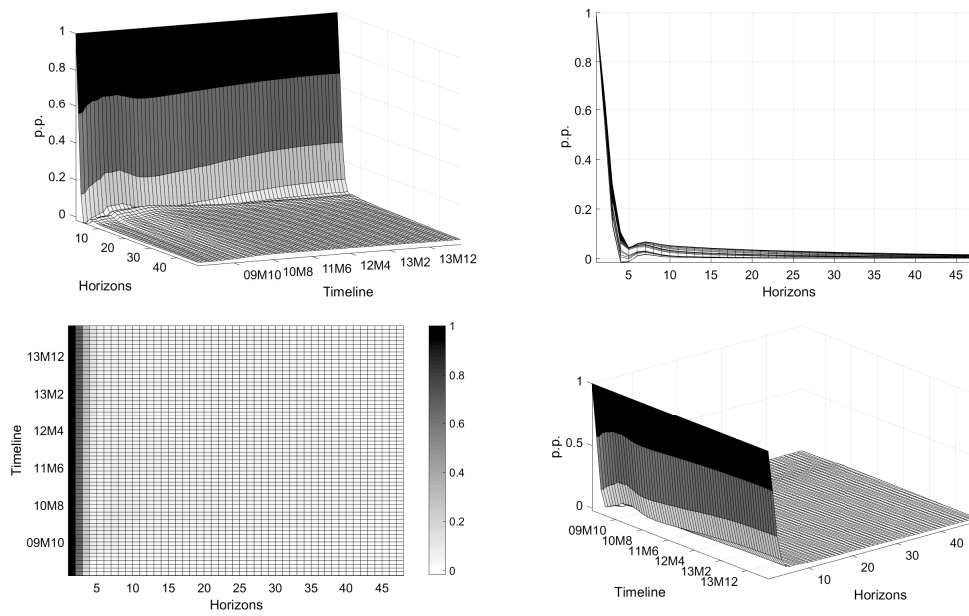


Figure B.6: TOP 1: CPI to QE shocks: Cholesky & BMS

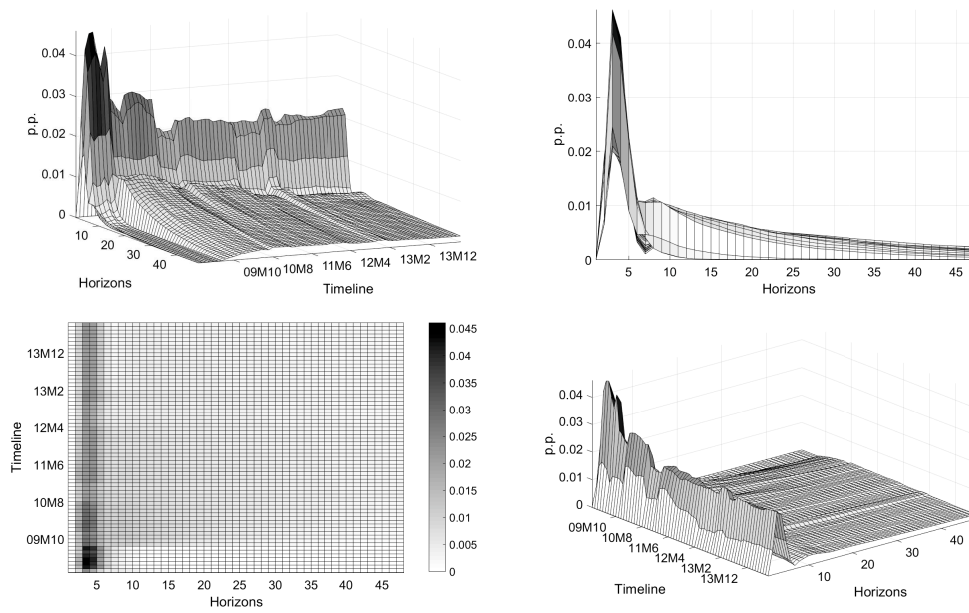


Figure B.7: TOP 1: GDP to QE shocks: Cholesky & BMS

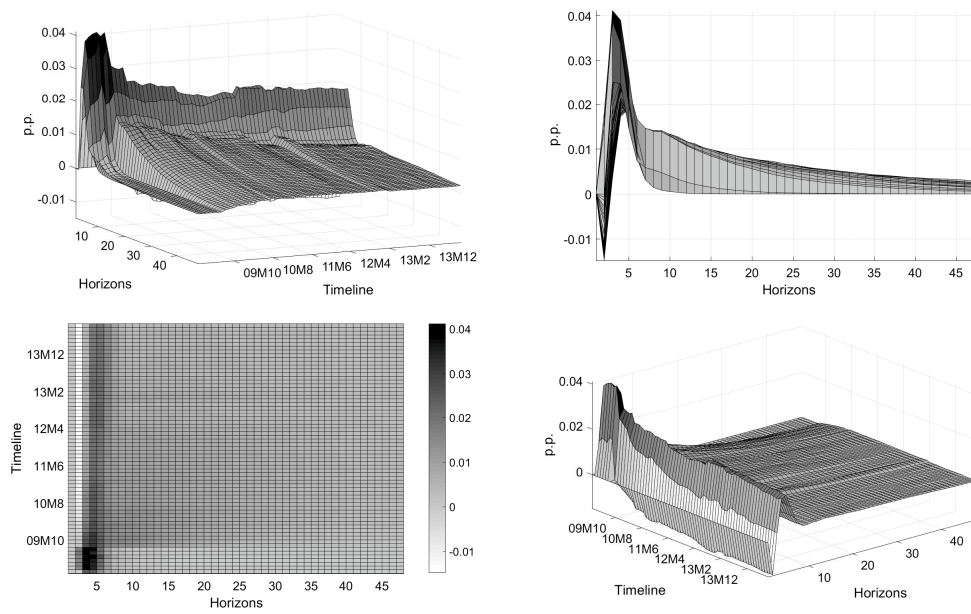


Figure B.8: TOP 1: QE to QE shocks: Cholesky & BMS

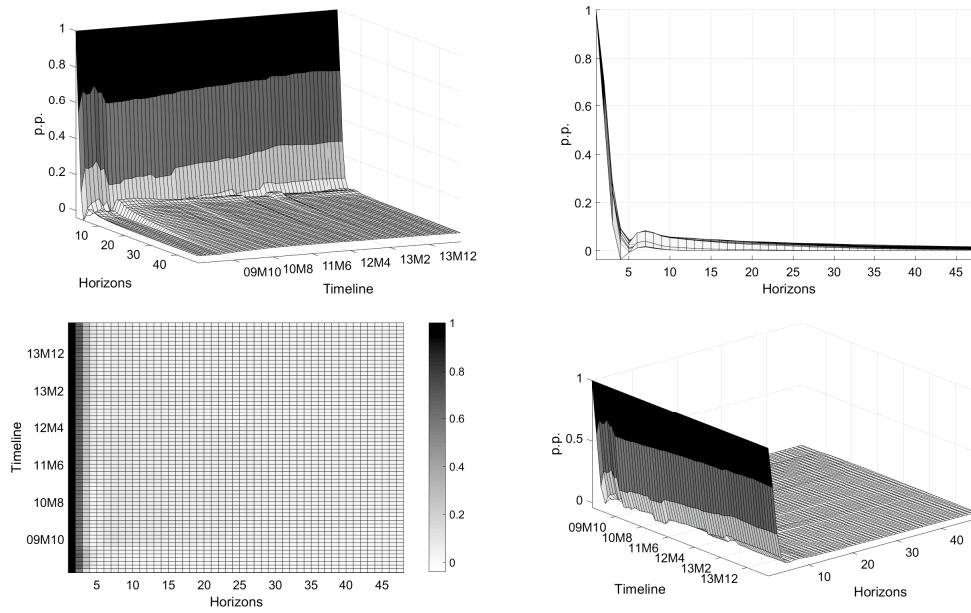


Figure B.9: TOP 1: CPI to QE shocks: Zero, Sign restrictions & BMA

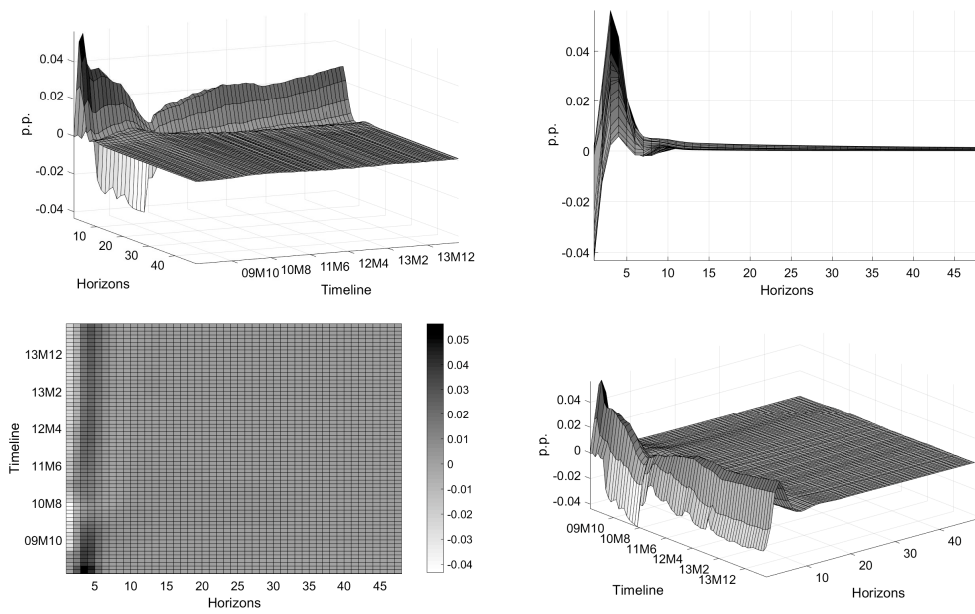


Figure B.10: TOP 1: GDP to QE shocks: Zero, Sign restrictions & BMA

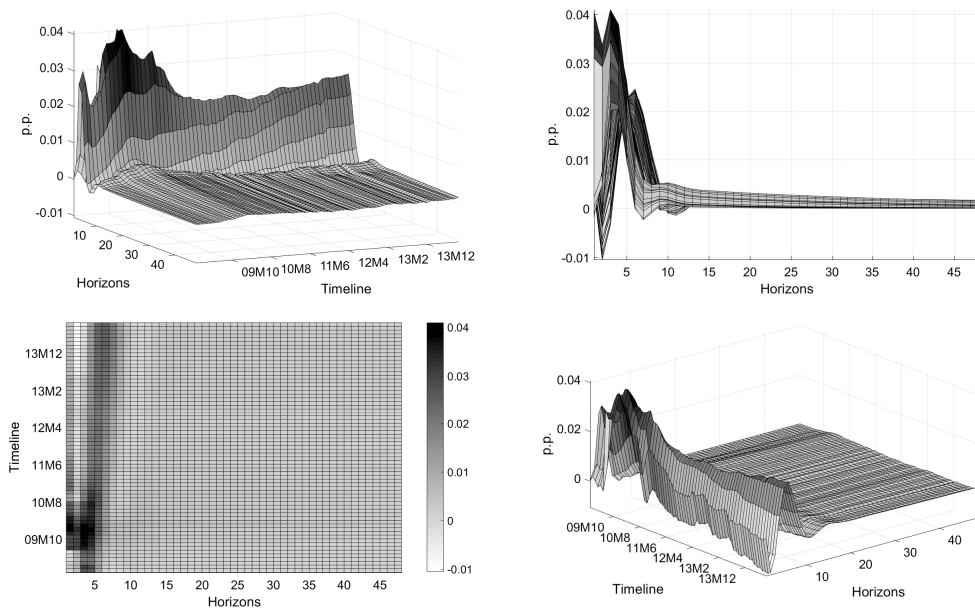


Figure B.11: TOP 1: QE to QE shocks: Zero, Sign restrictions & BMA

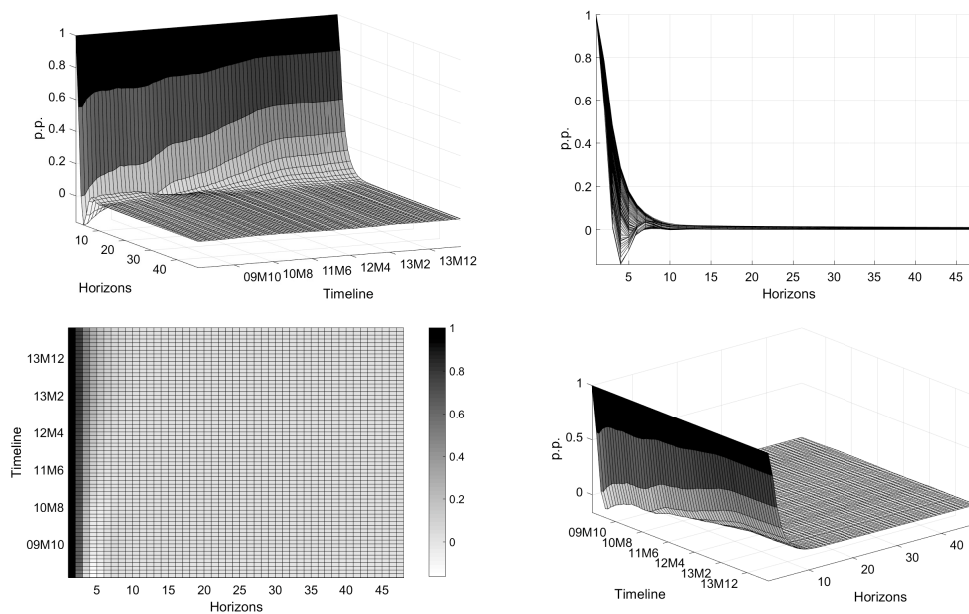


Figure B.12: TOP 1: CPI to QE shocks: Zero, Sign restrictions & BMS

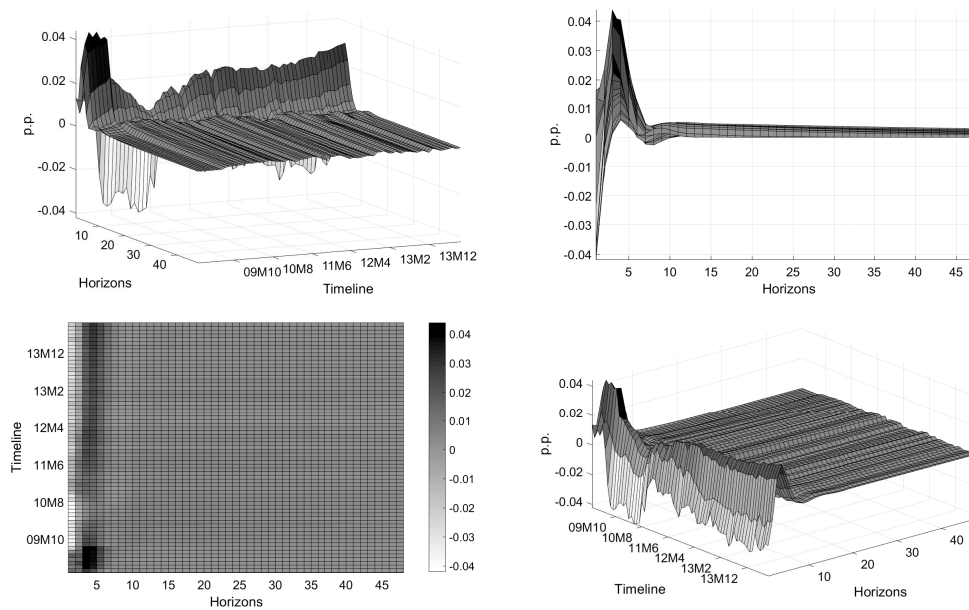


Figure B.13: TOP 1: GDP to QE shocks: Zero, Sign restrictions & BMS

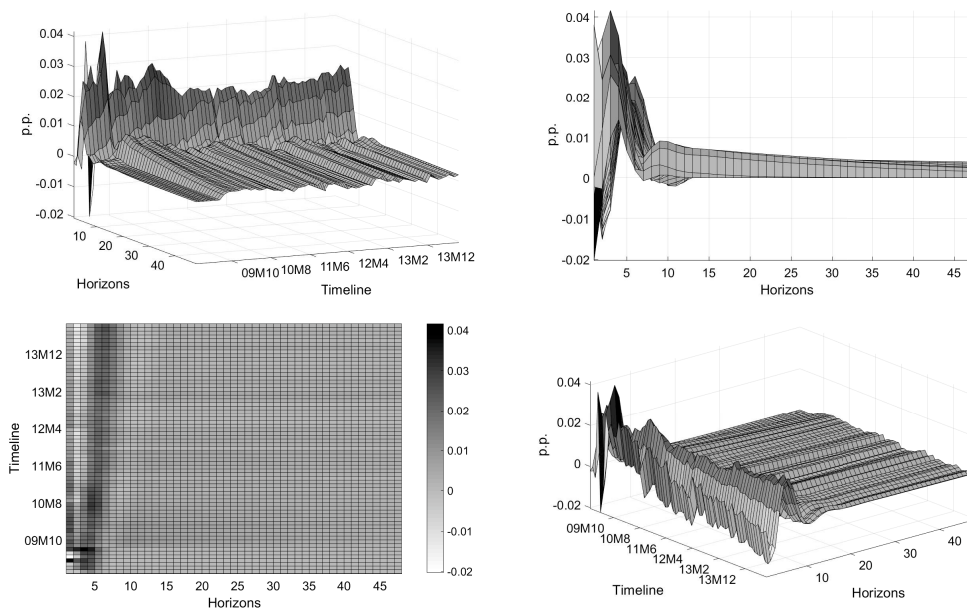


Figure B.14: TOP 1: QE to QE shocks: Zero, Sign restrictions & BMS

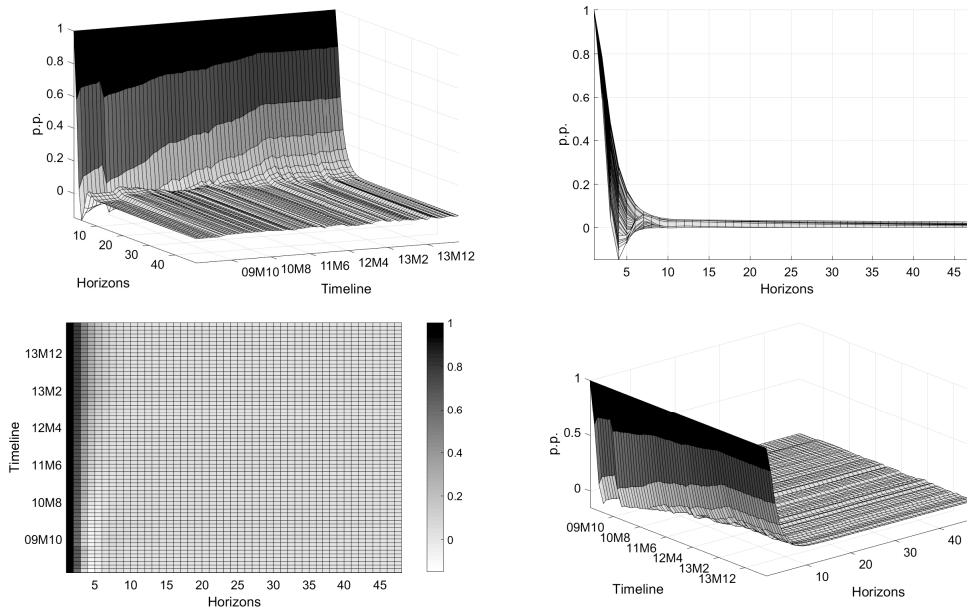
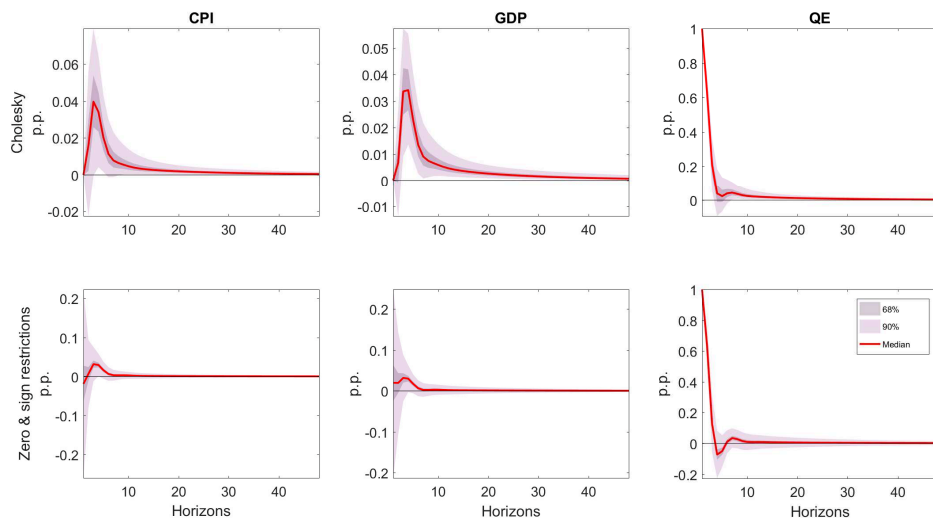


Figure B.15: TOP 1: Average IRF of QE structural shocks during QE1 with BMA



The 2D plots display the average of the median draws.

