



HAL
open science

Essays on macroeconomic theory

Elliot Aurissergues

► **To cite this version:**

Elliot Aurissergues. Essays on macroeconomic theory. Economics and Finance. Université Panthéon-Sorbonne - Paris I, 2018. English. NNT : 2018PA01E029 . tel-01972942

HAL Id: tel-01972942

<https://theses.hal.science/tel-01972942>

Submitted on 8 Jan 2019

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.



UNIVERSITÉ PARIS 1 PANTHÉON-SORBONNE
UFR de Sciences économiques
Paris Jourdan Sciences Economiques

THÈSE

Pour l'obtention du titre de Docteur en Sciences économiques
présentée et soutenue publiquement
le 28 septembre 2018 par
Elliot AURISSERGUES

Essays on Macroeconomic Theory

Sous la direction de :

M. Bertrand WIGNIOLLE, Professeur, Université Paris I Panthéon Sorbonne

M. Florin BILBIIE, Professeur, Université Paris I Panthéon Sorbonne

Membres du jury :

President du Jury

M. Jean Bernard CHATELAIN, Professeur, Université Paris I Panthéon Sorbonne

Rapporteurs

M. Edouard CHALLE, Professeur, École Polytechnique

M. Gerhard SORGER, Professeur, Université de Vienne

Examineurs

M. Roger GUESNERIE, Professeur émérite, Collège de France

M. Julien MATHERON, Docteur, Banque de France

Remerciements

Je remercie tout d'abord mon directeur de thèse principal Bertrand Wigniolle, ainsi que mon second directeur de thèse, Florin Bilbiie.

Je remercie les membres de mon jury de thèse pour les commentaires très précis et pertinents qu'ils ont pu faire durant la pré-soutenance ou par la suite et qui ont permis d'améliorer très substantiellement cette thèse.

Cette thèse n'aurait pu aboutir sans le soutien de plusieurs institutions, L'ENS Cachan qui a financé mon contrat doctoral initial, L'université Paris 1 Panthéon Sorbonne qui m'a accueilli pendant 5 ans et pour laquelle j'ai effectué ma mission d'enseignement pendant 3 ans et enfin l'université Paris 2 Panthéon Assas qui m'a offert un contrat d'ATER à plein temps pendant deux ans. A cet égard, je remercie les enseignants pour lesquels j'ai pu donner des travaux dirigés, en particulier Katheline Schubert et Antoine d'Autume à Paris 1, Alain Redslob et Jean Mercenier à Paris 2.

J'en profite également pour remercier les élèves que j'ai pu avoir durant ces cinq ans et en particulier les rares qui posaient parfois des questions.

Plus largement, mes remerciements vont à l'ensemble des professeurs et chercheurs avec lesquels j'ai pu interagir durant cette thèse ou auparavant. Je pense notamment à Nuno Coimbra, Gilles Saint Paul, Xavier Ragot, Gabriel Desgranges, Jean Olivier Hairault, Nicolas Dromel, Jean Imbs, Jordi Gali, Romain Rancière, Stéphane Gauthier, Cars Hommes, Henry Thompson, Francesco Furlanetto, Angelo Riva, Basile Grassi, Edouard Schaal ou encore Nicolas Drouhin. Je suis également reconnaissant pour toutes les remarques que j'ai pu recevoir sur mes différents papiers. Je pense par

exemple à Gaetano Gaballo, Emmanuel Carré, Egor Maslov, Russel Cooper, Allan Drazen, et à de nombreux autres participants aux conférences et séminaires auxquels j'ai présenté mes travaux .

Je souhaiterais tout particulièrement remercier deux de mes professeurs, tout d'abord le professeur Roger Guesnerie à la fois pour ses commentaires sur le premier chapitre de cette thèse et pour m'avoir permis de présenter le papier associé ce premier chapitre à la conférence de la Society for the Advancement of Economic Theory 2016 ainsi au séminaire qu'il organisait régulièrement au Collège de France et ensuite le professeur Alain Goergen qui a encadré mon mémoire de master 1 au sein de l'ENS cachan et a régulièrement commenté mon travail par la suite.

Cette thèse n'aurait jamais pu aboutir sans l'excellente ambiance qui a régné à la Maison des Sciences Économiques pendant 4 ans. Je remercie à cet égard mes collègues du bureau 219, Esther, Vincent, Antoine, Yvan, Mehdi, Sebastien, Anastasia, Ezgi, Idriss, et Bertrand. Je remercie également mes autres collègues doctorants, en premier lieu Pierre, mais aussi Sandra, Stefanija, Anna, Moutaz, Can, Mathias, Selim, Diane, Alessandra, Lorenzo, Kristel, Ilya, Mathieu B et S, Rudy, Guillaume C et A, Giulia, Johanna, Pierrick, Mathilde, Clément, Rémi, Pauline, ainsi que tout les doctorants que j'ai pu croiser dans les couloirs de la MSE ou dans des bars à proximité.

Ma gratitude va également à mes amis de l'ENS Cachan, en premier lieu à Charles qui fut un soutien moral constant au cours de ces cinq années mais aussi Mickael, Guillaume, Brice, Jeremy, Maxime. Je pense également à Wafa, Lionel, Sophie, Marine, Aurélie, Bruno, Jordan, Vincent, Nava, Marion S, Léonard, Pierre Louis.

Enfin, mes derniers et plus profonds remerciements vont à Aurélie ainsi qu'à ma famille pour leur soutien constant tant moral que matériel. Je remercie évidemment mes parents Christine et Eric, mon frère Alexis, ainsi que mes grands parents Jacqueline, Élise et Daniel. Je pense également à mon grand père paternel Jacques, qui nous a quitté en 2003.

Contents

Remerciements	iii
Contents	v
Introduction	1
1 Are consistent expectations better than rational expectations ?	11
1.1 Introduction	11
1.2 Literature	14
1.3 Framework	17
1.4 Long Run Equilibrium	20
1.4.1 Equilibrium definition	21
1.4.2 The Rational Learner Equilibrium	22
1.4.3 The Consistent Learner Equilibrium	27
1.5 Simulations	35
1.5.1 Overview	36
1.5.2 Summary of the model and the algorithm	38
1.5.3 Initialization	39
1.5.4 Baseline calibration	41
1.5.5 Experiments	42
1.6 Results	43
1.6.1 Summary	43
1.6.2 Main result	44
1.6.3 Exploration	46

1.6.4	The role of initialization	47
1.6.5	Path dependence	51
1.7	Extensions: misspecification and structural breaks	51
1.7.1	Intuition	51
1.7.2	Adding persistent unobservables	53
1.7.3	Adding structural breaks	53
1.7.4	Results	54
1.7.5	Why do rational learners misspecify their model	54
1.8	Discussion	55
1.9	Conclusion	57
2	Intertemporal Complementarity and the New Keynesian model	61
2.1	Introduction	61
2.2	Intratemporal Household choice	64
2.2.1	The household program	65
2.2.2	First order conditions	66
2.2.3	Separable preferences	67
2.2.4	Intertemporal nonseparability	69
2.3	Intertemporal Household choice	72
2.4	Implications for the New Keynesian model	74
2.4.1	Framework	75
2.4.2	The supply effect of the real interest rate	78
2.4.3	Monetary shocks	80
2.4.4	Demand Shock	87
2.5	A Bayesian estimation of the model	90
2.5.1	Baseline estimation	91
2.5.2	Inspecting the mechanism	93
2.5.3	Robustness	94
2.5.4	Quantitative implications of the estimated model	96

2.6	Discussion	97
2.7	Implications for the forward guidance puzzle	102
2.8	The model with varying wealth	104
2.9	Conclusion	108
3	Corporate Investment and Adverse Selection, A Reappraisal	115
3.1	Introduction	115
3.2	The adverse selection problem	117
3.2.1	Firms and lenders	117
3.2.2	Equilibrium Definition under private information	120
3.2.3	Characterization of the separating equilibrium	126
3.2.4	The scope of the equilibrium	137
3.3	The macroeconomic model	138
3.3.1	Extension of the equilibrium for time varying interest and dis- count rate	139
3.3.2	Dynamic equations for good and bad firms	141
3.3.3	Investment Dynamics	143
3.3.4	Steady state	144
3.4	Implications: Response to changes in borrowing cost	146
3.4.1	The determinants of corporate investment	146
3.4.2	Response of leverage to interest rate	147
3.4.3	Long run effects of changes in borrowing costs	148
3.4.4	Short run effects	148
3.4.5	Response to changes in opportunity cost	149
3.5	Conclusion	151
A	Appendix of chapter 1	157
A.1	Proofs	157
A.1.1	Proof of proposition 1.1	157

A.1.2	Proof of proposition 1.4	158
A.1.3	Proof of proposition 1.5	159
A.1.4	Proof of proposition 1.6	160
A.1.5	Proof for proposition 1.9	161
A.1.6	Proof for proposition 1.11	163
A.1.7	Proof for proposition 1.14	163
A.1.8	Proof of proposition 1.16	164
A.1.9	Proof of proposition 1.18	165
B	Appendix of chapter 2	167
B.1	Proofs	167
B.1.1	Proof of proposition 2.1	167
B.1.2	Proof of proposition 2.2	168
B.1.3	Proof of proposition 2.3	168
B.1.4	Proof of proposition 2.4	168
B.1.5	Proof of corollary 2	168
B.1.6	Proof of proposition 2.5	169
B.1.7	Proof of proposition 2.6	170
B.1.8	Proof of proposition 2.7	170
B.1.9	Proof of proposition 2.8	171
B.1.10	Proof of proposition 2.9	172
B.1.11	Proof of proposition 2.10	172
B.1.12	Proof of proposition 2.11	172
B.2	Wealth in the utility function and indivisible labor	173
B.3	Bayesian estimation	174
B.4	Generalized nonseparable preferences in the model	177
B.5	Adjustment cost for consumption in an overlapping generations model	178
B.6	Forward guidance computations	179
B.6.1	Forward guidance in the standard model	179

B.6.2	Forward guidance with wealth in the utility	181
B.7	Model with creditors and debtors	181
C	Appendix of chapter 3	185
C.1	The Adverse Selection Problem	185
C.1.1	Proof for proposition 3.2	185
C.1.2	Proof of proposition 3.4	186
C.1.3	Proof of proposition 3.5	187
C.1.4	Proof of proposition 3.13	189
C.2	The Macroeconomic Model	190
C.2.1	Proof of proposition 3.14	190
C.2.2	Model derivation	191
	List of Figures	195
	List of Tables	197
	Résumé français	199
	Abstract	218

Introduction

This thesis deals with several questions related to the business cycles literature. It is composed of three chapters. Each of them can be read independently. However, all of them may have some implications for the conduct of economic policy, and in particular the transmission of monetary policy.

Chapter 1 The first chapter is a theoretical contribution to the expectation literature. The modelling of expectations is a difficult but important issue for both macroeconomic theory and applied macroeconomics.

For an economic agent, many decisions involve forecasting future values of some important variables. For example, a firm willing to know if building a new plant will be profitable has to forecast future cash flows and future user costs of capital. An investor facing the choice between two stocks should anticipate the future earnings growth of both corporations. The importance of expectations for consumption, labor supply, price and wage setting is less intuitive but they probably matter too. Hence, at the aggregate level, most macroeconomic variables are affected in some extent by agents' forecasts about their future values.

Once it is agreed that expectations matter, the next question is how they are formed. To foresight the future value of a variable x , an obvious possibility is to use the *current* value of the x . An issue is that the current value may be a very noisy signal for the future one. A solution is to use more information by considering a moving average of past values. This is the idea behind adaptive expectations.

A drawback of this approach is that agents do not react to new relevant information. For example, a sudden financial crisis, a more conservative central banker, a populist government are likely to affect future values of output but an agent using adaptive expectations would not react to these changes because they have not affected past values of output. The Rational Expectations Hypothesis (REH), introduced by Muth(1961) and popularized by Lucas in various papers, assumes that agents are more forward looking. According to the REH, the expectation for a variable x is equal to the expected value of the variable x conditional on available information. Formally, denoting the expectation x^e , the information set I and using standard notation for the expected value operator, the expectation is given by

$$x^e = E(X | I)$$

In practice, in most models assuming the REH, the agent is supposed to have the same level of knowledge as the theorist and to solve the model in the same way. For example, in an optimal growth model with a representative agent living infinitely and a stochastic productivity, each agent is supposed to compute the expected path of productivity conditional on current information, to solve the macroeconomic model given this path, then to deduce the future path of capital stock, real wages and real interest rate before to solve his individual maximization problem.

Thus, the REH requires that agents have (very) high cognitive capabilities and a very good knowledge of the economic environment. They have to know the true model of the economy and true parameter values for this model. An individual should also know how other individuals form their expectations. These assumptions are extremely strong and the REH has been criticized on this ground (see for example for Guesnerie 2001, 2005). Several pieces of literature have explored alternative theories of expectations. One of them is the learning literature. Agents do not know true parameter values and learn them by estimating the rational expectation model with standard econometric

techniques. Consistent Expectations, defined by Hommes and Sorger (1998), are another one. Consistent expectations are close to adaptive expectations. The forecast of x depends on its past values. The difference with adaptive expectations is that the relation between the expectation and past values have to be consistent with observed autocorrelations.

My paper builds on these two pieces of literature. I study a simple economy populated by individuals learning and using different forecasting models. One of this model is the "rational" model in the sense that the endogenous variable is a function of exogenous variables. The second model borrows the concept of consistent expectations from Hommes and Sorger (1998). Both models are estimated at each period. Moreover, agents may switch from one model to another according to past forecasting performance. I define a long run equilibrium of this economy as a situation in which (i) agents have a stable estimation of both the rational and the consistent model (ii) there is only one type of agent remaining (iii) Forecast errors of the "dominant" model are lower than forecast errors of the other model. Thus, two equilibria may exist, the "rational" one in which only the rational model is used and the "consistent" one in which the consistent model dominates. My first contribution is to show that the consistent equilibrium *actually* exists for a large set of parameters. My second contribution is to perform simulations of the economy and to show that the economy may converge towards this consistent equilibrium.

Chapter 2 The second chapter explores an alternative specification of the household optimization problem. In many business cycles model, like the Real Business Cycles model or the New Keynesian model, household choose their consumption and labor

supply by solving an infinite horizon optimization program. The standard program is

$$\begin{aligned} \max_{C_t, L_t, A_{t+1}} \sum_{t=0}^{+\infty} \beta^t U(C_t, L_t) \\ \text{w.r.t. } \forall t \quad (1 + r_t)A_t + W_t = W_t L_t + C_t + A_{t+1} \end{aligned}$$

where C_t is the consumption of period t , L_t is the leisure, W_t is the wage and A_t is an asset delivering one unit of good in period t . U is the utility function. It is often assumed separable. For example, a popular functional form is

$$U(C_t, L_t) = \frac{C_t^{1-\sigma}}{1-\sigma} + \frac{C_t^{1-\theta}}{1-\theta}$$

Solving this program leads to two well-known first order conditions

$$\begin{aligned} C_t^{-\sigma} &= \beta(1 + r_t)C_{t+1}^{-\sigma} \\ L_t^{-\theta} &= W_t C_t^{-\sigma} \end{aligned}$$

The first is the Euler equation and the second is the first order condition for leisure. They form the core of the optimal growth model, one of the workhorses of modern macroeconomics. When assuming a representative agent, e.g all agents have the same utility function and the same endowment, these first order conditions hold for *aggregate* consumption and leisure.

I consider a slight modification of this model. Instantaneous utility is provided by consumption, leisure and wealth, assimilated to the variable A . The program becomes

$$\max_{C_t, L_t} \sum_{t=0}^{+\infty} \beta^t U(C_t, L_t, A_{t+1})$$

$$w.r.t \quad \forall t \quad (1 + r_t)A_t + W_t = W_t L_t + C_t + A_{t+1}$$

Introducing wealth in the utility function may seem counter-intuitive. *A priori*, Assets A do not provide utility by themselves but only through future consumption and leisure. There is, however, two reasons to study this specification. First, this may capture several motives for saving, like self insurance against unemployment risk or capital accumulation for retirement and inheritance, whereas keeping the convenience of the representative agent assumption. These factors seem quantitatively important in explaining actual saving behavior (see for example Gourinchas and Parker 2002). Modelling them explicitly is certainly better but also much harder to embed in a medium scale macroeconomic model. Second, a growing literature (Krishnamurthy and Vissing-Jorgensen 2012, Saez and Stantcheva 2016, Michailat and Saez 2018, Kumhof-Ranci re and Winant 2015 among others) has used this specification to explain stylized facts that the standard model is unable to reproduce, for example the response of consumption to forward guidance or the concentration of wealth at the top of the wealth distribution. Other explanations are possible and would be arguably more intellectually satisfying. On the other hand, Wealth in the utility function is probably the simplest way to solve these issues and once again one of the more practical to embed in a medium scale model.

My contribution to this literature is to study the case of nonseparability between consumption and wealth. Allowing for nonseparability (and more precisely complementarity) is a way to disentangle between the elasticity of intertemporal substitution and the income effect on labor supply. In the standard model, the former is the inverse of the latter. This "cross equation restriction" implies that desired hours worked of a

representative household have the same elasticity with respect to real wages growth as with respect to real interest rate. If we follow the identification assumption proposed by Gali (2011), it implies that the response of unemployment to a monetary policy shock is very large and is generated as much by the shift in labor supply as by the shift in labor demand. It also implies that the response of the natural (flexible price) interest rate¹ to a demand shock² is small, close to the actual response of monetary policy. As a consequence, there is a small difference between the response of output and the response of flexible price equilibrium output and hence the response of output gap is particularly small. Wealth in the utility function and complementarity between consumption and wealth (that I call "intertemporal complementarity") allow for a lower income effect on labor supply whereas keeping a similar intertemporal substitution effect, dampening the response of unemployment to monetary policy shock and enhancing the response of output gap to demand shocks. I document those propositions in both a simple illustrative model and a medium scale DSGE model.

Then, I estimate the model and in particular my two key parameters κ which is a discount rate in the Linear Euler equation and ν which is introduced by the nonseparability between consumption and wealth. I estimate a medium scale model using Bayesian techniques over US macroeconomic data between 1985 and 2007. I find a large estimate for both κ and ν . Data seem pointing towards a small intertemporal substitution effect and a moderate income effect on labor supply, probably explaining the high value found for ν . These findings are robust to many alternative specifications.

Chapter 3 Whereas the second chapter deals with household behavior, the third chapter focus on the behavior of corporations and in particular on their investment decisions.

According to the basic theory of capital, the capital stock is set by firms in order to

¹By natural interest rate, I mean the flexible price equilibrium interest rate and not the steady state interest rate.

²A public consumption or a private investment shock

equalize the marginal product of capital and the user cost of capital. Let me assume that firms maximize their profits $f(K) - rK$ where f is the production function, K the capital stock and r the user cost, the first order condition is $f'(K) = r$. The optimal capital stock is deduced from this expression and investment is the difference between current and future optimal capital stock. This basic theory is often complemented by introducing adjustment costs in the optimization program of the firm. Otherwise, realistic changes in real interest rate would cause a phenomenal jump in investment.

This neoclassical theory of investment makes the assumption that firms are not subject to any borrowing constraint. On the other hand, several pieces of evidence suggest that firms face financial frictions. For example, a significant correlation is observed at firm level between investment and cash flows or sales. The interpretation of this correlation is difficult due to the possible correlation between cash flows and future investment opportunities but it could support the existence of borrowing constraints. Moreover, the sensitivity to cash flows seems particularly high among financially distressed firms as shown to Fazzari, Hubbard and Petersen (1988). It is not clear however if financial frictions necessarily implies a higher sensitivity to cash flows among more constrained firms (see Kaplan and Zingales 1997). Another piece of evidence have been provided by Gilchrist and Zakrajsek (2012). They show that some shifts in the spread between corporate bonds and safe assets cannot be explained by increased riskiness of corporate bonds and suggest tighter financial constraints as a possible explanation. Studying transmission channels of monetary policy, Bernanke and Gertler (1995) cast some doubts on the user cost channel and argue for the importance of the credit and the balance sheet channel.

These various pieces of evidence have led economists to develop models of financial frictions and to embed them into macroeconomic models, this last trend being considerably amplified by the financial crisis of 2008. A popular friction is the costly state verification model introduced by Townsend (1979). In this model, lenders cannot ob-

serve the outcome of an investment project for free. A borrower may understate his net worth, default and run away with his assets unless the lender pay an auditing cost to assess the true state of the borrower. The auditing cost creates a wedge between the return on investment for borrowers and the return on loans for lenders. Bernanke, Gertler and Gilchrist (1999) introduces this friction in a calibrated medium scale New Keynesian model and shows this enhances the response of output and investment to monetary policy shocks. Christiano, Motto and Rostagno (2014) estimate a similar model with Bayesian techniques and suggest that increases in volatility of firms' capital return may drive the business cycles.

A disturbing aspect of this literature is the ability of borrowers to divert a very large fraction of their assets. If borrowers are identified with shareholders, which is usually the case in macroeconomic applications, it is equivalent to assume that shareholders of publicly listed corporations may declare their firm bankrupt and run away with firm's assets³. This seems difficult to believe. Despite the success of these models, for example in explaining the rise in corporate spread during the financial crisis, it seems important to pursue the exploration of alternative models.

My contribution belongs to the macroeconomic literature dealing with adverse selection on capital markets and in particular with its consequences for corporate investment. Adverse selection on credit markets has been introduced by Stiglitz and Weiss (1981). A recent paper by Kurlat (2015) shows that adverse selection may provide a microfoundation for the resaleability constraint introduced by Kiyotaki and Moore (2012). I consider a slightly different adverse selection problem in infinite horizon. My contribution is to provide an explicit formula for corporate investment. Investment is a linear function of retained earnings. I embed the relation in a dynamic model and explore implications, in particular for the response of investment to changes in borrowing costs.

³If they have not the possibility to divert, they have no incentives to understate their assets and thus there is no need for auditing

Bibliography

- [1] **Bernanke. B., and Gertler. M.** 1995. "Inside the black box: The Credit Channel of Monetary Policy Transmission", *NBER working papers*, 5146.
- [2] **Bernanke. B., Gertler. M., and Gilchrist S.** 1999. "The Financial Accelerator in a Quantitative Business Cycle Framework", *in Handbook of Macroeconomics*, North-Holland, 1341-1393.
- [3] **Christiano. L., Motto. R. and Rostagno. M.** 2014. "Risk Shocks", *American Economic Review*, 104, 27-65.
- [4] **Fazzari. S., Hubbard. G. and Petersen. B.** 1988. "Financing Constraints and Corporate Investment", *Brookings Papers on Economic Activity*, 141-195.
- [5] **Gali. J.** 2011. *Unemployment Fluctuations and Stabilization Policies*, MIT Press.
- [6] **Gilchrist. S. and Zakrajsek. E.** 2012. "Credit Spread and Business Cycles Fluctuations", *American Economic Review*, 102, 1692-1720.
- [7] **Gourinchas. P. O. and Parker. J.** 2002. "Consumption over the Life Cycle", *Econometrica*, 70(1), 47-89.
- [8] **Guesnerie, R.** 2001. *Assessing Rational Expectations: Sunspot Multiplicity and Economic Fluctuations*. MIT Press
- [9] **Guesnerie, R.** 2005. *Assessing Rational Expectations 2-Eductive stability in economics*. MIT Press.
- [10] **Hommes, C.** 2013. *Behavioral Rationality and Heterogeneous Expectations in Complex Economic Systems*. Cambridge University Press.

-
- [11] **Hommes, C. and Sorger. G.** 1998. "Consistent Expectations Equilibria." *Macroeconomic Dynamics* 2, 287-321.
- [12] **Kaplan. S. and Zingales. L.** 1997. "Do Investment-Cash Flow Sensitivities Provide Useful Measures of Financing Constraints?", *The Quarterly Journal of Economics*, 112, 169-215.
- [13] **Kiyotaki. N. and Moore. J.** 2008. "Liquidity, Business Cycles and Monetary Policy", *Mimeo*, Princeton University.
- [14] **Krishnamurthy. A. and Vissing Jorgensen. A.** 2012. "The Aggregate Demand for Treasury Debt", *Journal of Political Economy*, 120(2), 233-267.
- [15] **Kumhof. M. , Ranciere. R. and Winant. P.** 2015. "Inequality, leverage and crisis", *American Economic Review*, 105(3), 1217-1245.
- [16] **Kurlat. P.** 2013. "Lemons Market and the Transmission of Aggregate Shocks", *American Economic Review*, 103(4), 1463-1489.
- [17] **Michaillat. P. and Saez. E.** 2018. "Wealth in the Utility function and the New Keynesian model", *mimeo*.
- [18] **Muth, J.** 1961. "Rational expectations and the theory of price movements." *Econometrica* 29,315-335.
- [19] **Saez. E. and Stantcheva. S.** 2016. "Simpler Theory of Optimal Capital Taxation", NBER working papers.
- [20] **Stiglitz. J. and Weiss. A.** 1981. "Credit Rationing in Market with Imperfect Information", *American Economic Review*, 71, 393-410.
- [21] **Townsend. R.** 1979. "Optimal contracts and competitive markets with costly state verification", *Journal of Economic Theory*, 21, 265-293.

Chapter 1

Are consistent expectations better than rational expectations ?

1.1 Introduction

The Rational expectation hypothesis (REH thereafter) is the dominant theory of expectations in macroeconomics. This hypothesis remains a source of controversy. Agents are completely forward looking. Could it be a good approximation of the real world? Recent debates, for example about the discrepancy between predicted and actual outcome of forward guidance, have suggested that agents are less forward looking than the theory still assumes.

One of the simplest alternative to the REH remains the old adaptive expectation hypothesis. With the REH, agents use all available and relevant information about exogenous variables to forecast future values of endogenous variables. With the adaptive expectation hypothesis, agents use lagged endogenous variables in order to forecast future ones. Thus, they do not use all relevant available information contained in exogenous variables and use the irrelevant information contained in lagged endogenous variables. Consistent expectations are a refinement of adaptive expectations. They were introduced by Hommes and Sorger (1998) and then refined by Hommes in several

papers.¹ When forming consistent expectations, agents assume that endogenous variables follow an auto-regressive processes. Persistence coefficients for these processes have to be equal to true auto-correlations. Despite being extremely interesting, the concept has not been largely used. A possible explanation for this limited success is that a variant of adaptive expectations is still perceived as less "rational" than rational expectations.

Indeed, it is often taken for granted that forming rational expectations is the best way to forecast. This paper challenges this idea. I consider an economy populated by individuals with different forecasting models. The law of motion of this economy is given by a simple univariate linear model. An endogenous variable called x depends on its expectation, a persistent exogenous variable called y , and a random perturbation. There are two types of agents: rational learners and consistent learners. Rational learners believe that the endogenous variable is determined by the exogenous variable. Consistent learners believe that the endogenous variable follows a first order autoregressive process and thus depends on its lagged values. A consistent agent behaves like a VAR econometrician whereas rational learners can be viewed as "DSGE" economists.

In the first part of the paper, I study the long run behavior of the economy. I define long run equilibrium as a situation in which both types have a stable estimation of their respective model and in which one type is dominant (e.g the mass of the other type is 0). The model of the dominant type should deliver more accurate forecasts in average than the other model. I distinguish between the Rational Learner Equilibrium and the Consistent Learner Equilibrium. In the former (resp. latter), rational (resp. consistent) learners dominate and the rational (resp. consistent) forecast is more accurate. I demonstrate that for a large set of parameter values, both types of equilibrium exist.

In the second part of the paper, I simulate the economy in order to determine which

¹See Hommes 2013 for a summary

long run equilibrium prevails. Both types learn their model in the spirit of Bray and Savin (1986) and Evans and Honkapoja (2001). At each period, they estimate their model through econometric techniques and using data they have observed. Once they have estimated their model, they use it in order to forecast the value of the endogenous variable. Once the value of the endogenous variable has been determined, both types compare the accuracy of their respective model. A fraction of agents using the less accurate model switches to the other one. My simulations show that the consistent equilibrium is not a theoretical curiosity. For a large set of parameters, the economy converges towards the consistent equilibrium and not the rational one.

In a nutshell, in a world where DSGE economists and VAR econometricians compete, DSGE economists may lose even if their model is the true one under the rational expectation hypothesis. All agents end up using the misspecified VAR model.

This result may be viewed as surprising. The economy I consider is very simple. Rational learners have the right representation of it whereas consistent learners have a misspecified one because the endogenous variable is not directly related to its lagged values. Hence, consistent learners miss the information contained in the exogenous variable but also use the irrelevant information contained in lagged endogenous variable. Despite these disadvantages, consistent learners may still win. Two factors explain their success. The first factor is that the theory of consistent learners is self fulfilling in some extent because lagged endogenous variables may affect x through expectations. If many agents believe that the past value of x is relevant, a change in the past value of x modify their forecast for x , affecting the actual value of x , and hence justifying their forecasting model. The second factor is that the consistent model does not exploit causality but exploits correlation. The exogenous variable y is persistent. Thus, past values of y are both correlated to the past value of x and the current value of y . As the current value of y is one of the determinant of the current value of x , there is some correlation between x and its past value. In a nutshell, my results require both

a positive feedback from expectations of x to the actual value of x and a persistent exogenous variable.

The paper is organized as follows. The first section describes the relevant literature. The second presents the framework. The third provides the theoretical results. The fourth explains how I perform simulations of the economy. The fifth deals with results of simulations. The sixth considers an extension where rational learners also have a misspecified model. The seventh offers a discussion of the main results.

1.2 Literature

This paper is related to the very large literature dealing with expectations in macroeconomic models

Rational expectations and learning literature Rational expectations were introduced by Muth (1961). Forecasts are equal to expected values given by the model, conditional on available information. This implies that agents have substantial cognitive capabilities. They know the true model of the economy and correct parameter values. The learning literature relaxes this last assumption. Agents do not know parameters but learn their values by using econometric techniques. It is not trivial that agents are able to learn "true" parameter values. Expectations affect endogenous variables and when agents learn, they modify the way they forecast at each period, and hence the relation between endogenous and exogenous variables. A detailed analysis is provided in Evans and Honkapoja (2001). An important theorem is due to Bray and Savin (1986), later refined by Marcet and Sargent (1989). The theorem states that, if the increase in the forecast of an endogenous variable does not increase too much the actual value, the learning algorithm converges towards the "true" parameter values.

Consistent expectations Many authors have not been satisfied with the REH and have proposed alternatives. My paper heavily borrows from the consistent expectations literature. The literature was initiated by the seminal paper of Brock and Hommes (1997). The concept of consistent expectations equilibrium was defined by Hommes and Sorger(1998). A detailed exposition can be found in the textbook of Hommes (2012). My paper is related to that literature in several ways. First, I borrow the concept of consistent expectations. This concept departs from rational expectation by imposing that expectations should only be consistent with observed autocorrelations. The concept of Consistent Learner Equilibrium developed in this paper is close to the concept of Consistent Expectations Equilibrium. The CLE imposes the additional condition that forecasts using consistent expectations have to be more accurate in average than forecasts of an alternative model e.g. the Rational Learner model. Moreover, I follow the seminal paper of Brock and Hommes by simulating the economy and allowing agents to modify their forecasting model according to the past forecasting performance.

There are however several differences between my paper and previous papers from that literature. I am interested by the convergence towards one type of expectations, either rational expectations or consistent ones. This is why I introduce the concepts of long run consistent learner equilibrium. As compared to the usual CEE concept, I add the condition that consistent forecasts should be more accurate than rational ones in the long run. I provide conditions for the existence of this long run equilibrium. Brock and Hommes (1997, 1998) and Hommes (2012) focus more on the cyclical dynamic, or even the chaotic one, induced by the coexistence of several expectations regimes.

A second, less important, difference is the learning behavior. In this paper, both types learn their model using econometric techniques, like in the Evans and Honkapoja textbook on learning. The criteria by which agents switch from one forecasting model to the other is also different. In Brock and Hommes (1997), they use discrete choice

econometric techniques to choose between the two models whereas I apply a more intuitive criterion in order to make the convergence towards one model easier.

Self fulfilling prophecies Rational expectations were challenged by the sunspot literature initiated by Cass and Shell (1977) and refined by Azariadis and Guesnerie (1982). These two papers have shown that, in some class of models, exogenous variables completely unrelated to endogenous variable may affect them simply because agents believe they do. The idea of this paper is close. Lagged endogenous variable does not affect directly current ones but may through beliefs. The difference is that those lagged endogenous variable are correlated to current ones through the persistence of fundamental exogenous variables. Intuitively, they may play a role in a larger class of models whereas a pure sunspot needs strict conditions to emerge (see Guesnerie 2001 for a review).

Adaptive asset pricing Another paper closely related to mine is a recent paper by Adam and Marcet (2011). In this paper, they compare two learning strategies in an asset market. In the first one, agents learn the relation between price and current dividends. In the second one, agents learn the relation between current and past prices. They show that the second learning strategy offers a simple explanation to many asset pricing puzzles. By contrast, I look at the selection process between different learning strategies and do not compare the ability of two learning models to explain stylized facts. Moreover, in Adam and Marcet, both strategies converge to the rational expectation solution. This is because dividends follow a very simple process in which past dividends are a sufficient statistics to forecast future ones. Under rational expectations, past prices are function of past dividends. Thus, past prices are also a sufficient statistics for future assets prices. In my paper, the alternative strategy does not deliver rational expectations in the long run.

Evolutionary theory and economics This paper adopts an evolutionary viewpoint. The economy is initially composed of different type of agents with different behavioral rules. I study what behavioral rules survive in the long run. The evolutionary viewpoint has a long history in economics. Some intuition may be found in Schumpeter (1926) and in the Austrian school. Friedman (1953) has defended the rationality assumption by suggesting that "rational" agents will eliminate "irrational" ones in markets. The evolutionary viewpoint was formalized in a more rigorous way by Nelson and Winter (1982) and more recently by Saint Paul (2015).

1.3 Framework

Consider a model in which a macroeconomic variable x is given by the equation

$$x_t = \alpha + \beta y_t + \lambda x_t^E + u_t \quad (1.1)$$

α is a constant. β and λ are parameters. u_t is a white noise of standard deviation σ_u . The value of u_t is not observable by agents before they make their decision in period t .

y is an exogenous variable which follows an autoregressive process

$$y_t = \theta y_{t-1} + \epsilon_t \quad (1.2)$$

θ is a parameter and ϵ_t is a white noise of standard deviation σ_y

x_t^E is the expectation of the value of x in t . There is a mass one of agents. Each of them form a forecast. x_t^E is the aggregate forecast.

$$x_t^E = \int_0^1 x_t^e(i) di \quad (1.3)$$

$x_t^e(i)$ is the individual expectation formed at period t by agent i .

Agents observe current values of exogenous variables and lagged endogenous variables.

Formally, their information set can be summarized by

$$\Theta_t = \{y_T, x_{T-1}\}_{T=0}^{T=t}$$

Hypothesis 1.1. *Parameters β , α , λ , θ are positive. Parameters θ and λ are strictly inferior to one.*

$$0 \leq \lambda < 1$$

$$0 \leq \theta < 1$$

Parameters β and α are assumed to be positive for convenience only. Extending the analysis to negative values does not raise any specific issue. The assumption made for θ and λ is more important. θ is the persistence coefficient of an AR(1) process. A value between -1 and 1 is necessary to make sure that the process is stationary. The expectation feedback parameter λ has to be inferior to one. Otherwise, the standard learning algorithm does not converge.

I restrict my analysis to values between 0 and 1 for both λ and θ . Many of the results would hold for other values but positive values seem more relevant for the purpose of this paper. I discuss this choice in the last section.

Rational expectations solution The rational expectation for the model solution is given by

$$E(x_t) = \frac{\alpha}{1 - \lambda} + \frac{\beta}{1 - \lambda} y_t$$

The equilibrium value of the variable x under rational expectations follows

$$x_t = \frac{\alpha}{1-\lambda} + \frac{\beta}{1-\lambda}y_t + u_t$$

The "rational" forecast To form a rational expectation, an agent has to know the values of $\frac{\alpha}{1-\lambda}$ and $\frac{\beta}{1-\lambda}$. What would happen if he does not? The learning theory has been built to address this question.

Agents try to learn the values of the two parameters. In period t , using the history of variables x and y , they estimate the model

$$x_k = \pi_f y_k + \varphi_f + u_{f,k}$$

where k goes from period 0 (where observables start) to period $t - 1$.

After estimating the model, they make a forecast for x_t . This forecast is denoted by $x_{f,t}$.

$$x_{f,t} = \pi_{f,t} y_t + \varphi_{f,t}$$

I call these agents "rational learners".

The alternative forecast This learning strategy converges towards the "good" solution (according to economic theory). But, does it allow agents to make the more accurate forecasts ?

I explore the possibility of an alternative strategy providing better forecasts. Agents do not learn the "true" model but learn autocorrelation for endogenous variables. I label it as the consistent strategy. Individuals which adopt this strategy are called "consistent learners".

I now define the consistent learning and the consistent forecast.

At each period, using the history of variable x , consistent agents estimate the model.

$$x_k = \pi_o x_{k-1} + \varphi_o + u_{o,t}$$

In period t , their forecast for x is

$$x_{o,t} = \pi_{o,t} x_{t-1} + \varphi_{o,t}$$

It seems difficult to believe that the consistent strategy could deliver more accurate forecasts than the rational one for such a simple model. Indeed, it does not use all available information. In particular, it does not take into account contemporaneous innovation on y whereas it takes into account u_{t-1} , which is the past innovation on x and should not be relevant to forecast x_t . This reasoning is true if agents are fully informed about parameter values. But it is more complicated if they should first learn these values.

1.4 Long Run Equilibrium

In this section, I study the long run behavior of the economy. I define long run as a situation in which both types of agents have a stable estimation of their respective models and in which there is only one type of agent remaining.

There is a long run equilibrium if the dominant type makes more accurate forecasts in average than the other type. Thus, there are no incentives to deviate from the more accurate model.

Two results emerge. First, the situation in which rational learners dominate is always an equilibrium. Indeed, if there are only rational learners, their model is correctly specified and their average forecast error is equal to the standard deviation of the

white noise u . Consistent learners have a misspecified model and make larger errors in average. I label such a situation as a "Rational Learner Equilibrium" (RLE thereafter).

A more interesting result is that another equilibrium may exist, in which all agents are consistent learners. Intuitively, if all agents are consistent learners, rational learners learn a misspecified model because they do not take into account that there is another type. As a consequence, their forecasts are not necessarily more accurate than those of consistent learners.

I define the two equilibrium concepts in the first subsection before characterizing them in the second and the third subsections.

1.4.1 Equilibrium definition

First I define the two types of long run equilibrium. In following definitions and propositions, the limit of a sequence of random variables is a random variable towards which the sequence converges in probability. I define the two matrix of observable variables $Z_{o,T}$ and $Z_{f,T}$. These matrix have T columns and these columns are observations respectively for vectors $\begin{pmatrix} 1 \\ x_{t-1} \end{pmatrix}$ and $\begin{pmatrix} 1 \\ y_t \end{pmatrix}$. I also define the vector X_T which is the vector column for observations of x from 0 to T .

Definition 1.1. *A Long Run Rational Learner Equilibrium is a couple of vectors (π_o, φ_o) , (π_f, φ_f) for which*

1. $(\pi_o, \varphi_o) = \lim_{T \rightarrow \infty} \frac{1}{T} (Z'_{o,T} Z_{o,T})^{-1} (Z'_{o,T} X_T)$
2. $(\pi_f, \varphi_f) = \lim_{T \rightarrow \infty} \frac{1}{T} (Z'_{f,T} Z_{f,T})^{-1} (Z'_{f,T} X_T)$
3. $\forall T \quad x_T = \alpha + \lambda \varphi_f + (\beta + \lambda \pi_f) y_T + u_T$
4. $\lim_{T \rightarrow \infty} \sum_{k=0}^T \frac{1}{T} [(x_k - \varphi_f - \pi_f y_k)^2] < \lim_{T \rightarrow \infty} \sum_{k=0}^T \frac{1}{T} [(x_k - \varphi_o - \pi_o x_{k-1})^2]$

The consistent learner equilibrium is defined similarly

Definition 1.2. A Long Run Consistent Learner Equilibrium is a couple of vectors (π_o, φ_o) , (π_f, φ_f) , belonging to \mathbb{R}^2 for which

1. $(\pi_o, \varphi_o) = \lim_{T \rightarrow \infty} \frac{1}{T} (Z'_{o,T} Z_{o,T})^{-1} (Z'_{o,T} X_T)$
2. $(\pi_f, \varphi_f) = \lim_{T \rightarrow \infty} \frac{1}{T} (Z'_{f,T} Z_{f,T})^{-1} (Z'_{f,T} X_T)$
3. $\forall t \quad x_T = \alpha + \lambda \varphi_o + \beta y_T + \lambda \pi_o x_{T-1} + u_T$
4. $\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=0}^T [(x_k - \varphi_f - \pi_f y_k)^2] > \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=0}^T [(x_k - \varphi_o - \pi_o x_{k-1})^2]$

In a nutshell, an equilibrium exists if learning algorithms have converged, if there is only one type remaining and if this type makes more accurate forecasts in average. The last conditions states that there are no incentives to deviate from the dominant forecasting model and it is a stable situation because the two models remain unchanged and so their forecasting performance.

Thereafter, I will use the term equilibrium as a synonym of long run equilibrium.

1.4.2 The Rational Learner Equilibrium

In the two following sections, I will construct explicitly rational learner and consistent learner equilibrium. The difficulty is that in each case, the equation (1.1) giving the endogenous variable x depends on forecasting parameters (π_f, φ_f) and (π_o, φ_o) whereas these forecasting parameters depends on correlation between x_t , y_t and x_{t-1} which are determined by the equation (1.1). Thus, equilibrium equation and forecasting parameters should be computed simultaneously. Then, I have to check if forecast errors of rational (resp. consistent) learners are actually below forecast errors of consistent (resp. rational) learners. From forecasting parameters and implied forecast equations, I derive closed form expression for forecasts errors of both types in both equilibrium. I am able to give explicit condition for the existence of the two equilibrium.

Equilibrium equation and rational learners forecasts

I first construct the Rational Learner Equilibrium.

Proposition 1.1. *The equilibrium equation of the RLE is given by*

$$x_t = \alpha + \lambda\varphi_f + (\beta + \lambda\pi_f)y_t + u_t \quad (1.4)$$

Proof Under the RLE, there are only rational agents. Their forecast is given by $x_t^e = \pi_f y_t + \varphi_f$. Reintroducing into the equilibrium equation (1.1) gives the equality above \square

Parameters of the forecasting equation of rational learners π_f and φ_f can be computed by using standard formulas for univariate ordinary least squares.

Proposition 1.2. *Forecast parameters for rational agents are given by*

$$\pi_f = \frac{\beta}{1 - \lambda} \quad (1.5a)$$

$$\varphi_f = \frac{\alpha}{1 - \lambda} \quad (1.5b)$$

Proof The detailed proof is given in appendix A.1.1. It is however useful to give some hint. π_f is the outcome of an univariate linear regression : $\pi_f = \frac{Cov(x_t, y_t)}{V(y_t)}$. Computations give

$$Cov(x_t, y_t) = (\beta + \lambda\pi_f)V(y_t)$$

Thus, π_f is the root of the equation $\pi_f = \beta + \lambda\pi_f$, hence $\pi_f = \frac{\beta}{1 - \lambda}$.

The constant of the linear regression φ_f is deduced from the value of π_f \square

If I replace π_f and φ_f by their values in the equilibrium equation, then the law of

motion for the process x_t is given by

$$x_t = \frac{\alpha}{1-\lambda} + \frac{\beta}{1-\lambda}y_t + u_t \quad (1.6)$$

Forecast errors

I now compute forecast errors of rational agents. In period t , the rational forecast is $\pi_f y_t + \varphi_f$ and the actual outcome is x_t . The limit of the average of the squares of forecast errors is

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=0}^T [(x_k - \varphi_f - \pi_f y_k)^2]$$

I compute it using standard results on variance.

Proposition 1.3. *The squared forecast error of the rational learner is asymptotically given by σ_u^2*

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=0}^T [(x_k - \varphi_f - \pi_f y_k)^2] = \sigma_u^2 \quad (1.7)$$

Proof Using equations (1.6) and (1.5), I have

$$\forall k \quad x_k - \varphi_f - \pi_f y_k = u_k$$

Thus

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=0}^T [(x_k - \varphi_f - \pi_f y_k)^2] = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=0}^T u_k^2 = \sigma_u^2$$

□

The intuition behind the result is straightforward. If every agent is a rational learner, the rational learner model is correctly specified and errors are given by the white noise u .

This average error has to be compared with the average error of consistent learners.

Usual formulas for ordinary least squares also allows computing parameter values for the forecasting equation of consistent learners e.g. the vector (φ_o, π_o) .

Proposition 1.4. *Coefficients of the forecasting equation for consistent learners are*

$$\pi_o = \theta - \frac{\sigma_u^2 \theta}{\left(\frac{\beta}{1-\lambda}\right)^2 \frac{\sigma_y^2}{1-\theta^2} + \sigma_u^2}$$

$$\varphi_o = \frac{\alpha(1 - \pi_o)}{1 - \lambda}$$

Proof π_o is also the outcome of an univariate linear regression : $\pi_o = \frac{Cov(x_t, x_{t-1})}{V(x_{t-1})}$.

First, I compute the covariance

$$Cov(x_t, x_{t-1}) = \left(\frac{\beta}{1-\lambda}\right)^2 \theta V(y_t)$$

Then, I compute the unconditional variance of x_t

$$V(x_{t-1}) = V(x_t) = \left(\frac{\beta}{1-\lambda}\right)^2 V(y_t) + \sigma_u^2$$

Using standard results on autoregressive process, I have $V(y_t) = \frac{\sigma_y^2}{1-\theta^2}$ After computing π_o , I find the value given in proposition 1.4

The constant of the linear regression φ_o is given by $E(x_t) - \pi_o E(x_{t-1})$. As $E(x_t) = \frac{\alpha}{1-\lambda}$, I have $\varphi_o = \frac{\alpha}{1-\lambda}(1 - \pi_o)$.

Detailed computations are given in appendix A.1.2 □

Knowing π_o and φ_o allows me to compute forecast errors of consistent learners

Proposition 1.5. *The limit of the average forecast error of consistent learners is given*

by

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=0}^T [(x_k - \varphi_o - \pi_o x_{k-1})^2] = (1 + \pi_o^2) \sigma_u^2 + (1 + \pi_o^2 - 2\pi_o \theta) \left(\frac{\beta}{1 - \lambda} \right)^2 \frac{\sigma_y^2}{1 - \theta^2}$$

Proof The limit of the average of forecast errors squares is equivalent to a variance. Standard properties of variance and covariance may be used. The detailed proof can be found in appendix [A.1.3](#) □

An immediate corollary of this proposition is

Proposition 1.6. *For all vector of parameters $(\beta, \alpha, \sigma_u, \sigma_y, \theta, \lambda)$ belonging to the set defined in hypothesis [1.1](#), if the equilibrium equation is given by [\(1.4\)](#), forecasting parameters for rational learners are given by [\(1.5\)](#), forecasting parameters for consistent learners are given by proposition [1.4](#) and forecast errors of rational learners are lower in average than forecast errors of consistent learners*

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=0}^T [(x_k - \varphi_f - \pi_f y_k)^2] > \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=0}^T [(x_k - \varphi_o - \pi_o x_{k-1})^2]$$

Proof It follows immediately from proposition [1.5](#) and [1.3](#). See appendix [A.1.4](#) for details. □

Scope of the RLE

Having closed form expressions for the forecasting equation the and forecast error of both rational and consistent learners, I determine the set of parameters for which a RLE exists.

Proposition 1.7. *The vector (φ_f, π_f) defined by proposition [1.2](#), the vector (φ_o, π_o) defined by proposition [1.4](#) and the process x_t defined by equation [\(1.6\)](#) form a Rational Learner Equilibrium*

Proof Assume that (φ_f, π_f) is given by proposition 1.2. I show that all conditions stated in definition 1.1 are fulfilled.

The process for x_t is given by equation (1.4) fulfilling the third condition.

I showed in the proof of proposition 1.2 that the asymptotic least square estimator $\lim_{T \rightarrow \infty} \frac{1}{T} (Z'_{o,T} Z_{o,T})^{-1} (Z'_{o,T} X_T)$ is actually the vector (φ_f, π_f) found in proposition 1.2. The second condition is fulfilled.

I showed in the proof of proposition 1.4 that the asymptotic least square estimator $\lim_{T \rightarrow \infty} \frac{1}{T} (Z'_{o,T} Z_{o,T})^{-1} (Z'_{o,T} X_T)$ is the vector (φ_o, π_o) found in proposition 1.4. The first condition is fulfilled.

I showed in the proof of proposition 1.6 that forecast errors of rational learners are lower in average than forecast errors of consistent learners fulfilling the fourth condition \square

A corollary of this proposition is that a rational learner equilibrium always exists.

Corollary 1. *For all vector of parameters $(\beta, \alpha, \sigma_u, \sigma_y, \theta, \lambda)$ belonging to the set defined in hypothesis 1.1, a Rational Learner Equilibrium exists.*

This existence theorem states that the rational learner equilibrium always exists. This is an intuitive result. Let me assume that rational learners make forecasts using the vector (π_f, φ_f) defined in proposition 1.2. They actually make rational expectations. Thus, no alternative model may deliver more accurate forecast in average in the long run.

1.4.3 The Consistent Learner Equilibrium

The analysis of the CLE follows the same path, but finding forecasting parameters (π_o, φ_o) is more challenging than for the RLE.

Equilibrium equation and consistent forecasts

I first compute parameters for the equilibrium equation and for the forecasting equation of consistent learners.

Proposition 1.8. *The equilibrium equation is given by*

$$x_t = \alpha + \lambda\varphi_o + \beta y_t + \lambda\pi_o x_{t-1} + u_t \quad (1.8)$$

Proof Under consistent equilibrium, there are only consistent agents. Their forecast is given by $x_{o,t} = \pi_o y_t + \varphi_o$. Reintroducing into equilibrium equation (1.1) gives the equation above. \square

Then, I have to find π_o and φ_o . Several intermediate results are needed.