Figure B.16: TOP 1: Average IRF of QE structural shocks during QE1 with BMS

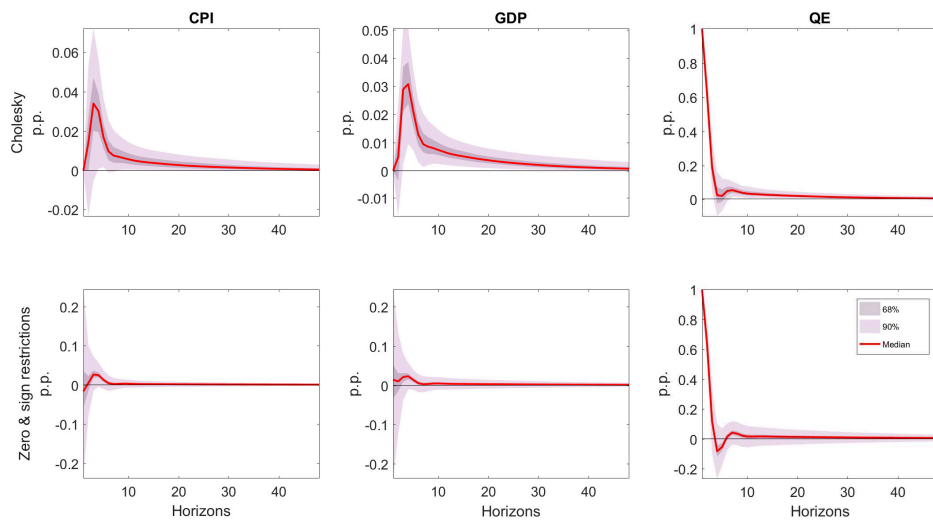


Figure B.17: TOP 1: Average IRF of QE structural shocks during QE2 with BMA

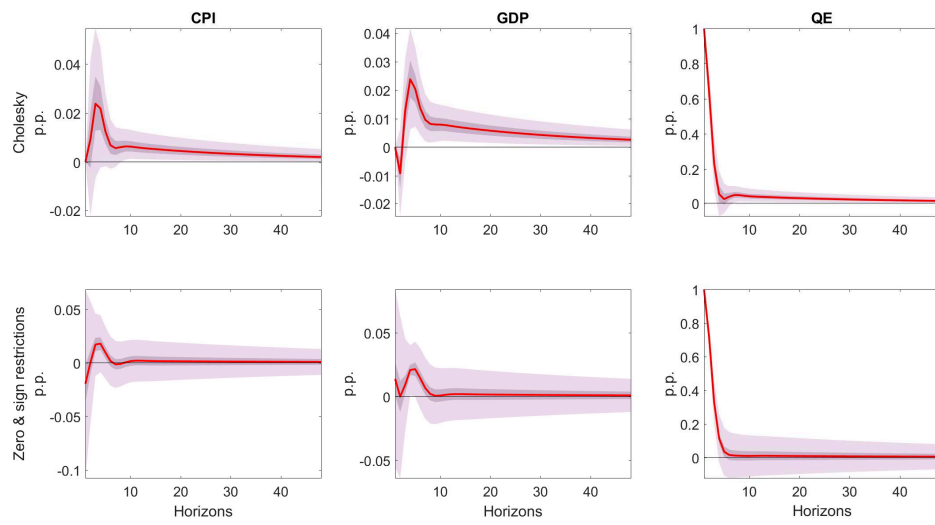


Figure B.18: TOP 1: Average IRF of QE structural shocks during QE2 with BMS

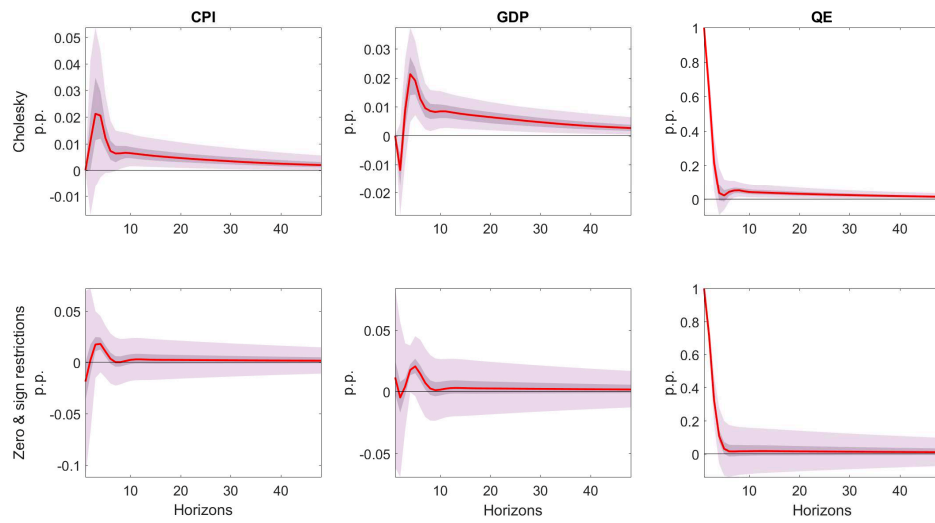


Figure B.19: TOP 1: Average IRF of QE structural shocks during QE3 with BMA

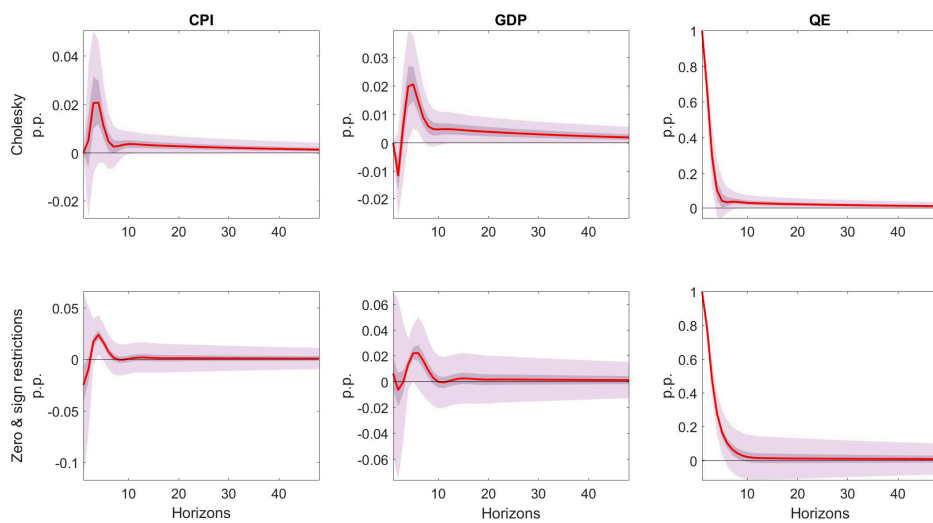


Figure B.20: TOP 1: Average IRF of QE structural shocks during QE3 with BMS

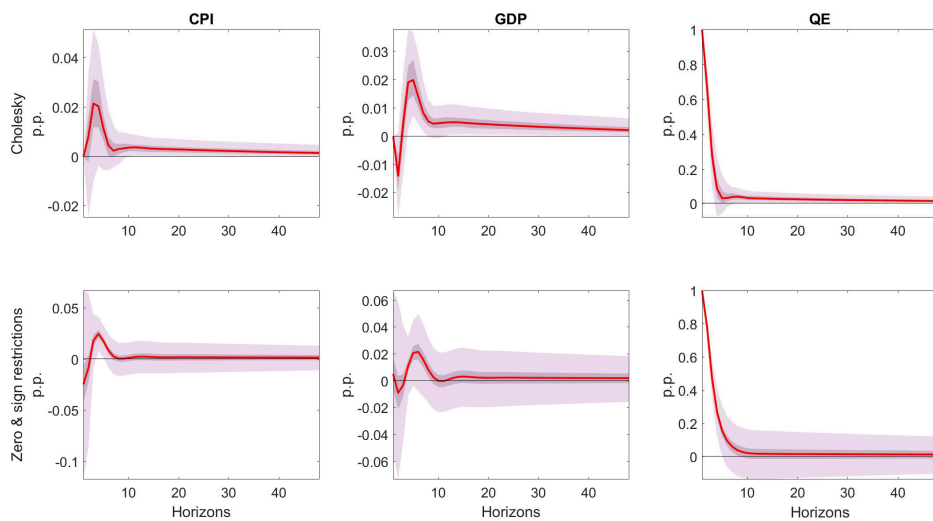


Figure B.21: TOP 1: Accumulated IRF of QE structural shocks during QE1 with BMA

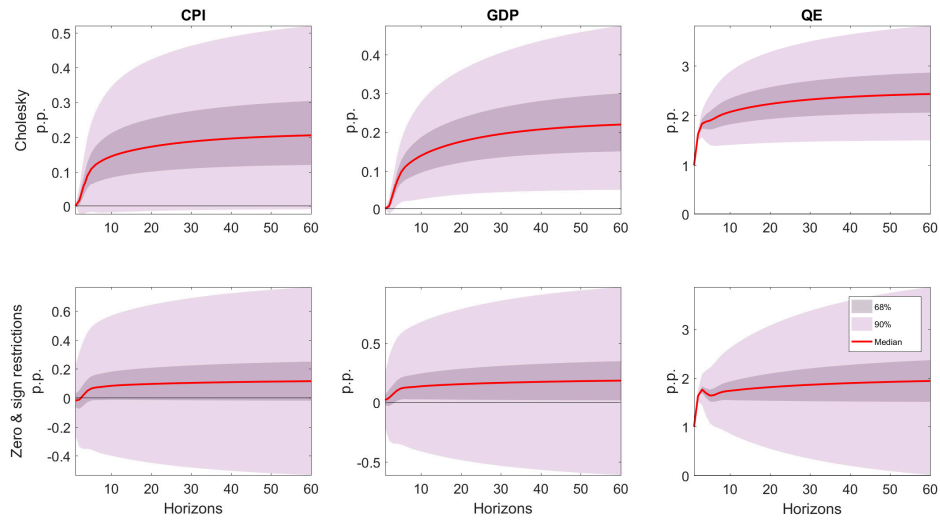


Figure B.22: TOP 1: Accumulated IRF of QE structural shocks during QE1 with BMS

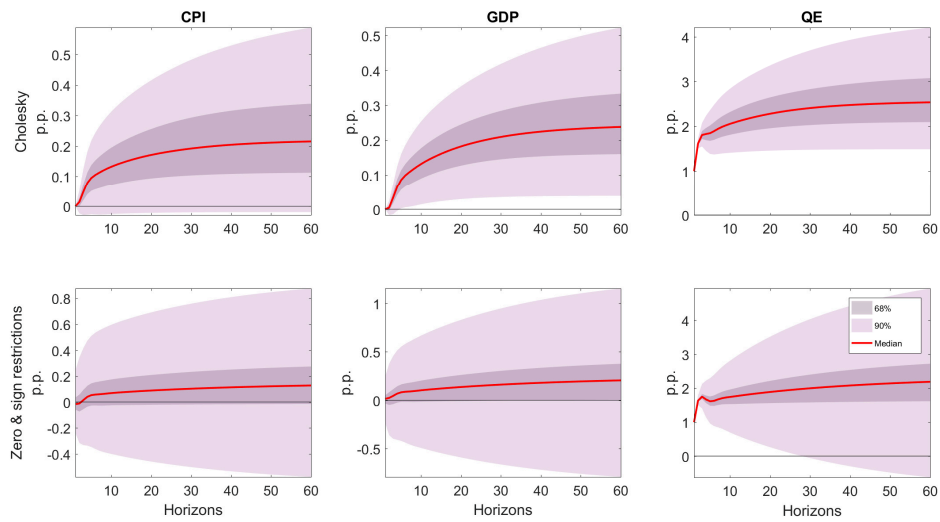


Figure B.23: TOP 1: Accumulated IRF of QE structural shocks during QE2 with BMA

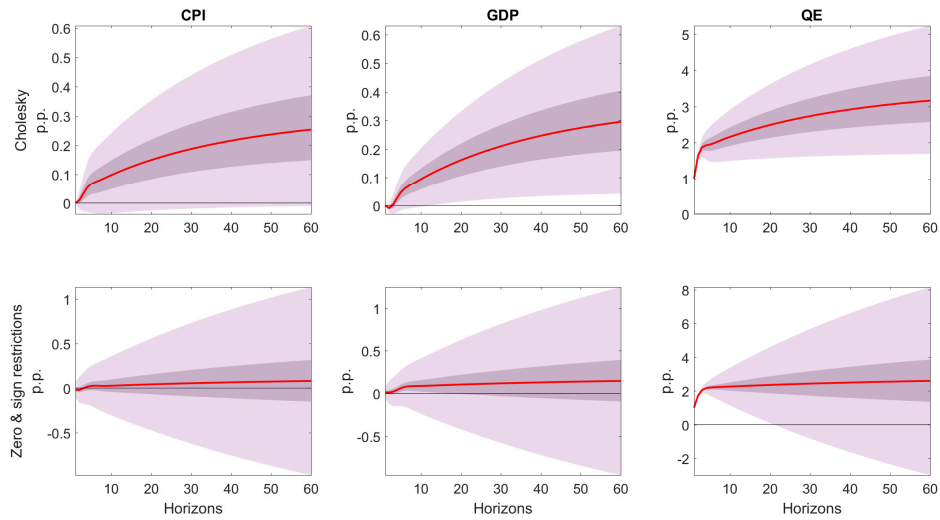


Figure B.24: TOP 1: Accumulated IRF of QE structural shocks during QE2 with BMS

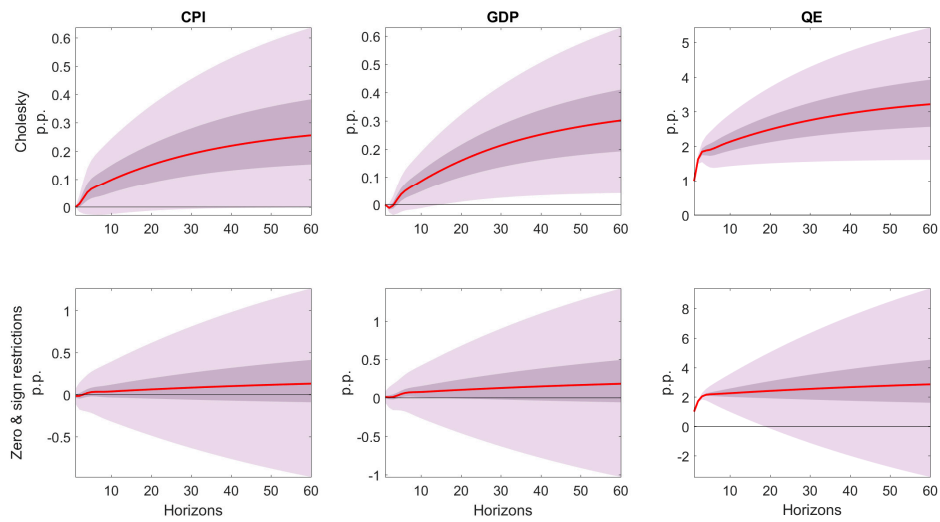


Figure B.25: TOP 1: Accumulated IRF of QE structural shocks during QE3 with BMA

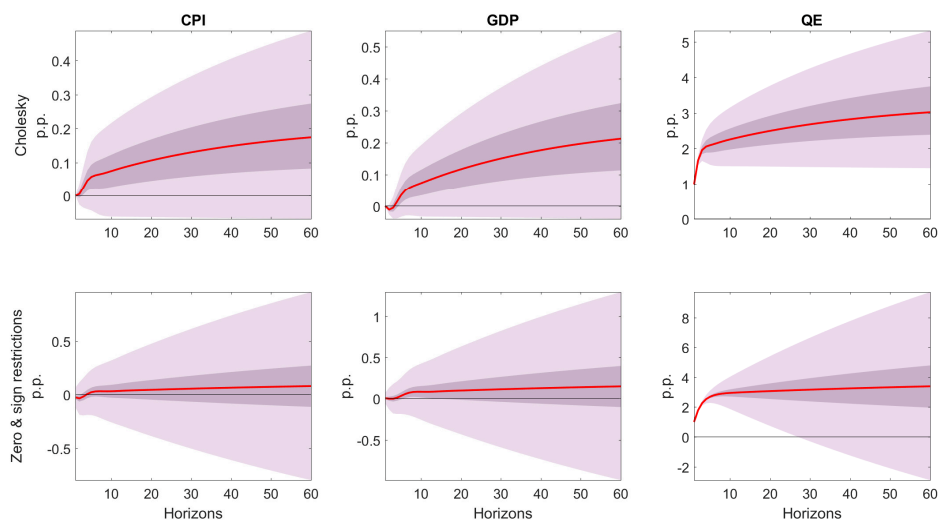


Figure B.26: TOP 1: Accumulated IRF of QE structural shocks during QE3 with BMS

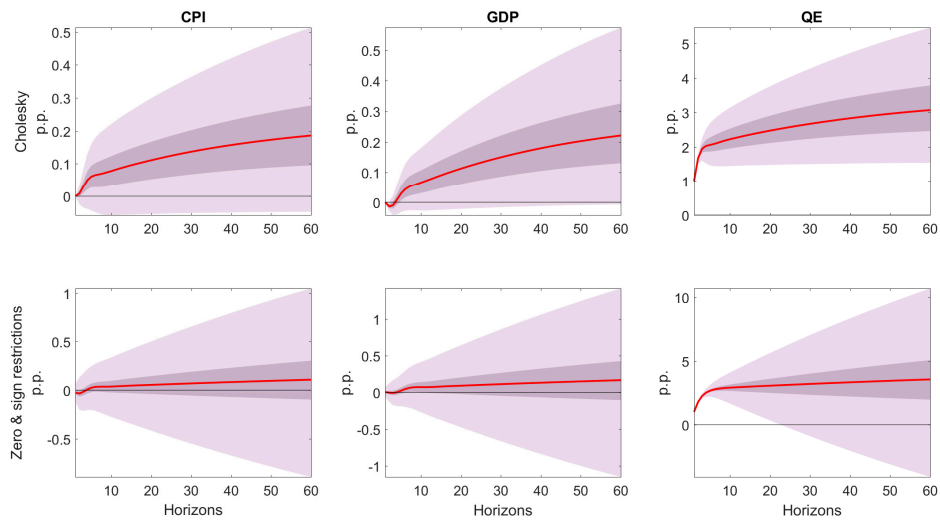


Figure B.27: VAR: CPI to QE shocks: Cholesky & BMA

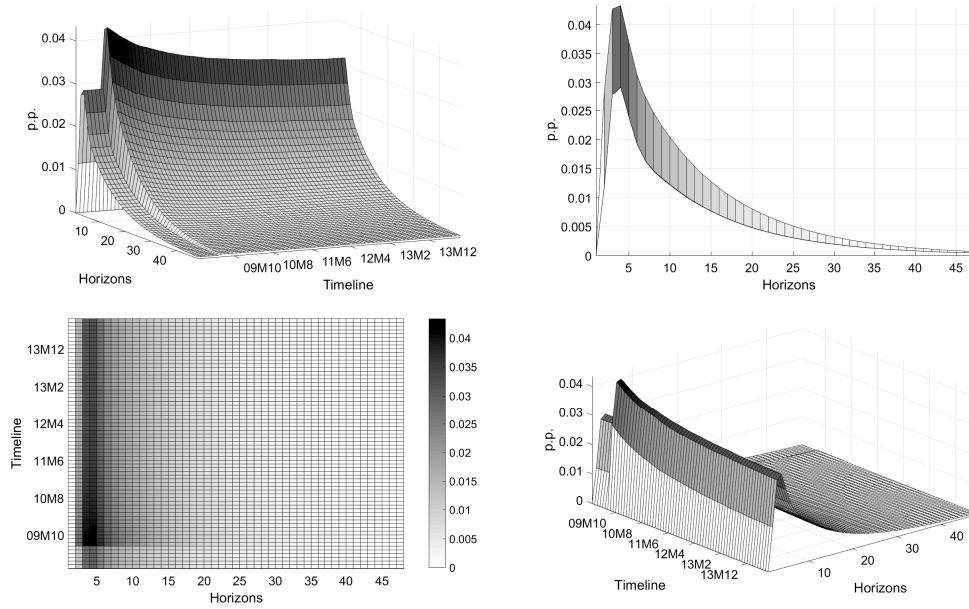


Figure B.28: VAR: GDP to QE shocks: Cholesky & BMA

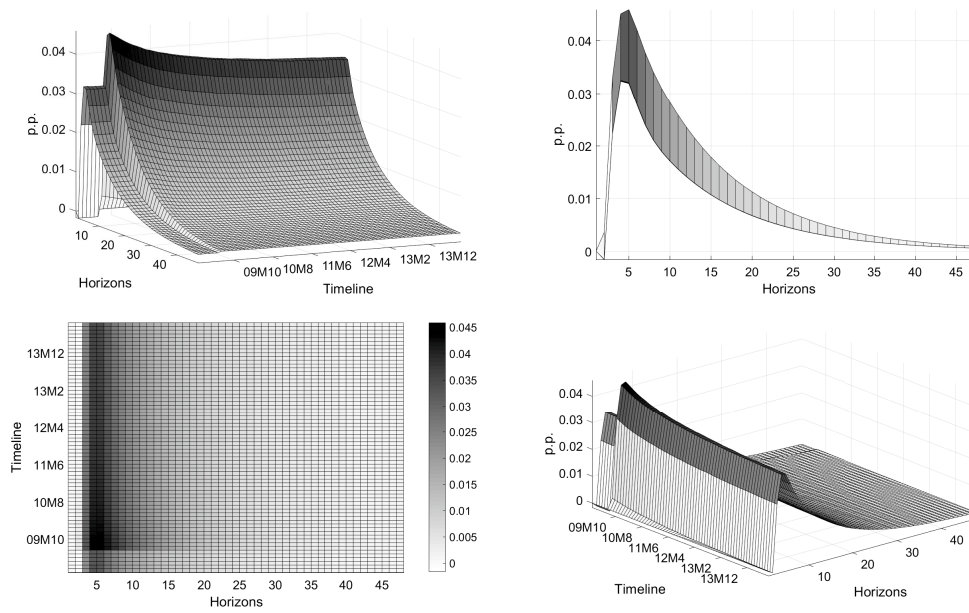


Figure B.29: VAR: QE to QE shocks: Cholesky & BMA

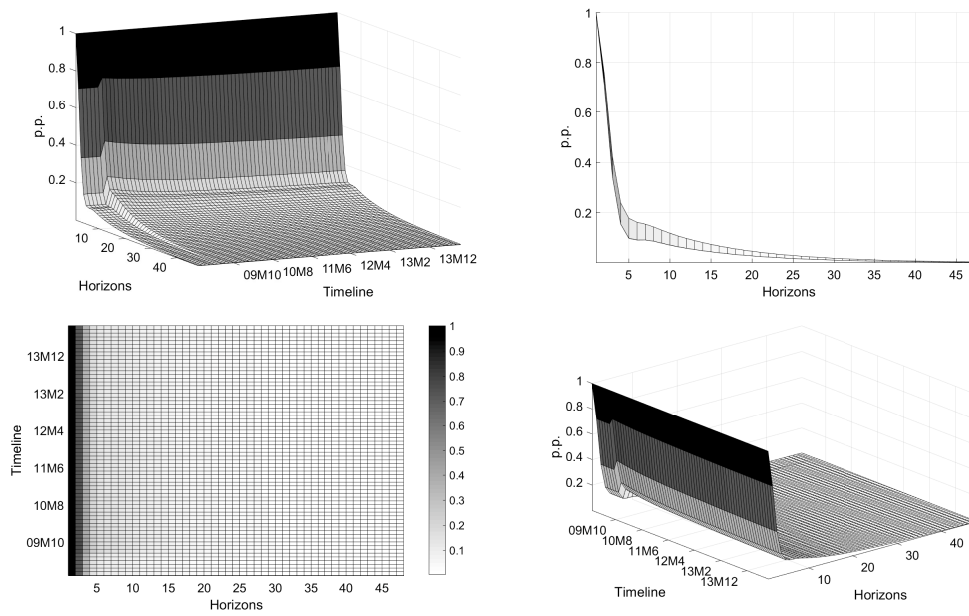


Figure B.30: VAR: CPI to QE shocks: Cholesky & BMS

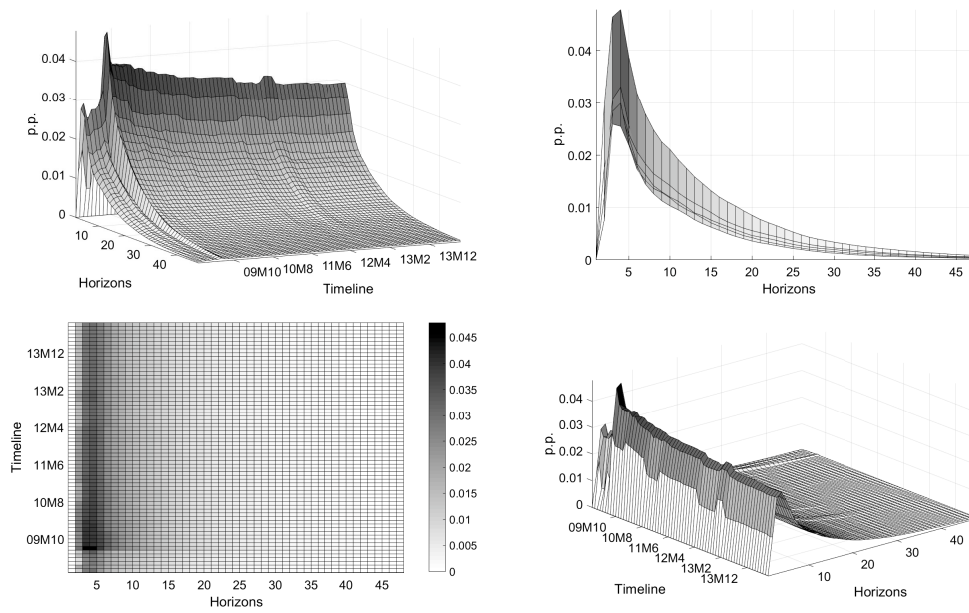


Figure B.31: VAR: GDP to QE shocks: Cholesky & BMS

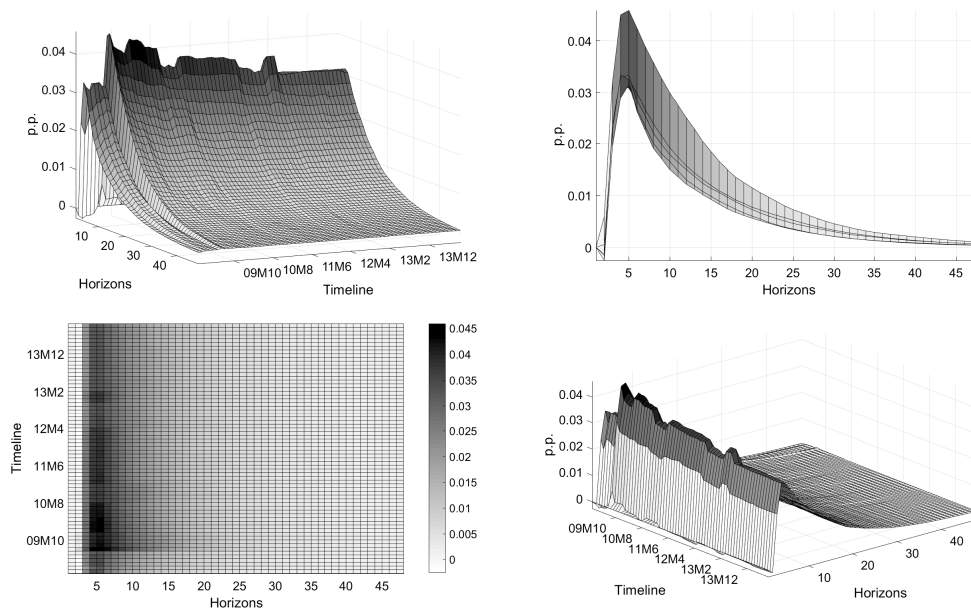


Figure B.32: VAR: QE to QE shocks: Cholesky & BMS

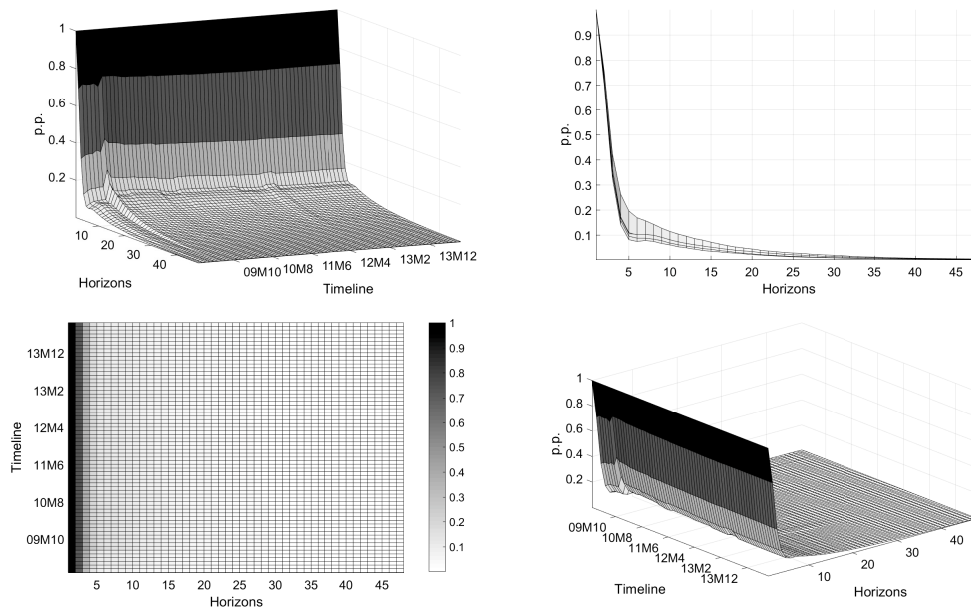


Figure B.33: VAR: CPI to QE shocks: Zero, Sign restrictions & BMA

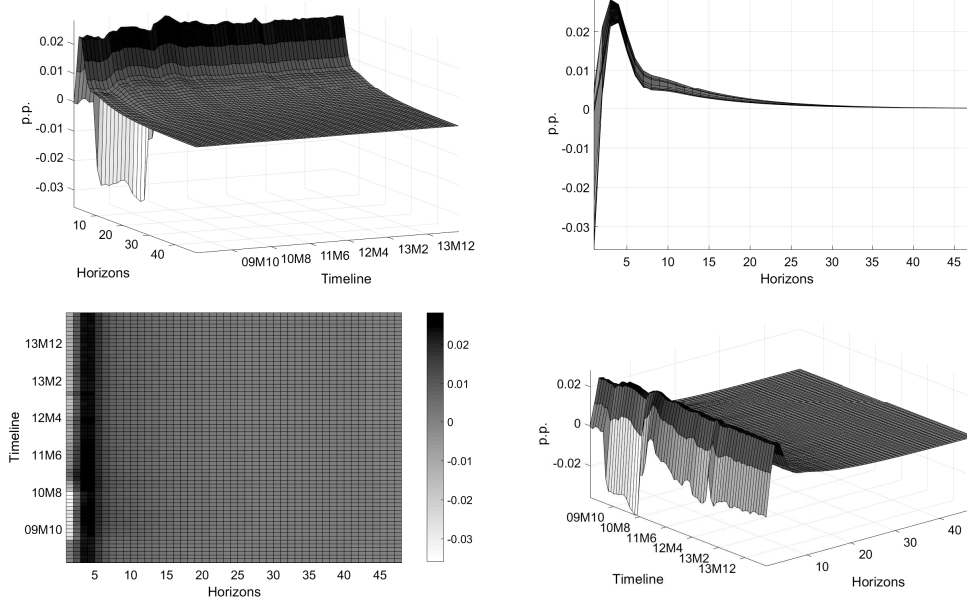


Figure B.34: VAR: GDP to QE shocks: Zero, Sign restrictions & BMA

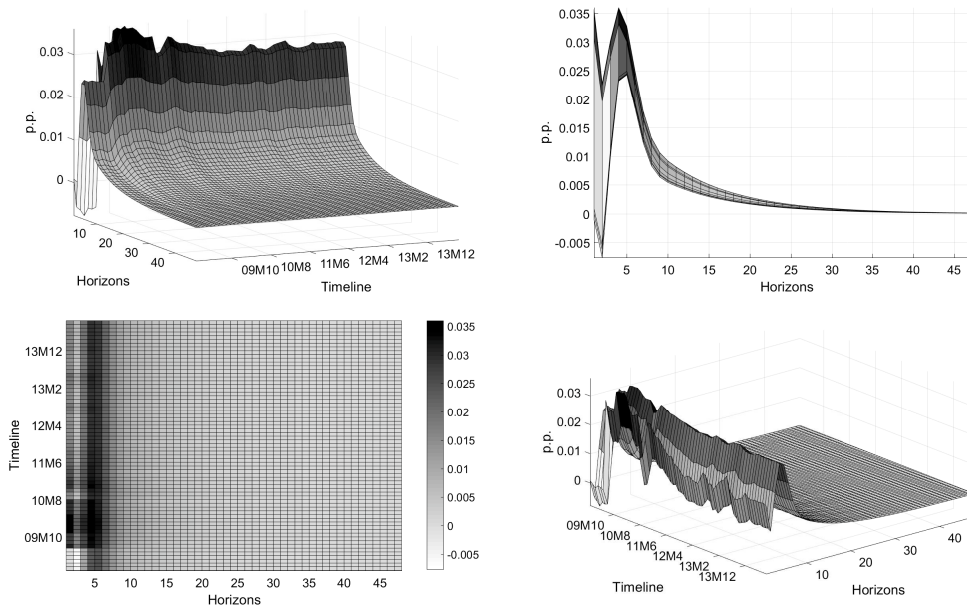


Figure B.35: VAR: QE to QE shocks: Zero, Sign restrictions & BMA

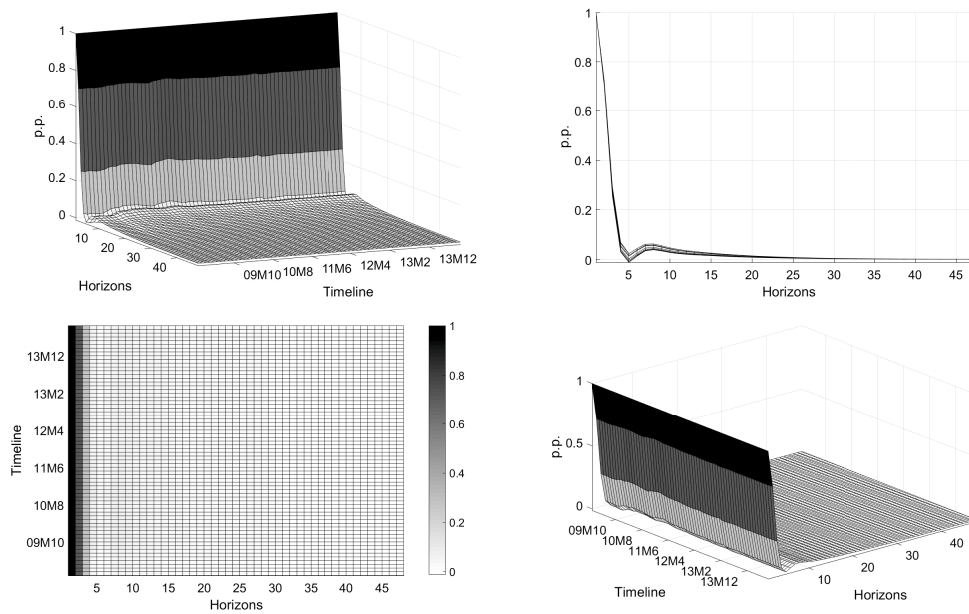


Figure B.36: VAR: CPI to QE shocks: Zero, Sign restrictions & BMS

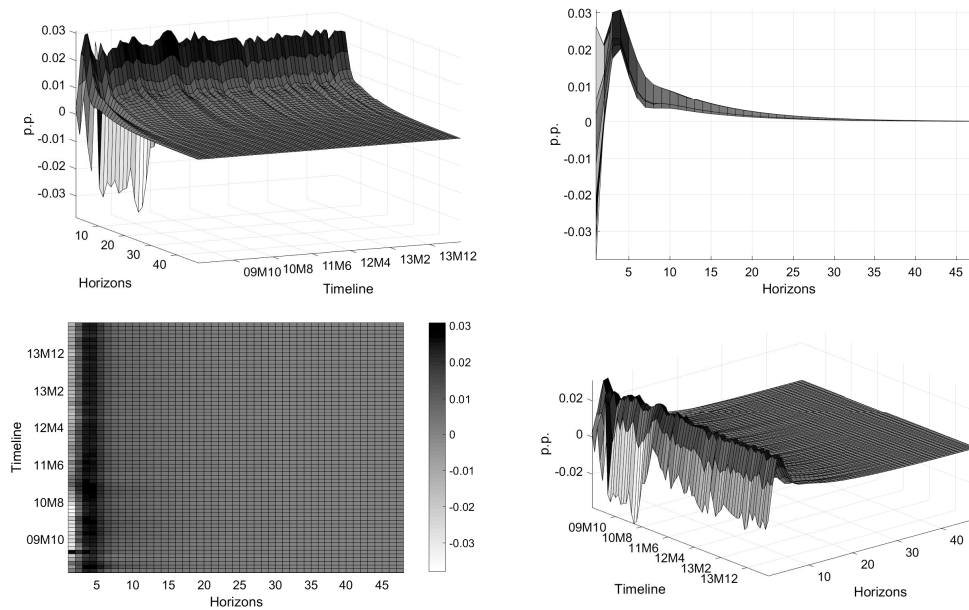


Figure B.37: VAR: GDP to QE shocks: Zero, Sign restrictions & BMS

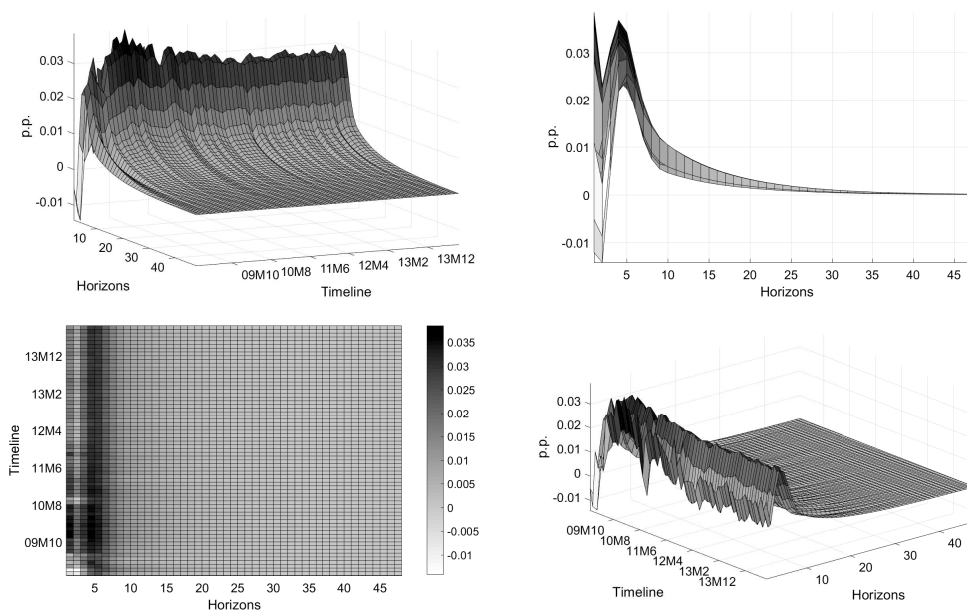


Figure B.38: VAR: QE to QE shocks: Zero, Sign restrictions & BMS

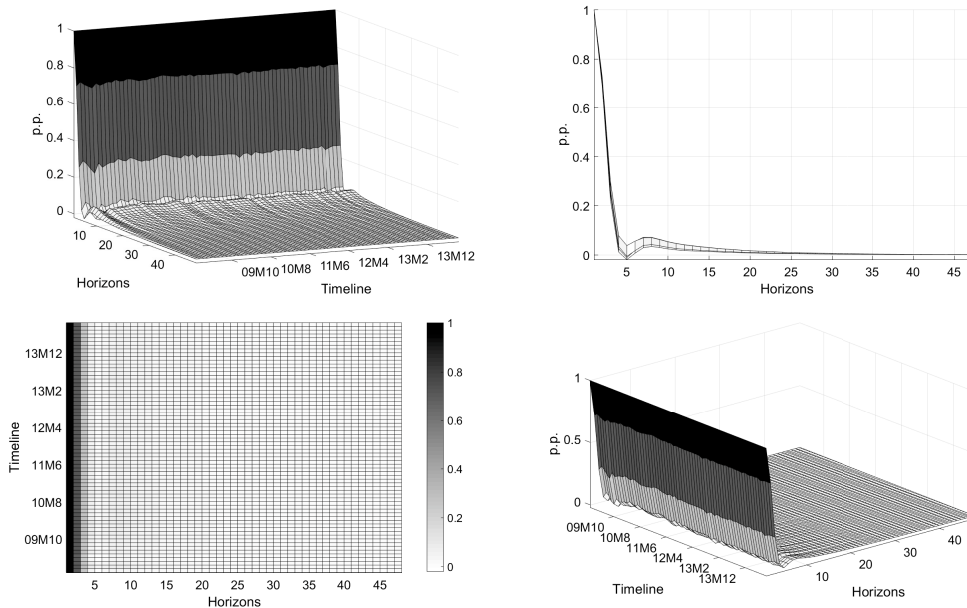


Figure B.39: VAR: Average IRF of QE structural shocks during QE1 with BMA

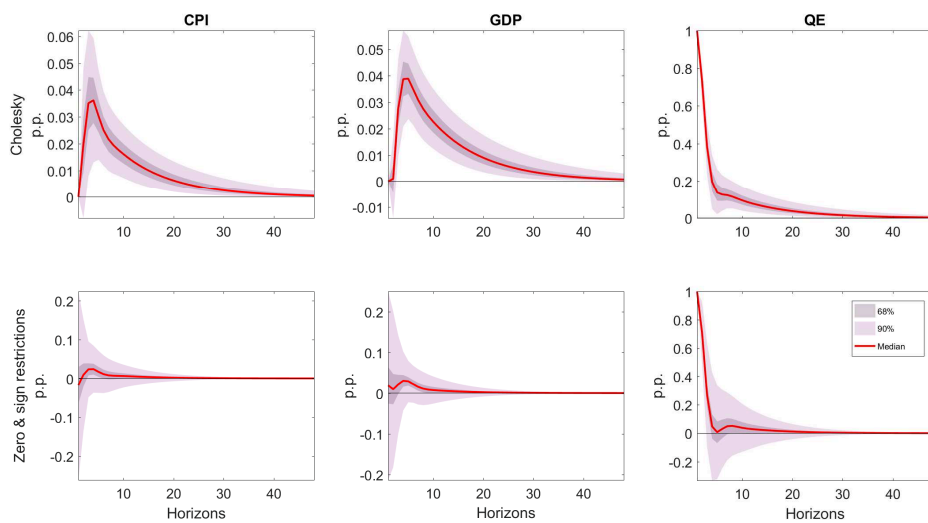


Figure B.40: VAR: Average IRF of QE structural shocks during QE1 with BMS

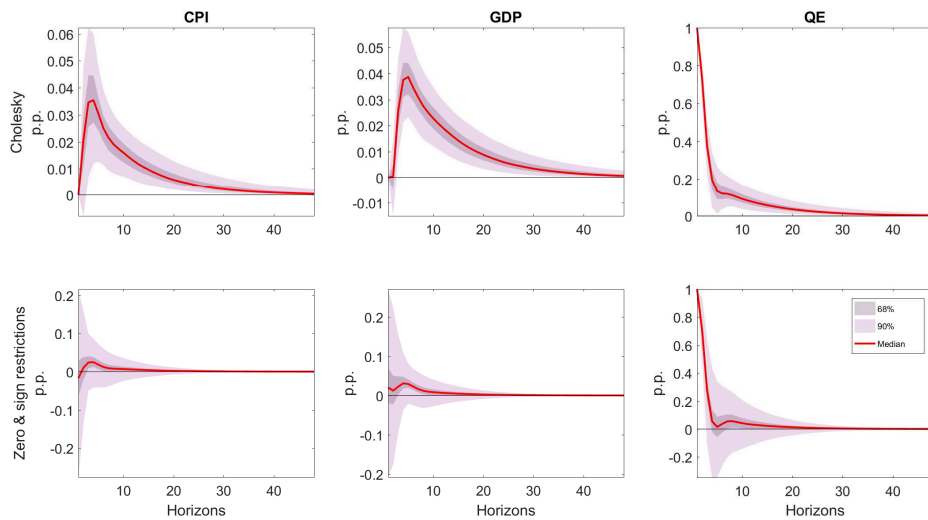


Figure B.41: VAR: Average IRF of QE structural shocks during QE2 with BMA

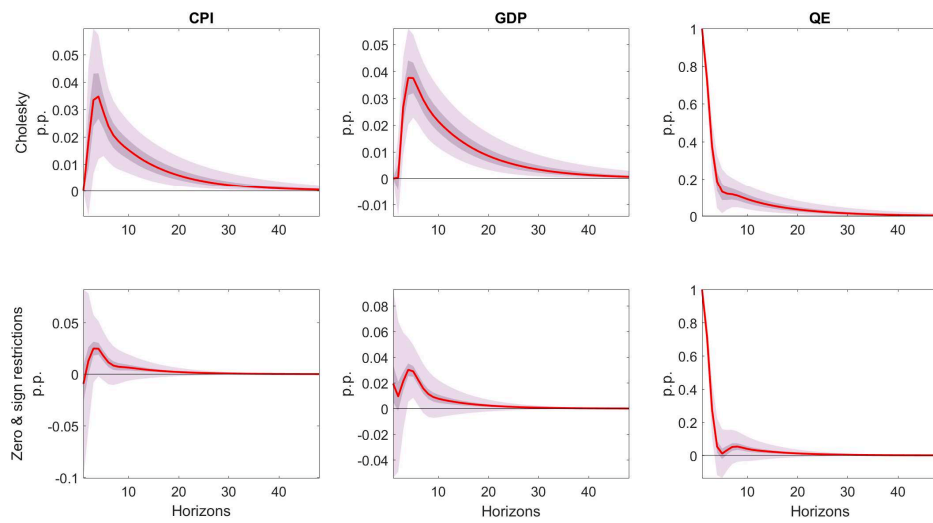


Figure B.42: VAR: Average IRF of QE structural shocks during QE2 with BMS

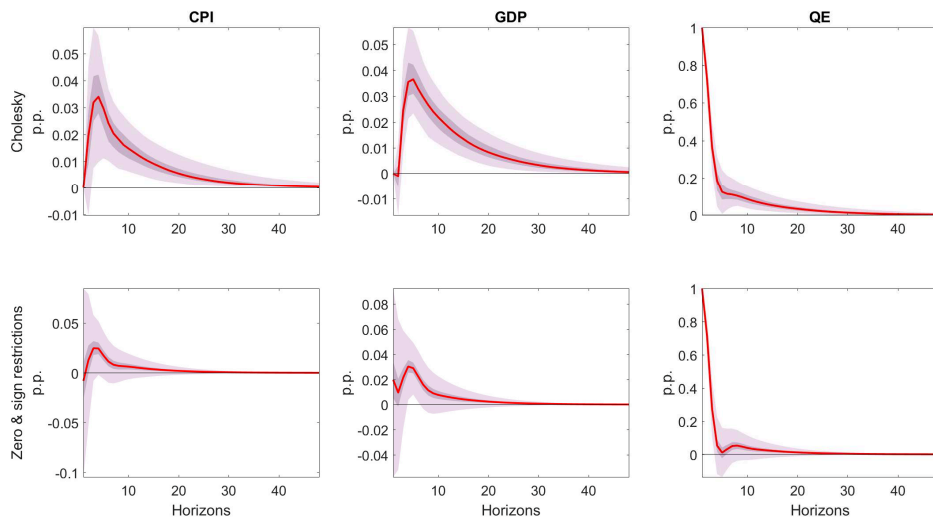


Figure B.43: VAR: Average IRF of QE structural shocks during QE3 with BMA

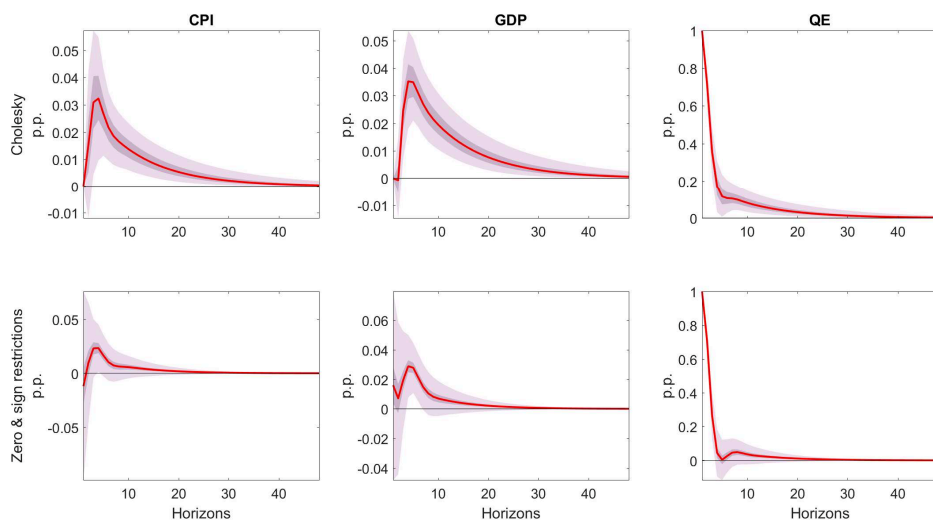


Figure B.44: VAR: Average IRF of QE structural shocks during QE3 with BMS

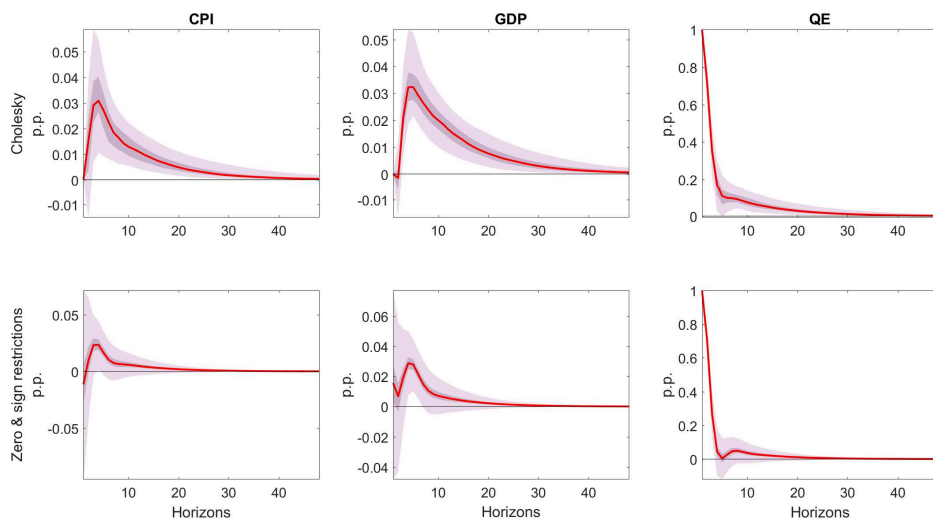


Figure B.45: VAR: Accumulated IRF of QE structural shocks during QE1 with BMA

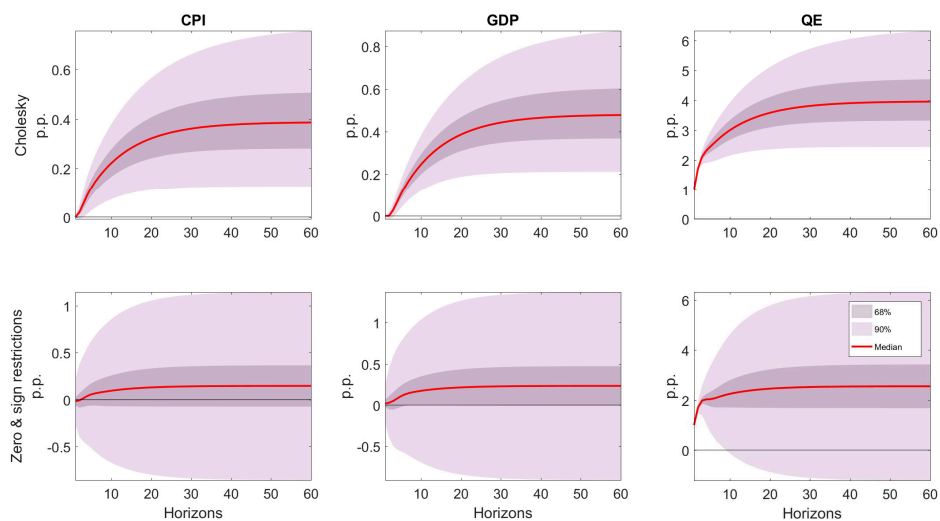


Figure B.46: VAR: Accumulated IRF of QE structural shocks during QE1 with BMS

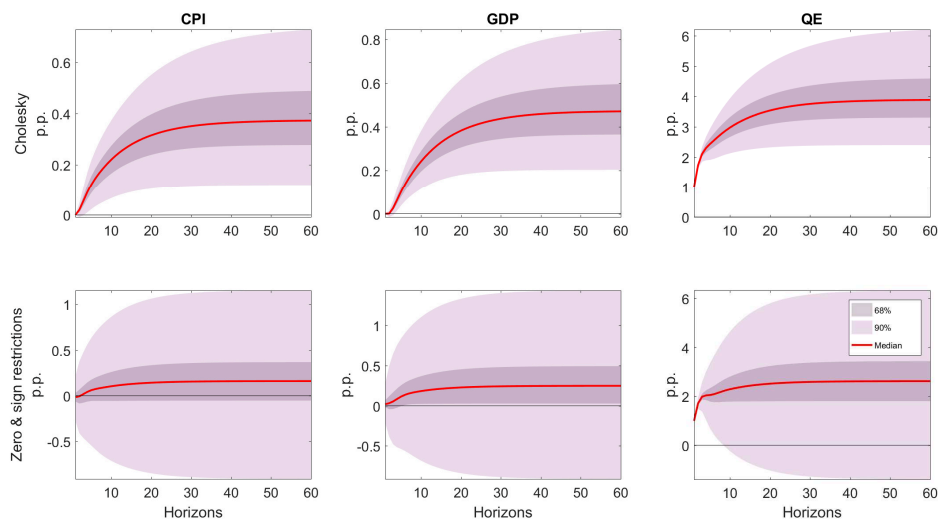


Figure B.47: VAR: Accumulated IRF of QE structural shocks during QE2 with BMA

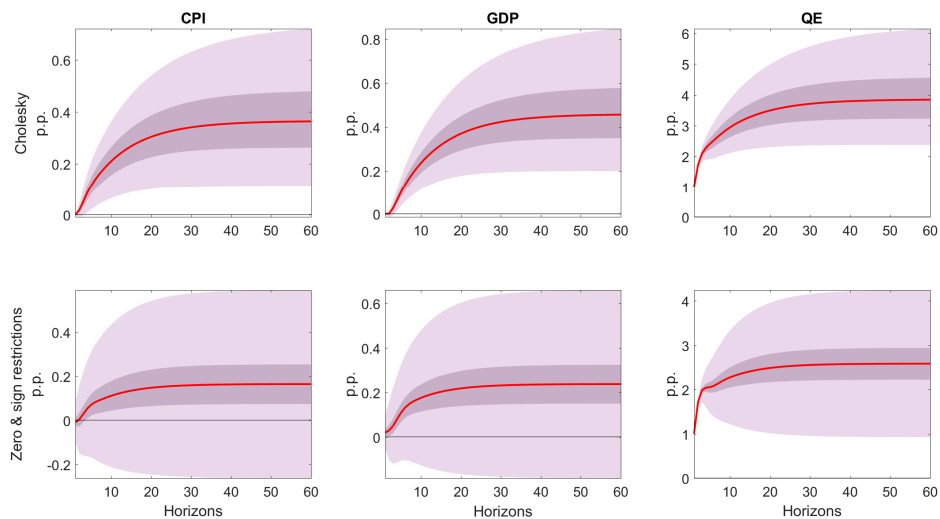


Figure B.48: VAR: Accumulated IRF of QE structural shocks during QE2 with BMS

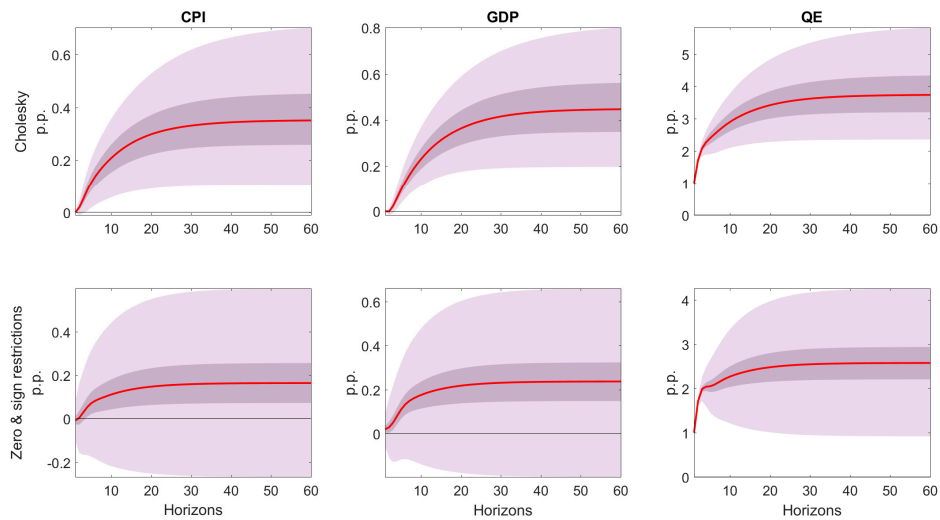


Figure B.49: VAR: Accumulated IRF of QE structural shocks during QE3 with BMA

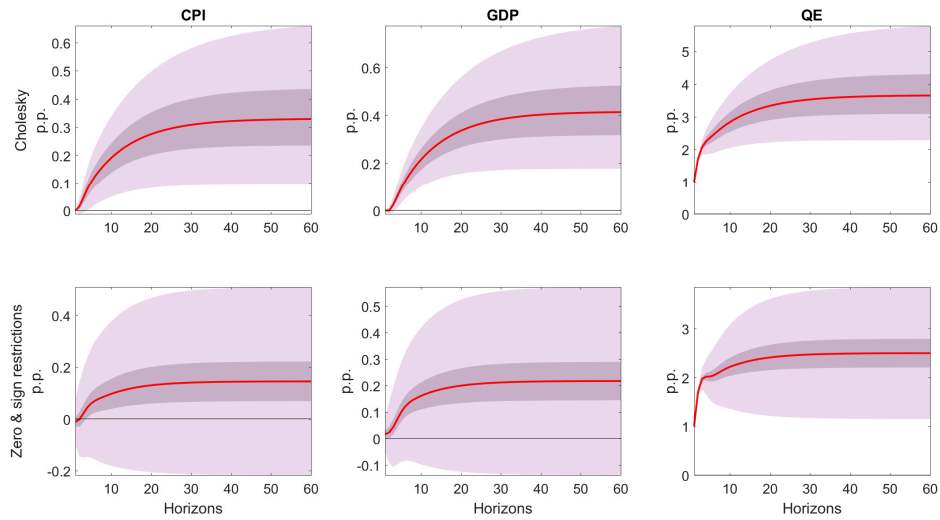
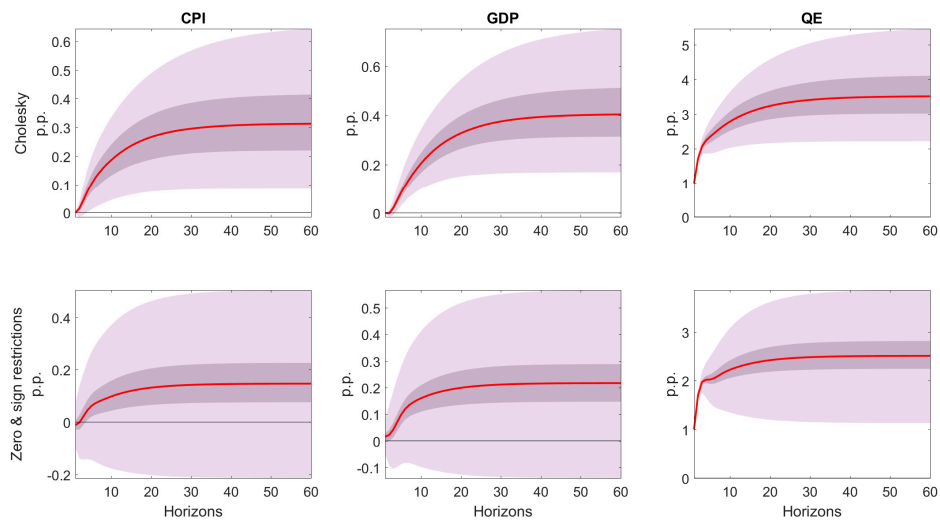


Figure B.50: VAR: Accumulated IRF of QE structural shocks during QE3 with BMS



B.2 Data

Table B.1: Data Chapter 2

Description & Acronyms	Units	Sources
Real Gross Domestic Product (GDP)	Billions	US B.E.A.
All Federal Reserve Banks: Total Assets (QE)	Millions	H.4.1 F.A.R.B.
Consumer Price Index for All Urban Consumers (CPI)	Index	US B.E.A.
Ten-year Treasury Term Premium (10Y TP)	Rates	Adrian et al. (2013)
Moody's Seasoned Baa Corporate Bond Yield (BAA)	Rates	Moody's
Ten-Year Treasury Constant Maturity Rate (10Y)	Rates	H.15 S.I.R.
Ten-year Expected Average Short-Term Rates (10Y exp. ST rates)	Rates	Adrian et al. (2013)
10-Year Breakeven Inflation Rate (10Y infl)	Rates	Federal Reserve St. Louis
Households and Nonprofit Organizations; Net Worth (HH NW)	Billions	Z.1 F.A.U.S.
Nonfinancial Corporate Business; Net Worth (Corp. NW)	Billions	Z.1 F.A.U.S.
St. Louis Fed Financial Stress Index (STL stress)	Index	Federal Reserve St. Louis
CBOE Volatility Index: VIX (CBOE VIX)	Index	CBOE Market Statistics

Initialisms: B.E.A. = Bureau of Economic Analysis, F.A.R.B. = Factors Affecting Reserve Balances, S.I.R. = Selected Interest Rates, F.A.U.S. = Financial Accounts of the United States.

B.3 Bayesian linear Gaussian state-space models

B.3.1 Bayesian filtering

The Bayesian statistics interpret the probabilities as a conditional measure of uncertainty under the assumption of rationality. Bayesian inference is opposed to the frequentist approach in the interpretation of probabilities. In the frequentist approach, a probability is a measure of the relative frequency of the occurrence of a random event, for an infinite number of trials. In the Bayesian philosophy, a probability is the measure of likelihood or the degree of certainty of the occurrence of that random conditional on evidence. In this environment, the posterior probability is the result of the normalized product between a prior knowledge and a likelihood function. Bayesian filtering is a recursive method to estimate the unknown probability distribution of a time-varying system for which an observer has noisy measurements. Mathematically, the state of this system is a dynamical random variable or a set of random variables. Measurements are data or a set of information about the evolution of that particular state. In the following, we present a Bayesian filtering method to solve for linear state-space models with Gaussian noises; that is, the so-called Kalman filter (Kalman (1960)).

A probabilistic state-space model is of the form of:

$$\begin{aligned}x_l &\sim p(x_l|x_{l-1}) \\ y_l &\sim p(y_l|x_l)\end{aligned}\tag{B.1}$$

where x_l describes the dynamic state of the system by a transition probability distribution $p(x_l|x_{l-1})$. The measurement model y_l is a conditional probability

distribution $p(y_l|x_l)$ between the current state x_l and a measure y_l . In addition, $l > 0$ refers to the time dynamic which can be, for instance, a discrete measure of time t . To allow for a Bayesian inference, the model is initialized with a prior probability distribution $p(x_0)$. The major issue of such dynamical system is that the computation of the full posterior distribution of the states is barely impossible. However, given some assumptions, one can easily compute the marginal distributions. In particular, the system is assumed to follow Markov sequences, as such:

- The states follow a first-order Markov sequence or a Markov chain chain (discrete time). The prediction of the future state depends only on the previous state. The latter is characterized by $p(x_l|x_{1:l-1}, y_{1:l-1}) = p(x_l|x_{l-1})$. In addition, the past is independent of the future given the present; $p(x_{l-1}|x_{l:T}, y_{l:T}) = p(x_{l-1}|x_l)$.
- The current measurement y_l given the current state x_l is conditionally independent of both, past measurements and states; that is, $p(y_l|x_{1:l}, y_{1:l-1}) = p(y_l|x_l)$.

The marginal posterior distribution of the state x_l given measurements $y_{1:l}$ is derived by a standard Bayes rule:

$$\begin{aligned}
p(x_l|y_{1:l}) &= \frac{p(y_{1:l}|x_l)p(x_l)}{p(y_{1:l})} \\
&= \frac{p(y_l, y_{1:l-1}|x_l)p(x_l)}{p(y_l, y_{1:l-1})} \\
&= \frac{p(y_l|y_{1:l-1}, x_l)p(y_{1:l-1}|x_l)p(x_l)}{p(y_l|y_{1:l-1})p(y_{1:l-1})} \\
&= \frac{p(y_l|y_{1:l-1}, x_l)p(x_l|y_{1:l-1})p(y_{1:l-1})p(x_l)}{p(y_l|y_{1:l-1})p(y_{1:l-1})p(x_l)} \\
&= \frac{p(y_l|x_l)p(x_l|y_{1:l-1})}{p(y_l|y_{1:l-1})} \tag{B.2}
\end{aligned}$$

where $p(y_l|y_{1:l-1}) = \int p(y_l|x_l)p(x_l|y_{1:l-1})dx_l$ is called the evidence. The posterior density is the product of the prior knowledge $p(x_l|y_{1:l-1}) = \int p(x_l|x_{l-1})p(x_{l-1}|y_{1:l-1})dx_{l-1}$ and the likelihood function $p(y_l|x_l)$.

B.3.2 Reminder: Multivariate Gaussian properties

Consider a random vector $\mathbf{Z} \in R^n$, Gaussian distributed with mean $\boldsymbol{\mu}$ and positive semi-definite covariance $\boldsymbol{\Sigma}$, such that $\mathbf{Z} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Assume that \mathbf{Z} is partitioned into two blocks (\mathbf{X}, \mathbf{Y}) . Accordingly:

$$\begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} \boldsymbol{\mu}_X \\ \boldsymbol{\mu}_Y \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_{XX} & \boldsymbol{\Sigma}_{XY} \\ \boldsymbol{\Sigma}_{YX} & \boldsymbol{\Sigma}_{YY} \end{pmatrix} \right) \tag{B.3}$$

Lemma 1: The marginal distributions of a joint Gaussian distribution are Gaussians. The marginal distributions are $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}_X, \boldsymbol{\Sigma}_{XX})$ and $\mathbf{Y} \sim \mathcal{N}(\boldsymbol{\mu}_Y, \boldsymbol{\Sigma}_{YY})$.

Proofs:

The marginal density of \mathbf{X} is of the form of:

$$f_{\mathbf{X}}(x) = \frac{1}{(2\pi)^{\frac{n+m}{2}} \begin{vmatrix} \boldsymbol{\Sigma}_{\mathbf{X}\mathbf{X}} & \boldsymbol{\Sigma}_{\mathbf{X}\mathbf{Y}} \\ \boldsymbol{\Sigma}_{\mathbf{Y}\mathbf{X}} & \boldsymbol{\Sigma}_{\mathbf{Y}\mathbf{Y}} \end{vmatrix}^{1/2}} \int_{\mathbf{Y} \in \mathbb{R}^m} \exp \left(-\frac{1}{2} \begin{pmatrix} x - \mu_{\mathbf{X}} \\ y - \mu_{\mathbf{Y}} \end{pmatrix}^T \begin{pmatrix} \boldsymbol{\Sigma}_{\mathbf{X}\mathbf{X}} & \boldsymbol{\Sigma}_{\mathbf{X}\mathbf{Y}} \\ \boldsymbol{\Sigma}_{\mathbf{Y}\mathbf{X}} & \boldsymbol{\Sigma}_{\mathbf{Y}\mathbf{Y}} \end{pmatrix}^{-1} \begin{pmatrix} x - \mu_{\mathbf{X}} \\ y - \mu_{\mathbf{Y}} \end{pmatrix} \right) dy \quad (\text{B.4})$$

Define the precision matrix \mathbf{P} such that it is the matrix inverse of the covariance matrix $\boldsymbol{\Sigma}$:

$$\mathbf{P} = \begin{pmatrix} \mathbf{P}_{\mathbf{X}\mathbf{X}} & \mathbf{P}_{\mathbf{X}\mathbf{Y}} \\ \mathbf{P}_{\mathbf{Y}\mathbf{X}} & \mathbf{P}_{\mathbf{Y}\mathbf{Y}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\Sigma}_{\mathbf{X}\mathbf{X}} & \boldsymbol{\Sigma}_{\mathbf{X}\mathbf{Y}} \\ \boldsymbol{\Sigma}_{\mathbf{Y}\mathbf{X}} & \boldsymbol{\Sigma}_{\mathbf{Y}\mathbf{Y}} \end{pmatrix}^{-1} = \boldsymbol{\Sigma}^{-1} \quad (\text{B.5})$$

By plugging \mathbf{P} into the marginal density (B.4) and expanding parentheses, one obtains:

$$f_{\mathbf{X}}(x) = \frac{1}{Q} \int_{\mathbf{Y} \in \mathbb{R}^m} \exp \left(-\frac{1}{2} (x - \mu_{\mathbf{X}})^T \mathbf{P}_{\mathbf{X}\mathbf{X}} (x - \mu_{\mathbf{X}}) - \frac{1}{2} (x - \mu_{\mathbf{X}})^T \mathbf{P}_{\mathbf{X}\mathbf{Y}} (y - \mu_{\mathbf{Y}}) - \frac{1}{2} (y - \mu_{\mathbf{Y}})^T \mathbf{P}_{\mathbf{Y}\mathbf{X}} (x - \mu_{\mathbf{X}}) - \frac{1}{2} (y - \mu_{\mathbf{Y}})^T \mathbf{P}_{\mathbf{Y}\mathbf{Y}} (y - \mu_{\mathbf{Y}}) \right) dy \quad (\text{B.6})$$

where Q are constant terms. Thereafter, by completing the squares:

$$f_{\mathbf{X}}(x) = \frac{1}{Q} \int_{\mathbf{Y} \in \mathbb{R}^m} \exp \left(-\frac{1}{2} (x - \mu_{\mathbf{X}})^T \mathbf{P}_{\mathbf{X}\mathbf{X}} (x - \mu_{\mathbf{X}}) + \frac{1}{2} (x - \mu_{\mathbf{X}})^T \mathbf{P}_{\mathbf{X}\mathbf{Y}} \mathbf{P}_{\mathbf{Y}\mathbf{Y}}^{-1} \mathbf{P}_{\mathbf{Y}\mathbf{X}} (x - \mu_{\mathbf{X}}) - \frac{1}{2} ((y - \mu_{\mathbf{Y}}) + \mathbf{P}_{\mathbf{Y}\mathbf{Y}}^{-1} \mathbf{P}_{\mathbf{Y}\mathbf{X}} (x - \mu_{\mathbf{X}}))^T \mathbf{P}_{\mathbf{Y}\mathbf{Y}} ((y - \mu_{\mathbf{Y}}) + \mathbf{P}_{\mathbf{Y}\mathbf{Y}}^{-1} \mathbf{P}_{\mathbf{Y}\mathbf{X}} (x - \mu_{\mathbf{X}})) \right) dy \quad (\text{B.7})$$

Factorize y and rewrite:

$$\begin{aligned}
f_{\mathbf{X}}(x) = & \frac{1}{Q} \int_{Y \in \mathbb{R}^n} \exp \left(-\frac{1}{2} ((y - \mu_Y) + \mathbf{P}_{\mathbf{Y}\mathbf{Y}}^{-1} \mathbf{P}_{\mathbf{Y}\mathbf{X}}(x - \mu_X))^T \mathbf{P}_{\mathbf{Y}\mathbf{Y}} ((y - \mu_Y) \right. \\
& \left. + \mathbf{P}_{\mathbf{Y}\mathbf{Y}}^{-1} \mathbf{P}_{\mathbf{Y}\mathbf{X}}(x - \mu_X)) \right) dy \cdot \exp \left(-\frac{1}{2} (x - \mu_X)^T \mathbf{P}_{\mathbf{X}\mathbf{X}} (x - \mu_X) \right. \\
& \left. + \frac{1}{2} (x - \mu_X)^T \mathbf{P}_{\mathbf{X}\mathbf{Y}} \mathbf{P}_{\mathbf{Y}\mathbf{Y}}^{-1} \mathbf{P}_{\mathbf{Y}\mathbf{X}} (x - \mu_X) \right)
\end{aligned} \tag{B.8}$$

Remembering that a multivariate Gaussian pdf is:

$$(2\pi)^{n/2} |\Sigma|^{1/2} = \int_{\mathbb{R}^n} \exp \left(-\frac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu) \right) \tag{B.9}$$

Therefore the marginal distribution writes:

$$f_{\mathbf{X}}(x) = \frac{1}{Q} \cdot (2\pi)^{n/2} |\mathbf{P}_{\mathbf{Y}\mathbf{Y}}|^{1/2} \exp \left(-\frac{1}{2} (x - \mu_X)^T (\mathbf{P}_{\mathbf{X}\mathbf{X}} - \mathbf{P}_{\mathbf{X}\mathbf{Y}} \mathbf{P}_{\mathbf{Y}\mathbf{Y}}^{-1} \mathbf{P}_{\mathbf{Y}\mathbf{X}}) (x - \mu_X) \right) \tag{B.10}$$

In which the mean appears, μ_X . The derivation of the covariance requires further efforts, like the inversion of the precision matrix. The use of Schur complement is thus helpful. Consider a linear system, such that:

$$\begin{aligned}
Ax + By &= e \\
Cx + Dy &= f
\end{aligned} \tag{B.11}$$

The second equation writes (if and only if D is invertible):

$$y = D^{-1}(f - Cx) \tag{B.12}$$

substituting out y in the first equation, yields to:

$$(A - BD^{-1}C)x = e - BD^{-1}f \quad (\text{B.13})$$

from which x and y are solved with:

$$\begin{aligned} x &= (A - BD^{-1}C)^{-1}(e - BD^{-1}f) \\ x &= (A - BD^{-1}C)^{-1}e - (A - BD^{-1}C)^{-1}BD^{-1}f \\ y &= D^{-1}(f - C(A - BD^{-1}C)^{-1}(e - BD^{-1}f)) \end{aligned} \quad (\text{B.14})$$

$$y = -D^{-1}C(A - BD^{-1}C)^{-1}e + (D^{-1} + D^{-1}C(A - BD^{-1}C)^{-1}BD^{-1})f$$

where $(A - BD^{-1}C)$ is the Schur complement of D . Hence, the system is solved by the following inverse matrix:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} (A - BD^{-1}C)^{-1} & -(A - BD^{-1}C)^{-1}BD^{-1} \\ -D^{-1}C(A - BD^{-1}C)^{-1} & D^{-1} + D^{-1}C(A - BD^{-1}C)^{-1}BD^{-1} \end{pmatrix} \quad (\text{B.15})$$

Accordingly, the inverse of the precision matrix writes:

$$\begin{aligned} \begin{pmatrix} \mathbf{P}_{XX} & \mathbf{P}_{XY} \\ \mathbf{P}_{YX} & \mathbf{P}_{YY} \end{pmatrix}^{-1} &= \begin{pmatrix} \boldsymbol{\Sigma}_{XX} & \boldsymbol{\Sigma}_{XY} \\ \boldsymbol{\Sigma}_{YX} & \boldsymbol{\Sigma}_{YY} \end{pmatrix} \quad (\text{B.16}) \\ &= \begin{pmatrix} (\mathbf{P}_{XX} - \mathbf{P}_{XY}\mathbf{P}_{YY}^{-1}\mathbf{P}_{YX})^{-1} & -(\mathbf{P}_{XX} - \mathbf{P}_{XY}\mathbf{P}_{YY}^{-1}\mathbf{P}_{YX})^{-1}\mathbf{P}_{XY}\mathbf{P}_{YY}^{-1} \\ -\mathbf{P}_{YY}^{-1}\mathbf{P}_{YX}(\mathbf{P}_{XX} - \mathbf{P}_{XY}\mathbf{P}_{YY}^{-1}\mathbf{P}_{YX})^{-1} & (\mathbf{P}_{YY} - \mathbf{P}_{YX}\mathbf{P}_{XX}^{-1}\mathbf{P}_{XY})^{-1} \end{pmatrix} \end{aligned}$$

Hence, the covariance of the marginal distribution of \mathbf{X} is $\boldsymbol{\Sigma}_{XX} = (\mathbf{P}_{XX} - \mathbf{P}_{XY}\mathbf{P}_{YY}^{-1}\mathbf{P}_{YX})^{-1}$ Q.E.D.

Lemma 2: the conditional distributions of a joint Gaussian are Gaussians:

$$\begin{aligned} \mathbf{X}|\mathbf{Y} &\sim \mathcal{N}\left(\mu_{\mathbf{X}} + \Sigma_{\mathbf{X}\mathbf{Y}}\Sigma_{\mathbf{Y}\mathbf{Y}}^{-1}(y - \mu_{\mathbf{Y}}), \Sigma_{\mathbf{X}\mathbf{X}} - \Sigma_{\mathbf{X}\mathbf{Y}}\Sigma_{\mathbf{Y}\mathbf{Y}}^{-1}\Sigma_{\mathbf{Y}\mathbf{X}}\right) \\ \mathbf{Y}|\mathbf{X} &\sim \mathcal{N}\left(\mu_{\mathbf{Y}} + \Sigma_{\mathbf{Y}\mathbf{X}}\Sigma_{\mathbf{X}\mathbf{X}}^{-1}(x - \mu_{\mathbf{X}}), \Sigma_{\mathbf{Y}\mathbf{Y}} - \Sigma_{\mathbf{Y}\mathbf{X}}\Sigma_{\mathbf{X}\mathbf{X}}^{-1}\Sigma_{\mathbf{X}\mathbf{Y}}\right) \end{aligned} \quad (\text{B.17})$$

Proofs:

The conditional distribution of \mathbf{Y} given \mathbf{X} is given by:

$$f_{\mathbf{Y}|\mathbf{X}}(y|x) = \frac{1}{Q'} \exp\left(-\frac{1}{2}\begin{pmatrix} x - \mu_{\mathbf{X}} \\ y - \mu_{\mathbf{Y}} \end{pmatrix}^T \begin{pmatrix} \Sigma_{\mathbf{X}\mathbf{X}} & \Sigma_{\mathbf{X}\mathbf{Y}} \\ \Sigma_{\mathbf{Y}\mathbf{X}} & \Sigma_{\mathbf{Y}\mathbf{Y}} \end{pmatrix}^{-1} \begin{pmatrix} x - \mu_{\mathbf{X}} \\ y - \mu_{\mathbf{Y}} \end{pmatrix}\right) \quad (\text{B.18})$$

where Q' is a normalization constant. By the substitution of the above precision matrix \mathbf{P} , one can obtain:

$$\begin{aligned} f_{\mathbf{Y}|\mathbf{X}}(y|x) &= \frac{1}{Q'} \exp\left(-\frac{1}{2}(x - \mu_{\mathbf{X}})^T \mathbf{P}_{\mathbf{X}\mathbf{X}}(x - \mu_{\mathbf{X}}) - \frac{1}{2}(x - \mu_{\mathbf{X}})^T \mathbf{P}_{\mathbf{X}\mathbf{Y}}(y - \mu_{\mathbf{Y}}) \right. \\ &\quad \left. - \frac{1}{2}(y - \mu_{\mathbf{Y}})^T \mathbf{P}_{\mathbf{Y}\mathbf{X}}(x - \mu_{\mathbf{X}}) - \frac{1}{2}(y - \mu_{\mathbf{Y}})^T \mathbf{P}_{\mathbf{Y}\mathbf{Y}}(y - \mu_{\mathbf{Y}})\right) \end{aligned} \quad (\text{B.19})$$

By completing the squares:

$$\begin{aligned} f_{\mathbf{Y}|\mathbf{X}}(y|x) &= \frac{1}{Q'} \exp\left(-\frac{1}{2}(x - \mu_{\mathbf{X}})^T \mathbf{P}_{\mathbf{X}\mathbf{X}}(x - \mu_{\mathbf{X}}) + \frac{1}{2}(x - \mu_{\mathbf{X}})^T \mathbf{P}_{\mathbf{X}\mathbf{Y}}\mathbf{P}_{\mathbf{Y}\mathbf{Y}}^{-1}\mathbf{P}_{\mathbf{Y}\mathbf{X}}(x - \mu_{\mathbf{X}}) \right. \\ &\quad \left. - \frac{1}{2}((y - \mu_{\mathbf{Y}}) + \mathbf{P}_{\mathbf{Y}\mathbf{Y}}^{-1}\mathbf{P}_{\mathbf{Y}\mathbf{X}}(x - \mu_{\mathbf{X}}))^T \mathbf{P}_{\mathbf{Y}\mathbf{Y}}((y - \mu_{\mathbf{Y}}) + \mathbf{P}_{\mathbf{Y}\mathbf{Y}}^{-1}\mathbf{P}_{\mathbf{Y}\mathbf{X}}(x - \mu_{\mathbf{X}}))\right) \end{aligned} \quad (\text{B.20})$$

Factor out anything that is not dependent on y :

$$f_{\mathbf{Y}|\mathbf{X}}(y|x) = \frac{1}{Q''} \exp \left(-\frac{1}{2} \left((y - \mu_{\mathbf{Y}}) + \mathbf{P}_{\mathbf{Y}\mathbf{Y}}^{-1} \mathbf{P}_{\mathbf{Y}\mathbf{X}}(x - \mu_{\mathbf{X}}) \right)^T \mathbf{P}_{\mathbf{Y}\mathbf{Y}} \left((y - \mu_{\mathbf{Y}}) + \mathbf{P}_{\mathbf{Y}\mathbf{Y}}^{-1} \mathbf{P}_{\mathbf{Y}\mathbf{X}}(x - \mu_{\mathbf{X}}) \right) \right) \quad (\text{B.21})$$

Hence, it is straightforward to see that the conditional density $f_{\mathbf{Y}|\mathbf{X}}(y|x)$ is a Gaussian density of mean $\mu_{\mathbf{Y}} + \mathbf{P}_{\mathbf{Y}\mathbf{Y}}^{-1} \mathbf{P}_{\mathbf{Y}\mathbf{X}}(x - \mu_{\mathbf{X}})$ and covariance $\mathbf{P}_{\mathbf{Y}\mathbf{Y}}^{-1}$. The above inversion of the precision matrix helps to derive the mean and covariance matrix, such that:

$$\begin{aligned} \mu_{\mathbf{Y}|\mathbf{X}} &= \mu_{\mathbf{Y}} + \mathbf{P}_{\mathbf{Y}\mathbf{Y}}^{-1} \mathbf{P}_{\mathbf{Y}\mathbf{X}}(x - \mu_{\mathbf{X}}) = \mu_{\mathbf{Y}} + \boldsymbol{\Sigma}_{\mathbf{Y}\mathbf{X}} \boldsymbol{\Sigma}_{\mathbf{X}\mathbf{X}}^{-1} (x - \mu_{\mathbf{X}}) \\ \boldsymbol{\Sigma}_{\mathbf{Y}|\mathbf{X}} &= \mathbf{P}_{\mathbf{Y}\mathbf{Y}}^{-1} = \boldsymbol{\Sigma}_{\mathbf{Y}\mathbf{Y}} - \boldsymbol{\Sigma}_{\mathbf{Y}\mathbf{X}} \boldsymbol{\Sigma}_{\mathbf{X}\mathbf{X}}^{-1} \boldsymbol{\Sigma}_{\mathbf{X}\mathbf{Y}} \end{aligned} \quad (\text{B.22})$$

Q.E.D.