First, the process governing the variable x is a first order autoregressive process whose persistence coefficient is $\lambda\pi_o$. I conjecture that this process is stationary

Conjecture 1.

$$\lambda\pi_o < 1$$

I verify that this conjecture is true in proposition 1.14.

Consistent learners believe that the variable x follows an autoregressive process. The challenge is to compute their long run estimation of the persistence coefficient π_o . I state the following proposition.

Proposition 1.9. *The asymptotic least square estimator of π_o is given by*

$$\frac{\text{Cov}(x_t, x_{t-1})}{V(x_{t-1})} = \frac{\beta^2(\lambda\pi_o + \theta)V(y(t)) + \lambda\pi_o(1 - \lambda\pi_o\theta)\sigma_u^2}{\beta^2(1 + \lambda\pi_o\theta)V(y(t)) + (1 - \lambda\pi_o\theta)\sigma_u^2} \quad (1.9)$$

Proof Detailed proof is given in appendix [A.1.5](#) □

The above expression for π_o itself contains π_o . Finding the true value of this parameter requires to solve the equation.

Proposition 1.10. *Possible values for π_o are given by the roots of the following equation whose unknown is denoted by π*

$$\pi = \frac{\beta^2 V(y(t))(\theta + \lambda\pi) + \lambda\pi(1 - \lambda\pi\theta)\sigma_u^2}{\beta^2 V(y(t))(1 + \lambda\pi\theta) + (1 - \lambda\pi\theta)\sigma_u^2} \quad (1.10)$$

For each possible value of π_o , φ_o is given by $\varphi_o = \frac{\alpha(1-\pi_o)}{1-\lambda}$

Proof If π_o is one of the roots of the equation (1.10), then the asymptotic least square estimator of π_o is actually equal to π_o .

The value of φ_o is deduced from π_o as in the previous proof. □

It is possible to show that the equation for π_o is actually a quadratic equation.

Proposition 1.11. *equation (1.10) is a quadratic equation and may be written*

$$A\pi^2 + B\pi + C = 0$$

where

$$A \equiv \lambda\theta[\beta^2 V(y_t) - \sigma_u^2(1 - \lambda)]$$

$$B \equiv (\beta^2 V(y_t) + \sigma_u^2)(1 - \lambda)$$

$$C \equiv -\beta^2\theta V(y_t)$$

Proof See proof in appendix [A.1.6](#) □

For a given vector of parameters $(\beta, \alpha, \sigma_u, \sigma_y, \theta, \lambda)$, there are several possible values for π_o . A CLE may exist for the two roots. Assume that π_o is equal to one of these two roots, proposition 1.9 shows that the asymptotic least square estimator for π_o is

actually equal to that root, confirming the belief of consistent learners. If for this value of π_o , consistent learners make lower forecast errors in average than rational learners, then a consistent equilibrium exists.

The next step is to compute forecast errors. I have to choose a value for π_o . I restrict the analysis to a parameter set for which equation (1.10) has a positive and a negative root. I focus on the positive root.

It does not mean that a CLE cannot exist for the negative root or outside the parameter set defined by hypothesis 1.2. I do not explore this possibility. I derive *sufficient* conditions not *necessary* ones. I discuss this choice in the last section.

The following assumption is made

Hypothesis 1.2.

$$\beta^2 \frac{\sigma_y^2}{1 - \theta^2} > \sigma_u^2 (1 - \lambda) \quad (1.11)$$

This hypothesis states that the parameter A defined in proposition 1.11 is positive. Intuitively, it means that the fraction of the variance of x due to innovations in u should not be too large as compared to the fraction of variance due to innovations in y . From this assumption, I deduce the next proposition.

Proposition 1.12. *Equation (1.10) has one positive and one negative real root. The negative root is $\frac{-B - \sqrt{B^2 - 4AC}}{2A}$. The positive root is*

$$\frac{-B + \sqrt{B^2 - 4AC}}{2A}$$

where, A , B and C are parameters defined in proposition 1.11

Proof Hypothesis 1.2 implies that the coefficient A is positive. Hypothesis 1.1 implies that C is negative and B is positive. Thus $B^2 - 4AC > B^2 > 0$. Thus, there are two roots. The root $\frac{-B - \sqrt{B^2 - 4AC}}{2A}$ is negative and the other one positive. \square

A closed form expression may be computed for my chosen value of π_o .

Proposition 1.13. π_o is given by

$$\pi_o = \frac{-B + \sqrt{B^2 - 4AC}}{2A}$$

where, A , B and C are parameters defined in proposition 1.11.

φ_o is given by

$$\varphi_o = \frac{\alpha(1 - \pi_o)}{1 - \lambda}$$

The reader should keep in mind this is one of the possible values for π_o but not the only one.

It also allows me to verify the conjecture 1.

Proposition 1.14. π_o is strictly inferior to one. A fortiori, given hypothesis 1.1

$$\lambda\pi_o < 1$$

Proof π_o is the root of the quadratic equation defined in proposition 1.11. I study the function associated to this quadratic equation and look at the zeros. I show that the value of the function is strictly positive in 1 and that the derivative is strictly positive for values superior to one. Thus, the function is strictly positive for all values superior to one and values of π for which the function is equal to zero are necessarily below 1. Detailed computations are given in appendix □

Forecast errors

Having an analytic expression for π_o , I compute forecast errors for consistent and rational learners. Before giving an explicit formula, I highlight some important intermediate results

Proposition 1.15. *The variance of x_t is given by*

$$V(x_t) = \frac{(1 - \lambda\pi_o\theta)\sigma_u^2 + (1 + \lambda\pi_o\theta)\beta^2V(y_t)}{(1 - \lambda^2\pi_o^2)(1 - \lambda\pi_o\theta)}$$

The covariance between y_t and x_t is

$$Cov(y_t, x_t) = \frac{\beta}{1 - \lambda\pi_o\theta}V(y_t)$$

Proof Formulas for $V(x_t)$ and $Cov(y_t, x_{t-1})$ are given in the proof of proposition 1.9.

The two formulas highlighted above are deduced from them. \square

Proposition 1.16. *Squared Errors of consistent learners are given by*

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=0}^T [x_k - \varphi_o - \pi_o x_{k-1}]^2 = \beta^2 V(y_t) + (1 - \lambda)^2 \pi_o^2 V(x_t) + \sigma_u^2 - 2(1 - \lambda)\pi_o\beta\theta Cov(y_t, x_t) \quad (1.12)$$

Proof I use the same line of reasoning as in the previous section. I show that the limit of the average of the squares of forecast errors is equivalent to a variance and then use properties of variance and covariance. The detailed proof can be found in appendix A.1.8 \square

Forecast errors of consistent learners have to be compared with those of rational agents. Following propositions deal with the rational learner estimation in the long run and their forecast errors.

Proposition 1.17. *If the variable x is given by equation (1.8), estimates for the vector*

(φ_f, π_f) are given by

$$\pi_f = \frac{\beta}{1 - \lambda\pi_o\theta}$$

$$\varphi_f = \frac{\alpha}{1 - \lambda}$$

Proof π_f is equal to

$$\pi_f = \frac{Cov(x_t, y_t)}{V(y_t)}$$

The formula for $Cov(x_t, y_t)$ is given in proposition 1.15. The value of π_f follows immediately and I deduce the value of φ_f from the value of π_f . \square

Then, I can compute forecast errors of rational learners.

Proposition 1.18. *Squared Errors of rational learners are given by*

$$\lim_{T \rightarrow \infty} \sum_{k=0}^T \frac{1}{T} [x_k - \varphi_f - \pi_f y_k]^2 = (\pi_f - \beta)^2 V(y_t) + (\lambda\pi_o)^2 V(x_t) + \sigma_u^2 - 2\lambda\pi_o(\pi_f - \beta)\theta Cov(x_t, y_t)$$

Proof The detailed proof can be found in appendix A.1.9 \square

Scope of the consistent equilibrium

Values found in the previous section for forecast errors of both type allows me to give an existence condition for the CLE

Proposition 1.19. *Under hypothesis 1.2, the equation (1.8), the vector (π_o, φ_o) defined in proposition 1.13, the vector (π_f, φ_f) defined in proposition 1.17 and the equilibrium equation (1.8) form a Consistent Learner Equilibrium if*

$$\begin{aligned}
& (\pi_f - \beta)^2 V(y_t) + (\lambda \pi_o)^2 V(x_t) + \sigma_u^2 - 2\lambda \pi_o (\pi_f - \beta) \theta Cov(x_t, y_t) \\
& \geq \beta^2 V(y_t) + (1 - \lambda)^2 \pi_o^2 V(x_t) + \sigma_u^2 - 2(1 - \lambda) \pi_o \beta \theta Cov(y_t, x_t)
\end{aligned} \tag{1.13}$$

Proof Let me assume that (φ_o, π_o) is given by proposition 1.13. I show that all the conditions stated in definition 1.2 are fulfilled.

The process for x_t is given by equation (1.8) fulfilling the third condition.

I showed in the proof of proposition 1.9 that the asymptotic least square estimator $\lim_{T \rightarrow \infty} \frac{1}{T} (Z'_{o,T} Z_{o,T})^{-1} (Z'_{o,T} X_T)$ is actually the vector (φ_o, π_o) found in proposition 1.13. The second condition is fulfilled.

I showed in the proof of proposition 1.17 that the asymptotic least square estimator $\lim_{T \rightarrow \infty} \frac{1}{T} (Z'_{f,T} Z_{f,T})^{-1} (Z'_{f,T} X_T)$ is the vector (φ_f, π_f) found in proposition 1.4. The first condition is fulfilled.

The fourth condition is fulfilled if the inequation stated in (1.13) is true as I showed in the proof of proposition 1.18 and 1.16. \square

To get a better idea of the scope of the consistent equilibrium, I compute numerically the difference of squared errors between rational and consistent learners for different values of parameters λ and θ . The calibration of the other parameters is given in table 1.1. The outcome is displayed in figure 1.1. The surface is red when rational forecasts are more accurate and in green when consistent forecast errors are. Thus, the green surface defines the scope of the consistent equilibrium. The figure shows that consistent equilibrium is not a curiosity and exists for a large set of parameters. The figure also shows that consistent forecasts tend to be more accurate when λ , the elasticity of x to expectations, or θ , the persistence coefficient of the exogenous variable, are high.

Variable	Value
σ_y	0.2
σ_u	0.1
α	1
β	1

TABLE 1.1 – Calibration Table for figure 1.1

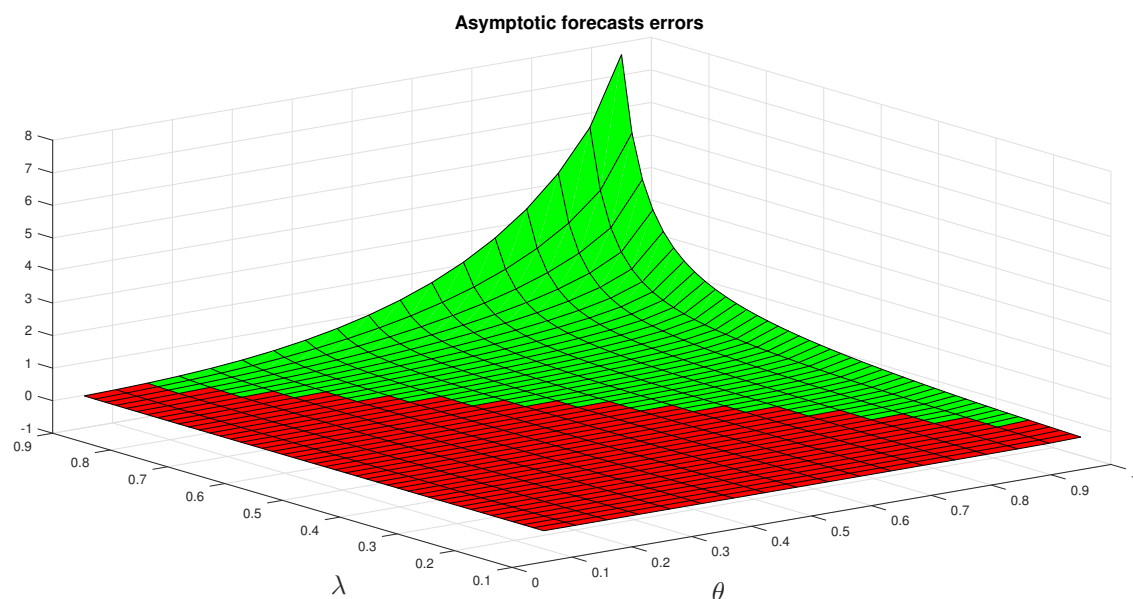


FIGURE 1.1 – Scope of multiple equilibrium

1.5 Simulations

Multiple long run equilibrium exist for a large set of parameter values. May the economy converge towards this consistent equilibrium ? I perform simulations to answer this question.

The economy is composed of three algorithms. At each period, both consistent and rational agents update their model using their forecast errors. In the meantime, they compare the forecasting performance of both models. When a model performs better, more agents adopt it. I describe these features in more details in the following section.

1.5.1 Overview

The rational learner algorithm As mentioned in the previous section, rational learners believe that the variable x can be forecast by estimating the equation

$$x_t = \pi_f y_t + \varphi_f + u_{f,t} \quad (1.14)$$

I now describe the recursive algorithm used by rational learners to estimate this equation.

I define the vector of exogenous variable $z_{f,t}$ and the vector of estimated parameters $\Phi_{f,t}$

$$z_{f,t} \equiv (1 \quad y_t)'$$

$$\Phi_{f,t} \equiv (\varphi_{f,t} \quad \pi_{f,t})'$$

I denote the covariance matrix by $R_{f,t}$. At each period, rational learners update the covariance matrix and their estimation of parameters using newly observed data. Formally, The law of motion for Φ_f and R_f are given by two recursive equations.

$$R_{f,t+1} = R_{f,t} + \frac{1}{t}(z_{f,t}z'_{f,t} - R_{f,t}) \quad (1.15a)$$

$$\Phi_{f,t+1} = \Phi_{f,t} + R_{f,t+1} \frac{1}{t} z_t \left(x_t - z'_t \Phi_{f,t} \right) \quad (1.15b)$$

At period t , the forecast of rational learner is

$$x_{f,t} = \pi_{f,t} y_t + \varphi_{f,t} \quad (1.16)$$

The consistent learner algorithm Consistent learners have a different strategy. They believe that the variable x is given by

$$x_t = \pi_o x_{t-1} + \varphi_o + u_{o,t} \quad (1.17)$$

Like rational learners, they try to learn the value of π_o and the value of φ_o . Both algorithms are similar.

I introduce the vector of exogenous variable $z_{o,t}$ and the vector of estimated parameters $\Phi_{o,t}$

$$z_{o,t} \equiv (1 \quad x_{t-1})'$$

$$\Phi_{o,t} \equiv (\varphi_{o,t} \quad \pi_{o,t})'$$

The variance covariance matrix is $R_{o,t}$ and the recursive estimation is given by

$$R_{o,t+1} = R_{o,t} + \frac{1}{t}(z_{o,t}z'_{o,t} - R_{o,t}) \quad (1.18a)$$

$$\Phi_{o,t+1} = \Phi_{o,t} + R_{o,t+1} \frac{1}{t} z_{o,t} (x_t - z'_{o,t} \Phi_t) \quad (1.18b)$$

At period t , the forecast of consistent learners is

$$x_{o,t} = \pi_{o,t} x_{t-1} + \varphi_{o,t} \quad (1.19)$$

Update of the share of consistent learners At period t , the fraction of consistent learners is γ_t and the fraction of rational learners is $1 - \gamma_t$

At the end of period t , agents observe forecasts of both types $x_{f,t}, x_{o,t}$ and the actual outcome x_t . They have infinite memory.

After t simulated periods, they compute the statistic²

$$\Delta_t = \frac{1}{t} \sum_{k=1}^t \sqrt{(x_{o,k} - x_k)^2} - \frac{1}{t} \sum_{k=1}^t \sqrt{(x_{f,k} - x_k)^2} \quad (1.20)$$

The statistic Δ is the average forecast error of the consistent learner minus the average forecasting error of the rational learner.

It should be noted that I use absolute values to compute errors whereas I used squares in the theoretical section. Squares allow for easier asymptotic computations but may give a too large advantage to consistent learners in simulations. Rational learners often make very large errors at the beginning of a simulation and squares would give too much weight to these initial errors. To show that simulation results are not explained by these large errors, I directly use absolute values.

If $\Delta < 0$, the consistent strategy is in average more accurate than the rational one until period t . A fraction μ of the rational learners shifts to the consistent learner strategy. Conversely, if the rational strategy has been more accurate in average, the same fraction shifts from the consistent strategy to the rational one. Thus, the evolution of γ is given by

$$\gamma_{t+1} = \gamma_t - \mu \gamma_t \mathbb{1}_{\{\Delta_t > 0\}} + \mu(1 - \gamma_t) \mathbb{1}_{\{\Delta_t < 0\}} \quad (1.21)$$

1.5.2 Summary of the model and the algorithm

The structure of the model can be summarized by nine equations.

A first block of equations is composed of equilibrium equations. It includes the two forecasting equations (1.16) and (1.19) and the equilibrium equation for x .

²I also consider the alternative statistic $\Delta_t = \frac{\sum_{k=0}^t \mathbb{1}_{\{(\bar{x}_{o,k} - x_k)^2 < (\bar{x}_{f,k} - x_k)^2\}}}{t} - 0.5$. Which is the number of times for which consistent learner strategy have delivered a more accurate forecasts than the rational learner strategy

$$x_t = \alpha + \beta y_t + \lambda \gamma_t x_{o,t} + \lambda(1 - \gamma_t)x_{f,t} + \zeta_t \quad (1.22)$$

Beside this static block, there are two dynamic blocks.

The first dynamic block includes the two recursive estimation algorithm (1.15a)-(1.15b) and (1.18a)-(1.18b), giving four equations. A second dynamic block gives the evolution of the share of consistent learners. These are equations (1.20) and (1.21)

Description of the economy algorithm I simulate this economy over a long period. The algorithm may be summarized by the following sequence of events

1. Using the model they have chosen, their past estimates of parameter values and the value of y_t , x_{t-1} , rational learners and consistent learners compute their forecasts for x_t
2. The equilibrium value of x_t is computed through equilibrium equation (1.22)
3. This value of x_t is compared with forecasts of both types.
4. If the model of rational learners (respectively consistent learners) under-performs, they switch to the other model with probability μ (respectively $1 - \mu$).
5. Once they have chosen their new model, they estimate it using the history of values for x and y

1.5.3 Initialization

The initial share of rational learners is γ_0 . I set it at 0.5. Half of the population is initially rational. An higher value would give an initial advantage to rational learners³. I perform simulations with different values for γ_0 . This robustness exercise suggests that the initial value of γ has few importance for the long run outcome.

³respectively a lower would give an advantage to consistent learners

A significant practical issue is the initialization of the two learning algorithms. I have to set priors for the two covariance matrix R_f and R_o and the two vectors of parameter estimates Φ_f and Φ_o .

I initialize R_f and R_o by using long run values for mean and variance. The long run values are those of the rational expectation equilibrium.

The matrix $R_{f,0}$ is given by

$$\begin{pmatrix} 1 & E(y) \\ E(y) & E(y^2) \end{pmatrix}$$

or

$$\begin{pmatrix} 1 & 0 \\ 0 & \frac{\sigma_y^2}{1-\theta^2} \end{pmatrix}$$

In a similar way, the matrix $R_{o,0}$ is given by

$$\begin{pmatrix} 1 & E(x) \\ E(x) & E(x^2) \end{pmatrix}$$

or

$$\begin{pmatrix} 1 & \frac{\alpha}{1-\lambda} \\ \frac{\alpha}{1-\lambda} & \left(\frac{\beta}{1-\lambda}\right)^2 \frac{\sigma_y^2}{1-\theta^2} + \sigma_u^2 + \left(\frac{\alpha}{1-\lambda}\right)^2 \end{pmatrix}$$

These long run values are similar for rational learners and consistent ones and thus do not give an advantage to one type. Setting other priors would be risky because the matrix has to be inverted. Bad initial values may lead to inconsistent or explosive estimations.

The long run outcome is sensitive to the initialization of Φ_f and Φ_o and especially to the initialization of φ_f and φ_o . There are several strategies to deal with this sensitivity issue. A first approach is to use values of the rational expectation equilibrium to set

initial guess for Φ_f and Φ_o . But, it gives a strong advantage to rational learners. I adopt a variant of this strategy. I center the prior around the rational expectation value but I allow for perturbations around it, possibly large ones. Perturbations are given by the vector of parameters $(\nu_0, \nu_1, \nu_2, \nu_3)$.

The vector $\Phi_{f,0}$ is equal to

$$\Phi_{f,0} = \begin{pmatrix} \frac{\mu}{1-\lambda}(1 - \nu_0) \\ \frac{\beta}{1-\lambda}(1 - \nu_1) \end{pmatrix} \quad (1.23)$$

The vector $\Phi_{o,0}$ is set in a similar way

$$\Phi_{o,0} = \begin{pmatrix} \frac{\mu(1-\theta)}{1-\lambda}(1 - \nu_2) \\ \theta(1 - \nu_3) \end{pmatrix} \quad (1.24)$$

I constrain the prior for the autoregressive coefficient to be compatible with a stationary autoregressive process $\theta(1 - \nu_3) < 1$.

1.5.4 Baseline calibration

I consider a baseline calibration to get a benchmark result. The calibration is summarized in Table 1.2⁴. The expectation feedback λ is set at 0.5. the persistence of the exogenous variable θ is set at 0.8. This can be considered as a high value but quarterly persistence of shocks in macroeconomic models is often closer to 0.9. I target a 5 percent standard deviation from the average value, in line with macroeconomic volatility. Innovations on y account for eighty percent of that volatility and white noise u count for twenty percent. The initial share of rational learners is 0.5. Parameters $(\nu_0, \nu_1, \nu_2, \nu_3)$ are all equal to the same value given by ν is set at 1. It means that both rational and consistent agents start their estimation by guessing a value of the

⁴with $p_1 = (1 - \theta) * \frac{\alpha}{1-\lambda}$ and $p_2 = \frac{\alpha}{1-\lambda}$.

constant equal to 0 and the value of the coefficient equal to 0.

Beside this baseline calibration, I consider two other calibrations. The first one favors rational learners, hence is called the "rational" calibration. The second one favors consistent learners. The main difference between the rational and the baseline calibration is the value of ν . In the rational calibration, it is set at 0, meaning that initial guesses for vectors (π_f, φ_f) and π_o, φ_o are given by their asymptotic values at the rational expectation equilibrium. The consistent calibration is characterized by an intermediate value for ν but a larger value for the expectation feedback parameter λ .

Variable	Baseline	"rational" calibration	"consistent" calibration
λ	0.5	0.5	0.75
ν	1	0	0.3
γ_0	0.5	0.5	0.5
θ	0.8	0.8	0.8
σ_y^2	$0.04p_1$	$0.04p_1$	$0.04p_1$
σ_u^2	$0.01p_2$	$0.01p_2$	$0.01p_2$
α	$\frac{1}{1-\lambda}$	$\frac{1}{1-\lambda}$	$\frac{1}{1-\lambda}$
β	1	1	1
μ	0.025	0.025	0.025

TABLE 1.2 – Calibration Table for the Simulated Economy

These different calibrations only aim to give a first look at the result. I perform simulations for a much larger set of parameters.

1.5.5 Experiments

I perform several experiments.

1. I first simulate the economy with the different calibrations displayed in table 1.2 to get some intuition about what happens in this economy.
2. The outcome seems dependent from deep parameter values. I explore more systematically the issue by performing simulation for many couple of parameters

(λ, θ) and then for many couple of parameters (σ_u, σ_y) . I represent the result in three dimensional figures.

3. For a given calibration, a simulation may converge towards different equilibrium. There is some history-dependence. To understand the extent of this history-dependence, I perform the same experiment as in the second part but I simulate several times for a given calibration.
4. The outcome seems sensitive to the initialization of the learning algorithms of consistent and rational learners. I simulate many times for different values of γ_0 and ν . I also simulate by using random draws for ν .

1.6 Results

1.6.1 Summary

Results can be summarized as follows

1. Consistent equilibrium is not a theoretical curiosity. Simulation converges towards it for a large set of parameters. This is the case of the baseline calibration in the first experiment.
2. The result of a simulation depends on deep structural parameters. The second experiment show that high values of the expectation feedback parameter λ and of the persistence coefficient θ favor consistent equilibrium.
3. Values of λ and θ for which the economy converges towards the CLE seem close to values for which consistent equilibrium exists. In other words, when the CLE exists, the economy has good chance to converge towards it.
4. The result of a simulation also depends on initialization parameters, in particular on the initial guess of rational learners and consistent learners for their respective

model. A guess close enough to rational expectation value allows rational agents to dominate. In some sense, the rational equilibrium is "locally stable" but not globally.

5. The initial guess of the two constants φ_f and φ_o seems much more important than the initial guess of π_f and π_o .
6. For some calibration, the outcome of the simulation is path dependent. For the same parameters, one simulation may converge towards the RLE and a second one towards the CLE. However, the set of parameters for which there is path dependence seems small.