Due to its importance in the Kalman filter, remind that a vector \mathbf{Z} is Gaussian distributed such that $\mathbf{Z} = \mathbf{A}\mathbf{S} + \varepsilon$ where $\varepsilon \sim \mathcal{N}(0, \boldsymbol{\Sigma}_{\varepsilon})$. Hence, the linear transformation of a Gaussian vector is also Gaussian distributed, $\mathbf{S} \sim \mathcal{N}(\mathbf{A}\boldsymbol{\mu}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^T + \boldsymbol{\Sigma}_{\varepsilon})$

B.3.3 Kalman Filter

A standard linear state-space model is:¹

¹In the two following subsections (“Kalman filter” and “Kalman smoother”), some notations and derivations are found in the book of Särkkä (2013). This book is highly recommended for further knowledge.

$$x_l = A_{l-1}x_{l-1} + q_{l-1} \quad (\text{B.23})$$

$$y_l = H_l x_l + r_l \quad (\text{B.24})$$

where $x_l \in \mathbb{R}^n$ is a state vector, $y_l \in \mathbb{R}^m$ is a measurement vector, q_{l-1} and r_l are Gaussian errors, $q_l \sim \mathcal{N}(0, Q_l)$ and $r_l \sim \mathcal{N}(0, R_l)$. A_{l-1} is a $n \times n$ transition matrix and H_l is a $m \times n$ measurement matrix. Given the Gaussian properties, the related probabilistic state-space model is:

$$p(x_l|x_{l-1}) \sim \mathcal{N}(x_l|A_{l-1}x_{l-1}, Q_{l-1}) \quad (\text{B.25})$$

$$p(y_l|x_l) \sim \mathcal{N}(y_l|H_l x_l, R_l) \quad (\text{B.26})$$

The Bayesian filtering of this linear state-space model is reduced to the following distributions:

$$p(x_l|y_{1:l-1}) \sim \mathcal{N}(x_l|\hat{m}_l, \hat{P}_l) \quad (\text{B.27})$$

$$p(x_l|y_{1:l}) \sim \mathcal{N}(x_l|m_l, P_l) \quad (\text{B.28})$$

$$p(y_l|y_{1:l-1}) \sim \mathcal{N}(y_l|H_l \hat{m}_l, H_l \hat{P}_l H_l^T + R_l) \quad (\text{B.29})$$

Hence, the parameters are recursively computed with a Kalman filter algorithm. The latter consists of the estimate of the marginal posterior distribution $p(x_l|y_{1:l}) \sim \mathcal{N}(x_l|m_l, P_l)$. The Kalman filtering consists of three steps:

- 1st step: Initialize the system with a prior distribution $p(x_0)$.
- 2nd step: Obtain the predictive distribution of the state x_l , given the prior

knowledge and the dynamic model (predictive step):

$$p(x_l|y_{1:l-1}) = \int p(x_l|x_{l-1})p(x_{l-1}|y_{1:l-1})dx_{l-1} \quad (\text{B.30})$$

- 3rd step: Obtain the posterior distribution of the state x_l according to the Bayes rule, given the prediction and measurements (update step):

$$p(x_l|y_{1:l}) = \frac{p(y_l|x_l)p(x_l|y_{1:l-1})}{\int p(y_l|x_l)p(x_l|y_{1:l-1})dx_l} \quad (\text{B.31})$$

The closed-form equations of the prediction step are:

$$\hat{m}_l = A_{l-1}m_{l-1} \quad (\text{B.32})$$

$$\hat{P}_l = A_{l-1}P_{l-1}A_{l-1}^T + Q_{l-1} \quad (\text{B.33})$$

The closed-form equations of the update step are:

$$i_l = y_l - H_l\hat{m}_l \quad (\text{B.34})$$

$$S_l = H_l\hat{P}_lH_l^T + R_l \quad (\text{B.35})$$

$$K_l = \hat{P}_lH_l^T S_l^{-1} \quad (\text{B.36})$$

$$m_l = \hat{m}_l + K_l i_l \quad (\text{B.37})$$

$$P_l = \hat{P}_l - K_l H_l \hat{P}_l \quad (\text{B.38})$$

where K_l is the so-called Kalman Gain. By *Lemma 1*, the marginal distribution of the predictive state is given by:

$$\begin{aligned}
p(x_{l-1}, x_l | y_{1:l-1}) &= p(x_l | x_{l-1}) p(x_{l-1} | y_{1:l-1}) \\
&= \mathcal{N}(x_l | A_{l-1} m_{l-1}, Q_{l-1}) \mathcal{N}(x_{l-1} | m_{l-1}, P_{l-1}) \quad (\text{B.39}) \\
&= \mathcal{N} \left(\begin{pmatrix} x_{l-1} \\ x_l \end{pmatrix} \middle| \begin{pmatrix} m_{l-1} \\ A_{l-1} m_{l-1} \end{pmatrix}, \begin{pmatrix} P_{l-1} & P_{l-1} A_{l-1}^T \\ A_{l-1} P_{l-1} & A_{l-1} P_{l-1} A_{l-1}^T + Q_{l-1} \end{pmatrix} \right)
\end{aligned}$$

Let $\hat{m}_l = A_{l-1} m_{l-1}$ be the one-step ahead prediction mean and the predicted covariance as $\hat{P}_l = A_{l-1} P_{l-1} A_{l-1}^T + Q_{l-1}$. The latter is derived with:

$$\begin{aligned}
\hat{P}_l &= \mathbb{E} \left[(x_l - \hat{m}_l)(x_l - \hat{m}_l)^T \right] \\
&= \mathbb{E} \left[(A_{l-1} x_{l-1} + q_{l-1} - A_{l-1} m_{l-1})(A_{l-1} x_{l-1} + q_{l-1} - A_{l-1} m_{l-1})^T \right] \\
&= \mathbb{E} \left[(A_{l-1}(x_{l-1} - m_{l-1}) + q_{l-1})(A_{l-1}(x_{l-1} - m_{l-1}) + q_{l-1})^T \right] \quad (\text{B.40}) \\
&= \mathbb{E} \left[(A_{l-1} e_{l-1} + q_{l-1})(A_{l-1} e_{l-1} + q_{l-1})^T \right] \\
&= A_{l-1} P_{l-1} A_{l-1}^T + Q_{l-1}
\end{aligned}$$

where $P_{t-1} = \mathbb{E}[e_{l-1} e_{l-1}^T]$ and $Q_{t-1} = \mathbb{E}[q_{l-1} q_{l-1}^T]$. In the same vein, the joint distribution between the current state x_l and the measurement y_l is:

$$\begin{aligned}
p(x_l, y_l | y_{1:l-1}) &= p(y_l | x_l) p(x_l | y_{1:l-1}) \\
&= \mathcal{N}(y_l | H_l x_l, R_l) \mathcal{N}(x_l | \hat{m}_l, \hat{P}_l) \quad (\text{B.41}) \\
&= \mathcal{N} \left(\begin{pmatrix} x_l \\ y_l \end{pmatrix} \middle| \begin{pmatrix} \hat{m}_l \\ H_l \hat{m}_l \end{pmatrix}, \begin{pmatrix} \hat{P}_l & \hat{P}_l H_l^T \\ H_l \hat{P}_l & H_l \hat{P}_l H_l^T + R_l \end{pmatrix} \right)
\end{aligned}$$

By *Lemma 2*, the conditional distribution of x_l given $y_{1:l}$ is:

$$p(x_l|y_{1:l}) = \mathcal{N}(x_l|m_l, P_l) \quad (\text{B.42})$$

where,

$$\begin{aligned} m_l &= \hat{m}_l + \hat{P}_l H_l^T S_l^{-1} (y_l - H_l \hat{m}_l) = \hat{m}_l + K_l i_l \\ P_l &= \hat{P}_l - \hat{P}_l H_l^T S_l^{-1} H_l \hat{P}_l = \hat{P}_l - K_l H_l \hat{P}_l \end{aligned} \quad (\text{B.43})$$

The Kalman filter is optimal, unbiased and of minimum variance. The optimality holds for various criteria such as the minimum mean squared error (MMSE), the maximum likelihood (ML) or the maximum a posteriori (MAP). The latter is a ML principle for which the prior is not restricted to be uniform. The MMSE and MAP rely on the posterior distribution $p(x_l|y_{1:l})$. The Gaussian properties ensure that the mean and covariance of $p(y_l|x_l)$ are given by:

$$\begin{aligned} \mathbb{E}[y_l|x_l] &= H_l x_l \\ \text{Cov}[y_l|x_l] &= R_l \end{aligned} \quad (\text{B.44})$$

In addition, the mean and covariance of $p(x_l|y_{1:l-1})$ are:

$$\begin{aligned} \mathbb{E}[x_l|y_{1:l-1}] &= \hat{m}_l \\ \text{Cov}[x_l|y_{1:l-1}] &= \text{Cov}[x_l - \hat{m}_l] = \hat{P}_l \end{aligned} \quad (\text{B.45})$$

Knowing that and using the previous derivations, one can infer that the posterior distribution $p(x_l|y_{1:l})$ is of the form of:

$$\begin{aligned}
p(x_l|y_{1:l}) &\sim \mathcal{N}(x_l|m_l, P_l) \\
&= Z \exp\left(-\frac{1}{2}\left((x_l - \hat{m}_l)^T \hat{P}_l^{-1}(x_l - \hat{m}_l)\right) - \frac{1}{2}\left((y_l - H_l x_l)^T R_l^{-1}(y_l - H_l x_l)\right)\right)
\end{aligned} \tag{B.46}$$

The MAP optimal condition is:

$$\left. \frac{\partial \log p(x_l|y_{1:l})}{\partial x_l} \right|_{x_l=m_l} = 0 \tag{B.47}$$

Using the following derivative conditions:

$$\begin{aligned}
\frac{\partial}{\partial \mu} (x - A\mu)^T \Sigma (x - A\mu) &= -2A^T \Sigma (x - A\mu) \\
\frac{\partial}{\partial x} (x - \mu)^T \Sigma (x - \mu) &= 2\Sigma (x - \mu)
\end{aligned} \tag{B.48}$$

It is straightforward to find that:

$$m_l = \frac{\hat{P}_l^{-1} \hat{m}_l + H_l^T R_l^{-1} y_l}{(H_l^T R_l^{-1} H_l + \hat{P}_l^{-1})} \tag{B.49}$$

Using the Woodbury identity $(A + UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1}$, the MAP (B.49) writes:

$$\begin{aligned}
m_l &= (\hat{P}_l - \hat{P}_l H_l^T (R_l + H_l \hat{P}_l H_l^T)^{-1} H_l \hat{P}_l) \times (\hat{P}_l^{-1} \hat{m}_l + H_l^T R_l^{-1} y_l) \\
&= (\hat{m}_l - \hat{P}_l H_l^T (R_l + H_l \hat{P}_l H_l^T)^{-1} H_l \hat{m}_l) \\
&\quad + (\hat{P}_l H_l^T R_l^{-1} y_l - \hat{P}_l H_l^T (R_l + H_l \hat{P}_l H_l^T)^{-1} H_l \hat{P}_l H_l^T R_l^{-1} y_l) \\
&= (\hat{m}_l - K_l H_l \hat{m}_l) + K_l y_l \\
&= \hat{m}_l + K_l (y_l - H_l \hat{m}_l) \\
&= \hat{m}_l + K_l i_l
\end{aligned} \tag{B.50}$$

where $K_l = \hat{P}_l H_l^T (R_l + H_l \hat{P}_l H_l^T)^{-1}$. By substituting out for y_l , the MAP m_l is:

$$m_l = \hat{m}_l + K_l (H_l x_l + r_l - H_l \hat{m}_l) \tag{B.51}$$

The definition of the covariance P_l then rewrites to:

$$\begin{aligned}
P_l &= \mathbb{E}[e_l e_l^T] = \mathbb{E}[(x_l - m_l)(x_l - m_l)^T] \\
&= \mathbb{E}[(x_l - (\hat{m}_l + K_l (H_l x_l + r_l - H_l \hat{m}_l)))(x_l - (\hat{m}_l + K_l (H_l x_l + r_l - H_l \hat{m}_l)))^T] \\
&= \mathbb{E}[((I - K_l H_l)(x_l - \hat{m}_l) - K_l r_l)((I - K_l H_l)(x_l - \hat{m}_l) - K_l r_l)^T] \\
&= (I - K_l H_l) \mathbb{E}[(x_l - \hat{m}_l)(x_l - \hat{m}_l)^T] (I - K_l H_l)^T + K_l \mathbb{E}[r_l r_l^T] K_l^T \\
&= (I - K_l H_l) \hat{P}_l (I - K_l H_l)^T + K_l R_l K_l^T \\
&= \hat{P}_l - K_l H_l \hat{P}_l - \hat{P}_l H_l^T K_l^T + K_l (H_l \hat{P}_l H_l^T + R_l) K_l^T \\
&= \hat{P}_l - K_l H_l \hat{P}_l - \hat{P}_l H_l^T K_l^T + K_l S_l K_l^T
\end{aligned} \tag{B.52}$$

knowing that the right multiplication of the Kalman gain by $S_l K_l^T$ gives:

$$K_l S_l K_l^T = \hat{P}_l H_l^T K_l \tag{B.53}$$

the last two terms of (B.52) cancels out, therefore the MAP P_l covariance writes:

$$P_l = \hat{P}_l - K_l H_l \hat{P}_l \quad (\text{B.54})$$

The Kalman filter is also a MMSE estimator. Indeed, the latter can be shown by computing the minimization of the trace of the covariance P_l , as follows:

$$\text{Tr}[P_l] = \text{Tr}[\hat{P}_l] - 2 \text{Tr}[K_l H_l \hat{P}_l] + \text{Tr}[K_l (H_l \hat{P}_l H_l^T + R_l) K_l^T] \quad (\text{B.55})$$

using the fact that $\text{Tr}[A] = \text{Tr}[A^T]$ and differentiating with respect to K_l :

$$\begin{aligned} \frac{d \text{Tr}[P_l]}{d K_l} &= -2(H_l \hat{P}_l)^T + 2K_l (H_l \hat{P}_l H_l^T + R_l) = 0 \\ (H_l \hat{P}_l)^T &= K_l (H_l \hat{P}_l H_l^T + R_l) \\ K_l &= \hat{P}_l H_l^T (H_l \hat{P}_l H_l^T + R_l)^{-1} \end{aligned} \quad (\text{B.56})$$

B.3.4 Kalman Smoother

The Kalman filter computes the posterior marginal distribution of the current state given measurements up to the time step l . The Kalman smoother is a backward recursive algorithm that estimates the posterior marginal distribution of the current state given measurements up to a time step $T > l$. In other words, the Kalman smoother makes use of future measurements to estimate the current state of the model. The Rauch, Tung and Striebel (RTS) smoother provides the corresponding closed-form equations for linear state-space Gaussian model. In probabilistic terms, the Kalman smoother translates into:

$$\begin{aligned}
p(x_l|y_{1:T}) &= \mathcal{N}(x_l|\tilde{m}_l, \tilde{P}_l) \\
&= p(x_l|y_{1:l}) \int \left[\frac{p(x_{l+1}|x_l)p(x_{l+1}|y_{1:T})}{p(x_{l+1}|y_{1:l})} \right] dx_{l+1} \quad (\text{B.57})
\end{aligned}$$

where $p(x_l|y_{1:l})$ is the filtered posterior marginal distribution. The Markovian properties of this model implies that $p(x_l|x_{l+1}, y_{1:T}) = p(x_l|x_{l+1}, y_{1:l})$. Hence,

$$\begin{aligned}
p(x_l|x_{l+1}, y_{1:T}) &= p(x_l|x_{l+1}, y_{1:l}) \\
&= \frac{p(x_{l+1}|x_l)p(x_l|y_{1:l})}{p(x_{l+1}|y_{1:l})} \quad (\text{B.58})
\end{aligned}$$

Consequently, the joint distribution of x_l and x_{l+1} given the measurements $y_{1:T}$ is:

$$\begin{aligned}
p(x_l, x_{l+1}|y_{1:T}) &= p(x_l|x_{l+1}, y_{1:T})p(x_{l+1}|y_{1:T}) \\
&= \frac{p(x_{l+1}|x_l)p(x_l|y_{1:l})p(x_{l+1}|y_{1:T})}{p(x_{l+1}|y_{1:l})} \quad (\text{B.59})
\end{aligned}$$

The closed-form equations of the RTS smoother are:

$$\hat{m}_{l+1} = A_l m_l \quad (\text{B.60})$$

$$\hat{P}_{l+1} = A_l P_l A_l^T + Q_l \quad (\text{B.61})$$

$$C_l = P_l A_l^T \hat{P}_l^{-1} \quad (\text{B.62})$$

$$\tilde{m}_l = m_l + C_l(\tilde{m}_{l+1} - \hat{m}_{l+1}) \quad (\text{B.63})$$

$$\tilde{P}_l = P_l + C_l(\tilde{P}_{l+1} - \hat{P}_{l+1})C_l^T \quad (\text{B.64})$$

where the mean m_l and the covariance P_l are filtered. By *Lemma 1*, the joint distribution of the current state x_l and the future state x_{l+1} given measurement $y_{1:l}$ is:

$$\begin{aligned}
p(x_l, x_{l+1} | y_{1:l}) &= p(x_{l+1} | x_l) p(x_l | y_{1:l}) \\
&= \mathcal{N}(x_{l+1} | A_l m_l, Q_l) \mathcal{N}(x_l | m_l, P_l) \\
&= \mathcal{N} \left(\begin{pmatrix} x_l \\ x_{l+1} \end{pmatrix} \middle| \begin{pmatrix} m_l \\ A_l m_l \end{pmatrix}, \begin{pmatrix} P_l & P_l A_l^T \\ A_l P_l & A_l P_l A_l^T + Q_l \end{pmatrix} \right)
\end{aligned} \tag{B.65}$$

from which we derive the predicted mean $\hat{m}_{l+1} = A_l m_l$ and the predicted covariance $\hat{P}_{l+1} = A_l P_l A_l^T + Q_l$. By *Lemma 2*, the conditional distribution $p(x_l | x_{l+1}, y_{1:k})$ is:

$$\begin{aligned}
p(x_l | x_{l+1}, y_{1:k}) &= p(x_l | x_{l+1}, y_{1:T}) \\
&= \mathcal{N}(x_l | \tilde{m}'_l, \tilde{P}'_l)
\end{aligned} \tag{B.66}$$

where,

$$\begin{aligned}
\tilde{m}'_l &= m_l + C_l(x_{l+1} - A_l m_l) \\
\tilde{P}'_l &= P_l - C_l A_l P_l \\
C_l &= P_l A_l^T \hat{P}_l^{-1} = P_l A_l^T (A_l P_l A_l^T + Q_l)^{-1}
\end{aligned} \tag{B.67}$$

The joint distribution of x_l and x_{l+1} given $y_{1:T}$ is:

$$\begin{aligned}
p(x_{l+1}, x_l | y_{1:T}) &= p(x_l | x_{l+1}, y_{1:T}) p(x_{l+1} | y_{1:T}) \\
&= \mathcal{N}(x_l | \tilde{m}'_l, \tilde{P}'_l) \mathcal{N}(x_{l+1} | \tilde{m}_{l+1}, \tilde{P}_{l+1}) \\
&= \mathcal{N} \left(\begin{pmatrix} x_{l+1} \\ x_l \end{pmatrix} \middle| \tilde{m}''_l, \tilde{P}''_l \right)
\end{aligned} \tag{B.68}$$

where:

$$\tilde{m}''_l = \begin{pmatrix} \tilde{m}_{l+1}, \\ m_l + C_l(\tilde{m}_{l+1} - A_l m_l) \end{pmatrix} \tag{B.69}$$

$$\tilde{P}''_l = \begin{pmatrix} \tilde{P}_{l+1} & \tilde{P}_{l+1} C_l^T \\ C_l \tilde{P}_{l+1} & C_l \tilde{P}_{l+1} C_l^T + \tilde{P}'_l \end{pmatrix} \tag{B.70}$$

By *Lemma 2*, the conditional distribution $p(x_l | y_{1:T})$ is therefore:

$$p(x_l | x_{l+1}, y_{1:k}) \sim \mathcal{N}(x_l | \tilde{m}_l, \tilde{P}_l) \tag{B.71}$$

with,

$$\tilde{m}_l = m_l + C_l(\tilde{m}_{l+1} - \hat{m}_{l+1}) \tag{B.72}$$

$$\tilde{P}_l = P_l + C_l(\tilde{P}_{l+1} - \hat{P}_{l+1})C_l^T \tag{B.73}$$

B.4 Kalman filter and smoother for multiple TVP-FAVAR

The essence of this algorithm using a dual Kalman filter and smoother (KFS) to estimate multiple TVP-FAVAR is described in the technical appendix of Koop & Korobilis (2014).² Multiple TVP-FAVAR models are written with:

$$x_{i,t} = z_{i,t}\lambda_{i,t} + \mu_{i,t} \quad \mu_{i,t} \sim \mathcal{N}(0, V_{i,t}) \quad (\text{B.74})$$

$$z_{i,t} = z_{i,t-1}\beta_{i,t} + \varepsilon_{i,t} \quad \varepsilon_{i,t} \sim \mathcal{N}(0, Q_{i,t}) \quad (\text{B.75})$$

$$\lambda_{i,t} = \lambda_{i,t-1} + \nu_{i,t} \quad \nu_{i,t} \sim \mathcal{N}(0, W_{i,t}) \quad (\text{B.76})$$

$$\beta_{i,t} = \beta_{i,t-1} + \eta_{i,t} \quad \eta_{i,t} \sim \mathcal{N}(0, R_{i,t}) \quad (\text{B.77})$$

where the i subscript denotes all the $2^n - 1$ possible combinations of $n \times 1$ vectors $x_{i,t}$ of observables. $z_{i,t} = \begin{pmatrix} f_{i,t} \\ y_t \end{pmatrix}$ contains an $s \times 1$ vector y_t of observables and $k \times 1$ vectors $f_{i,t}$ of unobserved factors. All errors $\mu_{i,t}$, $\varepsilon_{i,t}$, $\nu_{i,t}$ and $\eta_{i,t}$ are uncorrelated Gaussians disturbances with time-varying covariances. The matrices of coefficients $\beta_{i,t}$ and factor loadings $\lambda_{i,t} = ((\lambda_{i,t}^f)^T, (\lambda_{i,t}^y)^T)^T$ are specified by random walk equations. The time-varying covariances $V_{i,t}$ and $Q_{i,t}$ are computed using an exponential weighted moving average (EWMA) method; that is, a first-order infinite impulse response filter with exponentially decaying weights. The time-varying covariances $W_{i,t}$ and $R_{i,t}$ are estimated using a forgetting factors method à la Koop & Korobilis (2012).

These two methods depend on hyperparameters defined as $\kappa_i \in [0, 1]$, $i =$

²In this section, we use and complete the derivations of this technical appendix. Hence, we use similar notations for the ease of understanding.

$1, \dots, 4$, namely decay and forgetting factors. These hyperparameters control the degree time-variation in the parameters. They have similar interpretation: the lowest a κ is, the greatest the $t - 1$ observation and its squared residual receive weight in the estimation. In other words, the lowest a κ is, the highest is the corresponding time variation. The hyperparameters κ_1 and κ_2 are dedicated to the variation of $V_{i,t}$ and $Q_{i,t}$. In the case of $\kappa_1 = \kappa_2 = 1$, $V_{i,t}$ and $Q_{i,t}$ are constant, therefore the model specification is a FAVAR. The hyperparameters κ_3 and κ_4 are related to $W_{i,t}$ and $R_{i,t}$. In the case of $\kappa_3 = \kappa_4 = 1$, $W_{i,t} = R_{i,t} = 0$, hence the model specification is homoskedastic.

The initialization of the algorithm relies on the choice of various prior distributions, which are here uninformative:

$$\begin{aligned}
 f_0 &\sim \mathcal{N}(0, \Sigma_{0|0}^f), & \beta_0 &\sim \mathcal{N}(0, \Sigma_{0|0}^\beta), & \lambda_0 &\sim \mathcal{N}(0, \Sigma_{0|0}^\lambda), \\
 V_0 &\equiv I_{n \times n}, & Q_0 &\equiv I_{(s+k) \times (s+k)}.
 \end{aligned} \tag{B.78}$$

In a nutshell, the dual KFS corresponds to:

- 1st step: Estimate $\lambda_{i,t}$, $\beta_{i,t}$, $V_{i,t}$ and $Q_{i,t}$ given the prior conditions and conditional on the principal component estimates of the unobserved factors $f_{i,t}$; denoted by $\tilde{f}_{i,t}$. Hence, $\tilde{z}_{i,t} = \begin{pmatrix} \tilde{f}_{i,t} \\ y_t \end{pmatrix}$.
- 2nd step: Update $f_{i,t}$, conditional on the previous estimates of $\lambda_{i,t}$ and $\beta_{i,t}$, $V_{i,t}$ and $Q_{i,t}$.

DUAL KFS ALGORITHM FOR THE ESTIMATE OF MULTIPLE TVP-FAVAR:

- For $t = 1, \dots, T$ and $i = 1, \dots, I$, the filtering and the updating steps of the **first** KFS proceeds as follows:

- Estimate $\lambda_{i,t}$, $\beta_{i,t}$, $V_{i,t}$ and $Q_{i,t}$, given the prior conditions and the principal component estimates of the unobserved factors $\tilde{f}_{i,t}$:

$$\begin{aligned} p(\lambda_{i,t}|x_{i,1:t-1}) &\sim \mathcal{N}\left(\lambda_{i,t|t-1}, \Sigma_{i,t|t-1}^\lambda\right) \\ p(\beta_{i,t}|z_{i,1:t-1}) &\sim \mathcal{N}\left(\beta_{i,t|t-1}, \Sigma_{i,t|t-1}^\beta\right) \end{aligned} \quad (\text{B.79})$$

where the one-step ahead predicted means are $\lambda_{i,t|t-1} = \lambda_{i,t-1|t-1}$, $\beta_{i,t|t-1} = \beta_{i,t-1|t-1}$. The predicted covariances are estimated using the forgetting factor method:

$$\begin{aligned} \Sigma_{i,t|t-1}^\lambda &= \Sigma_{i,t-1|t-1}^\lambda + \hat{W}_{i,t} \\ \Sigma_{i,t|t-1}^\beta &= \Sigma_{i,t-1|t-1}^\beta + \hat{R}_{i,t} \\ \hat{W}_{i,t} &= \left(\frac{1}{\kappa_3} - 1\right) \Sigma_{i,t-1|t-1}^\lambda \\ \hat{R}_{i,t} &= \left(\frac{1}{\kappa_4} - 1\right) \Sigma_{i,t-1|t-1}^\beta \end{aligned} \quad (\text{B.80})$$

- Compute the estimates of $V_{i,t}$ and $Q_{i,t}$ using the EWMA method:

$$\begin{aligned} \hat{V}_{i,t} &= \kappa_1 V_{i,t-1|t-1} + (1 - \kappa_1) \hat{\mu}_{i,t} \hat{\mu}_{i,t}^T \\ \hat{Q}_{i,t} &= \kappa_2 Q_{i,t-1|t-1} + (1 - \kappa_2) \hat{\varepsilon}_{i,t} \hat{\varepsilon}_{i,t}^T \end{aligned} \quad (\text{B.81})$$

where, $\hat{\mu}_{i,t} = x_{i,t} - \tilde{z}_{i,t} \lambda_{i,t|t-1}$ and $\hat{\varepsilon}_{i,t} = \tilde{z}_{i,t} - \tilde{z}_{i,t-1} \beta_{i,t|t-1}$

– Update $\lambda_{i,t}$ and $\beta_{i,t}$ as follows:

$$\begin{aligned} p(\lambda_{i,t}|x_{i,1:t}) &\sim \mathcal{N}\left(\lambda_{i,t|t}, \Sigma_{i,t|t}^\lambda\right) \\ p(\beta_{i,t}|z_{i,1:t}) &\sim \mathcal{N}\left(\beta_{i,t|t}, \Sigma_{i,t|t}^\beta\right) \end{aligned} \quad (\text{B.82})$$

The closed-form updated estimates are:

$$\lambda_{i,t|t} = \lambda_{i,t|t-1} + \Sigma_{i,t|t-1}^\lambda \tilde{z}_{i,t}^T \left(\hat{V}_{i,t} + \tilde{z}_{i,t} \Sigma_{i,t|t-1}^\lambda \tilde{z}_{i,t}^T \right)^{-1} (x_{i,t} - \tilde{z}_{i,t} \lambda_{i,t|t-1}) \quad (\text{B.83})$$

$$\Sigma_{i,t|t}^\lambda = \Sigma_{i,t|t-1}^\lambda - \Sigma_{i,t|t-1}^\lambda \tilde{z}_{i,t}^T \left(\hat{V}_{i,t} + \tilde{z}_{i,t} \Sigma_{i,t|t-1}^\lambda \tilde{z}_{i,t}^T \right)^{-1} \tilde{z}_{i,t} \Sigma_{i,t|t-1}^\lambda \quad (\text{B.84})$$

$$\beta_{i,t|t} = \beta_{i,t|t-1} + \Sigma_{i,t|t-1}^\beta \tilde{z}_{i,t}^T \left(\hat{Q}_{i,t} + \tilde{z}_{i,t-1} \Sigma_{i,t|t-1}^\beta \tilde{z}_{i,t-1}^T \right)^{-1} (\tilde{z}_{i,t} - \tilde{z}_{i,t-1} \beta_{i,t|t-1}) \quad (\text{B.85})$$

$$\Sigma_{i,t|t}^\beta = \Sigma_{i,t|t-1}^\beta - \Sigma_{i,t|t-1}^\beta \tilde{z}_{i,t-1}^T \left(\hat{Q}_{i,t} + \tilde{z}_{i,t-1} \Sigma_{i,t|t-1}^\beta \tilde{z}_{i,t-1}^T \right)^{-1} \tilde{z}_{i,t-1} \Sigma_{i,t|t-1}^\beta \quad (\text{B.86})$$

– Update $V_{i,t}$ and $Q_{i,t}$ with:

$$\begin{aligned} V_{i,t|t} &= \kappa_1 V_{i,t-1|t-1} + (1 - \kappa_1) \hat{\mu}_{i,t|t} \hat{\mu}_{i,t|t}^T \\ Q_{i,t|t} &= \kappa_2 Q_{i,t-1|t-1} + (1 - \kappa_2) \hat{\varepsilon}_{i,t|t} \hat{\varepsilon}_{i,t|t}^T \end{aligned} \quad (\text{B.87})$$

where, $\hat{\mu}_{i,t|t} = x_{i,t} - \tilde{z}_{i,t} \lambda_{i,t|t}$ and $\hat{\varepsilon}_{i,t|t} = \tilde{z}_{i,t} - \tilde{z}_{i,t-1} \beta_{i,t|t}$

- For $t = T - 1, \dots, 1$ and $i = 1, \dots, I$, the smoothing step of the **first** KFS proceeds as follows:

– Smooth $\lambda_{i,t}$, $\beta_{i,t}$, using the RTS smoother:

$$\begin{aligned} p(\lambda_{i,t}|x_{i,1:T}) &\sim \mathcal{N}(\lambda_{i,t|T}, \Sigma_{i,t|T}^\lambda) \\ p(\beta_{i,t}|z_{i,1:T}) &\sim \mathcal{N}(\beta_{i,t|T}, \Sigma_{i,t|T}^\beta) \end{aligned} \quad (\text{B.88})$$

– The closed-form updated estimates are:

$$\lambda_{i,t|T} = \lambda_{i,t|t} + C_{i,t}^\lambda (\lambda_{i,t+1|T} - \lambda_{i,t+1|t}) \quad (\text{B.89})$$

$$\Sigma_{i,t|T}^\lambda = \Sigma_{i,t|t}^\lambda + C_{i,t}^\lambda (\Sigma_{i,t+1|T}^\lambda - \Sigma_{i,t+1|t}^\lambda) C_{i,t}^{\lambda T} \quad (\text{B.90})$$

$$C_{i,t}^\lambda = \Sigma_{i,t|t}^\lambda (\Sigma_{i,t+1|t}^\lambda)^{-1} \quad (\text{B.91})$$

$$\beta_{i,t|T} = \beta_{i,t|t} + C_{i,t}^\beta (\beta_{i,t+1|T} - \beta_{i,t+1|t}) \quad (\text{B.92})$$

$$\Sigma_{i,t|T}^\beta = \Sigma_{i,t|t}^\beta + C_{i,t}^\beta (\Sigma_{i,t+1|T}^\beta - \Sigma_{i,t+1|t}^\beta) C_{i,t}^{\beta T} \quad (\text{B.93})$$

$$C_{i,t}^\beta = \Sigma_{i,t|t}^\beta (\Sigma_{i,t+1|t}^\beta)^{-1} \quad (\text{B.94})$$

– Update $V_{i,t}$ and $Q_{i,t}$ with:

$$V_{i,t|t+1} = \kappa_1 V_{i,t|t} + (1 - \kappa_1) V_{i,t+1|t+1} \quad (\text{B.95})$$

$$Q_{i,t|t+1} = \kappa_2 Q_{i,t|t} + (1 - \kappa_2) Q_{i,t+1|t+1}$$

- For $t = 1, \dots, T$ and $i = 1, \dots, I$, the filtering and the updating steps of the **second** KFS proceeds as follows:

– Estimate $f_{it}(z_{i,t})$, given the smoothed estimates of $\lambda_{i,t}$, $\beta_{i,t}$, $V_{i,t}$ and $Q_{i,t}$:

$$p(z_{i,t}|x_{i,1:t-1}) \sim \mathcal{N}(z_{i,t|t-1}, \Sigma_{i,t|t-1}^z) \quad (\text{B.96})$$

where the predicted mean and covariance are:

$$\begin{aligned} z_{i,t|t-1} &= z_{i,t-1|t-1} \beta_{i,t-1|t-1} \\ \Sigma_{i,t|t-1}^z &= \beta_{i,t-1|t-1} \Sigma_{i,t-1|t-1}^z \beta_{i,t-1|t-1}^T + Q_{i,t-1|t-1} \end{aligned} \quad (\text{B.97})$$

– Update z_t as follows:

$$p(z_{i,t}|x_{i,1:t}) \sim \mathcal{N}(z_{i,t|t}, \Sigma_{i,t|t}^z) \quad (\text{B.98})$$

The closed-form updated estimates are:

$$z_{i,t|t} = z_{i,t|t-1} + \Sigma_{i,t|t-1}^z \lambda_{i,t}^T \left(V_{i,t} + \lambda_{i,t} \Sigma_{i,t|t-1}^z \lambda_{i,t}^T \right)^{-1} (x_{i,t} - z_{i,t|t-1} \lambda_{i,t}) \quad (\text{B.99})$$

$$\Sigma_{i,t|t}^z = \Sigma_{i,t|t-1}^z - \Sigma_{i,t|t-1}^z \lambda_{i,t}^T \left(V_{i,t} + \lambda_{i,t} \Sigma_{i,t|t-1}^z \lambda_{i,t}^T \right)^{-1} \lambda_{i,t} \Sigma_{i,t|t-1}^z \quad (\text{B.100})$$

- For $t = T - 1, \dots, 1$ and $i = 1, \dots, I$, the smoothing step of the **second** KFS proceeds as follows:

– Smooth $z_{i,t}$, using the RTS smoother:

$$p(z_{i,t}|x_{i,1:T}) \sim \mathcal{N}(z_{i,t|T}, \Sigma_{i,t|T}^z) \quad (\text{B.101})$$

– The closed form updated estimates are:

$$z_{i,t|T} = z_{i,t|t} + C_{i,t}^z (z_{i,t+1|T} - z_{i,t+1|t}) \quad (\text{B.102})$$

$$\Sigma_{i,t|T}^z = \Sigma_{i,t|t}^z + C_{i,t}^z \left(\Sigma_{i,t+1|T}^z - \Sigma_{i,t+1|t}^z \right) C_{i,t}^{zT} \quad (\text{B.103})$$

$$C_{i,t}^z = \Sigma_{i,t|t}^z \beta_{i,t}^T \left(\Sigma_{i,t+1|t}^z \right)^{-1} \quad (\text{B.104})$$

B.5 Identification of structural shocks

A single TVP-FAVAR (equations (B.74) and (B.75)) can be compactly written as:

$$\begin{aligned} \begin{pmatrix} x_t \\ y_t \end{pmatrix} &= \tilde{\lambda}_t \begin{pmatrix} f_t \\ y_t \end{pmatrix} + \tilde{\mu}_t \\ z_t = \begin{pmatrix} f_t \\ y_t \end{pmatrix} &= B_{t,1} \begin{pmatrix} f_{t-1} \\ y_{t-1} \end{pmatrix} + \dots + B_{t,p} \begin{pmatrix} f_{t-p} \\ y_{t-p} \end{pmatrix} + \varepsilon_t \end{aligned} \quad (\text{B.105})$$

where $\tilde{\mu}_t = (\mu_t^T, 0)^T$ and $\tilde{\lambda}_t = \begin{pmatrix} \lambda_t^f & \lambda_t^y \\ 0_{1 \times k} & 1 \end{pmatrix}$. Plugging the space into the state equation, solves for the VMA representation of the model:

$$\begin{aligned} \begin{pmatrix} x_t \\ y_t \end{pmatrix} &= \tilde{\lambda}_t B_t(L)^{-1} \varepsilon_t + \tilde{\mu}_t \\ &= \tilde{\Phi}_t(L) \eta_t \end{aligned} \quad (\text{B.106})$$

where $B_t(L) = I - \sum_{l=1}^p B_{l,t} L^l$ and L is the lag operator. Notice that the VMA representation of the state equation is sufficient for the identification of structural shocks only for z_t variables. The entire identification of the structural shocks for each i combination of x_t implies $2^n - 1$ restriction schemes. However, this computation is almost inconceivable for a mixture of zero and sign restrictions. For the sake of parsimony, we focus our identification of structural shocks to the z_t variables. Accordingly, we need the VMA form of an TVP-VAR:

$$z_t = \begin{pmatrix} f_t \\ y_t \end{pmatrix} = B_t(L)^{-1} \varepsilon_t \quad (\text{B.107})$$

More generally, a structural TVP-VAR is of the form of:

$$\mathcal{S}_{0,t}z_t = \sum_{l=1}^p z_{t-l}\mathcal{S}_{t,l} + \vartheta_t \quad (\text{B.108})$$

where p is the lag-length and $\bar{\mathcal{S}}_t = (\text{vec}(S_{t,1})^T, \dots, \text{vec}(S_{t,l})^T)^T$ is a $mp \times m$, $m = s + k$, matrix of structural parameters for $t = 1 \dots T$. The reduced-form is:

$$z_t = \sum_{l=1}^p z_{t-l}B_{t,l} + \varepsilon_t \quad (\text{B.109})$$

Let $\beta_t = (\text{vec}(B_{t,1})^T, \dots, \text{vec}(B_{t,l})^T)^T$, $\beta_t = \bar{\mathcal{S}}_t\mathcal{S}_{0,t}^{-1}$ and $\varepsilon_t = \mathcal{S}_{0,t}^{-1}\vartheta_t$. The vectors ϑ_t contain the desired uncovered structural shocks. By assuming that $\mathbb{E}_t[\vartheta_t\vartheta_t^T] = I$, the covariance of the state equation Q_t is:

$$\mathbb{E}_t[\varepsilon_t\varepsilon_t^T] = \mathcal{S}_{0,t}^{-1}\mathbb{E}_t[\vartheta_t\vartheta_t^T]\mathcal{S}_{0,t}^{T-1} \quad (\text{B.110})$$

$$Q_t = \left(\mathcal{S}_{0,t}\mathcal{S}_{0,t}^T\right)^{-1} \quad (\text{B.111})$$

For the system to be identified, the idea is to pin down a unique S_0 by the use of linear or nonlinear restrictions at the parameters point $(\mathcal{S}_0, \bar{\mathcal{S}})$. In our framework, we apply restrictions to the last point estimate of Q_t , at time T ; that is, the smoothed Q_T . The set of admissible restrictions applied on impulse responses are nonlinear on the structural parameter space. To tackle this issue and avoid heavy optimization routines, Rubio-Ramirez et al. (2010)(RWZ) delivers a transformation method for these restrictions to be linear in the structural parameter space. The latter forms a set of admissible $w \times m$ matrices, for $1 \leq w \leq mp + m$, $w > 0$. Let the $w \times w$ D_j be the admissible matrices, for $1 \leq j \leq m$, satisfying the restrictions if and only if, the following condition holds:

$$D_j f(\mathcal{S}_0, \bar{\mathcal{S}}) e_j = 0 \quad (\text{B.112})$$

where e_j is the j -th column of $I_{m \times m}$.³ The columns of $f(\cdot)$ are arbitrarily ordered in columns ranks $d_j = \text{rank}(D_j)$; such that, $d_1 \geq d_2 \geq \dots d_m$. Given that, RWZ develop theorems regarding the rank conditions and the existence of a unique rotation matrix P for structural VAR systems to be exactly and globally identified. Further, RWZ provide an algorithm to compute P that satisfies the identifying restrictions. The related Theorems 6 and 7 in RWZ are:

- Theorem 6:
 - “Consider an SVAR with admissible and strongly regular restrictions represented by R . The SVAR is exactly identified if and only if the total number of restrictions is equal to $m(m - 1)/2$ and the rank condition in Theorem 1 is satisfied for some $(\mathcal{S}_0, \bar{\mathcal{S}}) \in R$ ”
- Theorem 7:
 - “Consider an SVAR with admissible and strongly regular restrictions represented by R . The SVAR is exactly identified if and only if $q_j = m - j$ ”.⁴

In our paper, we first identify the structural shocks using triangular zero restrictions to the impulse responses. We second make use of a mix of zero and sign restrictions. In the first scheme, we use a lower Cholesky decomposition of the

³The reader is referred to the paper to RWZ for the regularity conditions, the admissibility of the linear transformation and the normalization rule applied to the D_j matrices.

⁴See more details about the conditions for the set of admissible matrices R in RWZ.

contemporaneous impulse responses to match the triangular zero restrictions.

The corresponding transformation mapping is:

$$f(\mathcal{S}_0, \bar{\mathcal{S}}) = \begin{pmatrix} \text{IR}_0 \\ \text{IR}_\infty \end{pmatrix} = \begin{matrix} & \text{AS} & \text{AD} & \text{QE} & \text{TC} \\ \text{CPI} & \begin{pmatrix} \times & 0 & 0 & 0 \end{pmatrix} \\ \text{GDP} & \begin{pmatrix} \times & \times & 0 & 0 \end{pmatrix} \\ \text{QE} & \begin{pmatrix} \times & \times & \times & 0 \end{pmatrix} \\ \text{FAC} & \begin{pmatrix} \times & \times & \times & \times \end{pmatrix} \\ \text{CPI} & \begin{pmatrix} \times & \times & \times & \times \end{pmatrix} \\ \text{GDP} & \begin{pmatrix} \times & \times & \times & \times \end{pmatrix} \\ \text{QE} & \begin{pmatrix} \times & \times & \times & \times \end{pmatrix} \\ \text{FAC} & \begin{pmatrix} \times & \times & \times & \times \end{pmatrix} \end{matrix} \quad (\text{B.113})$$

where IR_0 are the short-run impulse responses and $\text{IR}_\infty(\mathcal{S}_0, \bar{\mathcal{S}}) = \left(\mathcal{S}_0^T - \sum_{l=1}^p \bar{\mathcal{S}}_l^T \right)^{-1}$ the infinite horizon impulse responses. In these matrices, a “0” signifies an exclusion restriction and a “x” stands for an unrestricted shock. In plain english, these zero restrictions have the following meanings:

- The structural shocks of the estimated factor (TC) have no contemporaneous effects on the total of assets of all Federal Reserve banks (QE), GDP and CPI.
- The QE shocks have no contemporaneous effects on GDP and CPI.
- The aggregate demand shocks have no contemporaneous effects on CPI.

The corresponding D_j matrices are:

$$\begin{aligned}
 D_1 &= \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, & D_2 &= \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \\
 D_3 &= \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, & D_4 &= \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.
 \end{aligned}
 \tag{B.114}$$

Further, we reorder the columns so that the column ranks $d_4 = 3$, $d_3 = 2$, $d_2 = 1$, $d_1 = 0$ satisfy the ordering condition $d_1 \geq d_2 \geq d_3 \geq d_4$.

In addition, the total number of zero restrictions is 6; therefore, the order condition is satisfied. However, the order condition is necessary but not sufficient for that system to be globally identified. Indeed, the Theorem 1 of RWZ asserts that:

- Theorem 1:

- “Consider an SVAR with admissible restrictions represented by R . If $(\mathcal{S}_0, \bar{\mathcal{S}}) \in R$ and $M_j f(\mathcal{S}_0, \bar{\mathcal{S}})$ is of rank m for $1 \leq j \leq m$, then the SVAR is globally identified at the parameter point $(\mathcal{S}_0, \bar{\mathcal{S}})$ ”

The corresponding M_j matrices are computed as follows:

$$M_j f(\mathcal{S}_0, \bar{\mathcal{S}}) = \begin{pmatrix} Q_j & f(\mathcal{S}_0, \bar{\mathcal{S}}) \\ mp \times mp & mp \times m \\ [I & \mathbf{0}] \\ j \times j & j \times (m-j) \end{pmatrix} \quad (\text{B.115})$$

which are for these particular restrictions:

$$\begin{aligned}
 M_1 &= \begin{pmatrix} 0 & 0 & 0 & x \\ 0 & 0 & x & x \\ 0 & x & x & x \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \end{pmatrix}, & M_2 &= \begin{pmatrix} 0 & 0 & 0 & x \\ 0 & 0 & x & x \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \\
 M_3 &= \begin{pmatrix} 0 & 0 & 0 & x \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, & M_4 &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{B.116}
 \end{aligned}$$

The M_j are of full rank. Hence, this model is globally identified according to

Theorem 7. RWZ propose various algorithms to find the unique orthogonal matrix P in respect with the linear restrictions. In the case of exactly identified models, following the notations of Binning (2013), the related matrix of zero restrictions (B.113) is defined as:

$$f = \begin{matrix} & \text{AS} & \text{AD} & \text{QE} & \text{TC} \\ \text{CPI} & \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{array} \right) & & & \\ \text{GDP} & & & & \\ \text{QE} & & & & \\ \text{FAC} & & & & \\ \text{CPI} & & & & \\ \text{GDP} & & & & \\ \text{QE} & & & & \\ \text{FAC} & & & & \end{matrix} \quad (\text{B.117})$$

In this algorithm, the Cholesky decomposition of Q , so that $C = \text{chol}(Q)^T$ and $CC^T = Q$ is transformed using the rotation matrix P , to satisfy the restrictions. More precisely, the short and long-run impulse response are respectively, $\overline{IR}_0 = C$ and $\overline{IR}_\infty = \begin{bmatrix} I & -\beta \\ m \times m & \end{bmatrix}^{-1} C$. In the end, the algorithm computes $IR_0 = \overline{IR}_0 P$ and $IR_\infty = \overline{IR}_\infty P$. In the following algorithm, we define the rotation matrix as \bar{P} . The essence of the following algorithm is available in Binning (2013).

ALGORITHM: ZERO TRIANGULAR RESTRICTIONS:

- **For** $\tau = 1:t$
 - **for** $j = 1:m$
 - * Create an index for the zero entries of f :
 - $H = fe_{\cdot,j} == 0, \quad e = \underset{m \times m}{I}$
 - * Compute the D_j matrices, sorted in a descending order:
 - $D_j = \text{diag}(H)$
 - $r_j = \text{rank}(Q_j)$
 - $\text{ord} = \text{sort}(r_j, 'descend')$
 - * Create an index mapping the new and old orderings, indexed by i , such that:
 - $i(\text{ord}(j)) = j$
 - * Check if the rank conditions are satisfied: $r_i = m - i$.
 - **end for**
 - Compute $C = \text{chol}(Q)^T$, define: $\bar{P} = \underset{m \times m}{0}$ and $F = \begin{pmatrix} \bar{I}R_0 \\ \bar{I}R_\infty \end{pmatrix}$, where

$$\bar{I}R_0 = C \text{ and } \bar{I}R_\infty = \begin{bmatrix} I & \\ & -\beta \end{bmatrix}^{-1} C$$
 - **for** $i = 1:m$
 - * **if** $i = 1$
 - $\tilde{D}_i = D_i F$
 - * *else*
 - $\tilde{D}_i = \begin{pmatrix} D_i F \\ \bar{P}^T \end{pmatrix}$
 - * **end if**

- Perform a QR decomposition of D_i such that, $D_i f(\overline{IR}_0, \overline{IR}_\infty) \bar{P}_{.,i} = 0$.
 - * $[\bar{Q}, \bar{R}] = \text{qr}(\tilde{D}_i^T)$
 - * $\bar{P}_{.,i} = \bar{Q}_{.,m}$
 - * Reorder the columns of P to be consistent with the j ordering.
- **end for**
- Compute the structural impulse responses:
 - * $S_0 = C\bar{P}$
 - * $S_0 S_0^T = Q$
- **end for**

In the second *for* loop ($i = 1:m$), \tilde{D}_i is rank deficient ($m - 1$). Consequently, the unknown $m \times 1$ P_i orthogonal vectors are found using a QR factorization:

$$\tilde{D}_i^T = \bar{Q}\bar{R} = \underset{m \times m}{\bar{Q}} \begin{pmatrix} \bar{R}_1 \\ (m-1) \times (m-1) \\ 0 \\ 1 \times (m-1) \end{pmatrix} \quad (\text{B.118})$$

$$\tilde{D}_i^T = \begin{bmatrix} \bar{Q}_1 & \bar{Q}_2 \\ m \times (m-1) & m \times 1 \end{bmatrix} \begin{pmatrix} \bar{R}_1 \\ (m-1) \times (m-1) \\ 0 \\ 1 \times (m-1) \end{pmatrix} \quad (\text{B.119})$$

Given that $\text{rank}(\tilde{D}_i) = m - 1$, the solution is the last column of $\bar{Q}_{.,m}$. The system is:

$$\tilde{D}_i P_{.,i} = \bar{R}_1^T \bar{Q}^T P_{.,i} = \begin{bmatrix} \bar{R}_1^T & 0 \end{bmatrix} \begin{pmatrix} \bar{Q}_1^T \\ \bar{Q}_2^T \end{pmatrix} = 0 \quad (\text{B.120})$$

thus, $P_{.,i} = \bar{Q}_{.,m}$ and $\bar{Q}\bar{Q}_{.,m} = \begin{bmatrix} 0 & 1 \\ 1 \times m-1 & \end{bmatrix}^T$.

The second identification method is a mix of zero and sign restrictions. The main issue of zero triangular schemes is that such recursive systems are bold regarding theoretical wisdoms. In addition, the corresponding structural impulse responses might display puzzling signs. To tackle these issues, the use sign restrictions permits a matching between reduced-form IRF and expected theoretical signs. However, a system with sign restrictions is not identified according to the order and rank conditions. Consequently, the identification is handled using a large set of impulse responses that matches the sign restrictions. RWZ offers an algorithm to produce orthogonal matrices to compute IRF satisfying the sign restrictions. The algorithm follows the Theorem 9 of RWZ:

- Theorem 9:
 - “Let \tilde{X} be an $m \times m$ random matrix with each element having an

independent standard normal distribution. Let $\tilde{X} = \tilde{Q}\tilde{R}$ be the QR decomposition of \tilde{X} with the diagonal of \tilde{R} normalized to be positive.

Then \tilde{Q} has the uniform (or Haar) distribution.”

The QR decomposition is unique $\tilde{X} = \tilde{Q}\tilde{R}$ if and only if $\tilde{R} \in T(n)$, the set of invertible $n \times n$ upper triangular matrices with positive diagonal entries. According to the circular law, the distribution of \tilde{Q} thus converges to an uniform as n tends to infinity. In the notation of Binning (2013), the matrix of zero restrictions, we assume, in this mixture of restrictions is:

$$f = \begin{matrix} & \text{AS} & \text{AD} & \text{QE} & \text{TC} \\ \text{CPI} & \left(\begin{matrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{matrix} \right) & & & \\ \text{GDP} & & & & \\ \text{QE} & & & & \\ \text{FAC} & & & & \end{matrix} \quad (\text{B.121})$$

In plain english, these zero restrictions mean that:

- The structural aggregate supply shocks have no contemporaneous effects on the total of assets of all Federal Reserve banks (QE).
- The structural aggregate demand shocks have no contemporaneous effects on the total of assets of all Federal Reserve banks (QE).

The matrix of sign restrictions for a single horizon is:

$$f_{-}^{+} = \begin{matrix} & \text{AS} & \text{AD} & \text{QE} & \text{TC} \\ \text{CPI} & \left(\begin{matrix} + & + & \times & \times \\ - & + & \times & \times \\ \times & \times & + & \times \\ \times & \times & \times & \times \end{matrix} \right) \\ \text{GDP} \\ \text{QE} \\ \text{FAC} \end{matrix} \quad (\text{B.122})$$

In plain english, these sign restrictions mean that:

- CPI must respond positively to aggregate supply shocks.
- GDP must respond negatively to aggregate supply shocks.
- CPI must respond positively to aggregate demand shocks.
- GDP must respond positively to aggregate demand shocks.
- QE must respond positively to QE shocks.

ALGORITHM FOR ZERO AND SIGN RESTRICTIONS:

- **For** $\tau = 1:t$
 - **for** $j = 1:m$
 - * Create an index for the zero entries of f :
 - $H = fe_{.,j} == 0, \quad e = \begin{matrix} I \\ m \times m \end{matrix}$
 - * Compute the D_j matrices, sorted in a descending order:
 - $D_j = \text{diag}(H)$
 - $r_j = \text{rank}(Q_j)$
 - $\text{ord} = \text{sort}(r_j, 'descend')$
 - * Create an index mapping the new and old orderings, indexed by i , such that:
 - $i(\text{ord}(j)) = j$
 - * Check if the rank conditions are satisfied: $r_i = m - i$.
 - **end for**
 - Compute $C = \text{chol}(Q)^T$
 - **while** k_j draws
 - * $\tilde{X} = \text{normrnd}(0, 1)$
 $m \times m$
 - * $[\tilde{Q}, \tilde{R}] = \text{qr}(\tilde{X})$
 - * **for** $i = 1:m$
 - if $\tilde{R}_{i,i} < 0$
 - $\tilde{Q}_{.,i} = -\tilde{Q}_{.,i}$
 - **end if**
 - * **end for**
 - Define $F = \begin{pmatrix} \overline{IR}_0 \\ \overline{IR}_\infty \end{pmatrix}$, where $\overline{IR}_0 = C\tilde{Q}$ and $\overline{IR}_\infty = \begin{bmatrix} I \\ m \times m \end{bmatrix} - \beta \bigg]^{-1} C\tilde{Q}$
 - **for** $i = 1:m$

- * **if** $i = 1$
 - $\tilde{D}_i = D_i F$
- * *else*
 - $\tilde{D}_i = \begin{pmatrix} D_i F \\ \bar{p}^T \end{pmatrix}$
- * **end if**
- Perform a QR decomposition of D_i such that, $D_i f(\overline{IR}_0, \overline{IR}_\infty) \bar{P}_{.,i} = 0$.
 - * $[\bar{Q}, \bar{R}] = \text{qr}(\tilde{D}_i^T)$
 - * $\bar{P}_{.,i} = \bar{Q}_{.,m}$
 - * Reorder the columns of P to be consistent with the j ordering.
- **end for**
- Compute the structural impulse responses:
 - * $S_0 = C \bar{P}$
 - * $S_0 S_0^T = Q$
- if the sign restrictions are satisfied
 - * $k = k + 1$
- **end while**
- **end for**

Chapter 3

State-contingent Forward Guidance

Abstract

In this paper, the impacts of a state-contingent forward guidance are assessed in a DSGE model with search and matching frictions. Under an Odyssean perspective, the commitment is found to have relatively low effects in the economy. This result tackles the so-called *forward guidance puzzle*. In addition, the simulations suggest that a state-contingent forward guidance is mainly transmitted by shifts in the expectations of inflation.

Joint paper with: Julien Albertini¹ & Stéphane Moyen² – Preliminary draft.

¹GATE, University of Lyon 2, julien.Albertini@univ-lyon2.fr.

²Deutsche Bundesbank, stephane.moyen@bundesbank.de.

3.1 Introduction

The aftermath of the recent US financial crisis is characterized by a period of zero lower bound (ZLB, 2008M12-2015M11) in the Federal Funds rate. Nevertheless, projections of financial and economic conditions advocated for further monetary accommodation. As an emergency policy, the Federal Open Market Committee (FOMC) delivered several public statements about the future stance of the monetary policy. In particular, in December 12, 2012:

"[...] The Committee decided to keep the target range for the Federal Funds rate at 0 to 1/4 percent and currently anticipates that this exceptionally low range for the Federal Funds rate will be appropriate at least as long as the unemployment rate remains above 6-1/2 percent".

Along the lines of Campbell et al. (2012), this statement is an Odyssean forward guidance (FG). Phrased differently, a commitment to staying in ZLB, until the unemployment rate crosses an exit threshold is a state-contingent forward guidance (STFG).³ Hence,

To what extent is a STFG impacting the real economy?

Intriguingly, Boneva et al. (2018) (BHW) is the only study to answer to this question with the use of a DSGE model.⁴ Knowing the crucial role of DSGE models in the policy analysis, the gap is abyssal. In the study of BHW, a SCFG

³In the ZLB period, only two statements can be considered as STFG. The other one: "[...] will be appropriate to maintain the current target range for the federal funds rate well past the time that the unemployment rate declines below 6-1/2 percent, especially if projected inflation continues to run below the Committee's 2 percent longer-run goal." of December 18, 2013.

⁴The two-period model of Florez-Jimenez & Parra-Polania (2016) or the perfect foresight study of English et al. (2015) are the alternatives. The seminal papers related to FG under ZLB are Campbell et al. (2012) and Del Negro et al. (2012)

is found to have a temporary stimulus which outperforms a classical optimal monetary discretion (in terms of welfare). However, their use of a probabilistic mapping in the exit of STFG, is a bold modeling assumption. This *ad hoc* mechanism ensures the existence of an absorbing state for the trajectory of the paths; occasionally absent in the context of a *certain* threshold. This exit probability is an exponential function of the distance to a fixed threshold in the output gap and/or inflation. As such, the nominal interest rate lasts longer in ZLB, than otherwise required by an optimal discretion policy. This specification is likely to artificially amplify the effects of SCFG. Hence, it is reasonable to assert that these results remain greatly inconclusive.

This paper fills the gap in the literature by proposing a New-Keynesian DSGE model with a SCFG triggered by a monetary regime-switching system. The simulated scenario is a positive discount factor shock, large enough to cause a high unemployment and a ZLB. Private agents refer to an effective interest rate to take decisions. In a *normal* regime, this rate coincides with a notional interest rate. The notional rate follows a Taylor rule, that classically penalizes the deviations in inflation and in the employment rate. In addition, the modeling of an endogenous unemployment rate is a realistic support for the experiment of a SCFG. The model is solved using projection methods, which accounts for the nonlinearities implied by the regime-switching system.⁵

Our modeling assumption of the SCFG is an important novelty in the literature. The economy evolves in three endogenous monetary regimes. Apart from the *normal* regime presented above, the effective rate is stuck to zero in a ZLB regime without SCFG. The effective rate remains to zero, as long as the level of the notional rate reverts positive. Subsequently, enter from a SCFG

⁵The reader is referred to the technical appendix for further details about the solution method.

depends on an unemployment threshold. As an additional assumption, the entry is triggered if and only if a ZLB regime applied in the previous period. In contrary, exit from a SCFG is dependent of an exit unemployment threshold. This makes the possibility for the economy to last longer in ZLB, even if the notional interest rate advocates for a *normalization*. In the below experiments, the findings are multiple.

First, a SCFG is able to dampen a crash in the labor market. In particular, a SCFG stimulates the economy by refining the expectations in inflation. The underlying mechanisms work as follows. The credible commitment to an exit target convinces agents today, that the inflation will be higher in the future. This inflationary pressure diminishes the current real interest rate. Firms deliver job vacancies, in turn filled by workers. In addition, firms anticipate decline in relative prices; therefore compensate expected losses by increasing production today. Under these circumstances, workers claim for higher wages. For all of these reasons, the aggregate demand is boosted.⁶

Second, it is a widely held view that the impacts of SCFG are increasing in the duration of ZLB; the commitment to staying longer at the zero, inducing a faster recovery of the economy (*e.g.* in BHW). The upcoming results are contrasting. The SCFG effectively extends ZLB and accelerates the recovery in the labor market. However, in a simulated huge deflationary recession, the effects of a SCFG doubles, while the duration of the ZLB is not extended.⁷ From that point, any closed conclusion about the impact of duration would be too audacious, else is left for future research. However, it is clearly apparent that the

⁶More details about the mechanism are described in the subsection 3.3.2

⁷The impacts of SCFG are computed in percent change of output compared to simulations absent of SCFG. See results below.

highest proportion of the impacts of SCFG are due to the anticipation of agents to enter that regime. Conversely, the effects are squeezed by the anticipation about the liftoff from the ZLB. Indeed, the exit of a SCFG induces a convergence of the effective rate towards the level required by the Taylor rule. Hence, the effective and the real rates greatly increase, mitigating the cumulated stimulus.

Third, out of the anticipation effects, a SCFG is weakly impacting the economy; roughly 0.15 *p.p.* of cumulated output. This is an upper-bound estimate, since the SCFG regime is being perceived perfectly credible by agents.⁸ This result is a fair tackle to the *forward guidance puzzle*, described along the lines of Del Negro et al. (2012).

The rest of the paper is organized as follows. Section 2 is devoted to the presentation of model. Section 3 presents the calibration and the impulse responses, counterfactual and sensitivity analyses. Section 4 concludes.

3.2 The Model

The model is a New Keynesian DSGE model with search and matching frictions in the spirit of Mortensen & Pissarides (1994). In addition, it is featured with nominal price rigidities à la Rotemberg (1982), a monopolistic competition and a feedback Taylor rule for monetary policy. The flow in the labor force is evolving in the intensive margin; that is, the number of employed households. A Nash bargaining process determines the optimal wages. Last, the cycle is driven by exogenous deviations in the discount factor. The following subsections provide supplemental details.

⁸The SCFG is assumed to be a fully credible regime because the economy is absorbed to this regime until the exit threshold is crossed.

3.2.1 The Labor Market

The labor market is populated by identical employed or unemployed workers, and productive firms. Firms fill vacant jobs by publishing advertisements.⁹ The job search and the screening of workers is costly. By definition, the number of matches, m_t , is determined by a CES function of the form of:¹⁰

$$m_t = \left(s_t^{-\gamma} + v_t^{-\gamma} \right)^{-\frac{1}{\gamma}} \leq \min(s_t, v_t) \quad (3.1)$$

where $v_t \geq 0$ is the mass of job vacancies, and $s_t \geq 0$ the mass of searching workers. The labor force, L , is in a fixed supply. From that, we define $q_t = m_t/v_t$ as the probability of a filled vacancy and $f_t = m_t/s_t$ as the job finding probability. Accordingly, the tightness of the labor market is determined by $\theta_t = f_t/q_t$. It is necessary to clarify that the matching process (3.1) is an increasing and concave function in its arguments.

3.2.2 The Sequence of Events

For the purpose of introducing unemployment to the model, a fixed proportion of jobs are periodically destroyed at the exogenous rate ρ^x .¹¹ Accordingly, the number of job seekers is given by the following relation:

$$s_t = 1 - (1 - \rho^x)n_{t-1} \quad (3.2)$$

⁹The population of the model is normalized to 1; hence, the number of employed/unemployed agents and the employment/unemployment rate coincide.

¹⁰The use of a CES matching function ensures that the job finding and filling probabilities remain below 1.

¹¹This mechanism is borrowed from Hall (2005)

where n_t is the employment rate. Further, the law of motion of employment is defined by:

$$n_t = (1 - \rho^x)n_{t-1} + m_t \quad (3.3)$$

The equations (3.2) and (3.3) describe the sequence of events in the labor market. In the beginning of a period t , a number s_t of households are unemployed. Within the same period, the matching process m_t allows for these households to find another job. As such, this dynamic implies an unemployment rate, such that $u_t = 1 - n_t$. If an household remains unemployed, she receives fixed benefits in compensation to her situation.

3.2.3 The Representative Household

Identical households form a continuum of families, which are indexed by $i \in [0, 1]$. The labor incomes and unemployment benefits (UB) are pooled so that revenues are equally redistributed among the family members. By assumptions, households have preferences over a large variety of goods, which are indexed by $j \in [0, 1]$. Hence, the Dixit-Stiglitz demand system is:

$$c_{jt} = \left(\frac{p_{jt}}{p_t} \right)^{-\epsilon} c_t \quad (3.4)$$

$$p_t = \left[\int_0^1 p_{jt}^{1-\epsilon} dj \right]^{\frac{1}{1-\epsilon}} \quad (3.5)$$

where ϵ is a measure of substitutability, c_t the aggregate consumption and p_t the price index. The utility function of a representative household is defined by:

$$\max_{\Omega_t^H} \mathbb{E}_0 \sum_{t=0}^{\infty} \left(\prod_{k=0}^t \beta_k \right) \left[\frac{c_t^{1-\sigma}}{1-\sigma} + \ell(1 - n_t) \right] \quad (3.6)$$

where β_t is a discount factor shock and ℓ is the utility derived from leisure. Each household chooses the set of control variables, $\Omega_t^H = \{c_t, d_t, n_t\}_{t=0}^\infty$, by taking the state variables $\{p_t, w_t, i_t, f_t\}_{t=0}^\infty$, and the initial wealth (d_0), as given. The maximization of the utility function is subject to the law of motion of employment (3.3) and the following budget constraint:

$$p_t c_t + d_t = d_{t-1}(1 + i_{t-1}) + w_t n_t + (1 - n_t) b p_t + \Pi_t + T_t \quad (3.7)$$

in which d_t is a one-period domestic bonds bearing a nominal interest rate i_t . Let w_t be nominal wages, Π_t profits from the holding of firms, T_t lump-sum taxes and b fixed UB. The corresponding first-order conditions of an household's problem are given by:

$$\varphi_t = \lambda_t (w_t^R - b) - \ell + \mathbb{E}_t \beta_{t+1} (1 - \rho^x) (1 - f_{t+1}) \varphi_{t+1} \quad (3.8)$$

$$\lambda_t = c_t^{-1} \quad (3.9)$$

$$\lambda_t = (1 + i_t) \mathbb{E}_t \beta_{t+1} \lambda_{t+1} \frac{p_t}{p_{t+1}} \quad (3.10)$$

where $w_t^R = w_t / p_t$ is the real wage.

3.2.4 Firms

A continuum of j producers forms a monopolistically competitive market and produce the variety of goods. The labor force is the unique input of production. By assumption, the production function is linear in the employment rate:¹²

$$y_{jt} = n_{jt} \quad (3.11)$$

¹² $n_t = \int_0^1 n_{jt} dj$

The optimization problem of a firm consists of choosing the set of control variables, $\Omega_{jt}^F = \{v_{jt}, p_{jt}, n_{jt}\}_{t=0}^{\infty}$, by taking the state variables $\{p_t, w_{jt}, q_t\}_{t=0}^{\infty}$ as given. In addition, each producer maximizes the following intertemporal problem:

$$\max_{\Omega_t^p} \mathbb{E}_0 \sum_{t=0}^{\infty} \left(\prod_{k=0}^t \beta_k \right) \frac{\lambda_t}{\lambda_0} \Pi_{jt} \quad (3.12)$$

where $\Pi_{jt} = \left[\frac{p_{jt}}{p_t} y_{jt} - \frac{w_{jt}}{p_t} n_{jt} - \kappa v_{jt} - y_t \Gamma^\pi(p_{jt}) \right]$

subject to the production function (3.11) and the following evolution of the employment rate:

$$n_{jt} = (1 - \rho^x) n_{jt-1} + q_t v_{jt} \quad (3.13)$$

Along the specification of Rotemberg (1982), firms face quadratic adjustment costs to the price level, such that:

$$\Gamma^\pi(p_{jt}) = \frac{\psi}{2} \left(\frac{p_{jt}}{\pi p_{jt-1}} - 1 \right)^2 \quad (3.14)$$

where $\frac{p_{jt}}{p_t}$ are the relative price of goods. Let $\pi_t = p_t / p_{t-1}$ be the inflation rate, π its steady-state level and ψ the parameter that rules the inertia in the prices.