1.6.2 Main result

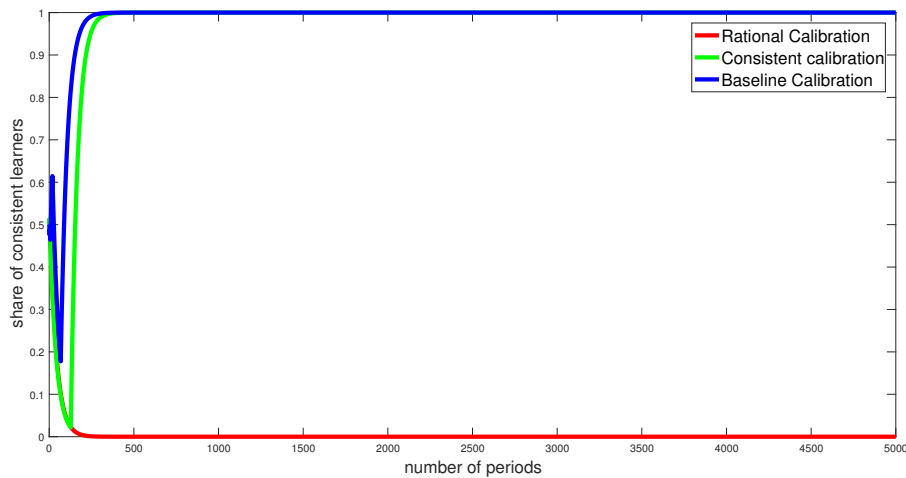


FIGURE 1.2 – Convergence for several calibrations

The main result of my simulations is that the economy does not always converge towards the Rational Learner Equilibrium and often converges towards the Consistent Learner Equilibrium. I illustrate the result by figures 1.2 and 1.3. Figure 1.2 repre-

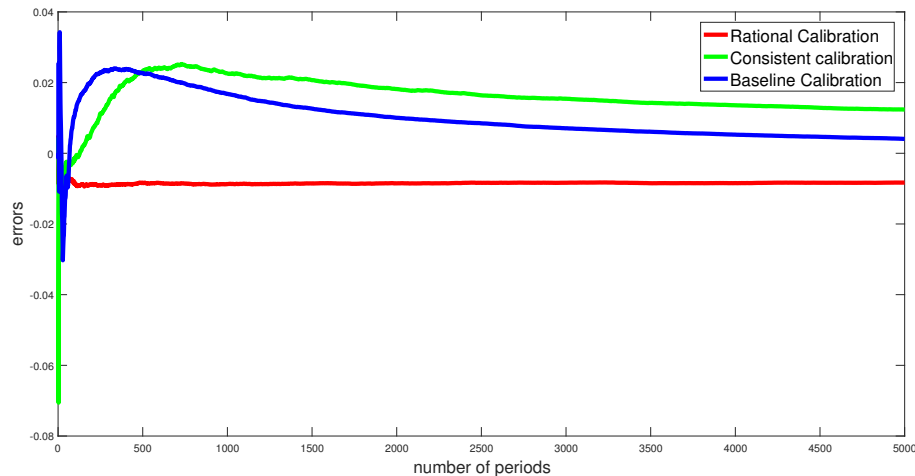


FIGURE 1.3 – Spread between rational and consistent errors for several calibrations

sents the evolution of the share of rational agents for the baseline, the rational and the consistent calibrations. Figure 1.3 displays the difference between average error of rational agents and average error of consistent agents (the error difference thereafter). A value higher than zero implies that rational errors are larger in average than consistent errors. The blue line corresponds to the baseline calibration. Initially, rational learners make smaller errors (by a very small margin) and their share increases but quickly the error difference is reversed and consistent learners become dominant. Error difference seems converging to a value superior to 0, showing that a longer simulation would not lead to a different outcome. The red line corresponds to the rational calibration. Rational agents dominate at the beginning, become the only type after fifty periods and then keep their advantage. The error difference is always inferior to zero, but remains quite small. The green line gives the evolution of the economy for the consistent calibration. It shows that convergence to the CLE may occur for initial parameters much closer to their REE value. The parameter ν is set at 0.3. Thus, components of vectors $\Phi_{o,0}$ and $\Phi_{f,0}$ are set around 70 percent of their REE value. The feedback parameter λ increases to 0.65 which remains a reasonable value. Similarly to the baseline calibration, rational agents have a short initial edge before

being crowded out by consistent agents. Error difference seems converging to a value above zero.

1.6.3 Exploration

Results from the baseline simulation may be particular cases whereas the bulk of the simulations converges towards the RLE. In this paragraph, I outline it is not an isolated result. I explore how changes in parameter values affect the outcome of the simulation. The outcome depends on many parameters : Four deep structural parameters $(\lambda, \theta, \sigma_y, \sigma_u)$ and two initialization parameters ν, γ_0 . Formally, I define the function $G : (\nu, \gamma_0, \lambda, \theta, \sigma_y, \sigma_u) \rightarrow \{0, 1\}$. The value of G is 0 (resp. 1) when the simulation run with parameter values $(\nu, \gamma_0, \lambda, \theta, \sigma_y, \sigma_u)$ converges to the consistent learner equilibrium (resp. the rational learner equilibrium). I fix four parameters at their baseline value and make two others varying. I consider successively variations of (λ, θ) , (σ_u, σ_y) . Results are displayed in figure 1.4 and 1.6. The surface is in green when the simulation converges towards the consistent learner equilibrium and in red if rational learners dominate.

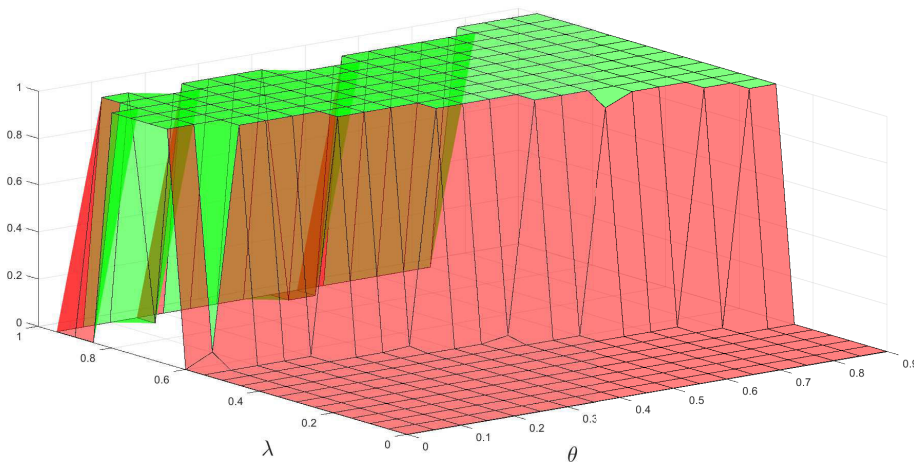
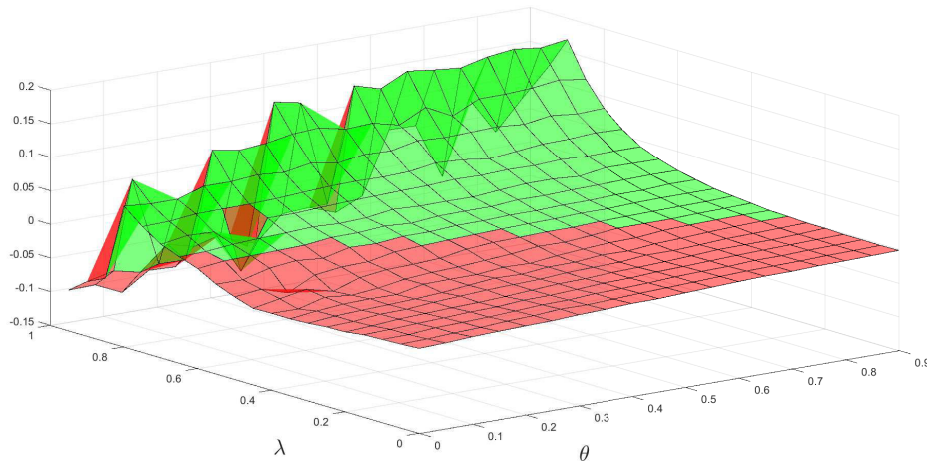


FIGURE 1.4 – Convergence with respect to λ and θ

FIGURE 1.5 – Error difference with respect to λ and θ

The green surface is important in all of these figures. Figure 1.4 shows that convergence to CLE occurs more often when the feedback parameter λ or the persistence coefficient θ are high. Figure 1.5 shows the error difference with respect to λ and θ . The exercise is an "experimental" counterpart to the theoretical result represented in figure 1.1. Experimental and theoretical green surface are quite close. When a consistent equilibrium exists, the economy seems having a high probability to converge towards the consistent equilibrium.

Figure 1.6 shows the sensitivity to standard deviation σ_u and σ_y . Larger standard deviations imply that consistent agents use a noisier information and miss a more important one. Surprisingly, they do not seem very important to determine the outcome of the simulation. Convergence for consistent equilibrium occurs for most values.

1.6.4 The role of initialization

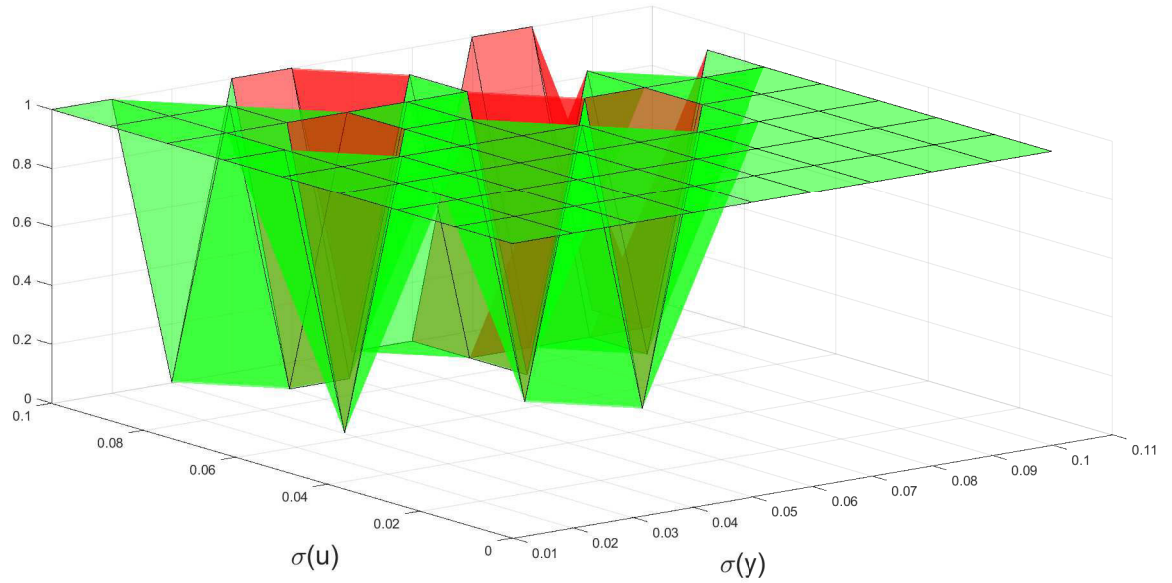
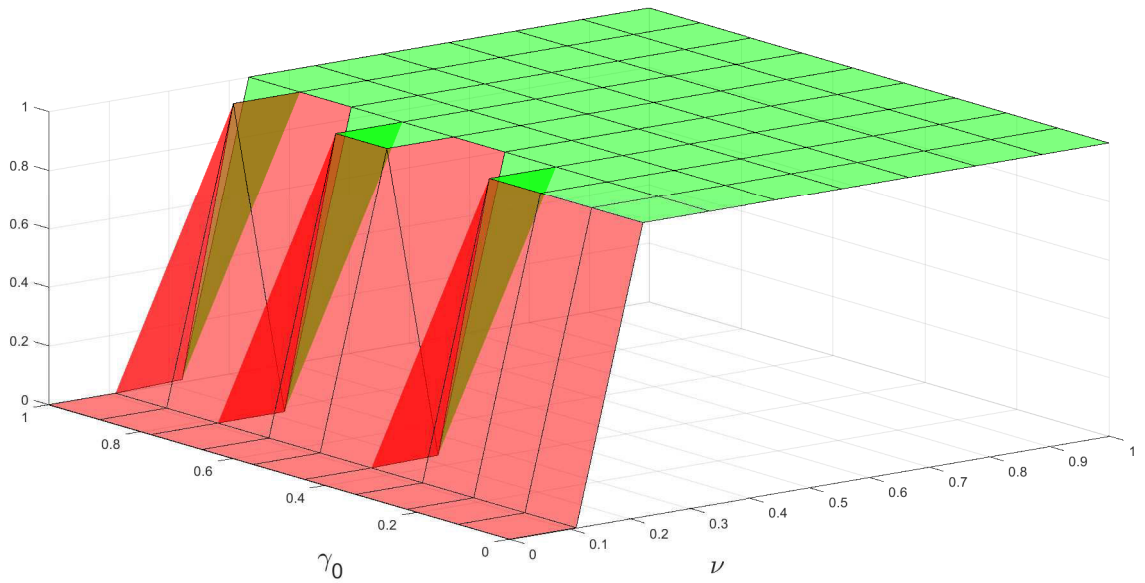
FIGURE 1.6 – Sensitivity to σ_y and σ_u FIGURE 1.7 – Sensitivity to γ_0 and ν

Figure 1.7 shows the sensitivity to initialization parameters. The outcome seems independent from the initial share of rational agents γ_0 , but seems affected by the value of ν . For a value of ν close to zero, implying an initialization close to REE value,

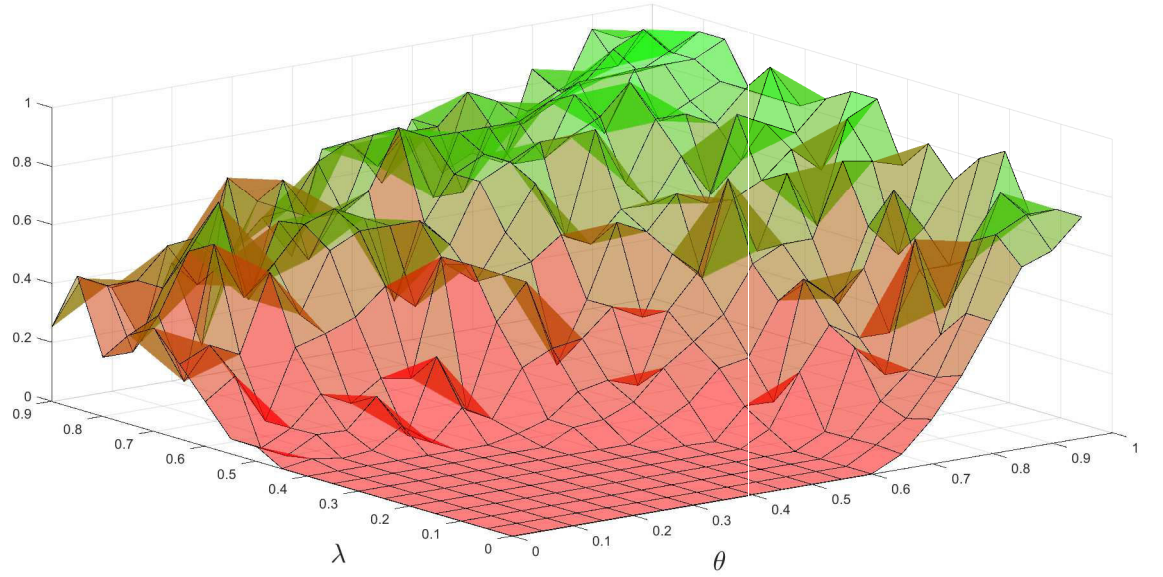


FIGURE 1.8 – Sensitivity to λ and θ when initialization parameters are random

convergence to REE is systematic. This result is reversed when initial parameters are set farther from their REE value.

As the outcome seems sensitive to initialization, I perform additional experiments. In figure 1.8, I represent the percentage of simulations which converges to the CLE with respect to λ and θ when initialization parameters are set randomly. $(\nu_0, \nu_1, \nu_2, \nu_3)$ are random variables and are drawn at each simulation. The result of the previous section is confirmed. Closer are λ and θ from one, higher is the probability to converge towards the consistent learner equilibrium .

I attempt to understand what explains this dependence to initialization. I separate the initial guess of φ_o and φ_f from the initial guess of coefficients π_o and π_f . Initial vectors $\Phi_{f,0}$ and $\Phi_{o,0}$ are respectively equal to $(\frac{\mu}{1-\lambda}(1-\nu_1), \frac{\beta}{1-\lambda}(1-\nu_2))$ and $(\frac{\mu(1-\theta)}{1-\lambda}(1-\nu_1), \theta(1-\nu_2))$. An "accurate" initial guess mean that the guessed value

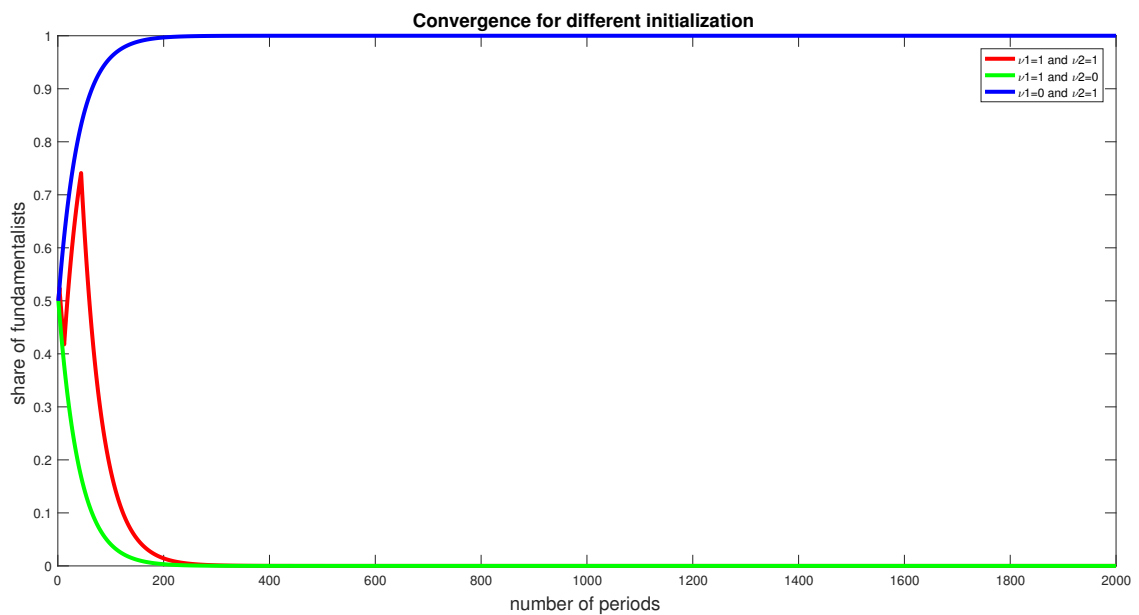


FIGURE 1.9 – Convergence for different initialization

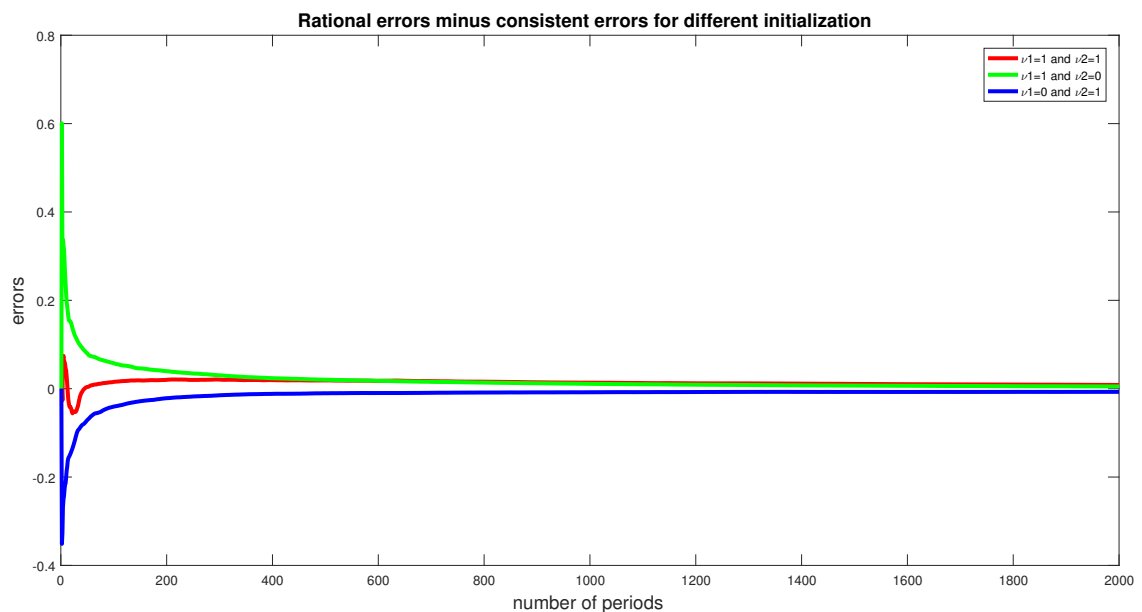


FIGURE 1.10 – Error spread for different initialization

corresponds to the Rational Expectation Equilibrium value. Figure 1.9 and 1.10 represents the evolution of the mass of rational agents and the error difference for three different calibrations of ν_1 and ν_2 . The blue line is the baseline initialization. Initial guessed values are equal to zero. Agents have a very inaccurate guess for both $(\varphi_o, \varphi_f$

and (π_f, π_o) . The economy converges to the consistent equilibrium. The green line represents a simulation in which agents have an accurate initial guess for π_f and π_o but start by believing φ_o and φ_f are equal to zero. Consistent learners still dominate in the long run. In the last simulation, agents have an accurate initial guess of the constant but an inaccurate guess for π_f and π_o . Unlike the two previous cases, the economy converges towards the RLE. Thus, an accurate initial guess of (π_f, π_o) is not sufficient to allow rational learners to dominate whereas an accurate initial guess of (φ_o, φ_f) is.

1.6.5 Path dependence

In what extent the outcome of two simulations may diverge whereas they have the same calibration ? I perform two experiments to answer this question. In the first one, I represent the percentage of simulations which converges to the RLE with respect to λ and θ . For each couple (λ, θ) , I sum the value of γ after 2000 periods and repeat the simulation one hundred times. I obtain the figure 1.11 which is not very different from figure 1.4. Surfaces of zeros and ones are nearly unchanged. However, at the frontier, there are values between 0 and 1, indicating path dependency.

1.7 Extensions: misspecification and structural breaks

1.7.1 Intuition

Until then, I have assumed that rational learners know the true structure of the economy. In practice, every model is misspecified in some extent. For example, rational learners may be unable to observe all relevant exogenous variables. In that case, they no longer learn the rational expectation solution of the model. However, agents still

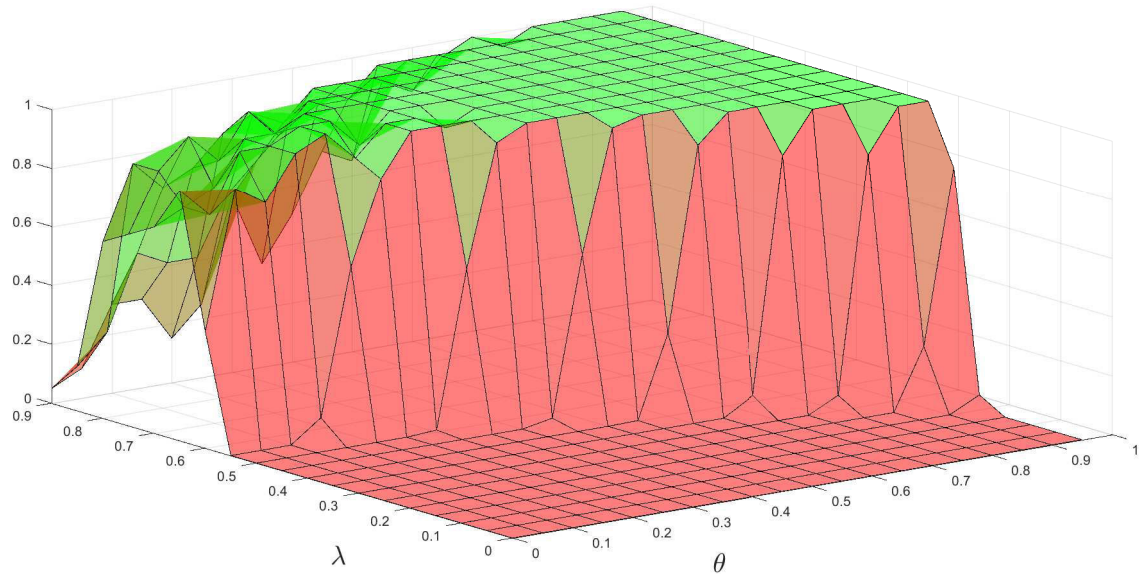


FIGURE 1.11 – Sensitivity to λ and θ : average mass of consistent learners after multiple simulations

face a choice between forecasting with exogenous variables or with lagged endogenous ones. This choice is important. With the rational expectations hypothesis, it is implicitly assumed that economic agents prefer the first option. I have shown in the previous section that the second option may be better even in the extreme case where the model with exogenous variables is correctly specified. The advantage of the consistent behavior could be bigger if the rational learner model is misspecified. Intuitively, past values of the endogenous variables may carry information about omitted variables in the exogenous model. I explore this intuition in the following section. I consider two forms of misspecification. The variable x may be affected by unobserved variables or the parameters of the equation (1.1) can be subject to structural breaks.