For the sake of parsimony, the optimal choices of price and vacancy are assumed to be symmetric among the firms. Hence, the first-order conditions

are given by:

$$q_t \mu_t = \kappa \quad (3.15)$$

$$\mu_t = mc_t - w_t^R + (1 - \rho^x) \mathbb{E}_t \beta_{t+1} \frac{\lambda_{t+1}}{\lambda_t} \mu_{t+1} \quad (3.16)$$

$$0 = (1 - \epsilon) + \epsilon mc_t - \psi \frac{\pi_t}{\pi} \left(\frac{\pi_t}{\pi} - 1 \right) + \mathbb{E}_t \beta_{t+1} \frac{\lambda_{t+1}}{\lambda_t} \psi \frac{\pi_{t+1}}{\pi} \left(\frac{\pi_{t+1}}{\pi} - 1 \right) \frac{y_{t+1}}{y_t} \quad (3.17)$$

where κ is the marginal cost of a recruitment, quantified by the Lagrangian multiplier μ_t . Further, plugging (3.16) into (3.15) refines the marginal cost/benefit of a job creation:

$$\frac{\kappa}{q_t} = mc_t - w_t^R + (1 - \rho^x) \mathbb{E}_t \beta_{t+1} \frac{\lambda_{t+1}}{\lambda_t} \frac{\kappa}{q_{t+1}} \quad (3.18)$$

Wage Setting

In equilibrium, the sum of the marginal values of filled jobs, for firms μ_t and workers φ_t , creates a surplus. It is thus greater than the sum of the marginal benefits of being unemployed or having vacant jobs. Consequently, the real wage is determined using a Nash bargaining process in which that surplus is shared between workers and employers. The optimal wage is valued by the solution of the following problem:

$$\max_{w_t} \left(\frac{\varphi_t}{\lambda_t} \right)^{1-\zeta} \mu_t^\zeta \quad (3.19)$$

The optimality condition of this problem is:

$$\zeta \frac{\varphi_t}{\lambda_t} = (1 - \zeta) \mu_t \quad (3.20)$$

where $\xi \in [0, 1]$ and $(1 - \xi)$ refer to the bargaining powers of the firms and workers, respectively. Using the definitions of φ_t in the equation (3.8) and μ_t in (3.15), the definition of the real wage writes:

$$w_t^R = (1 - \xi) \left(mc_t + \mathbb{E}_t \beta_{t+1} \frac{\lambda_{t+1}}{\lambda_t} (1 - \rho^x) \kappa \theta_{t+1} \right) + \xi \left(b + \frac{\ell}{\lambda_t} \right) \quad (3.21)$$

The important implication of the search and matching frictions is that the real wage is a weighted sum between the contribution of workers to the product (productivity and the amortization of the vacancy costs) and the *outside option* (UB and leisure) of being unemployed.

3.2.5 The Monetary and Fiscal Authorities

The notional interest rate i_t^* is adjusted according to a Taylor rule, as follows:¹³

$$1 + i_t^* = \left(\frac{\pi}{\beta} \left(\frac{\pi_t}{\pi} \right)^{\rho_\pi} \left(\frac{n_t}{n} \right)^{\rho_n} \right) \quad (3.22)$$

The fiscal authority finances the unemployment allowances $b(1 - n_t)$ by collecting lump-sum taxes T_t and issuing domestic bonds d_t . Accordingly, the balanced budget satisfies the following relation:

$$d_t + b(1 - n_t) = (1 + i_{t-1})d_{t-1} + T_t \quad (3.23)$$

3.2.6 Market Clearing

The aggregation of the individual profits Π_t yields to:

$$\Pi_t = p_t y_t - n_t w_t - p_t y_t \Gamma_t^\pi \quad (3.24)$$

¹³The implementation of the ZLB and the SCFG are explained below.

The combination of the fiscal budget (3.23) together with the households' budget constraint (3.7) and the aggregation of profits (3.24) determines the following aggregate resource constraint:

$$y_t \left[1 - \frac{\psi}{2} \left(\frac{\pi_t}{\pi} - 1 \right)^2 \right] = c_t + \kappa v_t \quad (3.25)$$

3.2.7 State-contingent Forward Guidance

The effective interest rate i_t is subject to a regime-switching system. Let $r_t = \{1, \dots, N\}$ be the definition of the different regimes. In this model, the economy switches between three specific monetary regimes: a *normal* regime ($r_t = 1$), and ZLB episodes, without ($r_t = 2$) and with SCFG ($r_t = 3$).

In a *normal* regime, the notional interest rate i_t^* and the effective nominal interest rate i_t , always coincide. In any ZLB situation, the effective interest rate remains constant as long as the level of the notional interest rate is negative. The beginning of a SCFG regime ($r_t = 3$) is always conditional in the occurrence of a ZLB regime ($r_t = 2$) in the previous period. In addition, it is function of an entry unemployment threshold. However, the end of SCFG does not depend of a ZLB regime, but only on an exit unemployment threshold. The regime-switching system writes:¹⁴

$$r_t = \begin{cases} 1 & \text{if } i_t^* > 0 & \text{and } u_t < \underline{u}(r_{t-1}) \\ 2 & \text{if } i_t^* \leq 0 & \text{and } u_t < \underline{u}(r_{t-1}) \\ 3 & \text{if } u_t \geq \underline{u}(r_{t-1}) & \text{and } r_{t-1} > 1 \end{cases} \quad (3.26)$$

Having defined what is meant by SCFG, the following lines detail the range of situations. In any case, a positive deviation in the level of the notional

¹⁴This system is borrowed from the technical related study Albertini & Moyen (2019).

rate always leads to the *normal* regime ($r_t = 1$). It is straightforward, but in this situation, the ZLB in the effective rate cannot apply. Conversely, as later experienced, a large and positive discount factor shock induces ZLB. In this situation, the level of the notional rate deviates below zero. The economy stays in a ZLB, as long as the notional rate reverts to a positive level. The endogenous switching between a ZLB without ($r_t = 2$) and with SCFG ($r_t = 3$) is determined by further conditions. The economy enters a SCFG regime ($r_t = 3$), if and only if, the unemployment crosses the entry threshold $\underline{u}(r_{t-1})$ and that the ZLB regime ($r_t = 2$) applied in the previous period $t - 1$. Finally, the return to the *normal* regime ($r_t = 1$) is triggered by the exit unemployment threshold, but is not function of the ZLB regime ($r_t = 2$). The economy is entirely absorbed in SCFG regime until the exit threshold is crossed. As such, this regime is fully credible.

In addition, the regime-switching system allows for the effective rate to differ from the Taylor rule out of the ZLB regimes. In other words, a ZLB can perfectly applies even if the notional interest rate had returned to a positive level. This elegant mechanism is important to fit the dynamic in the Federal Funds rate in late 2014. The effective interest rate is defined, in each of the regime, by:

$$i_t(r_t) = \begin{cases} i_t^* & \text{if } r_t = 1 \\ 0 & \text{if } r_t = 2 \\ 0 & \text{if } r_t = 3 \end{cases} \quad (3.27)$$

3.3 Quantitative Evaluations

3.3.1 Calibration

The calibration initializes the model on the bases of three targets of the US quarterly data: (i) the average unemployment rate, (ii) the average of the separation rate and (iii) the frequency of the ZLB episodes in the US history.¹⁵

Along the estimates of Fernández-Villaverde et al. (2015), the elasticity of substitution between goods is calibrated to, $\epsilon = 6$, which induces a steady-state markup of 20%. The equilibrium value of the discount factor shock is, $\beta = 0.996$.¹⁶ Together with, the monetary weights on inflation, $\rho_\pi = 1.50$, and employment, $\rho_n = 0.50$, the parameters related to the shock, $\rho_\beta = 0.85$ and $\sigma_\beta = 0.0025$, are borrowed from the estimates of Gust et al. (2012). These parameters help capturing the frequency of the ZLB episodes in the last century; that is, roughly 5% of time. The steady-state inflation, $\pi = 1.005$, fits the long-run inflation target of the FOMC, *i.e.*, 2%. All of these values imply an annual notional rate of, $i^* = 3.66\%$.

The steady-state unemployment rate is targeted to $u = 6\%$, matching the average of the civilian unemployment rate between 2000Q1-2019Q1. In line with Andolfatto (1996), the matching friction γ is adjusted so that the probability of filling a job is, $q = 0.9$. Following Pissarides & Petrongolo (2001), $\zeta = 0.50$, so that the bargaining power is even between workers and firms. In data, the average BLS monthly probability of job separation is 3.1% between 2000M1-2019M1. The latter implies that $\rho^x = 10\%$ in a quarterly frequency.

¹⁵Details about the computation of the steady-state values are given in the appendix.

¹⁶The discount factor shocks is in the form of an AR(1) process, such as: $\log \beta_t = \rho_\beta \log \beta_{t-1} + (1 - \rho_\beta) \log \beta + \sigma_\beta \epsilon_t^\beta$

Last, the vacancy costs are calibrated in respect with $\kappa v \simeq 0.01y$. The values of the parameters for the benchmark model are summarized in the Table 3.1.

Table 3.1: Parameters of the benchmark model

Parameters	Symbols	Values	Source
Discount factor	β	0.996	1.6% annual risk-free rate
Elast. of subst. between goods	ϵ	6.00	Target Markup of 20 percent
Steady state inflation	π	1.005	Target 2 percent net annual rate
Autocorr. coefficient	ρ_β	0.85	Albertini & Poirier (2015)
Std. of β_t	σ_β	0.0025	Target 5 percent at the ZLB
Vacancy posting costs	κ	0.09	Target $\kappa v \simeq 0.01y$
Leisure and home production	ℓ	0.277	Deduced from job creation
Matching frictions	γ	2.77	Target $q \simeq 0.9$
Separation rate	ρ^x	0.10	Target 0.032 monthly (BLS)
Worker bargaining power	ξ	0.50	Pissarides & Petrongolo (2001)
Price adjustment	ψ	210	Target observed deflation
Response to inflation	ρ_π	1.50	Gust et al. (2012)
Response to employment	ρ_n	0.50	Gust et al. (2012)

The unemployment thresholds $\underline{u}(r)$ for the regime-switching system are:

$$\underline{u}(r) = \{1, 0.075, 0.065\} \quad (3.28)$$

In the normal regime ($r_t = 1$), the value of the threshold is unimportant; hence fixed to a maximum value (1). The switching from ZLB ($r_t = 2$) to SCFG ($r_t = 3$) is calibrated to an unemployment rate of 7.5%. In the purpose of gauging the impacts of the FOMC statement of December 2012, the exit unemployment threshold is fixed to 6.5%.

3.3.2 State-contingent Forward Guidance Experiments

Impulse Response Analysis

In these experiments, the economy is hit by a large and positive discount factor shock. This unexpected shock is calibrated to produce a long ZLB episode, and to target the observed 10% deviations in the unemployment rate in 2009Q4 (in the regime of ZLB without SCFG). Three monetary specifications are considered to disentangle the effects of the ZLB itself, and those of SCFG: Firstly, agents always refer to the Taylor rule (equation (3.22)). Secondly, the effective interest rate is bounded by zero. In this specification, the ZLB binds as long as the level of the notional interest rate is negative. The monetary *normalization* occurs when the notional rate reverts to a positive level. Thirdly, the benchmark monetary specification is characterized by the above regime-switching system (3.26). Figure 3.1 depicts the impulse response functions in each of these specifications.

The green dashed lines display key responses of the first specification. In this specification, the effective and the notional rates always coincide and are dictated by a Taylor rule. In other words, no lower limit is imposed to the effective rate. As such, the contraction in the economy is dampened by the accommodation induced by the big decrease in the real interest rate. The unemployment rate peaks to 7.75% when reaching 10% in the ZLB without SCFG (blue solid lines with squares).

Under ZLB without SCFG, it is straightforward that the lower bound creates a large drop in the price level and causes a sizable unemployment rate. Nevertheless, it has been shown in Albertini & Poirier (2015) that a ZLB can indirectly dampen a deflation spiral by refining the expectations of agents. In

their model, a ZLB triggers an endogenous rise in UB. The implied negative elasticity between the unemployment duration and the UB, is nonlinear in the duration of the ZLB. Hence, the longer the duration, the lower the elasticity. In the present model, UB are assumed to be constant, which keeps that channel muted. As such, our model highlights the role of the shifts in expectations induced by a SCFG. Intuitively, the endogeneization of UB would reinforce the exposed results.

In the presence of monetary/fiscal stimulus, the effects of ZLB without SCFG must be understood as follows. On the one hand, the *standard* effects of ZLB cause a strong rise in the unemployment rate. On the supply side, the high discount rate squeezes the marginal value of a new hire to a firm.¹⁷ In addition, profits are eroded by deflation. Hence, firms are discouraged to employ workers. As a result, firms cut vacancies and negotiate wages downwards. On the demand side, workers suffer from the pressure in wages. Thus, the outside option becomes more attractive; therefore, households are less likely to search for a job. Overall, the output falls.

On the other hand, *expectation* effects can dampen the downturn of the labor market; the future actions of agents being contemporaneously changed by monetary/fiscal policy incentives. As such, the expectations in inflation rise. On the demand side, this inflationary pressure reinforces the bargaining power of workers (the right-hand side of equation (3.21)). Hence, workers claim for higher wages. On the supply side, the reduction of the real interest rate softens the hiring costs. In addition, the anticipation of the reduction in the relative prices causes a rise in production today. As a result, firms open vacancies and

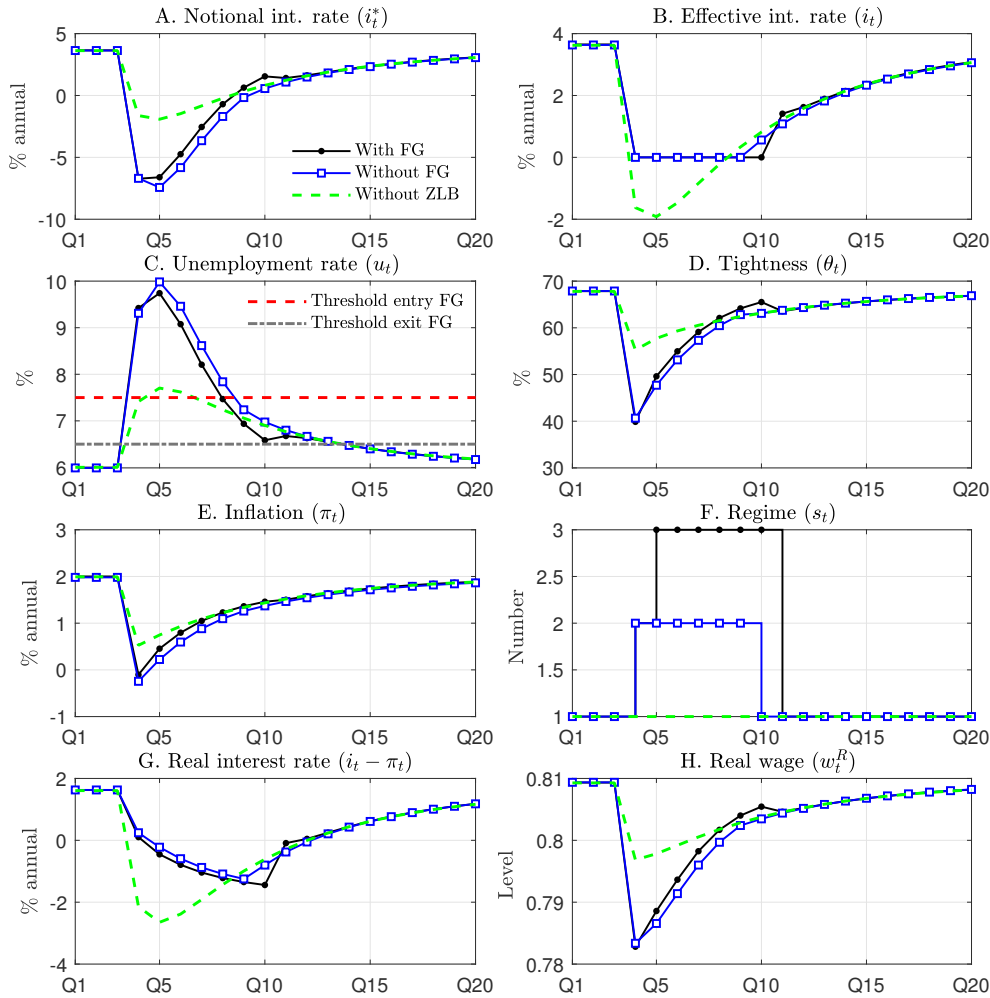
¹⁷See Hall (2017) as a supporting evidence on the link between discount rate and unemployment.

readjust wages. Thus, the economy receives a stimulus.

Here, in a ZLB without SCFG, the *expectation* effects are muted due to the absence of a reversal fiscal stimulus. However, a SCFG is likely to correct the expectations in inflation. The responses of the regime-switching system (black lines with dots) are consistent with this statement. Under SCFG, the ZLB is extended by a quarter. The effective rate stays longer to zero because of the commitment to the exit unemployment threshold. Meanwhile, the regime already switched to a *normalization* ($r_t = 1$). In that respect, Werning (2011) demonstrates that a credible extension of a ZLB, longer than advocated by the economic fundamentals (e.g. required by a Taylor rule), stimulates the economy. As exposed above, by assumption, the SCFG regime is perfectly credible. It is thus reasonable to assert that the observed stimulus is entirely explained by the refinement in the expectations of agents. Indeed, the only difference between a ZLB with or without SCFG, is the commitment to a level of unemployment. Under these circumstances, this promise to keep a zero nominal has apparent gains and accelerates the rebound of the economy (BHW find similar results). After ten quarters, the unemployment rate even lies below the economy without ZLB, by 0.5 *p.p.*

However, these benefits are softened by the transition to the *normalization*; *i.e.*, the regime-switching from SCFG to the standard Taylor rule ($r_t = 3 \rightarrow r_t = 1$). The surge in inflation is immediatly counterbalanced by the penalization of the rule. The latter provokes an upward readjustment of the real interest rate; hence, a rise in the unemployment rate. There are good reasons to believe that this reversal effect is dependent on the value of the thresholds. Thus, alternative entry and exit thresholds are experienced. Figures C.1 and C.2 in

Figure 3.1: Impacts of SCFG and ZLB



the appendix depict the results.¹⁸

The alternative exit scenarios are: A lower exit threshold, where $\underline{u}_3(r) = 6.3\%$, and a higher threshold $\underline{u}_3(r) = 6.8\%$. Apparent in the Figure C.1, the lower the exit threshold, the greater the effectiveness of SCFG. Conversely,

¹⁸Alternative exit thresholds must range in a small interval to avoid indeterminacy issues.

when the exit threshold is too high, the responses are similar to a regime of ZLB without SCFG. Henceforth, the entry and exit thresholds are calibrated to the same level $\underline{u}_2(r) = \underline{u}_3(r) = 6.5\%$. Lowering the entry threshold fosters the effects of SCFG (see the output gains in the Figure C.3 in the appendix).

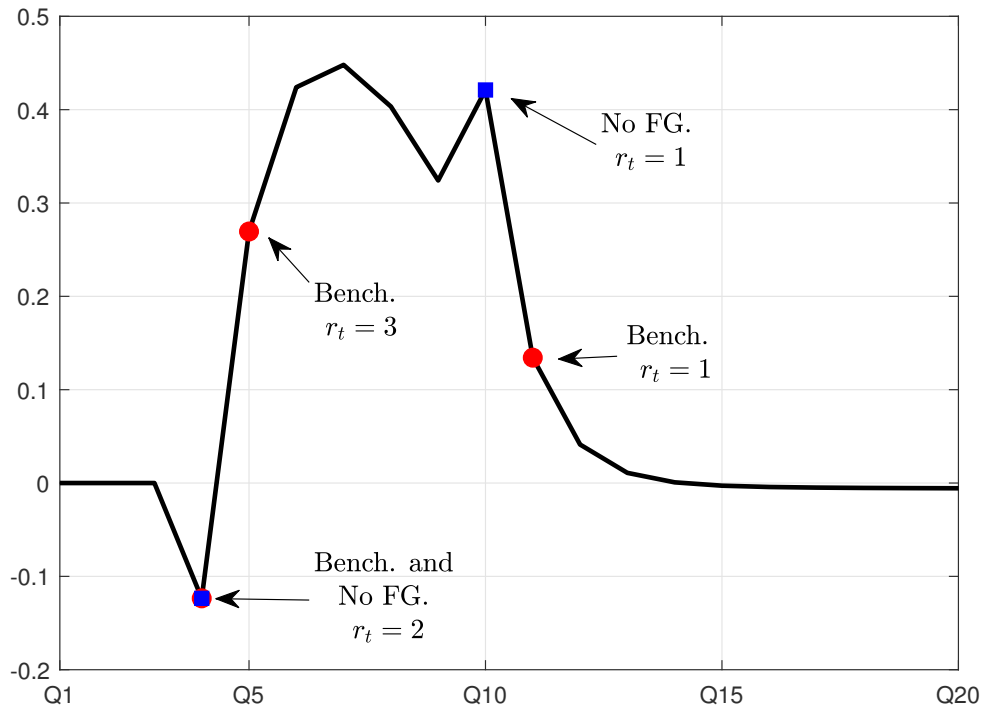
To sum up, a STFG stimulates the economy by refining the expectations in inflation. Phrased differently, the brighter the future, the higher the effectiveness. The inflationary spiral dampens the disruption in the labor market, under the condition that the exit unemployment threshold is low enough. Otherwise, the endogenous monetary *normalization* counterbalances the benefits of SCFG. As an interesting aside, the Figure 3.2 display the benefits of SCFG in terms of output gains:

$$\left(\frac{y_t^{FG} - y_t^{NFG}}{y_t^{NFG}} \right) \times 100 \quad (3.29)$$

which is the percent change between the output path when the SCFG is active (y_t^{FG}) and the output path, absent of SCFG (y_t^{NFG}).¹⁹ The benefits of the SCFG itself, cumulate between Q5-Q10. The effects are thus pretty weak, the difference being, in the end of period, of 0.15 *p.p.* In contrary, a great proportion of the benefits arises from the anticipation to enter a SCFG (see Q4). In that respect, the output gains have a kink in Q9, which highlights, again, the crucial role of the anticipation. In this case, the effects on output are lowered due to the upcoming switch to the monetary *normalization* ($r_t = 3 \rightarrow r_t = 1$). The kink tends to disappear in the increase of the exit threshold. Nevertheless, in that situation, the effective and the notional rates tend to coincide. As such, the impacts of SCFG converge to zero.

¹⁹These deviations are computed under the benchmark parametrization.

Figure 3.2: Effects of SCFG on output



The percent change in output is computed according to the equation (3.29). The blue squares refer to the second monetary specification with ZLB, absent of SCFG. In that specification, the economy is in ZLB between Q4 and Q11. The red dots characterizes the regime-switching in the benchmark specification. Here, ($r_t = 1$) in Q1-Q3 and Q11-Q20, ($r_t = 2$) in Q4 and ($r_t = 3$) in Q5-Q11.

Counterfactual analysis

The counterfactual experiment is based on the observed unemployment path between 1995Q1 and 2018Q1. This exercise questions what would have been the economy in the absence of: (i) the ZLB in the effective interest rate (ii) the SCFG. The methodology proceeds as follows. Using a Newton-Raphson algorithm, the path of the discount factor shock fits the time series of the unemployment rate. Subsequently, the path of the remaining endogenous variables

are simulated under the alternatives ((i),(ii)).

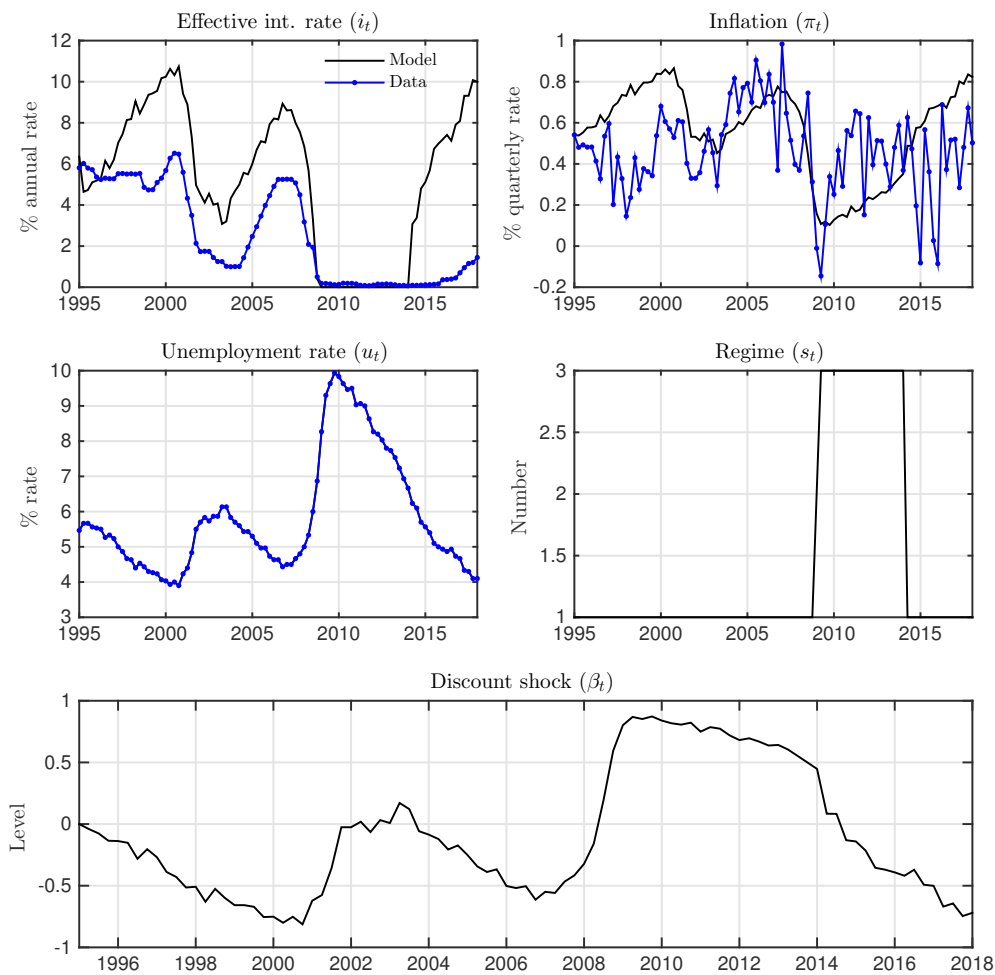
Simulations are depicted in the Figure 3.3. In most periods, the discount shock lies below zero, before abruptly and persistently increasing during the Great Recession. In the study of Hall (2017), measures of the discount rate have similar patterns. Despite the relative simplicity of the model, the fit between the simulations and data is fairly good. Indeed, the magnitude of the shifts in inflation is well matched; particularly the recent episode of disinflation. Hence, the inclusion of the employment rate in the Taylor rule, is an interesting modeling assumption. In general, the use of discount factor shocks has the main advantage to cause large unemployment swings and a reasonable deflation. As such, the model is able to reproduce the fast entry into the ZLB of the Great Recession, while capturing the observed disinflation path.

Figure 3.4 shows the simulations under the alternatives. The first scenario (blue dashed lines) overrides the presence of ZLB and SCFG. In line with the impulse responses experiments, a negative effective rate is highly beneficial for the economy. Interestingly, the path of the notional rate mimics estimates of shadow interest rates, see *e.g.* Wu & Xia (2016). Such monetary accommodation, would have lowered the unemployment rate by almost 2 *p.p.* in 2009Q4. In the second scenario (red solid lines with dots), the model enters in ZLB in 2009Q1, in line with the observed effective Federal Funds rate. Being completely endogenous, this similarity highlights the ability of the model to be realistic. Free of SCFG, the simulations overestimate the unemployment peak by 0.45 *p.p.*

Nevertheless, the simulations in the benchmark model fail to match the actual exit of the ZLB in 2015Q4 (see Figure 3.3). In addition, the effective rate presents a huge liftoff in the end of SCFG, which is off the beam compared to the observed sluggishness in the official nominal interest rate in late 2014. To

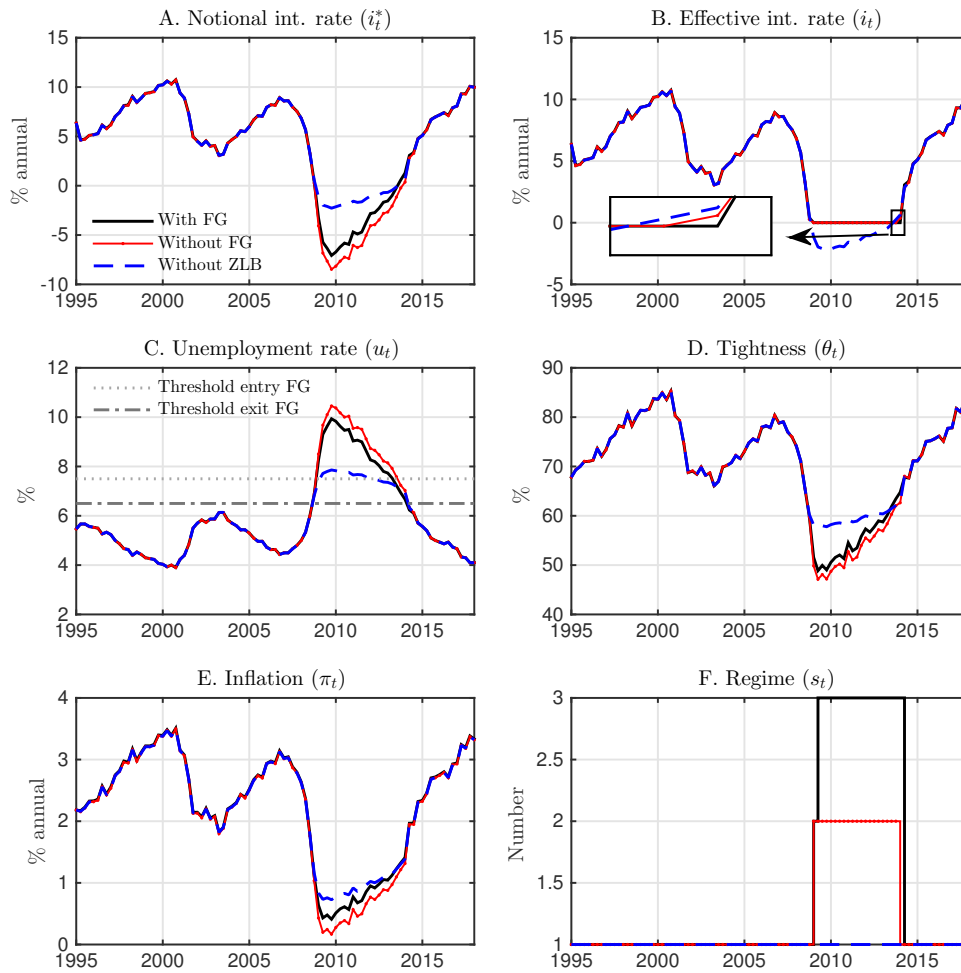
tackle this issue, the model would probably benefit of the inclusion of uncertainty in FG.

Figure 3.3: Extracted discount factor shocks



Shock to the discount factor is extracted using a Newton-Raphson method such that the model fits the observed unemployment rate over the period 1995Q1-2018Q1.

Figure 3.4: Counterfactual analysis



Using the smoothed shock series the model is simulated under two alternative scenarios: A counterfactual with ZLB, but without SCFG (red solid lines with dots). A counterfactual with ZLB and SCFG (black solid lines). The entry threshold is 7.5%. The exit threshold is 6.5%.

3.3.3 Sensitivity analysis

All of the results exposed above, mainly rely on: The response of inflation, the adjustment of the labor flows and the degree of labor market frictions. Hence, the results are confronted to changes in key parameters to isolate the main driver of the SCFG. The changes are summarized in the Table 3.2. The alterna-

Table 3.2: Alternative calibration

Parameters	Symbols	Bench.	C2	C3	C4
Vacancy posting costs	κ	0.09	0.09	0.19	0.14
Leisure and home production	ℓ	0.28	0.28	0.25	0.24
Separation rate	ρ^x	0.10	0.10	0.05	0.10
Price adjustment	ψ	210	120	210	210
Job finding rate	f	0.61	0.61	0.44	0.61

C2: Lower price adjustment cost. C3: Slower steady-state labor flows. C4: Higher vacancy costs.

tive calibration C2 involves a substantially lower adjustment cost to the prices ($\psi = 120$). In calibration C3, the separation rate is diminished to $\rho^x = 0.05$. Given that, the vacancy costs κ must increase while the job finding rate, f , must fall to 44% to ensure equilibrium. As a consequence, the long-run values of labor flows around which the economy fluctuates, are significantly lower. As such, the adjustments in the labor market are more sluggish. In calibration C4, the increase in the vacancy costs reinforces the labor market frictions. Thus, the utility of leisure, ℓ , must diminish to balance the wage equations.

Figures C.4 to C.6 in the appendix, display the impulse responses under the alternatives.²⁰ Figure C.7 complements the analysis by computing the output gains in each of the alternatives. In C2, the lower price adjustment costs induce

²⁰The discount factor shock has the benchmark calibration.

a much larger deflation, which in turn amplifies the impacts of SCFG. In this case, the benefits in terms of output are more than doubled. In C3, the effects of SCFG are also accentuated by the higher labor market rigidities. On the contrary, in C4, the impacts are narrowed by the rise in the vacancy costs.

In all of the alternatives, the duration of ZLB is no longer extended by a SCFG, compared to a specification of ZLB without SCFG. However, the larger the deflation, the greater the effectiveness. Hence, the main conclusion is that: the SCFG is mainly channeled through the *inflation expectations* channel of SCFG.

3.4 Conclusion

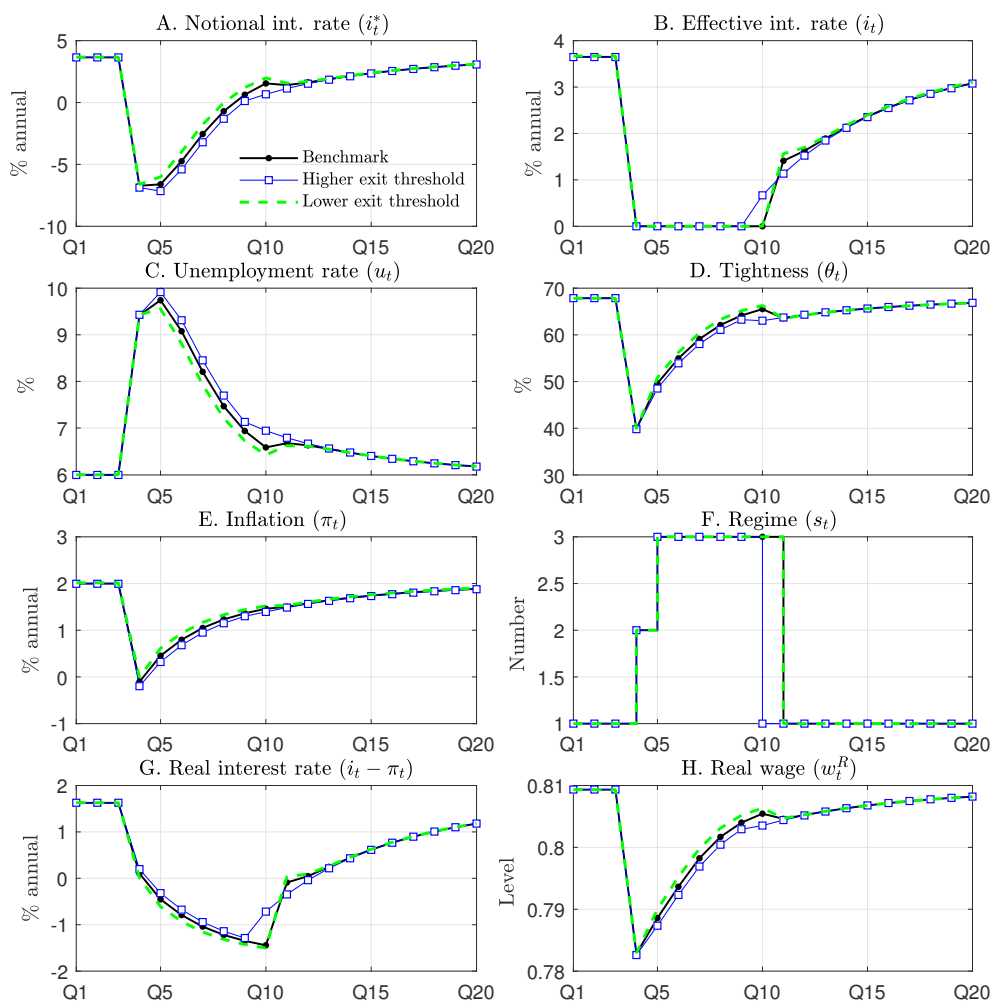
In this paper, the SCFG is analyzed by the use of a New-Keynesian DSGE model with search and matching frictions in the labor market. The model is subject to a monetary regime-switching system in which the SCFG is triggered by unemployment thresholds. The experiments analyze and gauge its effects in the economy.

The results suggest that a SCFG is mainly transmitted by the *inflation expectations* channel of SCFG. In other words, the stimulus to the economy today are caused by the conviction in future inflation. The anticipation of the entry in a SCFG regime takes the lion's share of the impacts. Conversely, the anticipation of the exit reduces the cumulated benefits. In addition, the fast liftoff from zero to a *normal* monetary regime further dampens the recovery. Under an accurate calibration, which is in fair fit with data, the SCFG is found to have a weak impact in the economy. Hence, these results open avenues on the method to tackle the theoretical *forward guidance puzzle*.

Appendix C

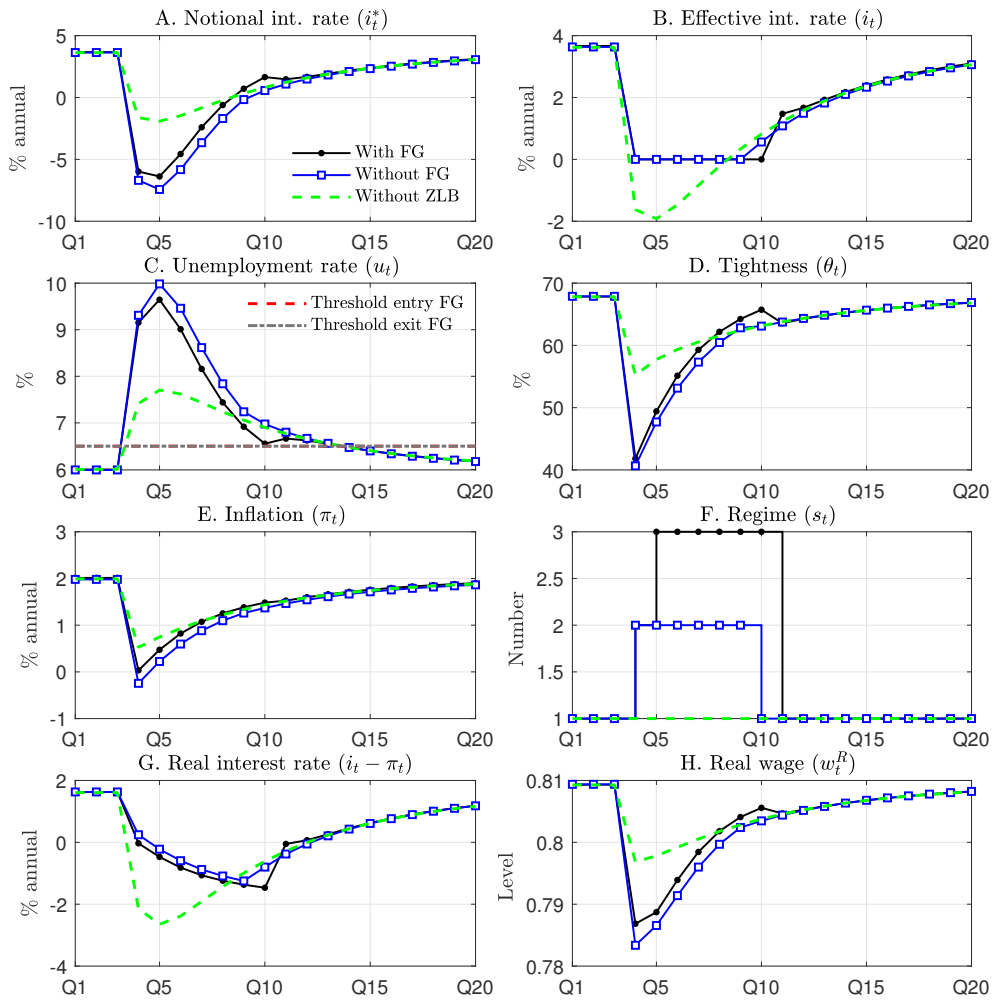
C.1 Additional Figures

Figure C.1: Alternative exit thresholds



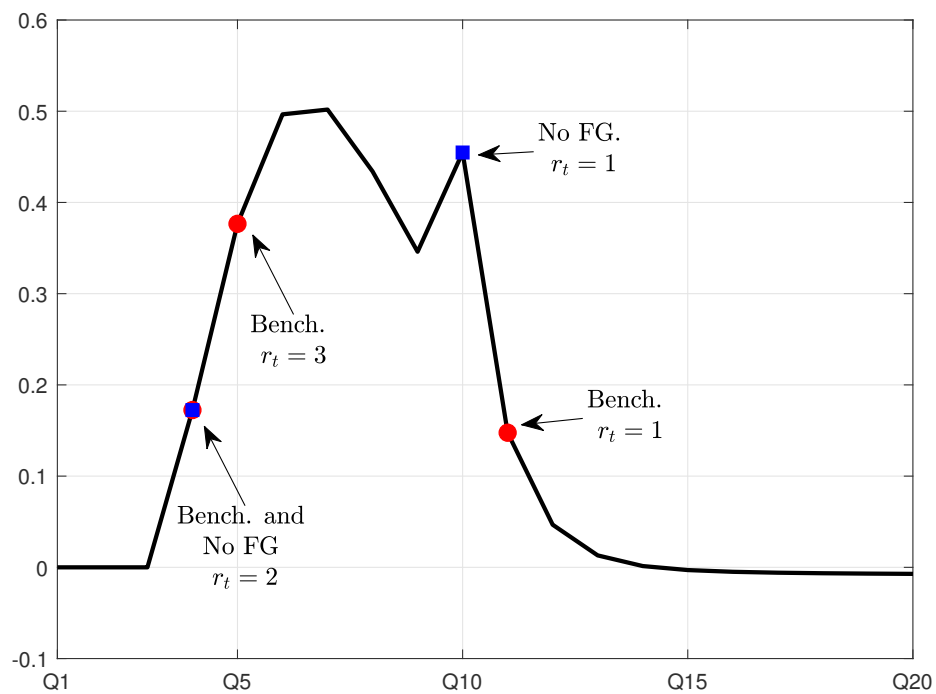
Benchmark: $\underline{u} = 6.5\%$. Higher exit threshold: $\underline{u} = 6.8\%$. Lower exit threshold: $\underline{u} = 6.3\%$.

Figure C.2: Alternative entry threshold



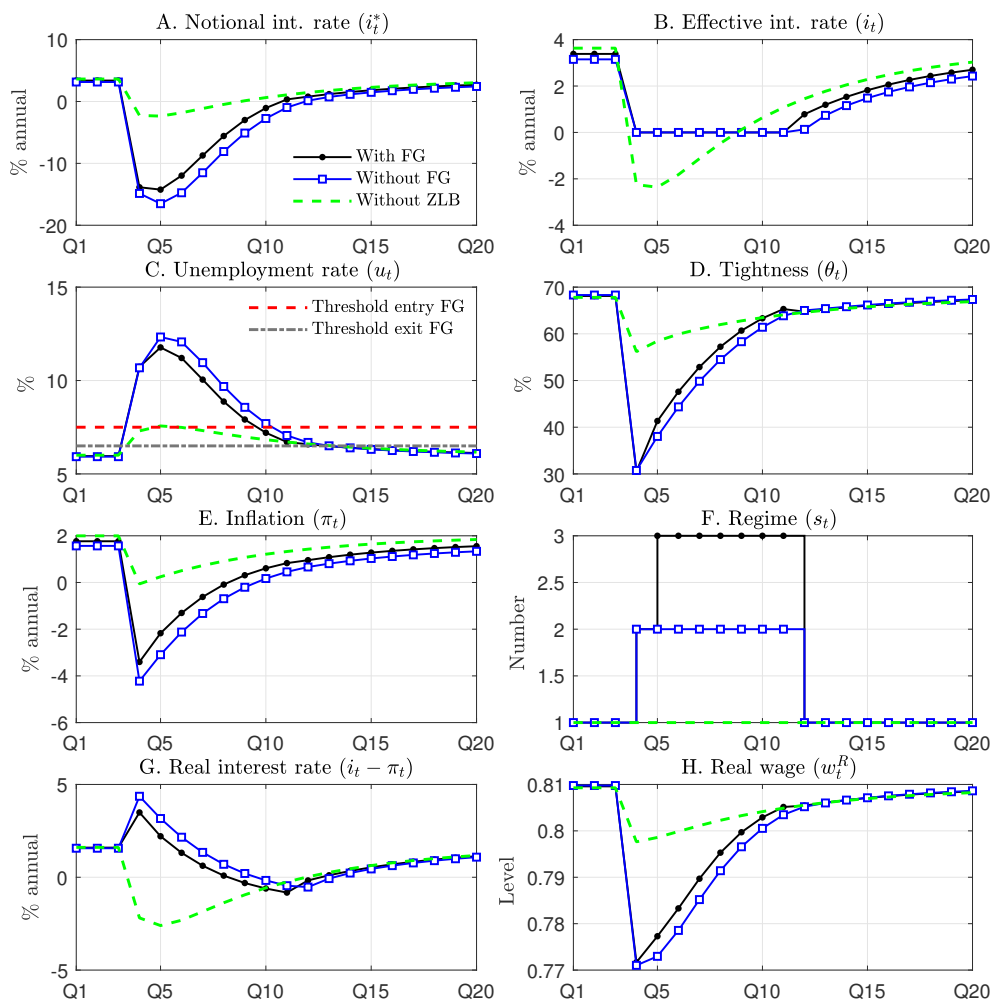
Entry and exit thresholds are the same $\underline{u}(2) = \underline{u}(3) = 6.5\%$.

Figure C.3: Effects of SCFG on output, alternative entry threshold



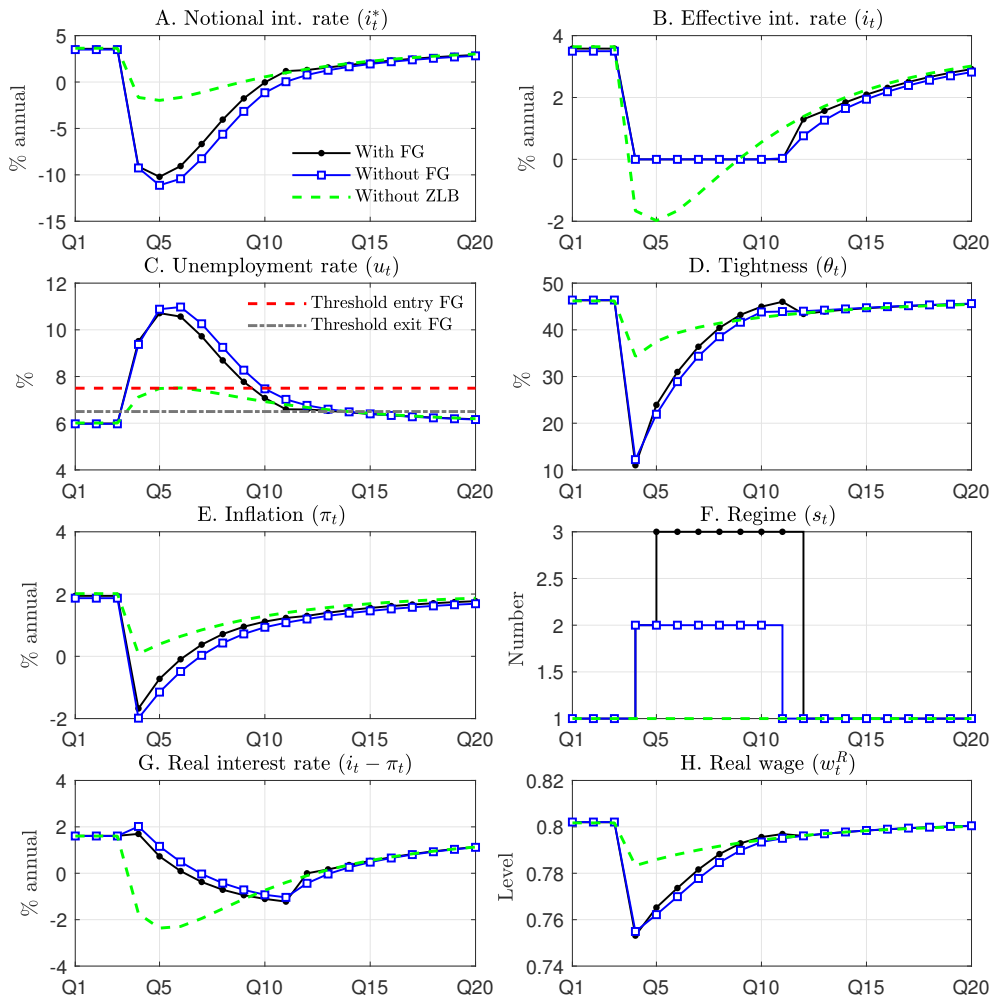
Entry and exit thresholds are the same, $\underline{u}(2) = \underline{u}(3) = 6.5\%$.

Figure C.4: Alternative calibration, C2



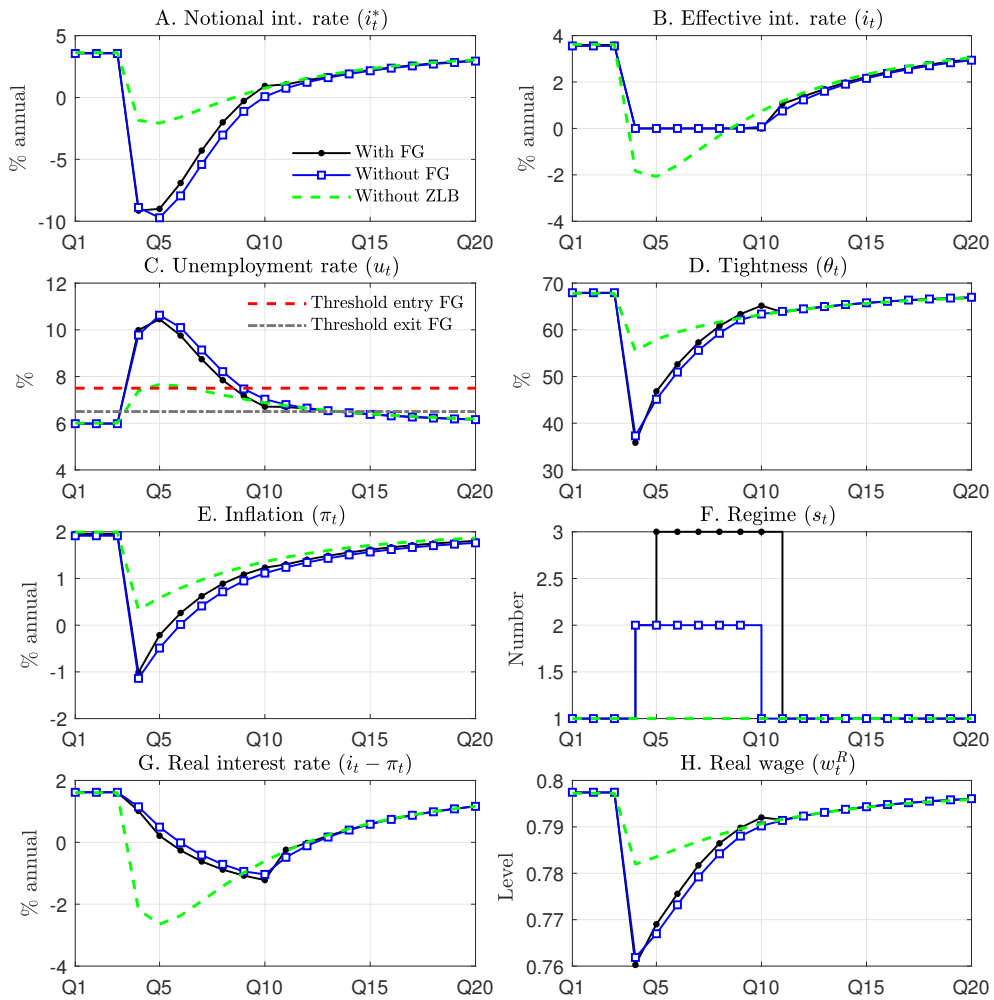
C2: lower price adjustment cost ($\psi = 120$).

Figure C.5: Alternative calibration, C3



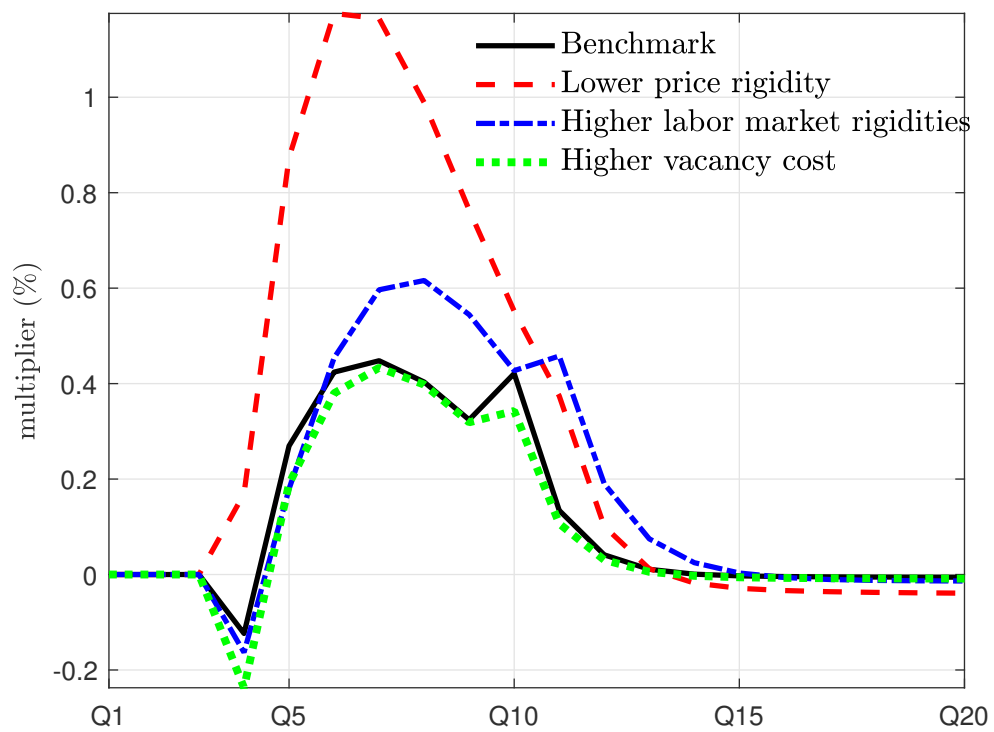
C3: Rigid labor market ($\rho^x = 0.05$ and $f = 0.44$).

Figure C.6: Alternative calibration, C4



C4: lower price adjustment cost ($\kappa = 0.14$).

Figure C.7: Effects of SCFG on output, alternative calibrations



C.2 Technical Appendix

C.2.1 Equations

Dynamic model

- Euler equations:

$$\lambda_t = c_t^{-1} \quad (\text{C.1})$$

$$\lambda_t = (1 + i_t) E_t \bar{\beta} \exp(\beta_{t+1}) \frac{\lambda_{t+1}}{\pi_{t+1}} \quad (\text{C.2})$$

- Law of motion of employment:

$$n_t = (1 - \rho^x) n_{t-1} + m_t \quad (\text{C.3})$$

- Hirings:

$$m_t = \left(s_t^{-\gamma} + v_t^{-\gamma} \right)^{-\frac{1}{\gamma}} \quad (\text{C.4})$$

- Job seekers:

$$s_t = 1 - (1 - \rho^x) n_{t-1} \quad (\text{C.5})$$

- Unemployment rate:

$$u_t = 1 - n_t \quad (\text{C.6})$$

- Job finding rate:

$$f_t = \frac{m_t}{s_t} \quad (\text{C.7})$$

- Job filling rate:

$$q_t = \frac{m_t}{v_t} \quad (\text{C.8})$$

- Tightness:

$$\theta_t = \frac{f_t}{q_t} \quad (\text{C.9})$$

- Free entry condition

$$\frac{\kappa}{q_t} = mc_t - w_t^R + (1 - \rho^x) E_t \bar{\beta} \exp(\beta_{t+1}) \frac{\lambda_{t+1}}{\lambda_t} \frac{\kappa}{q_{t+1}} \quad (\text{C.10})$$

- Output:

$$y_t = n_t \quad (\text{C.11})$$

- Real wage:

$$w_t^R = \xi \left(b + \frac{\ell}{\lambda_t} \right) + (1 - \xi) \left(mc_t + E_t \bar{\beta} \exp(\beta_{t+1}) (1 - \rho^x) \frac{\lambda_{t+1}}{\lambda_t} \kappa \theta_{t+1} \right) \quad (\text{C.12})$$

- New Keynesian Phillips Curve (NKPC):

$$\begin{aligned} 0 &= (1 - \epsilon) + \epsilon mc_t - \psi \frac{\pi_t}{\pi} \left(\frac{\pi_t}{\pi} - 1 \right) \\ &+ E_t \bar{\beta} \exp(\beta_{t+1}) \frac{\lambda_{t+1}}{\lambda_t} \psi \frac{\pi_{t+1}}{\pi} \left(\frac{\pi_{t+1}}{\pi} - 1 \right) \frac{y_{t+1}}{y_t} \end{aligned} \quad (\text{C.13})$$

- Notional interest rate:

$$1 + i_t^* = \frac{\pi}{\beta} \left(\frac{\pi_t}{\pi} \right)^{\rho_\pi} \left(\frac{n_t}{n} \right)^{\rho_n} \quad (\text{C.14})$$

- Market clearing:

$$y_t \left[1 - \frac{\psi}{2} \left(\frac{\pi_t}{\pi} - 1 \right)^2 \right] = c_t + \kappa v_t \quad (\text{C.15})$$

- Shocks:

$$\beta_{t+1} = \rho_\beta \beta_t + \sigma_\beta \varepsilon_{t+1}^\beta \quad (\text{C.16})$$

- Regime:

$$r_t = \begin{cases} 1 & \text{if } i_t^* > 0 & \text{and } u_t < \underline{u}(r_{t-1}) \\ 2 & \text{if } i_t^* \leq 0 & \text{and } u_t < \underline{u}(r_{t-1}) \\ 3 & \text{if } u_t \geq \underline{u}(r_{t-1}) & \text{and } r_{t-1} > 1 \end{cases} \quad (\text{C.17})$$

- Effective nominal interest rate:

$$i_t(r_t) = \begin{cases} i_t^* & \text{if } r_t = 1 \\ 0 & \text{if } r_t = 2 \\ 0 & \text{if } r_t = 3 \end{cases}$$

Variables and parameters

- The variables are:

$$\lambda_t, \pi_t, n_t, m_t, s_t, u_t, f_t, q_t, \theta_t, v_t, y_t, w_t^R, mc_t, i_t^*, c_t, \beta_t, r_t, i_t$$

- The endogenous state variables is:

$$n_t$$

- The exogenous state variables is:

$$\beta_t$$

- The control variables are:

$$\lambda_t, \pi_t, m_t, s_t, u_t, f_t, q_t, \theta_t, v_t, y_t, w_t^R, mc_t, i_t^*, c_t, r_t, i_t$$

- The parameters are:

$$\bar{\beta}, \ell, \epsilon, \bar{\pi}, \psi, \rho_i, \rho_n, \rho_\pi, b, \xi, \kappa, \gamma, \rho^x \bar{u}(r), \rho_\beta, \sigma_\beta$$

Steady state model

- Euler equations:

$$\lambda = c^{-1} \quad (\text{C.18})$$

$$(1+i) = \frac{\pi}{\bar{\beta}} \quad (\text{C.19})$$

- Law of motion of employment:

$$n = \frac{fs}{\rho^x} \quad (\text{C.20})$$

- Hirings:

$$m = (s^{-\gamma} + v^{-\gamma})^{-\frac{1}{\gamma}} \quad (\text{C.21})$$

- Job seekers:

$$s = 1 - (1 - \rho^x)n \quad (\text{C.22})$$

- Unemployment rate:

$$u = 1 - n \quad (\text{C.23})$$

- Job finding rate:

$$f = \frac{m}{s} \quad (\text{C.24})$$

- Job filling rate:

$$q = \frac{m}{v} \quad (\text{C.25})$$

- Tightness:

$$\theta = \frac{f}{q} \quad (\text{C.26})$$

- Free entry condition

$$\frac{\kappa}{q} = mc - w^R + (1 - \rho^x)\beta\frac{\kappa}{q} \quad (\text{C.27})$$

- Output:

$$y = n \quad (\text{C.28})$$

- Real wage:

$$w^R = (1 - \zeta) (mc + (1 - \rho^x)\bar{\beta}\kappa\theta) + \zeta \left(b + \frac{\ell}{\lambda} \right) \quad (\text{C.29})$$

- New Keynesian Phillips Curve (NKPC):

$$mc = \frac{\epsilon - 1}{\epsilon} \quad (\text{C.30})$$

- Taylor rule:

$$1 + i = \frac{\pi}{\beta} \quad (\text{C.31})$$

- Market clearing:

$$y = c + \kappa v \quad (\text{C.32})$$

- Regime:

$$r = 1 \quad (\text{C.33})$$

- Effective nominal interest rate:

$$i(r) = i^*$$

Remark. *Solving the model with global approximation methods does not involve an approximation around the deterministic steady state. It is however calculated in order to use perturbation methods as an initial guess for the the global approximation. Under the perturbation method, kinks are not taken into account (ZLB and regimes). Then, we impose the regime to be in state 1 (normal regime).*

Calibration

With $\pi = 1.005$ and $\bar{\beta} = 0.996$

$$i = \frac{\pi}{\bar{\beta}} - 1 = 0.9\%$$

In the benchmark we set the steady-state level of employment to 0.94 and $\rho^x = 0.10$.

$$m = \rho^x n = 0.094$$

$$s = 1 - (1 - \rho^x)n = 0.154$$

$$f = \frac{m}{s} = 0.6104$$

$$y = n = 0.94$$

Since $q = 0.9$ then

$$\begin{aligned} v &= \frac{m}{q} = 0.1044 \\ \theta &= \frac{f}{q} = 0.6782 \end{aligned}$$

γ is set to balance the hiring function given $m = 0.094$. Then $\gamma = 2.778$. We set $\epsilon = 6$ and $mc = (\epsilon - 1)/\epsilon \simeq 0.83$. We impose $\xi = 0.5$, $\kappa v/y = 1\%$ and $b = 0.5$. The remaining parameters is ℓ obtain from the job creation condition given the wage rate definition:

$$\ell = \frac{\lambda}{\xi} \left((mc - b)\xi - \frac{\kappa}{q}(1 - \beta(1 - \rho^x)) - (1 - \xi)\beta(1 - \rho^x)\kappa\theta \right) = 0.277$$

The remaining steady state values are pin down as follow:

$$\begin{aligned} c &= y - \kappa v = 0.93 \\ \lambda &= \frac{1}{c} = 1.0746 \\ w &= (1 - \xi)(mc + \beta\kappa\theta(1 - \rho^x)) + \xi(b + \ell/\lambda) = 0.822 \end{aligned}$$

C.2.2 Solution method

In order to explicitly take into account the non-linearity induced by the ZLB and the RS, we use the projection method developed by Albertini and Moyen (2019). It consists of approximating the policy rules of the system previously described using Chebyshev polynomials. We consider a third-order Chebyshev polynomial over a simulated grid. Our strategy departs from conventional projection methods because we approximate as many policy functions as we have regimes: the ZLB, the SCFG and *normal times*. We merge them in the algorithm to compute the expectations. We first present some numerical technics that will

be helpful for the understanding of the general algorithm.

General formulation of the solution

The general solution of this family of discrete-time rational expectations models with RS writes:

$$E_t(R(\mathbf{x}_t, r_t, \mathbf{y}_t, \mathbf{x}_{t+1}, r_{t+1}, \mathbf{y}_{t+1})) = 0 \quad (\text{C.34})$$

- r_t is today's regime
- $\mathbf{x}_t = \{x_t, z_t\}$ is the vector of continuous state variables. It evolves according to:

$$\mathbf{x}_{t+1} = \Gamma(\mathbf{x}_t, \varepsilon_{t+1}, r_t, \mathbf{y}_t) \quad (\text{C.35})$$

- $\{y_t\}_{t=0}^{\infty}$ is the sequence of optimal decision rules (controls).
- z_t denotes the vector of stationary and exogenous shocks. It satisfies the following representation:

$$z_{t+1} = \Phi(z_t, \varepsilon_{t+1}) \quad \text{with } z_0 \text{ given and } \varepsilon_{t+1} \sim \mathcal{N}(\mu, \Sigma)$$

where ε_{t+1} is a vector of i.i.d. structural disturbances with mean vector μ and covariance matrix Σ .

- x_t is a vector of continuous endogenous state variables¹ that evolves according to:

¹We distinguish continuous endogenous state variables from the regime r_t because the latter as a discrete support and must then be treated differently.

$$x_{t+1} = h(x_t, z_t, r_t, y_t) \quad \text{with } x_0 \text{ given} \quad (\text{C.36})$$

- The law of motion of the regime r_t takes the following form:

$$r_t = \Psi(x_t, z_t, r_{t-1}) \quad \text{with } s_0 \text{ given} \quad (\text{C.37})$$

- $R(\cdot)$ is the residual function defined by the FOC and the evolution of the state variables.

The solution of the above problem consists in expressing the set of control variables y_t as a function of the state variables:

$$y_t = g(\mathbf{x}_t; r_t) \quad (\text{C.38})$$

In most DSGE models, the function $g(\cdot)$ cannot be computed analytically, so does the residual function $R(\cdot)$. The standard approach of projection method is to consider Chebyshev polynomials as approximation functions. However, the regime has a discrete support $\{1, 2, \dots, N\}$ which involves non-differentiable policy rules. Spectral methods based on smooth approximation functions are likely to generate inaccuracy of the approximation at the kinks and in the area of the state-space that is discontinuous. In order to circumvent this difficulty, we consider regime-dependent decision rules. We denote by $\Theta(r) = \{\theta_1(r), \dots, \theta_m(r)\}$ the set of coefficients of the approximation function in regime r . The approximation function for (C.38) is:

$$y_t = P(\mathbf{x}_t; \Theta(r_t)) \quad (\text{C.39})$$

The general idea of the projection method is to find the set of coefficients $\Theta(\cdot)$ minimizing the residual function. We make this step through OLS regression. Formally it writes:

$$\hat{\Theta}(r_t) = \arg \min_{\Theta(r_t)} \int_{\mathbf{x}_t} R(\mathbf{x}_t; \Theta(r_t))^2 d\mathbf{x}_t \quad (\text{C.40})$$

where $R(\mathbf{x}_t; \Theta(r_t))$ is the residual equation given that y_t is approximated by a polynomial function of the state variables.

Compact form model

In order to simplify the algorithm, we combine several equations. By doing so, the competitive equilibrium is described by two control variables, namely: θ_t and π_t . This section describes the steps to obtain such a compact form of the competitive equilibrium. For the sake of clarity we do not show entirely the replacements of the variables as the expression would be too heavy. Variables in blue are intermediary variables.

Let first define the expectation functions:

$$\Phi_t^1 = E_t \bar{\beta} \exp(\beta_{t+1}) \frac{\lambda_{t+1}}{\pi_{t+1}} \quad (\text{C.41})$$

$$\Phi_t^2 = E_t \bar{\beta} \exp(\beta_{t+1}) \lambda_{t+1} \psi \frac{\pi_{t+1}}{\pi} \left(\frac{\pi_{t+1}}{\pi} - 1 \right) y_{t+1} \quad (\text{C.42})$$

$$\Phi_t^3 = E_t \bar{\beta} \exp(\beta_{t+1}) \lambda_{t+1} (1 - \zeta) (1 - \rho^x) \kappa \theta_{t+1} \quad (\text{C.43})$$

$$\Phi_t^4 = E_t \bar{\beta} \exp(\beta_{t+1}) \lambda_{t+1} (1 - \rho^x) \frac{\kappa}{q_{t+1}} \quad (\text{C.44})$$

- Using Equations (C.4) and (C.9), Equations (C.7) and (C.8) can be written as:

$$f_t = (1 + \theta_t^{-\gamma})^{-1/\gamma} \quad (\text{C.45})$$

$$q_t = (1 + \theta_t^\gamma)^{-1/\gamma} \quad (\text{C.46})$$

• Plugging the job seeker definition (Equation (C.5)) and the matching probabilities (C.45) and (C.46) into the law motion for employment (Equation (C.3)), one has:

$$n_t = (1 - \rho^x)n_{t-1} + (1 - (1 - \rho^x)n_{t-1})(1 + \theta_t^{-\gamma})^{-1/\gamma} \quad (\text{C.47})$$

• Given that $y_t = n_t$, $v_t = s_t\theta_t$ and using the market clearing condition (C.15), one has the following expression for λ_t :

$$\lambda_t = \left\{ n_t \times \left[1 - \frac{\psi}{2} \left(\frac{\pi_t}{\pi} - 1 \right)^2 \right] - \kappa(1 - (1 - \rho^x)n_{t-1})\theta_t \right\}^{-1} \quad (\text{C.48})$$

• The marginal cost can be obtained from the NKPC Equation (C.13):

$$mc_t = \frac{1}{\epsilon} \left[\epsilon - 1 + \psi \frac{\pi_t}{\pi} \left(\frac{\pi_t}{\pi} - 1 \right) \frac{\Phi_t^2}{n_t \lambda_t} \right] \quad (\text{C.49})$$

• The notional interest rate (C.14) can be rewritten as

$$1 + i_t^* = \frac{\pi}{\beta} \left(\frac{\pi_t}{\pi} \right)^{\rho\pi} \left(\frac{n_t}{n} \right)^{\rho n} \quad (\text{C.50})$$

• The effective nominal interest rate can then be computed as:

$$i_t(r_t) = \begin{cases} \frac{\pi}{\beta} \left(\frac{\pi_t}{\pi} \right)^{\rho\pi} \left(\frac{n_t}{n} \right)^{\rho n} & \text{if } r_t = 1 \\ 0 & \text{otherwise} \end{cases}$$

• The compact form of the competitive equilibrium is then given by the Euler equation (C.2) and the job creation condition (C.10) (after replacing the wage by its definition from Equation (C.12)):

$$\lambda_t = (1 + i_t(r_t))\Phi_t^1 \quad (\text{C.51})$$

$$\frac{\kappa}{(1 + \theta_t^\gamma)^{-1/\gamma}} = \xi \left(mc_t - b - \frac{\ell}{\lambda_t} \right) \frac{\Phi_t^3}{\lambda_t} + \frac{\Phi_t^4}{\lambda_t} \quad (\text{C.52})$$

Given expectation functions and initial value for the state variables (n_{t-1} and β_t), the decisions rule are given by Equations (C.51) and (C.52).