1.7.2 Adding persistent unobservables

A first misspecification is the existence of an unobserved exogenous variable. For example, the equation (1.1) becomes

$$x_t = \alpha + \beta y_t + \lambda x_t^E + u_t + v_t \quad (1.25)$$

with

$$v_t = \rho_v v_{t-1} + \epsilon_t^v$$

v is not observable by agents and rational learners continue to estimate the model $x_t = \varphi_f + \pi_f y_t$. Because v is persistent, past values of x carry information about the current value of v , giving an edge to consistent learners over rational learners.

1.7.3 Adding structural breaks

A second misspecification is that some parameters are not constant but time varying and follow for example a Markov chain.

The equation becomes

$$x_t = \alpha_t + \beta y_t + \lambda x_t^E + u_t + v_t \quad (1.26)$$

α_t is a random variable whose support is the vector $\{\alpha_l, \alpha_h\}$, where both α_l and α_h are real numbers.

α_t evolves according to a Markov chain. In the "h" state, the probability to remain in the high state is p_h whereas the probability to remain in the low state is p_l .

Rational learners still estimate the misspecified model $x_t = \varphi_f + \pi_f y_t$

1.7.4 Results

I perform simulations with both misspecification. I calibrate the Markov chain to have a structural break every 100 periods in average. α_l and α_h are three percent deviation from the average value of α . I set σ_v at the same level than σ_u but introduce a small persistence coefficient with $\rho_v = 0.3$. I display the convergence with respect to (θ, λ) in figure 1.12. The two sources of misspecification significantly enhance the dominance of the consistent learner equilibrium.

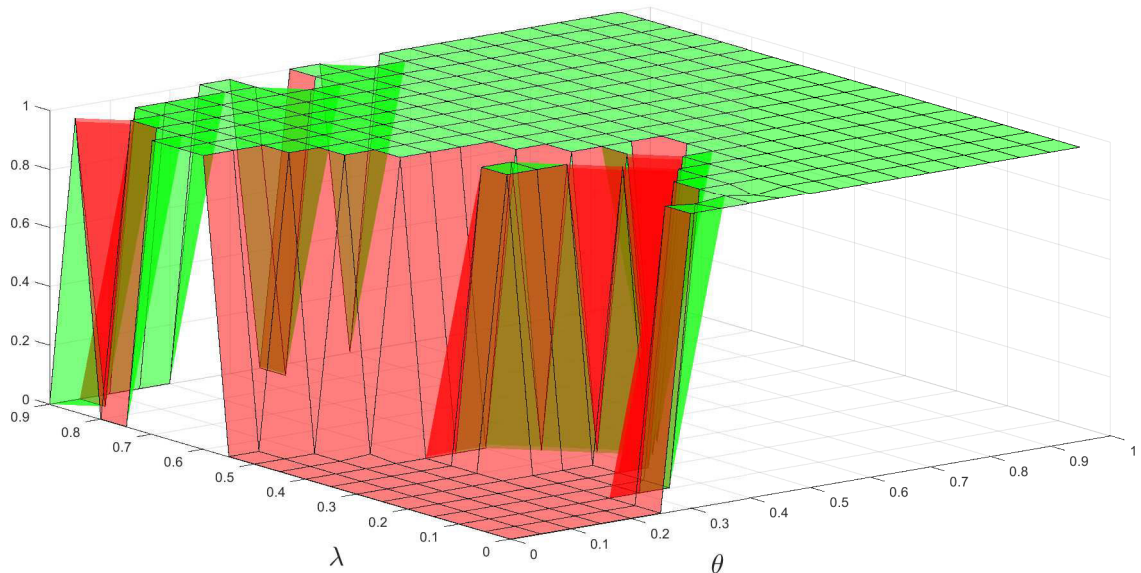


FIGURE 1.12 – Sensitivity of the response to λ and ν with structural breaks

1.7.5 Why do rational learners misspecify their model

It seems implausible to assume that rational learners do not detect the misspecification. Two reasons motivate this assumption.

First, I consider small deviations from the original model. For example, α_l and α_h are three percent deviation from the average value of α and the autoregressive coefficient

ρ_v is only 0.3.

Second, even if they detect a misspecification, they could have serious troubles to identify and estimate the true model. In case of structural breaks, they should estimate no less than five parameters $\pi_f, \alpha_l, \alpha_h, p_h, p_l$. The number is the same if there is an unobserved variable: $\sigma_v, \sigma_u, \rho_v, \varphi_f, \pi_f$. If both misspecification are present, They have eight parameters to estimate. In every case, they still observe two variables. The bottom line is that if agents have no other choice that forecasting with a misspecified model, the model closer to rational expectations is not necessarily the best option.

1.8 Discussion

I now provide an intuition for the existence and the "stability" of the consistent learner equilibrium (e.g. the convergence towards the CLE occurring in simulations). Then, I discuss the implications of several assumptions I made.

Inspecting the result In the first two sections, I compare two agents. Rational learners only use information contained in exogenous variables. If they have the right parameter values, there is a rational expectation equilibrium in which all agents are rational learners. Consistent learners use information contained in previous values of endogenous variables whereas endogenous variables are not directly related to their lagged values. If everyone makes forecasts in that way, the economy may have another equilibrium in which consistent learners forecast more accurately than rational learners. Simulations show that this equilibrium is a genuine possibility and not a mere theoretical curiosity. Why is such equilibrium possible ?

First, the rational learners' performance is lower than expected. Indeed, they do not take into account the existence of another type. If there is a non-zero mass of agents belonging to the other type, their model is misspecified. This may generate

large forecast errors in the long run. Moreover, this creates very large errors at the beginning of the simulation and makes the learning process of rational agents slower.

However, the more important point is that consistent learners perform better than expected. I highlight two reasons for that. First, the theory of consistent learners is self confirming. A positive expectation feedback parameter λ creates some strategic complementarity between agents. If they believe that x_t is positively correlated to x_{t-1} , they expect an higher x_t when x_{t-1} is high. Through expectation feedback, their high expectation for the value of x_t leads to a high actual value for x_t . This strategic complementarity effect is however not sufficient. I need another source of correlation between x_t and x_{t-1} . Otherwise, the theory of consistent learners would be a pure sunspot theory. In my model, a pure sunspot theory would need a λ equal to one to emerge whereas the consistent learner equilibrium may exist for values of λ well below one. The crucial difference between a pure sunspot and the consistent forecasting model is that whereas x_{t-1} does not affect directly x_t , they are positively correlated through the exogenous variable y_t . Indeed, y_t is an AR(1) process whose persistence parameter θ is positive. When y_{t-1} is high, both x_{t-1} and y_t are high, leading to a positive correlation between x_t and x_{t-1} . In other words, x_{t-1} carries some information about the value of x_t . Consistent learners miss some important information, e.g contemporaneous innovations on y and use an irrelevant information (e.g innovation on x_{t-1}) but they use indirectly some information contained in exogenous variable through their correlation with x_{t-1} . As a consequence, the process y and the persistence parameter θ are crucial for my result. With a θ equal to zero, consistent agents learn a pure sunspot and the consistent learner equilibrium would require a value of λ equal to one. Higher is θ , lower is the minimal value of λ for which the CLE emerges.

It should also be noted that the white noise u does not play any role in my results. It actually provides an advantage to rational learners because it reduces the correlation between x_t and x_{t-1} . I include it in the model to show that my results can be obtained

in a general setup.

Implications of some assumptions The role of complementarity and persistence also explains my first assumption in which I restrict the analysis to values of λ and θ between 0 and 1. Intuitively, a negative value for one of the two parameters would make more difficult the existence of a CLE. How restrictive are these two parametric assumptions? They seem reasonable if we have some macroeconomic application in mind. Exogenous disturbances are usually highly persistent in macroeconomic models. Strategic complementarity may also occur in those models, for example through aggregate demand if prices and wages are rigid or sticky.

My choice of focusing on positive values for θ and λ also explains why I choose to concentrate on the positive root of the equation giving possible values of π_o . A negative value for π_o would imply a negative correlation between x_t and x_{t-1} whereas a positive θ would suggest a positive correlation between the two. It is fairly possible that a consistent learner equilibrium may also emerge for a negative value of θ and a negative root for π_o . This paper does not explore this possibility.

1.9 Conclusion

In this paper I have shown that rational expectations may not be evolutionary dominant. Agents using a misspecified model can forecast more accurately, leading rational agents to adopt the misspecified model. This result suggests that the Consistent Expectations Hypothesis proposed by Hommes and Sorger (1998) is a genuine alternative to the Rational Expectation Hypothesis. This is particularly true when there is a large positive feedback effect from expectations to endogenous variables and when exogenous variables are persistent.

This result is obtained in a simple model. In more complex ones, learning the rational

expectation model is difficult. I show that it could be better to learn a misspecified model based on lagged endogenous variable rather than a misspecified model based on exogenous variables and thus closer to the rational expectation solution.

The interesting point is that large positive feedback from expectations or high persistence of exogenous variables are common in macroeconomics and finance. The next step of the research agenda is to verify that consistent expectations may be evolutionary dominant in a simple asset price model or a simple macroeconomic model.

Bibliography

- [1] **Adam, K. and A. Marcet.** 2011. "Internal rationality, imperfect market knowledge and asset prices." *Journal of Economic Theory* 146, 1224-1252
- [2] **Bray, M., and N. Savin.** 1986. "Rational expectations equilibria, learning, and model specification." *Econometrica* 54, 1129-1160
- [3] **Brock, W.A. and C.H. Hommes.** 1997. "A rational route to randomness." *Econometrica* 65, 1059-1095
- [4] **Bullard, J. and J. Suda.** 2011. "The stability of Macroeconomic systems with Bayesian learners." *Banque de France working papers*
- [5] **Cass, K. and A. Shell.** 1983. "Do sunspots matter ?" *Journal of Political Economy* 91, 193-227
- [6] **Evans, G., and S. Honkapoja.** 2001. *Learning and expectations in Macroeconomics*. Princeton University Press
- [7] **Friedman, M.** 1957. *A Theory of the Consumption Function..* Princeton University Press.
- [8] **Grandmont, M.** 1998. "Expectations Formation and Stability of Large Socioeconomic Systems." *Econometrica* 66, 741-781
- [9] **Guesnerie, R.** 2001. *Assessing Rational Expectations: Sunspot Multiplicity and Economic Fluctuations*. MIT Press
- [10] **Guesnerie, R.** 2005. *Assessing Rational Expectations 2-Eductive stability in economics*. MIT Press

-
- [11] **Guesnerie, R. and C. Azariadis** 1982. "Propheties creatrices et persistance des theories." *Revue Economique* 33, 787-806
- [12] **Guesnerie, R. and M. Woodford** 1991. "Stability of Cycles with Adaptive Learning Rules." in *Equilibrium Theory and Applications* ed. by W. Barnett, B. Cornet, C. d'Aspremont, J. Gabszewicz, and A. Mas-Colell. Cambridge University Press.
- [13] **Hommes, C.** 2013. *Behavioral Rationality and Heterogeneous Expectations in Complex Economic Systems*. Cambridge University Press.
- [14] **Hommes, C. and G. Sorger** 1998. "Consistent Expectations Equilibria." *Macroeconomic Dynamics* 2, 287-321
- [15] **Marcet, A., and J.P. Nicolini.** 2003. "Recurrent Hyperinflation and Learning." *American Economic Review* 93(5), 1476-1498
- [16] **Marcet, A., and T. Sargent.** 1989. "Convergence of least squares learning mechanisms in self referential linear stochastic models." *Journal of economic Theory* 48,337-368
- [17] **McGough, B.** 2003. "Statistical learning with time varying parameters." *Macroeconomic Dynamics* 7, 119-139
- [18] **Muth, J.** 1961. "Rational expectations and the theory of price movements." *Econometrica* 29,315-335
- [19] **Nelson, R. and S.G. Winter** 1982. *An Evolutionary Theory of Economic Change*. Harvard University Press.
- [20] **Saint Paul, G.** 2014. "How Darwinian should an economy be?" PSE working papers.
- [21] **Woodford, M.** 1990. "Learning to believe in sunspots." *Econometrica* 58, 277-307.

Chapter 2

Intertemporal Complementarity and the New Keynesian model

2.1 Introduction

This paper explores an alternative specification of the utility function and its theoretical implications for the new keynesian model.

A recent strand of literature has introduced wealth in the utility function. Krishnamurthy and Vissing-Jorgensen (2012) use bonds in the utility function and infer demand functions to explain the behavior of the bond market. Michaillat and Saez (2014, 2018) explore consequences for the New Keynesian model, especially at the zero lower bound. Saez and Stantcheva (2017) look at implications for optimal capital taxation. Kumhof, Ranciere and Winant (2015) use it to match profiles of income distribution in a model of financial crisis. Michau (2017) introduces a similar specification in a model with downward nominal wage rigidity and zero lower bound on nominal interest rate. A related class of models comes from the overlapping generations literature. Yaari (1964) introduced the joy of giving specification to explain bequests. Under this specification, households care about the amount of wealth their children will inherit but do not care about the consumption of their children, a model refined by Abel and Warshawsky (1988).

My contribution is to study the case of nonseparability between consumption and future wealth. The utility function of a representative agent depends on three variables, consumption, leisure and next period wealth. Allowing for wealth in the utility function and nonseparability introduces two new parameters in the model. The first parameter, denoted κ , governs the discount rate in the linear Euler equation. The second parameter ν reflects the degree of complementarity between consumption and future wealth. A positive value of ν allows me to obtain a low elasticity of intertemporal substitution along a moderate income effect on labor supply whereas, under the standard specification, the former is the inverse of the latter. I call such complementarity between consumption and future wealth the intertemporal complementarity. Disentangling the income effect and the intertemporal substitution effect has important implications for the model. The elasticity of hours worked with respect to real wages becomes different from the elasticity with respect to the real interest rate, modifying the response of real wages and unemployment to a monetary policy shock and the response of output gap to a demand shock.

I extend the analysis to a medium scale DSGE model. I focus on implications for labor market variables, especially real wages and unemployment. Following Galí(2011), I introduce sticky wages and identify unemployment as the difference between the desired labor supply, given by the first order condition of a family behaving competitively, and the effective labor demand. In the standard model, an expansionary monetary policy shock generates a relatively large response of real wages and a very large response of unemployment. This large response of unemployment is caused by the shift in labor demand but also by a large shift in labor supply. Introducing intertemporal complementarity allows me to reduce substantially this response of labor supply.

Then, I provide an estimation of parameters κ and ν . I estimate the medium scale model using Bayesian techniques. I find a large value for the parameter κ , suggesting a substantial discount rate in the Euler equation. I also find a large and positive value

for the parameter ν , supporting intertemporal complementarity. This high value of ν is the consequence of the difference it creates between the income effect on labor supply and the intertemporal substitution effect. When the elasticity of intertemporal substitution is the inverse of the income effect, the estimation gives a low EIS and a large income effect, leading to implausible fluctuations in labor supply. A positive parameter ν allows for a low EIS and a moderate income effect providing a more plausible serie for labor supply. High values for ν and κ seem a robust result. I reestimate the model with several alternative specifications, like alternative prior for κ , habits consumption, a longer sample, and labor force participation instead of hours worked in observables. Outcomes are consistent with my baseline estimation.

Substituability between consumption and leisure time is an alternative way to separate the income effect and the intertemporal substitution effect. This has been extensively analyzed by Bilbiie (2009) to explain the response of consumption to fiscal policy shocks. Some degree of substituability seems plausible but the evolution of consumption at retirement provides an upper bound to it (See Kimball and Shapiro 2008 for some quantitative exercise). Compatibility with balanced growth is also a concern. It seems interesting to complement this approach by exploring an alternative specification focusing on the intertemporal choice.

Various explanations have been put forward to justify the inclusion of wealth in the utility function. Wealth can provide an important social status leading consumers to have a preference for it. Alternatively, it may capture several saving motivations. Under the standard specification, a representative agent only saves for consumption smoothing. Households also save to insure themselves against negative income shock, to increase their income at retirement or to hand their estate to their children. A model integrating explicitly all these motivations would be better. However, the cost in complexity would be very high. Wealth in the utility function may generate in some extent a similar behavior for aggregate consumption and leisure whereas keeping the

convenience of the representative agent framework.

Several pieces of literature have recently cast doubts on the standard model of intertemporal choice. The model implies that theoretical responses to expected monetary policy shocks are much larger than their empirical counterparts (Del Negro et al. 2013). A simple way to solve the forward guidance puzzle is to introduce a discount in the linear Euler equation (see McKay, Nakamura and Steinsson 2016 and Gabaix 2017). Such discount rate emerges immediately when wealth is in the utility function. It is interesting to note that intertemporal complementarity increases the discount in the linear Euler equation. Heterogeneous agents models suggest that a substantial part of the response of consumption to monetary policy could come from an indirect effect, through the increase of the income of "hand to mouth" households, and not from the direct effect through intertemporal substitution (Auclert 2017, Kaplan, Moll and Violante 2017).

The paper is organized as follows. The first section explores the household choice in a simple optimization problem in finite horizon with wealth in the utility. The second section extends the results to infinite horizon. The third gives several implications for the macroeconomic model using both a simple model to highlight intuition and a medium scale model to confirm these insights in a more "realistic" environment. I estimate the model in the fourth. I discuss some assumptions and implications in the fifth. I examine additional consequences for forward guidance in the sixth. The seventh section introduces time varying wealth in the model.

2.2 Intratemporal Household choice

When wealth enters into the utility function, households have two motives to accumulate it. First, because it increases the "income" of the next period. Second, because it increases their *current* utility. To get a better intuition of the household's behavior, it

is useful to start by only considering this second motive. To do so, I consider an household which only cares about its current utility whose wealth is one of the arguments. I label this model as the "Wealth Targeting Model".

2.2.1 The household program

I consider the optimization problem of a consumer who does not care about future utility streams but whose current utility function accepts its future wealth as an argument. At period t , the objective function of the consumer is

$$U(C_t, L_t, A_{t+1}) \tag{2.1}$$

under the budget constraint

$$Q_t A_{t+1} + W_t L_t + C_t = A_t + W_t + \Pi_t \tag{2.2}$$

where C is the consumption, L is leisure time, W is real wage, and Π are profits distributed by firms. A is an asset which gives the right to receive one unit of consumption good at the next period. A does not provide utility through a continuation value but directly provides some utility, hence there is a positive demand for assets even if the optimization program of the household is purely static. Households still have a choice to make between current consumption and future assets.

To buy one unit of this asset, the consumer should pay a price Q . This price is the inverse of the interest factor.

$$Q_t = \frac{1}{1 + rr_t} \tag{2.3}$$

where rr is the real interest rate. Q_t is the price of future consumption goods.

For the moment, I identify wealth with safe bonds. The point is that consumers may

buy in period t a "promise" on final good of period $t + 1$. The amount of this promise enters into the utility function. I consider alternative interpretations and some of their consequences in the section dedicated to the model with varying wealth. Until then, I will use "wealth" and "assets" as synonyms.

2.2.2 First order conditions

I now solve for first order conditions.

Proposition 2.1. *The utility function reaches its local maximum under the budget constraint if the following first order conditions are fulfilled*

$$U_C(C_t, L_t, A_{t+1}) = \Lambda_t \quad (2.4a)$$

$$U_L(C_t, L_t, A_{t+1}) = W_t \Lambda_t \quad (2.4b)$$

$$U_A(C_t, L_t, A_{t+1}) = Q_t \Lambda_t \quad (2.4c)$$

Where U_c (resp. U_L and U_A) is the first derivative of the utility function with respect to consumption (resp. leisure and wealth). Λ_t is the Lagrange multiplier.

The problem is a basic consumer choice problem. whose solution is straightforward. A proof is given in appendix B.1.1. Now, suppose that consumption C , wealth A and leisure L reach steady state values. I can linearize conditions from proposition 2.1 around the steady state. I defined c_t , a_t and l_t and λ_t as percentage deviation from their steady state value. More generally, small letters will denote percentage deviation from steady state or deviation from steady state.

Proposition 2.2. *A linear approximation of the system of equations defined in propo-*

sition 2.1 is

$$\frac{U_{CC}C}{U_C}c_t + \frac{U_{CAA}A}{U_C}a_{t+1} + \frac{U_{CL}L}{U_C}l_t = \lambda_t \quad (2.5a)$$

$$\frac{U_{CL}C}{U_L}c_t + \frac{U_{LAA}A}{U_L}a_{t+1} + \frac{U_{LLL}L}{U_L}l_t = w_t + \lambda_t \quad (2.5b)$$

$$\frac{U_{CAC}C}{U_A}c_t + \frac{U_{AAA}A}{U_A}a_{t+1} + \frac{U_{LAL}L}{U_A}l_t = q_t + \lambda_t \quad (2.5c)$$

Terms U_C (resp U_L, U_A), and U_{CC} (and other similar terms), denote the steady state value of first and second order derivatives of the utility function.

Proof The linear approximation is a first order Taylor expansion of first order conditions defined in proposition 2.1. Detailed computations are given in appendix B.1.2 \square

2.2.3 Separable preferences

Before considering the case of intertemporal complementarity, I show that, under separable preferences, first order conditions defined in proposition 2.2 are related to the first order conditions of the standard model.

Hypothesis 2.1. *Preferences are separable. Cross derivative of the utility function are equal to zero :*

$$U_{CL} = 0$$

$$U_{AL} = 0$$

$$U_{CA} = 0$$

To allow a proper comparison with the standard model, I combine first order conditions with several general equilibrium conditions. I assume that the supply of assets is fixed. Thus, the deviation from steady state is equal to zero. I also introduce a relation between leisure time and hours worked.

Hypothesis 2.2. *the asset supply equation is given by*

$$A_{t+1} = \bar{A} \quad (2.6)$$

where A is a constant, hence the percentage deviation from steady state is

$$a_{t+1} = 0 \quad (2.7)$$

This assumption states that asset supply is not sensitive to the asset price and thus to the demand of assets by consumers. In the last section of the paper, I relax this assumption and study the model with a varying asset supply.

With assumptions 2.1 and 2.2, The system of linear equations considered in proposition 2.2 becomes

$$\frac{U_{CCC}}{U_C} c_t = \lambda_t \quad (2.8a)$$

$$\frac{U_{LLL}}{U_L} l_t = w_t + \lambda_t \quad (2.8b)$$

$$q_t + \lambda_t = 0 \quad (2.8c)$$

To obtain more friendly equations, I substitute leisure with hours worked and the price of assets with real interest rate

Hours worked are given by

$$l_t = \eta n_t \quad (2.9)$$

where η is the ratio between the steady state working time and the steady state leisure time¹.

¹Often denoted $\frac{N}{1-N}$

The deviation from the steady state real interest rate rr_t is directly related with the percentage deviation from the steady state bond price q_t

$$rr_t = -q_t \quad (2.10)$$

It is convenient to make notations easier by introducing parameters $\sigma = \frac{U_{CC}C}{U_C}$ and $\theta = \frac{U_{LL}L}{U_L}$. I obtain two equations for labor supply and consumption.

Proposition 2.3. *Labor supply and consumption equation are*

$$\theta\eta n_t = w_t - \sigma c_t \quad (2.11)$$

$$\sigma c_t = -rr_t \quad (2.12)$$

Both equations are derived from the system 2.8. Labor supply equation (2.11) is common with the standard model for intertemporal choice. The difference lies in equation (2.12). Instead of having an equation for consumption growth, I have an equation for consumption *levels* with respect to interest rate. The parameter σ governs both the intertemporal substitution effect and the income effect on labor supply. This feature is shared with the standard model of intertemporal choice.

2.2.4 Intertemporal nonseparability

I now allow the cross derivative between wealth and consumption to be different from zero. Hypothesis 2.1 becomes

Hypothesis 2.3.

$$U_{CL} = 0$$

$$U_{AL} = 0$$

$$U_{CA} \neq 0$$

A positive cross derivative between consumption and wealth implies that assets and consumption are complements in the sense of Edgeworth, whereas a negative cross derivative means the two are substitutes.

I keep the separability assumption for leisure. $U_{AL} = 0$, $U_{CL} = 0$. This is a strong assumption but the goal is to keep a tractable model and to focus on intertemporal choice. My analysis follows the analysis made by Bilbiie (2009) for the nonseparability between consumption and leisure.

Nonseparability between consumption and assets allows me to disentangle the consumption elasticity to interest rate from the income effect on labor supply.

Proposition 2.4. *The system from proposition 2.2 becomes*

$$\frac{U_{CC}C}{U_C}c_t = \lambda_t \quad (2.13a)$$

$$\frac{U_{LL}L}{U_L}l_t = w_t + \lambda_t \quad (2.13b)$$

$$\frac{U_{CA}C}{U_A}c_t = q_t + \lambda_t \quad (2.13c)$$

Let me define the parameter $\nu \equiv \frac{U_{CA}C}{U_A}$.

Corollary 2. *The intertemporal substitution effect is governed by $\sigma + \nu$ whereas the income effect is is governed by the parameter σ*

$$\theta\eta n_t = w_t - \sigma c_t \quad (2.14)$$

$$(\sigma + \nu)c_t = -rr_t \quad (2.15)$$

Intuition Some intuition may be given for this result. The sensitivity of leisure with respect to interest rate is equal to $\frac{\sigma}{\sigma + \nu}$. A large and positive ν is obtained if $U_{CA} > 0$ and thus complementarity between present consumption and future assets. A fall in real rates implies that the marginal utility of assets should rise relative to the marginal utility of consumption and relative to the marginal utility of leisure. With separable

preferences, consumption increases and thus reduces the marginal utility of consumption. With nonseparable preferences, if consumption and assets are complements, the rise in consumption decreases the marginal utility of consumption and increases the marginal utility of assets. Thus, a much lower rise in consumption may achieve the equality between the relative price of future consumption goods and the marginal rate of substitution between consumption and assets. A similar line of reasoning explains the smaller sensitivity of leisure to real interest rate. The increase in the marginal utility of assets implies a lower fall in the marginal utility of leisure and thus a lower rise in leisure.

Concavity and Noninferiority requirements Parameters σ and ν cannot be calibrated freely. They should respect concavity requirements for the utility function:

Proposition 2.5. *The utility function U is concave if*

$$U_{CC} \leq 0$$

$$U_{LL} \leq 0$$

$$U_{AA} \leq 0$$

$$U_{AA}U_{CC} - U_{CA}^2 \geq 0$$

The last condition implies that U_{CA} cannot be "too large" with respect to U_{CC} and thus ν should not be "too large" with respect to σ , except if U_{AA} is large enough. With our particular asset supply function, U_{AA} can be calibrated freely, allowing a low value for $\frac{\sigma}{\sigma+\nu}$.

Whereas not compulsory, it also seems reasonable to impose that assets and consumption are not inferior goods whose demand decreases when income rises. A positive value of ν (i.e. assets and consumption are complements) is however a sufficient condition for noninferiority.

Generalized nonseparable preferences In the previous paragraph, I focus on nonseparability between consumption and assets. It is useful to consider the general case with several forms of nonseparability. Notations are burdensome and I relegate computations to appendix B.4. The last equation of the appendix B.4 gives the consumption equation when U_{CL} , U_{AL} and U_{CA} are different from zero. Combining complementarity between consumption and assets, complementarity between leisure and assets and substitutability between consumption and leisure reduces the consumption elasticity to real interest rate further without affecting the income effect on labor supply.

2.3 Intertemporal Household choice

The optimization problem I extend the analysis of the previous section in a more standard setup. Wealth still enters in the utility function but households care about future utility. They maximize

$$\sum_{T=t}^{+\infty} \beta^{T-t} E_t U(C_T, L_T, A_{T+1}) \quad (2.16)$$

Budget constraint is the same as in the previous section. It is important at this stage to note that I do not make any assumption about the value of the parameter β . In the standard model, $\frac{1}{\beta} - 1$ is equal to the steady state real interest rate. It necessarily implies a value of β close to one. Where wealth enters into the utility function, this equality no longer holds and β can be calibrated with more freedom as I show in proposition 2.7. I now derive first order conditions.

Proposition 2.6. *First order conditions for the optimization problem are*

$$U_C(C_t, L_t, A_{t+1}) = \Lambda_t \quad (2.17a)$$

$$U_L(C_t, L_t, A_{t+1}) = w_t \Lambda_t \quad (2.17b)$$

$$U_A(C_t, L_t, A_{t+1}) + \beta E_t U_C(C_{t+1}, L_{t+1}, A_{t+2}) = Q_t \Lambda_t \quad (2.17c)$$

Proof The problem is close to the standard problem. The solution follows the same steps. See appendix [B.1.7](#) for details. \square

Compare to the intratemporal problem of the previous section, the only change is the forward looking term in the first order condition for wealth.

Steady state and linearization Under the standard specification, the discount rate β is constrained to be the inverse of the interest factor. This restriction no longer holds with wealth in the utility function

Proposition 2.7. *At the steady state, there is a wedge between the discount rate and the inverse of the interest factor*

$$\beta = Q - \frac{U_A}{U_C} \quad (2.18)$$

Proof It follows immediately from computing the steady state of the system of recursive equations given by the first order conditions above. Computations are given in appendix [B.1.8](#). \square

I now linearize first order conditions around the steady state. I combine them with the asset supply equation (3.2). Asset supply is still fixed, implying $a_{t+1} = 0$. I keeps the nonseparability assumption of the previous section

Proposition 2.8. *Under hypothesis 2.3, first order conditions for leisure and consumption becomes*

$$\begin{aligned}\frac{U_{LLL}}{U_L}l_t &= w_t + \frac{U_{CCC}}{U_C}c_t \\ (1 - \frac{\beta}{Q})\frac{U_{CAC}}{U_A}c_t + \frac{\beta}{Q}\frac{U_{CCC}}{U_C}E_t c_{t+1} &= q_t + \frac{U_{CCC}}{U_C}c_t\end{aligned}$$

I denote $\kappa = 1 - \frac{\beta}{Q}$. Other notations are unchanged.

Proposition 2.9. *The system from proposition 2.8 can be rewritten*

$$\theta\eta n_t = w_t - \sigma c_t \tag{2.19a}$$

$$(\sigma + \kappa\nu)c_t = -rr_t + (1 - \kappa)\sigma E_t c_{t+1} \tag{2.19b}$$

Those equations are extremely close to those of the standard model but have two new parameters κ and ν . The term κ introduces a discount in the Euler equation for consumption. It reduces the elasticity of consumption with respect to **future** real interest rate values. As in the wealth targeting model, The term ν amplifies this discounting and as in the wealth targeting model, dampens the response of consumption to current real interest rate without modifying the income effect.

2.4 Implications for the New Keynesian model

I now study implications of intertemporal complementarity (IC thereafter) for the New Keynesian model. The more interesting property of IC is to relax the cross equation restriction between the income effect and the intertemporal substitution effect. For a given intertemporal substitution effect, I can obtain a lower income effect on labor supply. I show it has important consequences for responses to monetary policy shocks and demand shocks in the New Keynesian model (NK model thereafter). I use a very simple version of the NK model to derive those implications. I also verify in what

extent they are still relevant in a medium scale model. The two models are presented in the first subsection. The second subsection is dedicated to monetary shocks and the third to demand shocks.

2.4.1 Framework

I display equations of the simple model in table 2.1. There are four behavioral equations: the consumption equation, the labor supply equation, the monetary policy rule, and the Philips curve. Consumption is given by equation (2.19b). It becomes the standard Euler equation if κ and ν are equal to zero. The monetary policy rule is unusual but simple. The nominal interest rate is equal to the expected inflation rate plus a disturbance². The economic interpretation is that the central bank sets directly the real interest rate. It is not a realistic feature but aims at providing a better intuition by focusing on households' behavior. Indeed, whereas the New Keynesian Philips Curve is still there, it is no longer relevant for output and real variables in general. It only determines the path of inflation whose effects on real variables are neutralized by the response of the central bank.³⁴ Aggregate demand on the good market is equal to $\varphi c_t + d_t$. φ is the steady state consumption over output ratio. d_t is an exogenous shock directly expressed in terms of GDP percentage points. It encompasses all other components of aggregate demand including private investment and public consumption⁵. The aggregate supply equation is the reduced form of the usual New Keynesian Philips Curve derived from Calvo Pricing. β_π is the coefficient associated with ex-

²Both the nominal interest rate and the expected inflation are in deviation from their steady state values

³This simple model can be viewed as an IS LM version of the New Keynesian model whereas the standard model with a Taylor rule would be as an AS-AD version

⁴This simple version does not grant determinacy. Our results are derived by assuming there are no sunspots. It is however easy to restore determinacy. For example, a monetary policy rule $r_t = \pi_{t+1} + \phi_y y_t$, ϕ_y being positive and possibly very small, would be sufficient.

⁵To remain compatible with the fixed supply of assets assumption, the increase (resp. decrease) in public consumption has to be tax- financed. Capital stock and investment do not appear in the model, but it is possible to add a fixed cost to firm. An increase in this fixed cost would increase aggregate demand without affecting the supply of assets

pected inflation and γ_p the coefficient associated with the marginal cost. On the labor market, I consider two variants of the model, a flexible wage variant in which the labor market is walrasian and a rigid wage variant in which the real wage does not depart from its steady state value and in which the difference between desired and effective hours worked is assimilated to the unemployment rate.

This very simple NK model is useful to provide some intuition. However, it is better to verify if results hold in a medium scale NK model. Equations are displayed in table 2.2. I introduce a more conventional monetary policy rule along wage stickiness and wage and price indexation. I rely on simulations to compute impulse response functions. The calibration used to obtain these IRFs is displayed in table 2.3. Frisch elasticity is equal to one. The consumption output ratio is set at 0.65, targeting the average value on US data between 1985 and 2007. Price and wage indexation parameters are both calibrated at 0.25 which is consistent with values found in estimated models. I set γ_p and γ_w at 0.1. The value of γ_p is high for reduced form estimation but is consistent with a yearly frequency for price changes. Coefficients for expected inflation β_π and β_w are set at 0.985.

Equation	Interpretation
$y_t = \alpha n_t$	Production Function
$y_t = \varphi c_t + d_t$	Market clearing on good market
$r_t - E_t \pi_{t+1} = r r_t$	Accounting equation
$\mu_t = w_t + n_t - y_t$	Marginal cost equation
$\pi_t = \gamma_p \mu_t + \beta_\pi E_t \pi_{t+1}$	Philips Curve
$r_t = E_t \pi_{t+1} + e_t$	Monetary policy rule
$n_t^s = \frac{1}{\theta \eta} (w_t - \sigma c_t)$	Desired Hours worked
$u_t = n_t - n_t^d$	Unemployment equation
$(\sigma + \kappa \nu) c_t = -r r_t + (1 - \kappa) \sigma E_t c_{t+1}$	Consumption equation
$w_t = 0$	Rigid wage model
$u_t = 0$	Flexible wage model

TABLE 2.1 – Simple Model

Identifying unemployment as the difference between desired and effective hours worked is a debatable assumption. Unemployment is an extensive margin phenomenon whereas the difference between desired and effective hours worked is an intensive margin phe-

Equation	Interpretation
$y_t = \alpha n_t$	Production Function
$y_t = \varphi c_t + d_t$	Market clearing on good market
$r_t - \pi_{t+1} = r r_t$	Accounting equation
$\mu_t = w_t + n_t - y_t$	Marginal cost equation
$\pi_t = \frac{\beta_\pi}{1+\beta_\pi\tau_p} E_t \pi_{t+1} + \frac{\tau_p}{1+\beta_\pi\tau_p} \pi_{t-1} + \gamma_p \mu_t$	Philips Curve
$r g_t = \phi_\pi \pi_t + e_t$	Monetary policy rule
$r_t = \lambda r_{t-1} + (1 - \lambda) r g_t$	Effective nominal interest rate
$n_t^s = \frac{1}{\theta \eta} (w_t - \sigma c_t)$	Desired Hours worked
$u_t = n_t - n_t^d$	Unemployment equation
$\pi_{w,t} = \frac{\beta_w}{1+\beta_w\tau_w} E_t \pi_{w,t+1} + \frac{\tau_w}{1+\beta_w\tau_w} \pi_{w,t-1} - \gamma_w u_t$	Wage Philips curve
$w_t = w_{t-1} + \pi_{w,t} - \pi_t$	Real wage equation
$(\sigma + \kappa \nu) c_t = -r r_t + (1 - \kappa) \sigma E_t c_{t+1}$	Consumption equation

TABLE 2.2 – Medium scale Model

Variable	value	Interpretation
α	0.68	Labor coefficient
ϕ_π	1.5	Inflation coefficient in MP rule
φ	0.65	consumption-output ratio
η	1	Hours-leisure ratio
θ	1	Inverse Frisch elasticity
β_π	0.985	Coefficient for the expected term in inflation equation
β_w	0.985	Coefficient for the expected term in wage equation
τ_p	0.25	Price indexation
τ_w	0.25	Wage indexation
γ_w	0.1	Wage Philips curve coefficient
γ_p	0.1	Philips curve coefficient
λ	0.7	Nominal rate persistence

TABLE 2.3 – calibration

nomenon. However, Gali (2011) considers a model of indivisible labor in which households member differ by their labor disutility but have a common level of consumption. It shows that the reduced form for labor force participation is the same as the reduced form for desired hours worked in the classical model. In appendix B.2, I show that the Gali's framework is compatible with a utility function whose wealth is one of the argument.

Wealth in the utility function has also some implications for the supply block. The derivation of the New Keynesian Philips curve for prices and wages is mostly unaffected

but the relevant discount rate may be subject to some debate. Should firms (resp. "trade unions") discount profit streams (resp. utility streams) using the safe real interest rate or the pricing kernel $\beta \frac{U_c(C_{t+1})}{U_c(C_t)}$? If firms and trade unions maximize the utility of their shareholders (resp. members) and if only bonds enter in the utility function, the latter is the relevant one, leading to potentially large discount not only in the consumption equation but also in inflation and wage inflation equation. However, I want to focus on implications of WIU and IC on households' choice. Thus, I choose to allow for different discount rate in consumption, inflation and wage inflation equation. In the section dedicated to the estimation of the model, I check the robustness of my findings to this assumption.

2.4.2 The supply effect of the real interest rate

To better understand the effects of IC, it is useful to start by considering the standard model, when κ and ν are equal to 0. The same parameter governs the income effect on labor supply and the elasticity of intertemporal substitution. A first consequence of this cross equation restriction is that the real interest rate has the same impact on labor supply as the real wage growth. Indeed, consider consumption and desired hours worked equation from table 2.1 with κ and ν equal to zero

$$\begin{aligned}\sigma c_t &= -rr_t + \sigma E_t c_{t+1} \\ -\theta \eta n_t^s &= w_t - \sigma c_t\end{aligned}$$

Combining the two equations allows me to derive an euler equation for hours.

$$\theta \eta (E_t n_{t+1}^s - n_t^s) = E_t w_{t+1} - w_t - rr_t \quad (2.20)$$

The elasticity of desired hours worked with respect to real wage growth is always equal to the elasticity with respect to the real interest rate. Real interest rate matters here because of the income effect on labor supply. But, it is important to keep in mind that the effect of the real interest rate on hours worked does not depend on the parameter σ which governs the income effect. A naive view of the problem would state that real interest rate affects labor supply because it affects consumption. Then, reducing the sensitivity of consumption to real interest rate would lower the impact of the real interest rate. Equation 2.20 shows it is misleading. Real interest rate affects labor supply because it affects the *marginal utility of consumption* and under the standard specification, the real interest rate is equal to the growth rate of the marginal utility of consumption whatever the value of the parameter σ is.

Now, let me consider the case where κ and ν are positive. This supply effect of the real interest rate vanishes

Proposition 2.10. *If the utility function accepts wealth as an argument and if U_{CA} is positive, the elasticity of labor supply with respect to real interest rate is lower than the elasticity of labor supply to real wages.*

Indeed, deriving the labor supply equation for κ and ν different from zero and dropping expected terms for more clarity, I get

$$\theta\eta n_t^s = w_t + \frac{\sigma}{\sigma + \kappa\nu} r r_t \quad (2.21)$$

The elasticity of hours with respect to wages is given by $\frac{1}{\theta\eta}$ whereas the elasticity of hours with respect to real interest rate is $\frac{\sigma}{\sigma + \kappa\nu} \frac{1}{\theta\eta}$. If ν is positive, the elasticity with respect to real interest rate is lower. A more detailed proof is given in appendix B.1.11

2.4.3 Monetary shocks

Real wages in the flexible wage variant

The restriction on the hours equation has an important consequence for the relation between the real wage and the real interest rate in the flexible wage variant of the model.

I derive the New Keynesian labor demand conditional to a monetary policy shock. In the New Keynesian model, the production is determined by the demand in the short run. Firms collect orders and use production function to determine the amount of labor they need to satisfy these orders.⁶

I first derive the labor demand in the standard model. I combine the consumption equation, the market clearing condition and the production function of table 2.1. To make notations more friendly, I assume that the monetary shock is perfectly anticipated and that there is no demand shock. It allows me to drop the expectation operator and the term d_t .

$$\frac{\sigma\alpha}{\varphi}(n_{t+1} - n_t) = rr_t \quad (2.22)$$

I combine equation (2.20) which represents labor supply and equation (2.22) which represents labor demand. The equilibrium value of real wage growth appears and depends on real interest rate. I break down the real wage equation between a demand

⁶This New Keynesian labor demand is different from the neoclassical labor demand which in the NK model is more relevant in the long run. In the short run, the neoclassical labor demand is replaced by the marginal cost (or the inverse markup) equation and the Philips Curve which together with the monetary policy rule determine the real interest rate. In the simple model of this section, the real interest rate is directly set by the central bank and does not react to changes in inflation and thus changes in marginal cost. Because of this particular assumption, the supply block of the model is unimportant. The marginal cost equation may also act as a labor demand equation if the marginal cost is kept constant. For my purpose, it seems more relevant to use the New Keynesian labor demand. Moreover, this "alternative" labor demand is difficult to interpret because the marginal cost is jointly determined with hours by firms and not an exogenous variable to their labor demand decision.

effect and a supply effect. The demand effect is defined as the change of real wages following a change in real interest rate when the labor supply curve is held constant. The supply effect is the change when the labor demand curve is held constant. I give a formal definition

Definition 2.1. *Consider the system of labor supply and demand*

$$\begin{aligned}\theta\eta(n_{t+1}^s - n_t^s) &= w_{t+1} - w_t - rr_t \\ \frac{\sigma\alpha}{\varphi}(n_{t+1} - n_t) &= rr_t\end{aligned}$$

The demand effect is the response of real wages following a change in real interest rate when the labor supply curve does not shift (e.g in an "imaginary" world in which $\theta\eta(n_{t+1}^s - n_t^s) = w_{t+1} - w_t$). The supply effect is the difference between the total effect and the demand effect.

Real wages are given by

$$w_t = - \left(\underbrace{1}_{\text{Supply effect}} + \underbrace{\frac{\theta\eta\varphi}{\sigma\alpha}}_{\text{Demand effect}} \right) rr_t + w_{t+1} \quad (2.23)$$

Real interest rate affects the growth of real wages through the two channels. The demand effect represents the traditional keynesian channel. An expansionary monetary policy increases aggregate demand stimulating labor demand and thus real wages. In addition, a supply effect arises. A fall in real interest rate leads workers to substitute current leisure to future one and thus to reduce their labor supply, stimulating wage growth. A counter intuitive result is that this supply effect on real wages *neither depends on the Frisch elasticity of the labor supply nor on the elasticity of intertemporal substitution.*

To better understand the intuition behind the response of real wages, I give a stylized representation of the labor market in figure 2.1

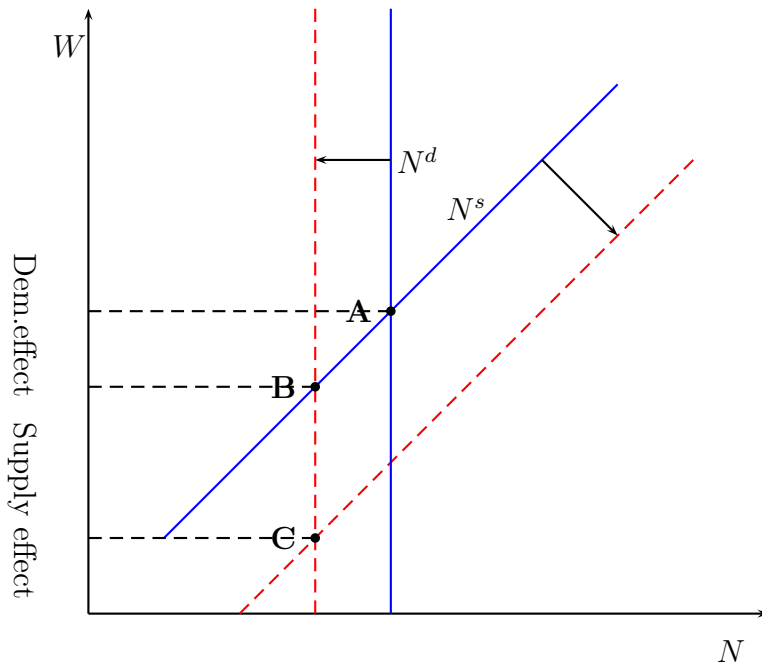


FIGURE 2.1 – Real wages following a rise in real rate

In the short run, the labor demand is given by equation (2.22). It is vertical and does not depend on real wages. Labor supply increases with real wages for a given level of real interest rate. The point A is the initial equilibrium. A rise of real rate has two effects on the figure. It depresses aggregate demand, shifting the labor demand curve to the left, reducing equilibrium real wages. This is the standard Keynesian channel. If only this channel is at play, the economy moves to the point B .

A second effect is the supply effect of the interest rate. The rise in the interest rate pushes the labor supply curve to the right, further lowering the equilibrium real wages, moving equilibrium to point C .

Consider now the case of intertemporal complementarity. Positive values for ν and κ alleviate the supply effect of the interest rate and thus lower the response of real wages to change in real interest rates.

Proposition 2.11. *The elasticity of equilibrium real wages with respect to real interest is decreasing with the value of the parameter ν governing intertemporal complementar-*

ity and with the value of the parameter κ which governs the discount rate in the euler equation.

Indeed, computing the equilibrium real wages on the labor market when κ and ν are different from zero gives

$$w_t = - \left(\underbrace{\frac{\sigma}{\sigma + \kappa\nu}}_{\text{Supply effect}} + \underbrace{\frac{\sigma}{\sigma + \kappa\nu} \frac{\theta\eta\varphi}{\sigma\alpha}}_{\text{Demand effect}} \right) rr_t + \frac{\sigma}{\sigma + \kappa\nu} (1 - \kappa) w_{t+1} \quad (2.24)$$

Consider a temporary change in the real interest rate, the response of the real wage is lower when κ or σ are greater. Following a rise in real interest rate, High values for κ and ν lower the response of leisure and thus the increase of labor supply. The rightward shift of labor supply is less important, limiting the fall in real wages. The demand effect is affected in a similar way. $\sigma + \kappa\nu$ governs the intertemporal effect of substitution. Higher values of κ and σ imply a lower impact of real interest rate on consumption and thus on aggregate demand.

What is interesting is that IC allows a lower supply effect for a given demand effect. The response of consumption to a certain path of real interest rate is mainly determined by the discount rate κ and the inverse of the consumption elasticity to real interest rate $\sigma + \kappa\nu$. Assuming both are given, a higher value of ν would lead to a lower value of σ lowering the supply effect whereas keeping the response of consumption to real interest rate roughly unchanged.

Proposition 2.12. *For a given value of $\sigma + \kappa\nu$ and a given value of κ , a larger value of ν (e.g a larger intertemporal complementarity) lowers the supply effect without affecting the demand effect.*

The proof immediately follows from the demand and supply effects highlighted in equation (2.24)

Unemployment in the rigid wage variant

The response of unemployment in the rigid wage model is very similar to the response of real wages in the flexible wage model. Shifts in labor demand and supply affects unemployment instead of real wages. Figure 2.2 provides some intuition. The equilibrium for labor and real wages is given by the intersection of the labor demand curve and the real wage curve whereas the difference between the labor demand and the labor supply for this real wage gives the unemployment rate. Unemployment is initially equal to zero. Following a rise in real interest rate, labor demand shifts to the left and labor supply to the right. Real wages remain at the same level, causing a rise in unemployment, coming from both the demand and the supply effect.

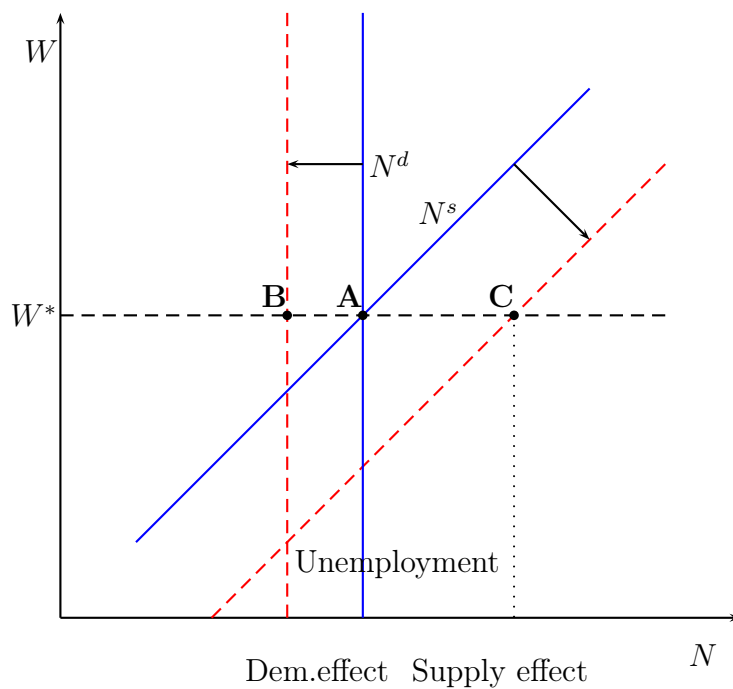


FIGURE 2.2 – Unemployment following a rise in real rate

Computations confirm the graphical intuition. Unemployment is given by

$$u_t = - \left(\underbrace{\frac{\sigma}{\sigma + \kappa\nu} \frac{1}{\theta\eta}}_{\text{Supply effect}} + \underbrace{\frac{\sigma}{\sigma + \kappa\nu} \frac{\varphi}{\sigma\alpha}}_{\text{Demand effect}} \right) rr_t + \frac{\sigma}{\sigma + \kappa\nu} (1 - \kappa) u_{t+1} \quad (2.25)$$

The medium scale model

Are these insights still relevant in a medium scale model? Figure 2.3 represents impulse responses (IRFs thereafter) to a monetary policy shock for real wages, output, unemployment and expected real interest rate when κ and ν are equal to zero. The shock is an unexpected one percent decrease in the nominal rate in annual value. On impact, the real rate falls by 0.4 percent and output increases by the same amount. Real wages slightly underreacts on impact but displays a hump shaped response with a peak around 0.4 percent too. More striking is the response of unemployment. Unemployment falls by 1.2 points on impact, nearly three times the response of output. It suggests that the supply effect of monetary policy is still sizable in the medium scale model.