Some useful notations and operators

Throughout the algorithm we will use the following notations and operators.

- x defines x as a matrix with l lines and c columns.
- $x_{[l \times c]}$ defines the state vector (excluding the regime).
- $\mathbf{x}_t = \{\beta_t, n_{t-1}\}$ defines the state vector (excluding the regime).
- $r = 1, \dots, R$ denote the range of regime.
- \mathcal{I}_n is a $[n \times 1]$ column vector whose elements are each equal to one.
- $\mathbb{1}_{\{cond\}}$ is an indicator variable taking the value 1 if the condition is satisfied and zero otherwise.
- $d = 1, \dots, D$ is the approximation order of the Chebyshev polynomial.
- $k = 1, \dots, K$ is the index of continuous state variables.
- $h = 1, \dots, H$ is the range quadrature nodes and weight used for integration.
- \otimes stands for the Kronecker product.
- \circ stands for the Hadamar product (or Schur product).

Chebyshev functions

In order to approximate the unknown functions (the policy rules) we use Chebyshev polynomials.

One-dimensional approximation

Consider first that \mathbf{x}_t is a one-dimensional state vector. The function $\varphi(\mathbf{x}_t)$ ensures that \mathbf{x}_t fits into the Chebyshev domain $[-1, 1]$:

$$\varphi(\mathbf{x}_t) = 2 \frac{\mathbf{x}_t - a}{b - a} - 1$$

Chebyshev polynomial basis of order $d \in [2, \mathbf{D}]$ are built according to the following recursion:

$$T_{d+1}(\mathbf{x}_t) = \mathbf{x}_t T_d(\mathbf{x}_t) - T_{d-1}(\mathbf{x}_t)$$

with $T_0(\mathbf{x}_t) = 1$ and $T_1(\mathbf{x}_t) = \mathbf{x}_t$. Applying the trigonometric identities, the d -th member of the polynomial is:

$$T_d(\mathbf{x}_t) = \cos(d \arccos(\mathbf{x}_t))$$

Let $P(\cdot)$ be an approximation function of a one-dimensional state variable \mathbf{x}_t . It writes:

$$P(\mathbf{x}_t; \Theta) = \sum_{d=0}^{\mathbf{D}} \theta_d T_d(\varphi(\mathbf{x}_t))$$

where $\Theta = \{\theta_1, \dots, \theta_{\mathbf{D}}\}$ are the parameters of the Chebyshev function.

Multidimensional approximation

When the number of state variables is higher than one ($\mathbf{K} > 1$), we have to build a multidimensional Chebyshev polynomial basis. The \mathbf{K} -fold tensor product basis for the function with \mathbf{K} variables ($\mathbf{x}_t = \{x_t^1, \dots, x_t^{\mathbf{K}}\}$) is built by taking all possible \mathbf{K} -term products of T_{d_k} . The tensor product basis \mathcal{B}^T is defined as:

$$\mathcal{B}^T_{[1 \times (1+d \mathbf{K})^d]} = \left\{ \prod_{k=1}^{\mathbf{K}} T_{d_k}(\varphi(x_t^k)) \mid d_k = 0, \dots, \mathbf{D}; k = 1, \dots, \mathbf{K} \right\}$$

The problem of the tensor product basis is that some of the basis have an order higher than D which increase the number of elements. In the complete base approach, all products of polynomial terms must have an order not higher than D . Denote by \mathcal{B}^C the complete polynomial basis with $\mathcal{B}^C \subseteq \mathcal{B}^T$. It is given by:

$$\mathcal{B}_{[1 \times \mathbf{b}]}^C = \left\{ \prod_{k=1}^{\mathbf{K}} T_{d_k}(\varphi(x_t^k)) \mid \sum_{d_1=0}^D \dots \sum_{d_{\mathbf{K}}=0}^D \left(\sum_{j=1}^{\mathbf{K}} d_j \right) \leq D \right\}$$

$$d_k = 0, \dots, D$$

$$k = 1, \dots, \mathbf{K}$$

where \mathbf{b} is the number polynomial basis such that:

$$\mathbf{b} = \sum_{d_1=0}^D \dots \sum_{d_{\mathbf{K}}=0}^D \mathbb{1}_{\{\sum_{j=1}^{\mathbf{K}} d_j \leq D\}} \quad \text{for } k = 1, \dots, \mathbf{K}$$

The multidimensional approximation function writes:

$$P(\mathbf{x}_t; \Theta) = \mathcal{B}^C \circ \Theta$$

where $\Theta = \{\theta_1, \dots, \theta_{\mathbf{b}}\}^\top$ is a $[\mathbf{b} \times 1]$ vector of Chebyshev coefficients.

Piecewise approximation

In the algorithm described latter we use piecewise approximation. The general idea is that the decision rule (consumption) is approximated by a collection of approximation functions, each of them belongs to a particular regime. Consequently, they are as many polynomials as there are regimes². We adopt

²This method approximates the solution more accurately than if we use a single Chebyshev polynomial. It allows to manage the RS in the policy rules. It is important to note that the two policy rules are linked when the expectations are computed. It means that agents take into

simplified notations. We denote by $P(\mathbf{x}_t; \Theta(r))$ the approximation function of the control variable when the regime is r with $\Theta(r) = \{\theta_1(r), \dots, \theta_m(r)\}^\top$ the vector of coefficients that belongs to the regime r . This notation mean that the decisions rules, summarized by the coefficient $\Theta(r)$, are actually time-varying. Therefore it is excluded from the state vector \mathbf{x}_t . The “Aggregate” policy rule is defined as:

$$P(\mathbf{S}_t; \Theta) = \begin{cases} P(\mathbf{x}_t; \Theta(1)) & \text{if } r_t = 1 \\ P(\mathbf{x}_t; \Theta(2)) & \text{if } r_t = 2 \\ \vdots & \vdots \\ P(\mathbf{x}_t; \Theta(R)) & \text{if } r_t = N \end{cases} \quad (\text{C.53})$$

with $\Theta = \{\Theta(1), \dots, \Theta(R)\}$ being a set of Chebyshev coefficients.

Numerical integration

Euler equations involve two important difficulties: (1) the presence of the aggregate shock and (2) the RS that may occur in the next period. Since the regime changes according to the aggregate shock, the evaluation of the expectation functions must take into account potential RS. To simplify notation, we drop time subscript t from present variable and use x' to define the next period state variables. Consider the expectation functions:

$$\Phi'_\iota = E\mathcal{E}(\mathbf{x}', P(\mathbf{x}'; \Theta(r'))) \quad \iota = 1, \dots, 4$$

account the endogenous RS process when they take their decisions.

where $\mathcal{E}(\cdot)$ stands for the expectation function, E the expectation operator and ι is the subscript for each of the four expectation functions. Next period state variables are obtained using the following law of motion:

$$\mathbf{x}' = \Gamma(\mathbf{x}, r, \varepsilon')$$

Since the next period regime r depends on unemployment (or one minus employment) one has the following evolution for the regime:

$$r' = \psi(r, \mathbf{x}') \equiv \psi(r, \Gamma(\mathbf{x}, r, \varepsilon'))$$

The Euler equation can then be rewritten as:

$$\Phi'_i = \int_{\varepsilon'} \mathcal{E} \left(\Gamma(\mathbf{x}, r, \varepsilon'), P(\Gamma(\mathbf{x}, r, \varepsilon'); \Theta(\psi(r, \Gamma(\mathbf{x}, r, \varepsilon')))) \right)$$

In order to solve the numerical integration problem, we use Gauss-Hermite quadratures. It consists in evaluating the integral at different nodes and summing the evaluations using particular weights. Let \mathbf{e}_h be the h -th nodes and ω_h the h -th weights of the Gauss-Hermite quadratures with $h = 1, \dots, H$, H being the number of quadrature nodes³.

Let Ξ be an approximation function of the integral (the right-hand-side of the previous equation). With Gauss-Hermite quadratures, one has:

$$\Xi'_i = \frac{1}{\sqrt{\pi}} \sum_{h=1}^H \omega_h \mathcal{E} \left(\Gamma(\mathbf{x}, r, \sigma_z \sqrt{2} \mathbf{e}_h), P(\Gamma(\mathbf{x}, r, \sigma_z \sqrt{2} \mathbf{e}_h); \Theta(\psi(r, \mathbf{x}, \sigma_z \sqrt{2} \mathbf{e}_h))) \right)$$

³With multiple shock, one must use a tensor product.

Newton algorithm

Once we obtain a value for the expectation functions, the objective is to pin down the policy rules. By doing so, we will be able to project the policy rules on the state-space in order to define the polynomial coefficients. Obtaining a value for the policy rules given expectation functions is not an easy task in models that are non-linear. For that purpose we rely on a Newton algorithm. The algorithm consists in solving the system of two equations (C.51) and (C.52) in two unknowns (θ_t and π_t) given initial value for the state variables (n_{t-1} and β_t) and given a value for the expectations ($\Phi_t^j, j = 1, \dots, 4$). Formally, we have:

$$\theta_t, \pi_t = \arg \min_{\theta_t, \pi_t} \mathcal{F} \left(\theta_t, \pi_t; \Xi^j, \mathbf{x}_t, r_t \right)$$

with \mathcal{F} being given by:

$$\mathcal{F} = \left[\begin{array}{c} -\lambda_t + (1 + i_t(r_t))\Phi_t^1 \\ -\frac{\kappa}{(1+\theta_t^\gamma)^{-1/\gamma}} + \zeta \left(mc_t - b - \frac{\ell}{\lambda_t} \right) \frac{\Phi_t^3}{\lambda_t} + \frac{\Phi_t^4}{\lambda_t} \end{array} \right] \quad (\text{C.54})$$

Generalized Stochastic Simulation Algorithm

The presentation of the algorithm is borrowed from Albertini and Moyen (2019) and adapted to our New Keynesian model with search and matching frictions and regime switching (ZLB and SCFG). As mentioned previously the compact form of the model is characterized by two control variables: θ_t and π_t . The algorithm aim at finding invariant policy functions of the form $\theta_t = P(\beta_t, n_{t-1}; \Theta_1(r_t))$ and $\pi_t = P(\beta_t, n_{t-1}; \Theta_2(r_t))$. Θ_1 and Θ_2 correspond to the set of Chebyshev polynomial coefficient that belong to the tightness and inflation respectively.

Step 1 - Initialization Choose the order of the Chebyshev polynomial \mathbf{D} . Denote by i the i -th iteration. At this stage, $i = 0$. Set a convergence criteria \underline{h} . Initialize the coefficients $\Theta(1)^i, \Theta(2)^i, \Theta(3)^i$ using OLS regression on a stochastic simulation computed from the solution of a perturbation method⁴.

Step 2 - Stochastic simulations Compute a stochastic simulation over $T=40000$ periods using the same sequence of shock $\{\varepsilon_t^\beta\}_{t=0}^T$ as in **Step 1**. Given initial condition of n_0, β_0 (deterministic steady state), $r_0 = 1$ and the sequence of shocks, the steps for the stochastic simulation are as follow. For $t = 2, \dots, T$ do:

a. Initial guess on policy rules:

$$\tilde{\theta}_t = P(\beta_t, n_{t-1}; \Theta_1(r_{t-1}))$$

$$\tilde{\pi}_t = P(\beta_t, n_{t-1}; \Theta_2(r_{t-1}))$$

b. Unemployment and notional interest rate:

$$\begin{aligned} \tilde{u}_t &= 1 - (1 - \rho^x)n_{t-1} - [1 - (1 - \rho^x)n_{t-1}](1 + \tilde{\theta}_t^{-\gamma})^{-1/\gamma} \\ \tilde{i}_t &= \left(\frac{\pi}{\bar{\beta}} \left(\frac{\tilde{\pi}_t}{\pi} \right)^{\rho\pi} \left(\frac{1 - \tilde{u}_t}{n} \right)^{\rho n} \right) - 1 \end{aligned}$$

c. Regime:

$$r_t = 1 \times \mathbb{1}_{\{\tilde{i}_t > 0\}} \mathbb{1}_{\{\tilde{u}_t < \underline{u}(r_{t-1})\}} + 2 \times \mathbb{1}_{\{\tilde{i}_t \leq 0\}} \mathbb{1}_{\{\tilde{u}_t < \underline{u}(r_{t-1})\}} + 3 \times \mathbb{1}_{\{r_{t-1} > 1\}} \mathbb{1}_{\{\tilde{u}_t > \underline{u}(r_{t-1})\}}$$

d. New policy rules:

$$\theta_t = P(\beta_t, n_{t-1}; \Theta_1(r_t))$$

$$\pi_t = P(\beta_t, n_{t-1}; \Theta_2(r_t))$$

⁴This step is achieved thanks to the Dynare software. We consider a perturbation method of order 2 and assume that the regime is constant and equal to 1.

e. *Unemployment and notional interest rate:*

$$u_t = 1 - (1 - \rho^x)n_{t-1} - [1 - (1 - \rho^x)n_{t-1}](1 + \theta_t^{-\gamma})^{-1/\gamma}$$

$$i_t = \left(\frac{\pi}{\bar{\beta}} \left(\frac{\pi_t}{\pi} \right)^{\rho\pi} \left(\frac{1 - u_t}{n} \right)^{\rho n} \right) - 1$$

Step 3 - Representative points [Maliar and Maliar (2015), section 2.2.2]. Set the parameter ε_n , $r = 1, \dots, R$. Given the time series previously calculated, define $X_r = \{\beta_t, n_{t-1} \mid r_t = r\}_{t=0}^T$, $r = 1, \dots, R$ as a set of points for which the regime is equal to r . Denote by \mathcal{P}_r an empty set such that $\mathcal{P}_r = \{\emptyset\}$. For $r = 1, \dots, R$, do:

a. Select $x_j \in X_r$. Compute $\Delta(x_j, x_\ell)$, $j \neq \ell$ to all x_j in X_r with the following formula

$$\Delta(x_j, x_\ell) = \sqrt{\sum_{k=1}^K (PC_j^k - PC_\ell^k)^2}$$

$\Delta(x_j, x_\ell)$ is the euclidian distance between principal components (PC) of X_r normalized to unit variance.

b. Eliminate from X_r all x_j for which $\Delta(x_j, x_\ell) < \varepsilon_i$

c. Add x_j to \mathcal{P}_r and eliminate it from X_r

d. Go back to step a. until $X_r = \{\emptyset\}$

Step 4 - Expectations Set Gauss-Hermite quadratures nodes \mathbf{e} and weights $\boldsymbol{\omega}$. We consider $H = 10$ quadrature nodes. Denote by $x_{jr}^\beta \in \mathcal{P}_r$ and $x_{jr}^n \in \mathcal{P}_r$ the j -th elements (discount shock and employment respectively) of the set \mathcal{P}_r in regime r . Furthermore, denote by v'_{jrh} the j -th elements of the next period variable $v = \beta, n, r, \theta, \pi$ in regime r at quadrature node h . The quadrature rule expands variables to account for potential

future values of the shock. Then, for each regime ($r = 1, \dots, R$), each element $j = 1, \dots, J_r$ from \mathcal{P}_r and each quadrature nodes $h = 1, \dots, H$:

a. Compute guess for next period variables:

$$\begin{aligned}\beta'_{jrh} &= \rho\beta\beta_{jrh} + \sigma_z \mathbf{e}_h \sqrt{2} \\ \tilde{\theta}'_{jrh} &= P(\beta'_{jrh}, n_{jrh}; \Theta_1(r)) \\ \tilde{\pi}'_{jrh} &= P(\beta'_{jrh}, n_{jrh}; \Theta_2(r))\end{aligned}$$

b. Compute next period unemployment and notional interest rate:

$$\begin{aligned}\tilde{u}'_{jrh} &= 1 - (1 - \rho^x)n_{jrh} - [1 - (1 - \rho^x)n_{jrh}](1 + (\tilde{\theta}'_{jrh})^{-\gamma})^{-1/\gamma} \\ \tilde{i}'_{jrh} &= \left(\frac{\pi}{\beta} \left(\frac{\tilde{\pi}_t}{\pi} \right)^{\rho\pi} \left(\frac{1 - \tilde{u}'_{jrh}}{n} \right)^{\rho n} \right) - 1\end{aligned}$$

c. Regime:

$$r'_{jrh} = 1 \times \mathbb{1}_{\{\tilde{i}'_{jrh} > 0\}} \mathbb{1}_{\{\tilde{i}'_{jrh} < \underline{u}(r)\}} + 2 \times \mathbb{1}_{\{\tilde{i}'_{jrh} \leq 0\}} \mathbb{1}_{\{\tilde{i}'_{jrh} < \underline{u}(r)\}} + 3 \times \mathbb{1}_{\{r_{t-1} > 1\}} \mathbb{1}_{\{\tilde{i}'_{jrh} > \underline{u}(r)\}}$$

d. New policy rules:

$$\begin{aligned}\theta'_{jrh} &= P(\beta'_{jrh}, n_{jrh}; \Theta_1(r'_{jrh})) \\ \pi'_{jrh} &= P(\beta'_{jrh}, n_{jrh}; \Theta_2(r'_{jrh}))\end{aligned}$$

e. Compute intermediary variables used for the expectation functions:

$$\begin{aligned} n'_{jrh} &= (1 - \rho^x)n_{jrh} + [1 - (1 - \rho^x)n_{jrh}](1 + (\theta'_{jrh})^{-\gamma})^{-1/\gamma} \\ v'_{jrh} &= \theta'_{jrh}[1 - (1 - \rho^x)n_{jrh}] \\ c'_{jrh} &= n'_{jrh} \left[1 - \frac{\psi}{2} \left(\frac{\pi'_{jrh}}{\pi} - 1 \right) \right] - \kappa v'_{jrh} \\ \lambda'_{jrh} &= 1/c'_{jrh} \\ q'_{jrh} &= (1 + (\theta'_{jrh})^\gamma)^{-1/\gamma} \end{aligned}$$

f. Compute the expectation functions at each node h :

$$\Phi^1_{jrh} = \bar{\beta} \exp(\beta'_{jrh}) \frac{\lambda'_{jrh}}{\pi'_{jrh}} \quad (\text{C.55})$$

$$\Phi^2_{jrh} = \bar{\beta} \exp(\beta'_{jrh}) \lambda'_{jrh} \psi \frac{\pi'_{jrh}}{\pi} \left(\frac{\pi'_{jrh}}{\pi} - 1 \right) n'_{jrh} \quad (\text{C.56})$$

$$\Phi^3_{jrh} = \bar{\beta} \exp(\beta'_{jrh}) \lambda'_{jrh} (1 - \xi) (1 - \rho^x) \kappa \theta'_{jrh} \quad (\text{C.57})$$

$$\Phi^4_{jrh} = \bar{\beta} \exp(\beta'_{jrh}) \lambda'_{jrh} (1 - \rho^x) \frac{\kappa}{q'_{jrh}} \quad (\text{C.58})$$

g. For each expectation functions $\iota = 1, \dots, 4$, sum over h using weights ω :

$$\Xi'_{jr} = \frac{1}{\sqrt{\pi}} \sum_{h=1}^H \omega_h \Phi'_{jrh}$$

Step 5 - Estimate coefficients The new value of the coefficients is calculated using OLS. The dependent variables is pinned down by the expectation terms and the explanatory variables is calculated from the basis function of the Chebyshev Polynomial. For $r = 1, \dots, R$

a. Pined down the estimate of today's policy rules using Newton algorithm to solve

Equation (C.54):

$$\hat{\theta}_{jr}, \hat{\pi}_{jr} = \arg \min_{\theta_{jr}, \pi_{jr}} \mathcal{F}(\theta_{jr}, \pi_{jr}; \Xi_{jr}^t, \mathbf{x}_{jr}, r)$$

b. Regress the estimate of consumption on the basis functions to get a new value for the coefficients:

$$\begin{aligned} \Theta_1(r)^{i+1} &= \left(\mathcal{B}^C(\mathcal{P}_r)^\top \mathcal{B}^C(\mathcal{P}_r) \right)^{-1} \left(\mathcal{B}^C(\mathcal{P}_r)^\top \hat{\theta}_r \right) \\ \Theta_2(r)^{i+1} &= \left(\mathcal{B}^C(\mathcal{P}_r)^\top \mathcal{B}^C(\mathcal{P}_r) \right)^{-1} \left(\mathcal{B}^C(\mathcal{P}_r)^\top \hat{\pi}_r \right) \end{aligned}$$

Step 7 - Convergence

a. Define the policy rule criteria

$$b_r = \frac{\|\Theta(r)^{i+1} - \Theta(r)^i\|}{\|\Theta(r)^i\|}$$

The final convergence criteria at iteration i is

$$b^i = \max(b_1, \dots, b_R)$$

b. Check if $b^i < \underline{b}$. If it is the case, stop the algorithm.

c. Otherwise, if $b^i \geq \underline{b}$, set $i = i + 1$ and go back to **Step 2**.

d. To achieve a convergence in the algorithm we use a smoothing parameter for the update of the policy rules:

$$\Theta(r)^{i+1} = \gamma^i \Theta(r)^i + (1 - \gamma^i) \Theta(r)^{i+1}$$

where γ^i depends on the iteration with the following values:

$$\gamma^i = \begin{cases} \gamma_1 & \text{if } i < i_{\min} \\ \gamma_2 & \text{if } i_{\min} \leq i < i_{\max} \\ \gamma_3 & \text{if } i \geq i_{\max} \end{cases}$$

Bibliography

- Adrian, T., Crump, R. K. & Moench, E. (2013), 'Pricing the term structure with linear regressions', *Journal of Financial Economics* **110**(1), 110–138.
- Akerlof, G. A. (1978), The market for "lemons": Quality uncertainty and the market mechanism, in 'Uncertainty in Economics', Elsevier, pp. 235–251.
- Albertini, J. & Moyon, S. (2019), 'Solving regime-switching dsge model with projection methods', *Forthcoming* .
- Albertini, J. & Poirier, A. (2015), 'Unemployment benefit extension at the zero lower bound', *Review of Economic Dynamics* **18**(4), 733–751.
- Andolfatto, D. (1996), 'Business cycles and labor-market search', *American Economic Review* **86**(1), 112–32.
- Andrade, P., Gaballo, G., Mengus, E. & Mojon, B. (2015), Forward guidance and heterogeneous beliefs, Working papers 573, Banque de France.
URL: <https://ideas.repec.org/p/bfr/banfra/573.html>
- Bauer, M. D. & Rudebusch, G. D. (2013), 'The signaling channel for federal reserve bond purchases', *International Journal of Central Banking*, *forthcoming* .
- Baumeister, C. & Benati, L. (2013), 'Unconventional Monetary Policy and the Great Recession: Estimating the Macroeconomic Effects of a Spread Compression at the Zero Lower Bound', *International Journal of Central Banking* **9**(2), 165–212.
- Bernanke, B. S. (2012), Opening remarks: monetary policy since the onset of the crisis, in 'Proceedings: Economic Policy Symposium Jackson Hole', pp. 1–22.
- Bernanke, B. S., Gertler, M. & Gilchrist, S. (1999), 'The financial accelerator in a quantitative business cycle framework', *Handbook of macroeconomics* **1**, 1341–1393.

- Binning, A. (2013), 'Underidentified svar models: A framework for combining short and long-run restrictions with sign-restrictions'.
- Boneva, L., Harrison, R. & Waldron, M. (2018), 'Threshold-based forward guidance', *Journal of Economic Dynamics and Control* **90**(C), 138–155.
- Borio, C. & Zabai, A. (2018), 'Unconventional monetary policies: a re-appraisal', *Research Handbook on Central Banking* p. 398.
- Caballero, R. & Farhi, E. (2017), 'The safety trap', *The Review of Economic Studies* **85**(1), 223–274.
- Campbell, J. R., Evans, C. L., Fisher, J. D., Justiniano, A., Calomiris, C. W. & Woodford, M. (2012), 'Macroeconomic effects of federal reserve forward guidance', *Brookings Papers on Economic Activity* pp. 1–80.
- Carlstrom, C. T., Fuerst, T. S. & Paustian, M. (2015), 'Inflation and output in new keynesian models with a transient interest rate peg', *Journal of Monetary Economics* **76**, 230–243.
- Christensen, J. H. & Rudebusch, G. D. (2012), 'The response of interest rates to us and uk quantitative easing', *The Economic Journal* **122**(564), F385–F414.
- Christiano, L. J., Motto, R. & Rostagno, M. (2014), 'Risk shocks', *American Economic Review* **104**(1), 27–65.
- Christiano, L., Motto, R. & Rostagno, M. (2010), 'Financial factors in economic fluctuations'.
- Curdia, V. & Woodford, M. (2011), 'The central-bank balance sheet as an instrument of monetary policy', *Journal of Monetary Economics* **58**(1), 54–79.
- Del Negro, M., Giannoni, M. P. & Patterson, C. (2012), 'The forward guidance puzzle', *FRB of New York Staff Report* (574).
- Dell'Ariccia, G., Laeven, L. & Suarez, G. A. (2017), 'Bank leverage and monetary policy's risk-taking channel: evidence from the united states', *the Journal of Finance* **72**(2), 613–654.
- English, W. B., Lopez-Salido, J. D. & Tetlow, R. J. (2015), 'The federal reserve's framework for monetary policy: Recent changes and new questions', *IMF Economic Review* **63**(1), 22–70.

- Fernández-Villaverde, J., Gordon, G., Guerrón-Quintana, P. & Rubio-Ramirez, J. F. (2015), 'Nonlinear adventures at the zero lower bound', *Journal of Economic Dynamics and Control* **57**, 182–204.
- Florez-Jimenez, M. L. & Parra-Polania, J. A. (2016), 'Forward guidance with an escape clause: when half a promise is better than a full one', *Applied Economics* **48**(15), 1372–1381.
- Friedman, M. & Schwartz, A. (1963), 'A monetary history of the united states', *Princeton University Press* .
- Gagnon, J., Raskin, M., Remache, J., Sack, B. et al. (2011), 'The financial market effects of the federal reserve's large-scale asset purchases', *international Journal of central Banking* **7**(1), 3–43.
- Gambacorta, L., Hofmann, B. & Peersman, G. (2014), 'The effectiveness of unconventional monetary policy at the zero lower bound: A cross-country analysis', *Journal of Money, Credit and Banking* **46**(4), 615–642.
- Gertler, M. & Karadi, P. (2011), 'A model of unconventional monetary policy', *Journal of monetary Economics* **58**(1), 17–34.
- Gertler, M. & Karadi, P. (2013), 'Qe 1 vs. 2 vs. 3...: A framework for analyzing large-scale asset purchases as a monetary policy tool', *international Journal of central Banking* **9**(1), 5–53.
- Gertler, M. & Karadi, P. (2015), 'Monetary policy surprises, credit costs, and economic activity', *American Economic Journal: Macroeconomics* **7**(1), 44–76.
- Gertler, M., Kiyotaki, N. & Queralto, A. (2012), 'Financial crises, bank risk exposure and government financial policy', *Journal of Monetary Economics* **59**, S17–S34.
- Guerrieri, L. & Iacoviello, M. (2017), 'Collateral constraints and macroeconomic asymmetries', *Journal of Monetary Economics* **90**, 28–49.
- Gürkaynak, R. S., Sack, B. & Swanson, E. (2005), 'The sensitivity of long-term interest rates to economic news: Evidence and implications for macroeconomic models', *American economic review* **95**(1), 425–436.
- Gust, C., Lopez-Salido, D. & Smith, M. (2012), 'The empirical implications of the interest-rate lower bound', *CEPR Discussion Papers* **9214**.

- Hall, R. E. (2005), 'Employment fluctuations with equilibrium wage stickiness', *American Economic Review* **95**(1), 50–65.
- Hall, R. E. (2017), 'High discounts and high unemployment', *American Economic Review* **107**(2), 305–30.
- Hamilton, J. D. & Wu, J. C. (2012), 'The effectiveness of alternative monetary policy tools in a zero lower bound environment', *Journal of Money, Credit and Banking* **44**, 3–46.
- Hancock, D. & Passmore, W. (2011), 'Did the federal reserve's mbs purchase program lower mortgage rates?', *Journal of Monetary Economics* **58**(5), 498–514.
- Hoeting, J. A., Madigan, D., Raftery, A. E. & Volinsky, C. T. (1999), 'Bayesian model averaging: a tutorial', *Statistical science* pp. 382–401.
- Iacoviello, M. (2005), 'House prices, borrowing constraints, and monetary policy in the business cycle', *American economic review* **95**(3), 739–764.
- Jouvanceau, V. (2019), New evidence on the effects of quantitative easing, GATE, Working Papers 1912.
- Joyce, M., Miles, D., Scott, A. & Vayanos, D. (2012), 'Quantitative easing and unconventional monetary policy—an introduction', *The Economic Journal* **122**(564), F271–F288.
- Kalman, R. E. (1960), 'A new approach to linear filtering and prediction problems', *Journal of basic Engineering* **82**(1), 35–45.
- Kiyotaki, N. & Moore, J. (1997), 'Credit cycles', *Journal of political economy* **105**(2), 211–248.
- Koop, G. & Korobilis, D. (2012), 'Forecasting inflation using dynamic model averaging', *International Economic Review* **53**(3), 867–886.
- Koop, G. & Korobilis, D. (2013), 'Large time-varying parameter vars', *Journal of Econometrics* **177**(2), 185–198.
- Koop, G. & Korobilis, D. (2014), 'A new index of financial conditions', *European Economic Review* **71**, 101–116.

- Krishnamurthy, A. & Vissing-Jorgensen, A. (2011), 'The effects of quantitative easing on interest rates: channels and implications for policy'.
- Martin, A., McAndrews, J., Palida, A. & Skeie, D. (2015), 'Federal Reserve Tools for Managing Rates and Reserves'.
- Meeks, R., Nelson, B. & Alessandri, P. (2017), 'Shadow banks and macroeconomic instability', *Journal of Money, Credit and Banking* **49**(7), 1483–1516.
- Mishkin, F. S. (1982), 'Does anticipated monetary policy matter? an econometric investigation', *Journal of Political Economy* **90**(1), 22–51.
- Modigliani, F. (1977), 'The monetarist controversy or, should we forsake stabilization policies?', *American Economic Review* **67**(2), 1–19.
URL: <https://EconPapers.repec.org/RePEc:aea:aecrev:v:67:y:1977:i:2:p:1-19>
- Modigliani, F. & Sutch, R. (1966), 'Innovations in interest rate policy', *The American Economic Review* **56**(1/2), 178–197.
- Mortensen, D. T. & Pissarides, C. A. (1994), 'Job creation and job destruction in the theory of unemployment', *Review of Economic Studies* **61**(3), 397–415.
- Nelson, B., Pinter, G. & Theodoridis, K. (2018), 'Do contractionary monetary policy shocks expand shadow banking?', *Journal of Applied Econometrics* **33**(2), 198–211.
- Pissarides, C. A. & Petrongolo, B. (2001), 'Looking into the black box: A survey of the matching function', *Journal of Economic Literature* **39**(2), 390–431.
- Raftery, A. E., Kárný, M. & Ettler, P. (2010), 'Online prediction under model uncertainty via dynamic model averaging: Application to a cold rolling mill', *Technometrics* **52**(1), 52–66.
- Rotemberg, J. J. (1982), 'Monopolistic price adjustment and aggregate output', *The Review of Economic Studies* **49**(4), 517–531.
- Rubio-Ramirez, J. F., Waggoner, D. F. & Zha, T. (2010), 'Structural vector autoregressions: Theory of identification and algorithms for inference', *The Review of Economic Studies* **77**(2), 665–696.
- Särkkä, S. (2013), *Bayesian filtering and smoothing*, Vol. 3, Cambridge University Press.

- Smets, F. & Wouters, R. (2007), 'Shocks and frictions in us business cycles: A bayesian dsge approach', *American economic review* **97**(3), 586–606.
- Swanson, E. T. (2011), 'Let's twist again: a high-frequency event-study analysis of operation twist and its implications for $qe2$ ', *Brookings Papers on Economic Activity* **2011**(1), 151–188.
- Tobin, J. (1969), 'A general equilibrium approach to monetary theory', *Journal of money, credit and banking* **1**(1), 15–29.
- Townsend, R. M. (1979), 'Optimal contracts and competitive markets with costly state verification', *Journal of Economic theory* **21**(2), 265–293.
- Vayanos, D. & Vila, J.-L. (2009), A preferred-habitat model of the term structure of interest rates, Technical report, National Bureau of Economic Research.
- Weale, M. & Wieladek, T. (2016), 'What are the macroeconomic effects of asset purchases?', *Journal of Monetary Economics* **79**, 81–93.
- Werning, I. (2011), Managing a liquidity trap: Monetary and fiscal policy, NBER Working Papers 17344, National Bureau of Economic Research, Inc.
- Williamson, S. D. (2012), 'New monetarist economics: Understanding unconventional monetary policy', *Economic Record* **88**, 10–21.
- Williamson, S. D. (2016), 'Scarce collateral, the term premium, and quantitative easing', *Journal of Economic Theory* **164**, 136–165.
- Williamson, S. D. (2018), 'Interest on reserves, interbank lending, and monetary policy', *Journal of Monetary Economics* .
- Woodford, M. & Eggertsson, G. (2003), 'The zero bound on interest rates and optimal monetary policy', *Brookings Papers on Economic Activity* **1**, 139–233.
- Woodford, M. et al. (2012), 'Methods of policy accommodation at the interest-rate lower bound', *The Changing Policy Landscape* **185**, 288.
- Wu, J. C. & Xia, F. D. (2016), 'Measuring the macroeconomic impact of monetary policy at the zero lower bound', *Journal of Money, Credit and Banking* **48**(2-3), 253–291.