Figure 2.4 represents the same IRFs in three different cases, corresponding to different values of σ and ν . κ is set at 0.5 and $\sigma + \kappa\nu$ is set at 1.5 in each case, hence parameters governing the demand effect are roughly unchanged. The main difference between the three experiments is the income effect on labor supply. The solid line is the "standard case" with ν equal to 0 and σ equal to 1.5. The dotted line is the "IC case". It displays responses obtained with a significant level of intertemporal complementarity (ν equal to 2 and σ equal to 0.5). The dashed line is an intermediate case. The response of the real interest rate is quite similar across the different calibrations. The

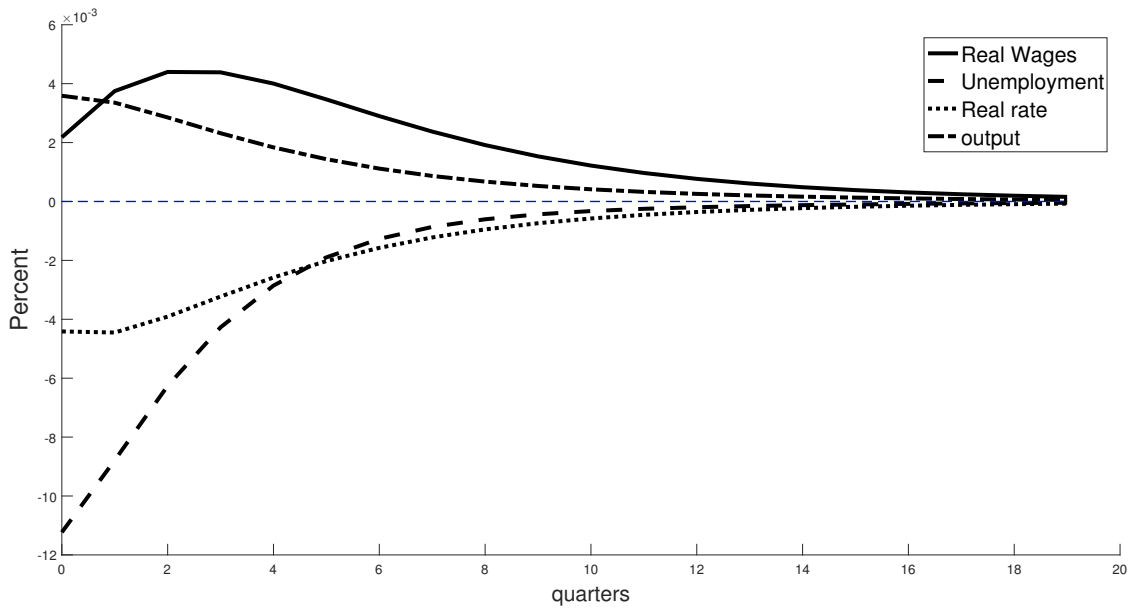


FIGURE 2.3 – Responses to monetary policy shocks

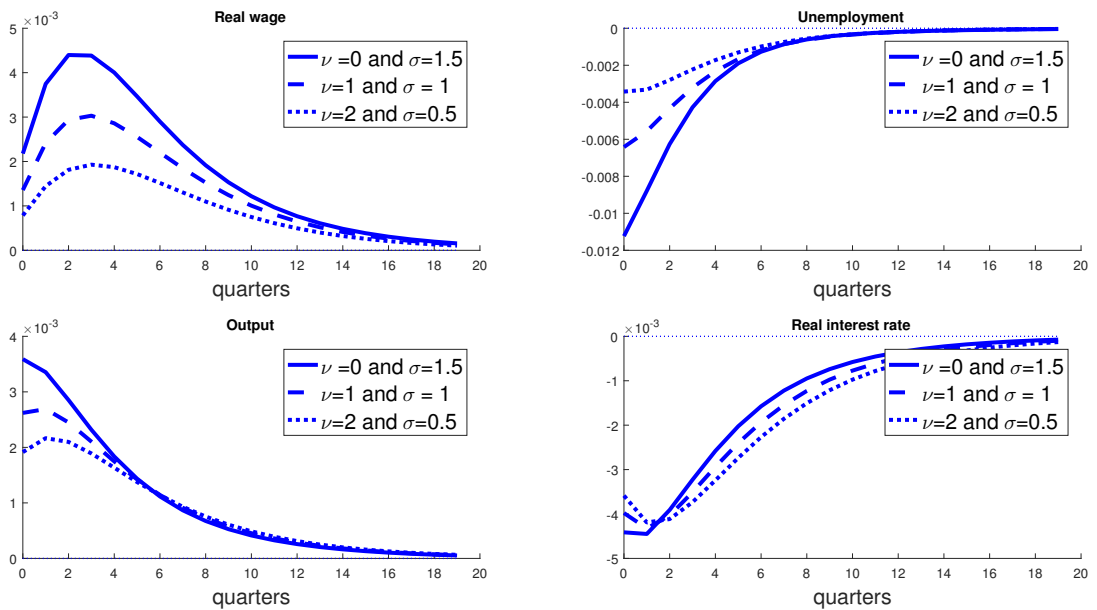


FIGURE 2.4 – Responses for several calibration

response of output slightly falls when ν increases (because a lower σ also lowers the discount rate in the Euler equation). Responses of real wages and unemployment are much more affected. They are roughly divided by three when the response of output is divided by 1.5. Indeed, the lower value of σ reduces the income effect on labor supply,

dampening the response of labor supply to a change in real interest rate, and thus the response of real wages or the response of unemployment. In the meantime, the higher value of ν keeps the demand effect at a similar level across the three experiments. This interpretation is supported by figure 2.5. It represents IRFs of labor demand and labor supply after a monetary policy shock in the standard case and in the IC case. In the standard case, the labor demand increases and the labor supply decreases, roughly by the same magnitude. The rise in unemployment is caused equally by the demand and the supply effect. In the IC model, the increase in unemployment is mainly caused by the increase in labor demand. The response of labor supply is ambiguous.

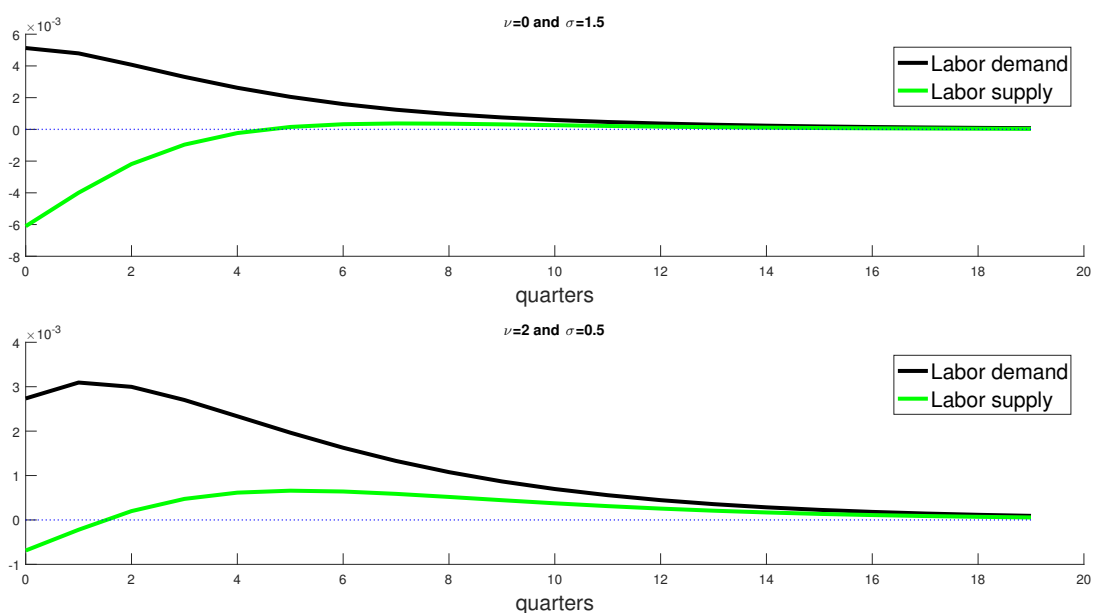


FIGURE 2.5 – Labor supply and demand after a monetary policy shock

2.4.4 Demand Shock

Real wages and Unemployment after a demand shock In this section, I study the response of real wages and unemployment after a demand shock. There are several reasons to be interested in them. First, private investment and public consumption

are probably important drivers of the business cycles. Second, in the conventional (keynesian) wisdom, effects of demand and monetary shocks on output, real wages and unemployment are similar. This conventional wisdom is only partially true in the NK model. A contractionary demand shock is equivalent to a contractionary monetary policy shock minus the supply effect.

Indeed, consider for example the response of real wages to a demand shock in the simple model⁷.

$$w_t = -\frac{\theta\eta}{\alpha}(d_t - d_{t+1}) + w_{t+1} \quad (2.26)$$

An exogenous demand shock increases output, labor demand and thus real wages. In the meantime, real interest rate and thus consumption are unaffected keeping the labor supply curve unchanged. Obviously, it is an extreme result due to the very specific monetary policy rule of the simple model. In practice, real interest rates react to demand shocks.

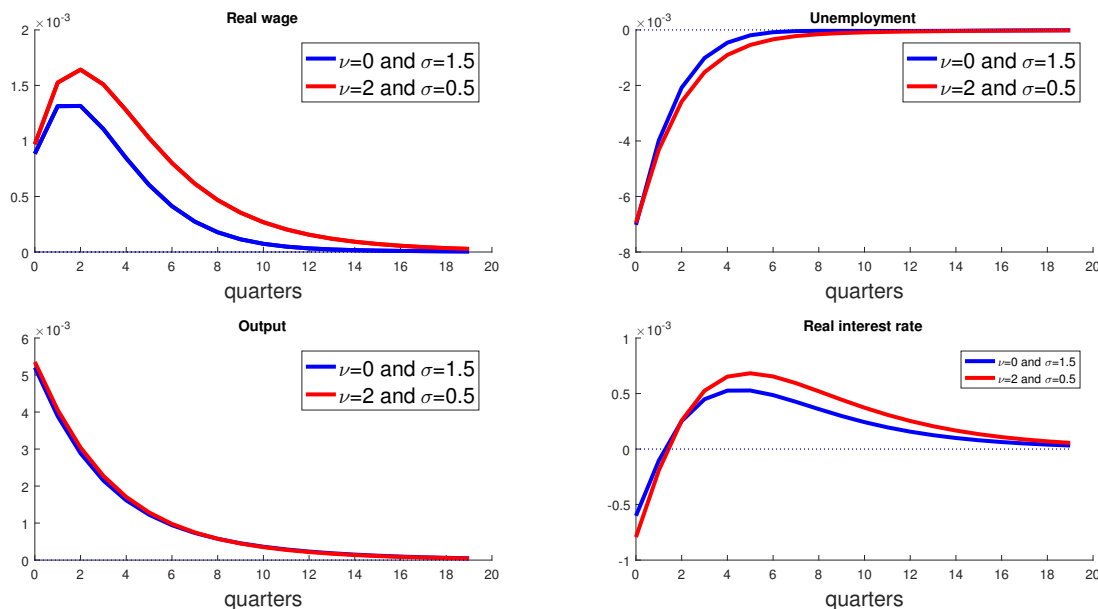


FIGURE 2.6 – Response of Real wages to demand shocks

⁷As in the previous section, I assume that the demand shock is the only shock is perfectly anticipated, allowing me to drop the expectation term and the monetary shock.

Figure 2.6 shows that the result is actually quite robust in the medium scale model. I represent impulse responses of output, real rate, unemployment and real wages to a demand shock in the standard and in the IC case. IRFs are nearly unaffected by the different values of ν and σ .

Response of the output gap to demand shocks Whereas ν and σ are unimportant for the response of output, real wages and unemployment, they actually matter for the response of output gap. Figure 2.7 displays the response of output, output gap, actual and natural interest rate following a demand shock in the standard case and in the IC case. In the standard model, the response of the output gap to a demand shock is nearly three times smaller than the response of output. The output gap has nearly disappeared after four quarters. This small response is not because the response of output is small but because the response of flexible price output is large. It is a consequence of the supply effect of real interest rate. Because of it, in the flexible price equilibrium, a small increase in the real interest rate may increase labor supply and decrease consumption enough to reestablish the equality between aggregate demand and the flexible price output. In other words, the response of the natural interest rate⁸ to a demand shock is small, close to the the response of the real rate in the sticky price equilibrium. This one is initially negative because of nominal interest rate smoothing, but the effects of the smoothing vanished after four quarters and the real rate becomes only marginally different from the natural one. As a consequence, the response of output is close to the response of natural output and the output gap is small. In the model with IC, the response of the natural interest rate to a demand shock is much larger and farther away from the actual response of monetary policy. The response of output gap is larger and more persistent.

⁸Here, the natural interest rate is the flexible price equilibrium real interest rate and not the steady state real interest rate.

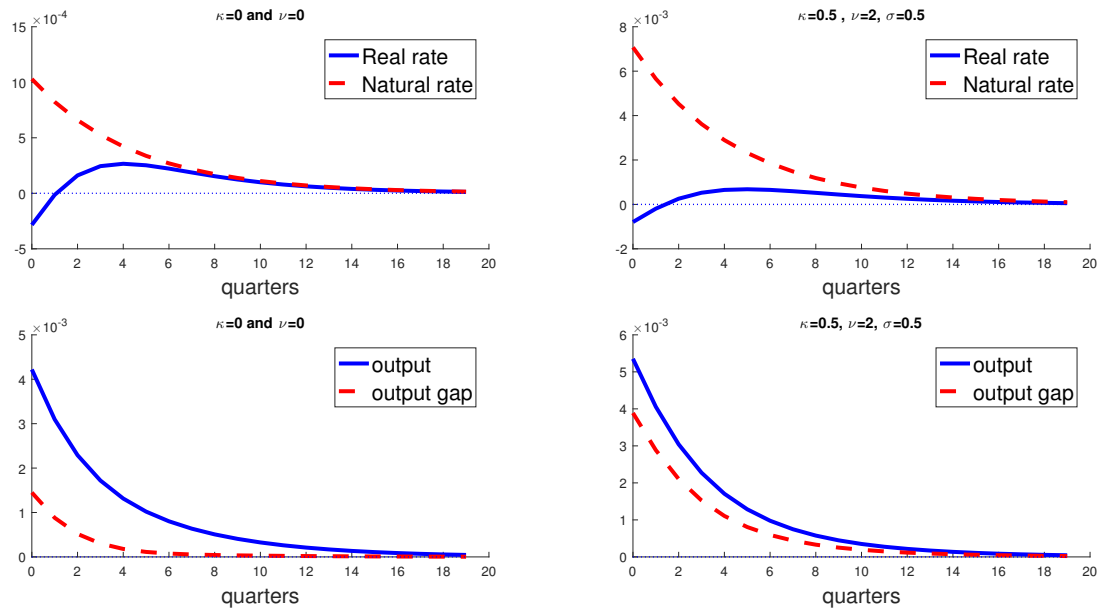


FIGURE 2.7 – Response of output gap and output to demand shocks

2.5 A Bayesian estimation of the model

In this section, I estimate parameter κ and ν . My approach is to perform a bayesian estimation of the medium scale model and to focus on parameters κ and ν .

The model I estimate is a variant of the model displayed in table 2.2. I add five shocks: a productivity shock, a markup price shock, a markup wage shock, a labor disutility shock and a discount factor shock. The complete specification can be consulted in table B.1. Parameters γ_p and γ_w (e.g Philips Curve coefficients) are estimated directly. I use quarterly data for seven macroeconomic variables : Real GDP, Real compensation per Hour in Nonfarm business sector, GDP deflator, Hours worked by all persons in Nonfarm business sector, Real Personal Consumption Expenditures, the effective FED funds rate, and the Unemployment rate. Following Smets and Wouters(2007), observables are first differences for the log of each of these variables, except for unemployment and FED funds rate, which are simply detrended. The sample contains 91 data points from the 1985:2 to 2007:4 . This dataset is small but the risk of structural breaks would be higher over a longer sample. Ours goes from the "great inflation"

to the "great recession". A stable relation between macroeconomic variables seems reasonable over that period. I perform a robustness test with a longer sample.

Variable	distribution	mean	std
κ	beta	0.5	0.2
ν	normal	0	1.5
σ	normal	1	1.5
θ	gamma	2	0.25
γ_p	normal	0.25	0.05
γ_w	normal	0.25	0.05
ϕ_π	normal	1.5	0.25
ϕ_y	normal	0.12	0.25
τ_p	beta	0.5	0.15
τ_w	beta	0.5	0.15

TABLE 2.4 – Priors for parameters

2.5.1 Baseline estimation

Priors are given in table 2.4. For usual parameters, they follow the literature. My prior on κ is a beta distribution whose mean is 0.5 and whose variance is 0.2. This choice excludes a value equal to zero. Thus, it is not possible to recover the standard model but it is still possible to be very close from it. I also estimate the model with a more conservative prior for κ . Results are described in the robustness subsection. The parameter ν is initially supposed to follow a Gaussian distribution centered around a zero mean with a large standard deviation set at 1.5.

Posterior estimates are displayed in table 2.5. Figure 2.8 represents prior and posterior distributions of κ and ν . They show a substantial shift between the prior and the posterior. Both parameters are positive and large. The estimated discount rate κ is equal to 0.8 whereas the estimated value of ν is as high as 2. It means that the coefficient for the expected term in the consumption equation is roughly equal to 0.05. σ is close to 1 suggesting a still sizable income effect, whereas the inverse of the elasticity of consumption to real interest rate $\sigma + \kappa\nu$ is close to 3. Other parameters

are in line with the literature, except for the marginal cost coefficient in the Philips curve which is close to zero.

It is interesting to compare these results with an estimation of the standard model. I set κ and ν equal to 0 and reestimate the model. Results are displayed in table 2.6. The main change is the mean estimate for σ . It is equal to 2.77 instead of 1.17. The estimate of σ in the standard model probably captures a small apparent response of consumption to changes in real interest rate. In my baseline estimation, this small response leads to a high value for the parameter ν whereas the value of σ is determined by the apparent income effect on labor supply. It is worth noting that the estimate of γ_p is also very low when estimating the standard model.

Variable	mean	mode	inf	sup
κ	0.797	0.839	0.644	0.961
ν	2.13	2.01	1.03	3.19
σ	1.17	1.11	0.627	1.69
θ	2.31	2.26	1.9	2.7
γ_p	0.0004	0.0005	-0.0006	0.0013
γ_w	0.264	0.284	0.178	0.374
ϕ_y	0.463	0.44	0.113	0.788
ϕ_π	1.47	1.46	1.07	1.86
τ_p	0.287	0.315	0.131	0.435
τ_w	0.126	0.0956	0.0378	0.21

TABLE 2.5 – Posteriors for the baseline estimation

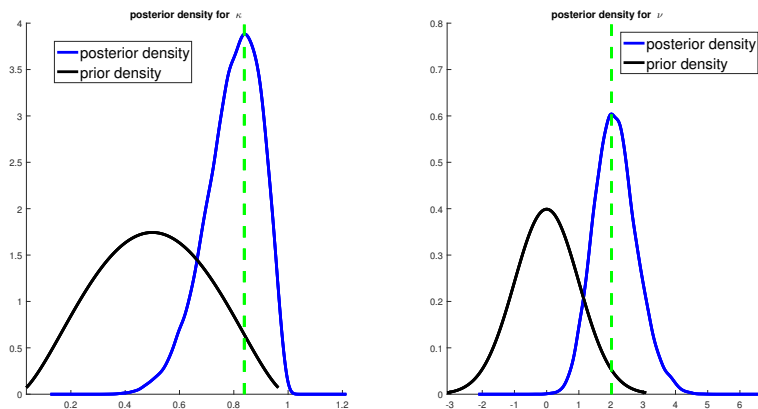


FIGURE 2.8 – Posterior and prior distribution for κ and ν

Variable	mean	mode	inf	sup
σ	2.77	2.63	2.16	3.37
θ	2.58	2.51	2.14	2.98
γ_p	0.0042	0.0019	-0.0015	0.0109
γ_w	0.206	0.247	0.0513	0.318
ϕ_y	1.24	1.27	0.952	1.57
ϕ_π	1.75	1.66	1.23	2.3
τ_p	0.374	0.356	0.153	0.592
τ_w	0.138	0.104	0.0471	0.225

TABLE 2.6 – Posteriors for the standard model estimation

2.5.2 Inspecting the mechanism

To better understand these results, it is useful to consider the figure 2.9. This figure represents the observed labor supply, the "predicted" one and the labor disutility shock for the standard model and the model with wealth in the utility function and intertemporal complementarity. The "predicted" labor supply is the value predicted by the model, given observations of wages and consumption and without taking into account labor disutility shock. It is equal to $\frac{1}{\theta\eta}(w_t - \sigma c_t)$ where w_t and c_t are observed. For the standard model, the predicted labor supply seems very weakly correlated with the actual one and strongly negatively correlated with the shock. The labor disutility shock is large not because this is necessary to explain large changes in the *observed* labor supply but because this explains a wide discrepancy between *predicted* changes in labor supply and *observed* ones. In other words, with the standard specification and the estimated value for σ , the model predicts large changes in labor supply because of observed changes in consumption and wages. These changes are not observed in data and the model has to create a large labor disutility shock in order to explain that the observed labor supply does not change.

With intertemporal complementarity, the main change for the predicted labor supply is a much lower estimated value for σ . The serie seems more reasonable. The predicted value still seems unable to explain the actual one but does not induce an artificially large labor disutility shock. Data points to moderate values for σ . But, such values

are not compatible with the observed response of consumption to real interest rate. Introducing the parameter ν allows to combine a moderate income effect and a small elasticity of consumption with respect to real interest rate.

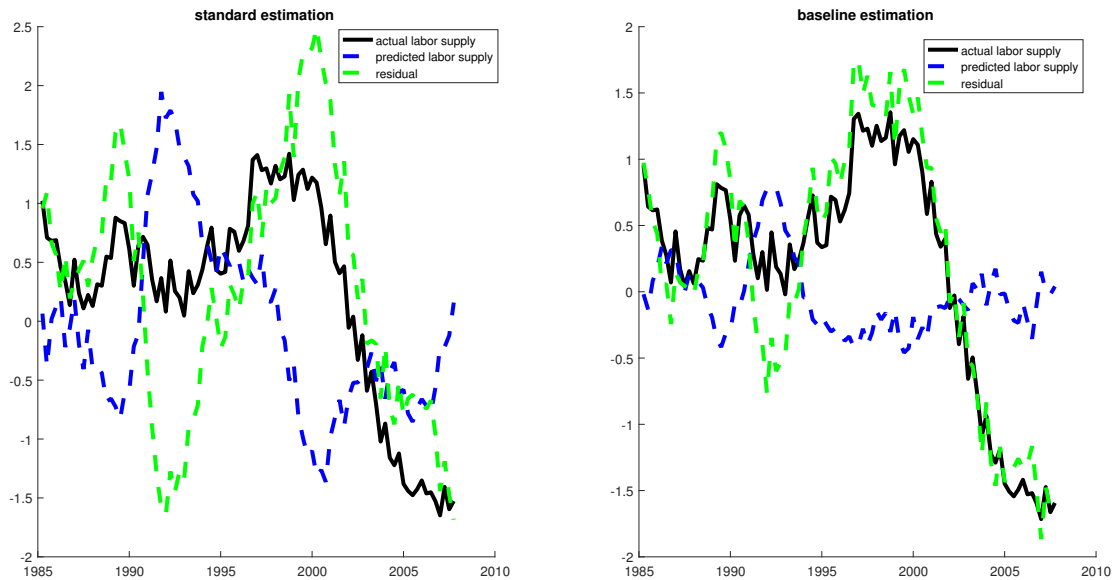


FIGURE 2.9 – Predicted and actual labor supply

2.5.3 Robustness

I now perform several robustness exercises. Posterior mean and confidence interval of κ , ν and σ are displayed in table 2.7 for all these exercises.

In the first one, I assume a different prior for κ . κ still follows a beta distribution but whose mean is 0.2 and whose standard deviation is 0.1. Results show that the posterior estimate of κ is lower than in the baseline estimation but still very high around 0.55.

In a second one, I introduce some form of habits consumption. Indeed, not adding a lag in the consumption equation may introduce some bias. I choose external habits.

Utility is provided by $c_t - hC_{t-1}$ where c_t is individual consumption and C_{t-1} aggregate consumption. h is calibrated at 0.7. Posterior mean of κ and ν are slightly lower than in the baseline estimation. The more affected parameter is σ whose posterior mean collapses.

In the third one, I estimate the model again by using a longer sample from 1954:3 to 2007:4. The estimate κ is still at 0.8 whereas ν is larger than in the baseline estimation.

In a fourth "experiment", I replace hours worked by civilian labor force participation among the observables. Results for ν and κ are roughly equivalent. Interestingly the value of σ is lower than in the baseline estimation, around 0.4, suggesting that income effect on labor supply is lower with labor force participation being the measure of labor supply.

In the fifth exercise, the discount rate in price and wage Philips curve is directly related to households' discount rate. Results seems unaffected. I also perform an estimation in which β_π and β_w are estimated. It leads to very low values for both of them (0.13 and 0.27 respectively) but also to a more reasonable value for γ_p , close to 0.08 compatible with a yearly frequency for price change.

The value of γ_p is very small in my baseline estimation (and in the standard one). Eventually, I estimate the model with γ_p and γ_w calibrated at 0.025. Whereas I obtain a larger estimate for σ , I still find large and positive values for κ and ν .

	Post. mean for κ	Post. mean for ν	Post. mean for σ
Alter. prior for κ	0.556 [0.4,0.7]	2.607 [1.56,3.8]	1.14 [0.64,1.7]
Habits consumption	0.652 [0.41,0.91]	1.67 [0.75,2.5]	0.196 [0.03,0.35]
Long sample	0.86 [0.76,0.98]	3.542 [2.37,4.51]	0.75 [0.4, 1.1]
Labor force part. as observable	0.82 [0.69,0.96]	2.80 [1.41,4.07]	0.43 [0.09, 0.81]
High disc. rate for NKPC	0.79 [0.60,0.94]	2.11 [1.11,2.94]	1.12 [0.65, 1.52]
γ_p and γ_w calibrated	0.7252 [0.53,0.93]	1.47 [0.4, 2.6]	1.49 [0.92,2.01]

TABLE 2.7 – Posterior mean of key parameters for different variants

2.5.4 Quantitative implications of the estimated model

In this paragraph, I look at IRFs generated by the estimated model. Figure 2.10 represents responses of output, expected real interest rate, unemployment and real wages. The fall in unemployment is in line with the rise of output suggesting that changes in unemployment are mostly explained by the demand side. The magnitude of the response of output is relatively low as compared to the response of the real rate. However, it should be kept in mind that this is only the response of consumption. Residential investment is a part of the exogenous demand shock d_t and thus is not sensitive to the real interest rate in the model, whereas in reality it represents a substantial part of the response of aggregate demand. The response of real wages is quite large, reflecting the high value of γ_w (e.g frequent price changes). Figure 2.11 displays response of labor demand and supply. Interestingly, labor supply increases following the expansionary monetary policy shock. In the calibrated exercise of section 3, labor supply was decreasing in the standard model and flat with intertemporal complementarity. I conjecture that this positive response of labor supply is related to the relatively large response of real wages. The fall in real rates pushes the labor supply curve to the left but labor supply also moves along the labor supply curve with the rise in real wages. Figure 2.12 shows the response of output and output gap following a demand shock. The two responses are quite close. It is worth noting that the demand shock generates a substantial rise in the natural interest rate.

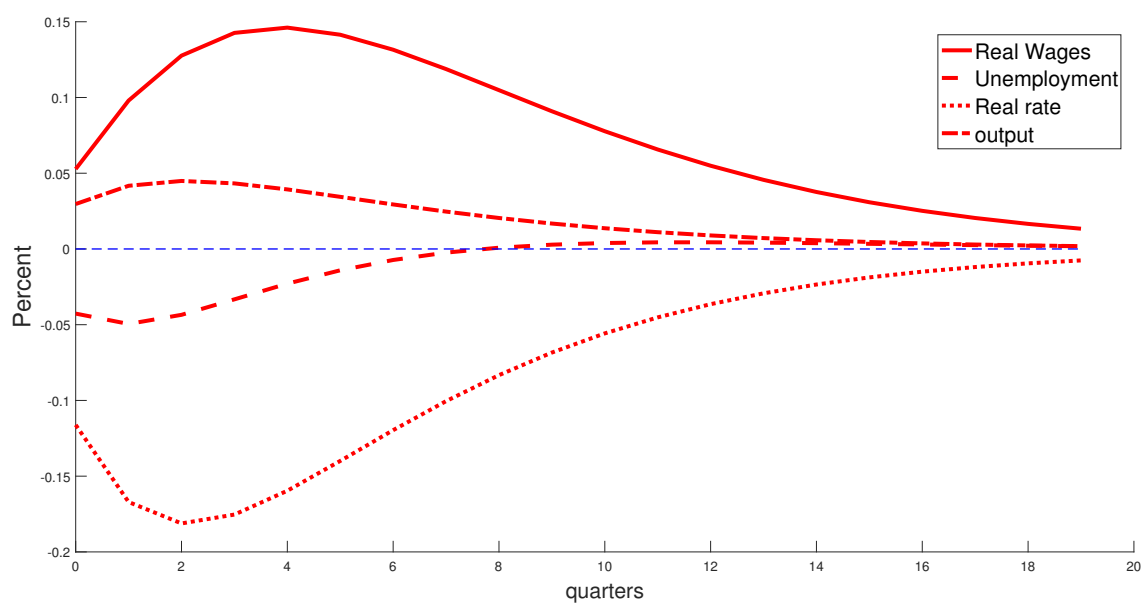


FIGURE 2.10 – Impulse responses to Monetary Policy shocks

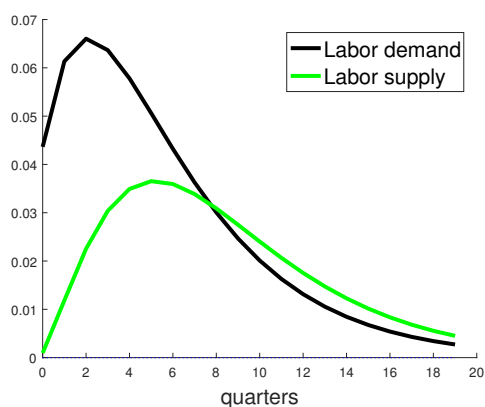


FIGURE 2.11 – Responses of labor supply and demand to Monetary Policy Shocks

2.6 Discussion

In this section, I discuss various pieces of literature which may support the specification proposed in this paper.

Labor market response to Monetary Policy Shocks Wealth in the utility function and intertemporal complementarity have important implications for the behavior of labor market variables following a monetary policy shocks. A substantial literature

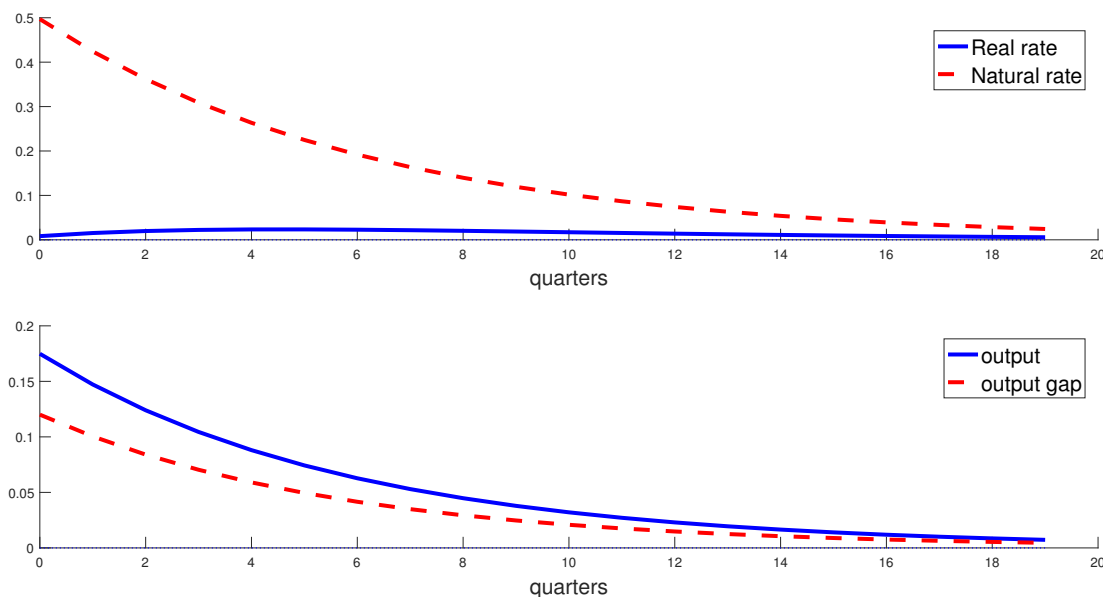


FIGURE 2.12 – Response of output gap to demand shocks

has dealt with the issue.

Intertemporal complementarity could help to explain the cyclical behavior of labor force participation. Pieces of evidence from VAR models collected by Christiano, Trabandt and Walentin (2010) suggest that labor force is mildly procyclical conditionally to a monetary policy shock. By contrast, labor supply, if identified in the way proposed by Galí (2011), responds positively to a contractionary monetary policy shock in the standard model because of the income effect on labor supply. Reducing the income effect is necessary to reconcile the model and the data. WIU and IC are a possibility to do so.

Simple New Keynesian model also fails to match the response of real wages to monetary policy shocks. Sims and Zha (1998) finds that a very persistent increase of nominal rate by 0.4 percent in annual value have not a significant impact on average real wages when looking at US postwar data. Christiano, Eichenbaum and Evans (1997) find relatively similar results. A less persistent 0.7 percent increase in nominal interest rate rises real wages by 0.1 percent. Both papers focus on US data, but estimations

across countries do not support a strongly procyclical response of real wages. Peersman and Smets (2001) finds that response of real wages is small in most countries of the euro area. Normandin (2006) finds a similar result for United Kingdom and Canada. These last two papers also find a counter-cyclical response of real wages for several countries. Wage rigidity provides a simple explanation for this mild procyclicality of real wages conditional to monetary policy shocks. However, the degree of wage stickiness needed to match data could be lower with an alternative specification of the utility function disentangling between the income effect on labor supply and the elasticity of intertemporal substitution.

Empirical evidence on euler equation A large empirical literature has dealt with the Euler equation. Hall(1988) finds no evidence of intertemporal substitution, a result confirmed for example by Yogo (2004). Another disappointing result was the negative correlation found between the FED funds rate and the real rate implied by consumption growth, found by Canzoneri et al. (2007) for several widely used consumption models. More positive results came from several papers by Attanasio and Weber (1993, 1995, 2010). Using microeconomic data on individual consumption and introducing controls for demographics and labor supply, they find a larger elasticity of intertemporal substitution. Results from these various papers are conflicting but pieces of evidence accumulated by the literature do not clearly endorse the standard model.

More importantly, such tests do not really allow the econometrician to choose between the standard model and a model with a discounted Euler equation, or even between the standard model and our wealth targeting model. Indeed, an equation giving consumption levels as a decreasing function of real interest rate also implies a positive correlation between consumption growth and real interest rate if changes in real interest rate are positively auto correlated. Conversely, under the same condition, an equation giving consumption growth as an increasing linear function of real interest

rate implies a negative correlation between real interest rate and consumption levels.

Indeed, consider a consumption equation in level like in the wealth targeting model.

$c_t = -\varphi r r_{t+1}$. Consumption growth becomes $c_{t+1} - c_t = -\varphi(rr_{t+2} - rr_{t+1})$, implying $cov(c_{t+1} - c_t, rr_{t+1}) = \varphi(1 - \rho_r)$ where ρ_r is the autocorrelation coefficient of real interest rate.

Consider now a consumption equation in growth $c_{t+1} - c_t = \sigma r r_{t+1}$. Correlation between consumption levels and real interest rate is equal to $cov(c_t, rr_{t+1}) = cov(c_t, \frac{c_{t+1}-c_t}{\sigma}) = \frac{\rho_c - 1}{\sigma}$ where ρ_c denotes consumption autocorrelation

Most macroeconomic shocks are positively auto correlated and persistence coefficients are often large. As a consequence, both consumption and real interest rate are very persistent at business cycles frequencies. Thus, a model with discounted Euler equation and the standard model have similar predictions for the sign of the two correlations.

Income effect on labor supply and low elasticity of intertemporal substitution Intertemporal complementarity disentangles between the income effect on labor supply and the intertemporal substitution effect.

A moderate but non negligible income effect is supported by survey directly asking to participants their labor supply response after an exogenous change in income (typically a lottery prize) (Kimball and Shapiro 2008). This result was confirmed by Cesarini et al. (2015). A moderate but not very small income effect is also supported the relative stability of hours worked in the long run despite large changes in real wages.

The small elasticity of consumption with respect to change in real interest rate also comes from a variety of observations. Direct estimations of the elasticity of intertemporal substitution have usually found a very low value for the EIS, sometimes close to zero (Hall 88, Yogo 2004). These small values are compatible with VAR evidence. The response of consumption to a positive monetary policy shocks is small given the response of the real interest rate (Bernanke and Gertler 95). Moreover, a large fraction

of this response could be generated by indirect effects of monetary policy rather than by the direct effect on intertemporal substitution (Auclert 2017, Kaplan, Moll and Violante 2017).

In the standard model, under separable preferences, these two facts are hard to reconcile. The parameter governing the income effect of labor supply is also the inverse of the intertemporal elasticity of substitution. A small EIS implies a very large income effect.

With intertemporal complementarity, the income effect is related to the parameter σ and $\sigma + \kappa\nu$ is the equivalent of the EIS. The complementarity between consumption and assets, measured by ν allows the elasticity of consumption to interest rate to be reduced substantially whereas keeping a moderate value for the income effect.

Precautionary saving and adjustment cost in consumption Another argument can be drawn in favor of complementarity between current consumption and future wealth. I show in appendix B.5 that it naturally arises when habits consumption and precautionary savings are combined. Consider an agent living two periods. He works and consumes in period one. In period 2, He does not work but uses assets accumulated in period one to consume. The utility at period 2 is affected by habits. It is not given by period two consumption but by the difference between the consumption of period two and a fraction of the consumption of period one. It is possible to rewrite the decision problem of period one by replacing the period two utility function by an indirect function depending on assets and period one consumption. The cross derivative of this indirect utility function is positive, indicating complementarity.

2.7 Implications for the forward guidance puzzle

A related issue is the forward guidance puzzle. Michailat and Saez (2018) study consequences of the wealth in the utility function for forward guidance at the zero lower bound. They keep a separable form for the utility function. In this section, I show that intertemporal complementarity may also help to solve the puzzle. I give a formal characterization of forward guidance by computing the response of output and inflation to an expected monetary policy shock in the New Keynesian model. Whereas it does not encompass all forms of forward guidance, the experiment clearly shows the overreaction of output. Then, I show analytically that intertemporal complementarity dampens the response.

A formal characterization The response to an expected shock on interest rate depends on the duration between the announcement and the realization of the shock but also depends on the contemporary reaction of monetary policy with respect to inflation and output gap. I choose the lower computational burden. I compute output and inflation multipliers with respect to an expected shock on nominal interest rate with two additional assumptions. First, the expected shock occurs in period $t + 1$. Second, the nominal interest rate in period t is kept constant by the central bank and do not react either to inflation or output gap. The underlying idea is that multipliers for other forward guidance shocks are linked to multipliers for this simple case.

The multipliers for output and inflation in response to such expected shock are denoted by \mathcal{M}_y and \mathcal{M}_π . I compute them relatively to multipliers associated with a contemporaneous monetary policy shock. I denote these multipliers Ψ_y and Ψ_π . I use a baseline New Keynesian model (see details in appendix [B.6](#))

$$\mathcal{M}_y = \left(1 + \frac{\psi(\theta\eta + \sigma\alpha + 1 - \alpha)}{\sigma(1 - \beta_\pi\rho)\alpha} \right) \Psi_y \geq \Psi_y$$

$$\mathcal{M}_\pi = \Psi_\pi \left(\beta_\pi + \frac{\psi}{\alpha\sigma}(\theta\eta + \sigma\alpha + 1 - \alpha) \right)$$

Current output overreacts. Its response to the future shock is always superior to the response to a current shock. Inflation is very likely to overreact as well. The response of inflation is superior to β_π which is usually close to 1.

Forward guidance and the supply effect of interest rate It is worth noting that the supply effect of interest rate enhances a forward guidance shock. Indeed, the term $\sigma\alpha$ at the numerator of the expression of \mathcal{M}_y is a consequence of the supply effect. In a imaginary world in which labor supply would only depend on wages and not on consumption whereas the Euler equation for consumption remains unchanged, the expression for \mathcal{M}_y would be $\left(1 + \frac{\psi(\theta\eta+1-\alpha)}{\sigma(1-\beta_\pi)\alpha} \right) \Psi_y$. There is still a substantial overreaction but lower than in the standard model.

Forward guidance and intertemporal complementarity I now compute the response to a forward guidance shock in the WIU model. Consumption is given by equation (2.19b). Computing the output multiplier leads to

$$\mathcal{M}_y = \left(\frac{(1 - \kappa)\sigma}{\varphi} + \frac{\psi(\theta\eta + \sigma\alpha + 1 - \alpha)}{\varphi(1 - \beta_\pi\rho)\alpha} \right) \Psi_y$$

where $\varphi \equiv \sigma + \kappa\nu$. Note that $\frac{(1-\kappa)\sigma}{\varphi}$ is simply the discount rate in the linear euler equation. The multiplier with respect to expected shocks is not always superior to the multiplier with respect to current monetary policy shock. There is no longer systematic overreaction of current output. A high value of φ dampens the response to forward

guidance announcements not only by increasing the discount in the euler equation but also by diminishing the effects of expected inflation.

2.8 The model with varying wealth

Until then, I have identified wealth with safe assets and I have supposed that the supply of bonds is not varying. In this section, I allow for a varying supply and for different interpretations. It may alter significantly the response of leisure and consumption in the wealth targeting model. I conjecture that similar issues may arise in the wealth in the utility model. I assume separable preferences to make computations easier.

The system of equations defined in proposition 2.3 may be rewritten

$$-\sigma c_t = -\gamma a_{t+1} - q_t \quad (2.27a)$$

$$-\theta l_t = w_t - q_t - \gamma a_{t+1} \quad (2.27b)$$

where parameters σ and θ are usual and γ is defined by

$$\gamma = -\frac{U_{AAA}}{U_A}$$

In the first order conditions above, a represents the amount of wealth *desired* by households for a given real wage and a given real interest rate. At this stage, they should not be viewed as equilibrium values even if obviously those two equations holds at equilibrium.

How wealth reacts to changes in interest rate deeply modifies the response of leisure

and consumption. If wealth is an increasing function of q_t , the two responses are enhanced whereas they are dampened if wealth is a decreasing function of q_t .

I now give three examples of wealth which gives different outcomes for monetary policy shocks.

A broad definition of wealth The first example is the closest to the standard model. The wealth is defined as the sum of the financial wealth and the labor wealth

$$A_t = F_t + \Omega_t$$

Where, F_t is the financial wealth and Ω_t is the labor wealth defined by the recursive equation

$$\Omega_t = W_t + Q_t \Omega_{t+1}$$

If the previous equation is iterated forward, the labor wealth is the discounted sum of future real wages

$$\Omega_t = \sum_{T=0}^{+\infty} \left[\prod_{k=0}^T Q_{t+k} \right] W_{t+T}$$

It is easy to see this model is very close to the standard model in many respect. Indeed, the budget constraint is nothing else than the usual intertemporal budget constraint. $\Omega_{t+1} + F_{t+1}$ is the income at period $t + 1$ and thus is equal to the discounted sum of consumption and leisure spending. The intertemporal first order condition relates current consumption C_t with future wealth $\Omega_{t+1} + A_{t+1}$. In period $t + 1$, consumption C_{t+1} will be a function of this wealth. Thus, you recover an equation linking current consumption with future ones and real interest rate.

Predictions with respect to monetary policy effects are also close. Assuming that supply of financial assets is null and linearizing of labor wealth gives

$$a_{t+1} = \omega_{t+1}$$

$$\omega_t = (1 - \beta)w_t + \beta(\omega_{t+1} + q_t)$$

A persistent fall in real interest rate will increase interest factors q and the future wealth ω_{t+1} . It will cause a rise in consumption and in leisure, implying a strong response of real wage to a current shock and strong responses to forward guidance.

Fixed public debt The second example leads to a very different conclusion. I assume that wealth is only financial and take the form of public debt. Responses to monetary policy are affected by fiscal policy.

I assume that the government has a very simple fiscal policy rule. Public debt, denoted B_t is fixed, equal to \bar{B} . Market clearing for public debt implies

$$Q_t A_{t+1} = \bar{B}$$

The linearized equation is simply

$$a_{t+1} = -q_t$$

Consumption and leisure may be expressed with respect to real wages and interest rate

$$\sigma c_t = (1 - \gamma)q_t$$

$$\theta l_t = (1 - \gamma)q_t - w_t$$

If $\gamma > 1$, both leisure and consumption becomes an increasing function of real interest rate. Obviously, under such parameters, effect of current and future monetary policy shocks on output are reversed.

A two agent framework The third example shows how preferences heterogeneity may affect response to monetary policy. I consider a model with two types of agents. Agents differ by the elasticity of their wealth to interest rate for a given marginal utility of consumption and by their Frisch elasticity of leisure. The first type is called "debtors" and the second type "creditors". Behavioral equations are more complicated, but aggregation and linearization remain straightforward (see appendix B.7). Combining them leads to a three equation system for leisure, consumption and financial assets of creditors (also equal to the financial liabilities of debtors). Asset distribution affects both leisure and consumption and makes their response to change in real interest rates ambiguous.

The asset distribution equation also introduces an endogenous amplification mechanism in the model. A shift in creditors' assets in period t will affect labor supply in period $t + 1$. Following an expansionary monetary policy shocks, creditors lower their savings and increase their labor supply on impact. At the next period, they have fewer assets and thus reduce their labor supply and their consumption through income effect.

2.9 Conclusion

In this paper, I refined wealth in the utility function model by considering a more general specification of the utility function. I show that complementarity between consumption and future wealth may have important implications for the New Keynesian model. It disentangles the income effect on labor supply and the intertemporal substitution effect, allowing for a lower response of labor supply to monetary policy shocks and a larger response of output gap to a demand shock. Then, I estimate the model using bayesian methods. I find a large value for both ν which governs the intertemporal complementarity and κ which governs the discount factor in the Euler equation. These findings probably reflects the ability of the model to provide a more plausible time serie for labor supply than the standard model. They are robust to several alternative specifications.

Bibliography

- [1] **Abel. A.B. and Warshawsky. M.** 1988. "Specification of the Joy of Giving: Insights from Altruism", *The Review of Economics and Statistics*, 70(1), 145-149.
- [2] **Attanasio. O. and Davis. S.J.** 1996. "Relative Wage Movements and the Distribution of Consumption", *Journal of Political Economy*, 104, 1227-1262.
- [3] **Auclert. A.** 2017. "Monetary policy and the redistribution channel", *The Review of Economics and Statistics*, mimeo.
- [4] **Barakchian. S.M. and Crowe. C.** 2013. "Monetary policy matters : Evidence from new shocks data", *Journal of Monetary Economics*, 60, 950-966.
- [5] **Bernanke. B., and Gertler. M.** 1995. "Inside the black box: The credit channel of Monetary Policy Transmission", *NBER working papers*, 5146.
- [6] **Bernanke. B., and Mihov. I.** 1998. "Measuring Monetary Policy", *The Quarterly Journal of Economics*, 113, 869-902.
- [7] **Bilbiie. F.** 2011. "Nonseparable Preferences, Frisch Labor Supply, and the Consumption Multiplier of Government Spending: One Solution to a Fiscal Policy Puzzle", *Journal of Money, Credit and Banking*, 43, 221-251.
- [8] **Bilbiie. F.** 2009. "Nonseparable Preferences, Fiscal Policy Puzzles, and Inferior Goods", *Journal of Money, Credit and Banking*, 41, 443-450.
- [9] **Bilbiie. F.** 2017. "The New Keynesian Cross: Understanding Monetary Policy with Hand-to-Mouth Households", mimeo.

-
- [10] **Camerer. C., Babcock. L., Loewenstein.G., Thaler. R.** 1997. "Labor supply of New York city cabdrivers : one day at a time", *Quarterly Journal of Economics*, 407-441.
- [11] **Campbell. J. and Mankiw. G.** 1990. "Permanent Income, Current Income, and Consumption" , *Journal of Business Economics and Statistics*, 8, 265-279.
- [12] **Canzoneri. M. B., Cumby. R. E. and Diba. B. T.** 2007. "Euler equations and money market interest rates: A challenge for monetary policy models", *Journal of Monetary Economics*, 79(1), 1863-1881.
- [13] **Cesarini. D., Lindqvist. E., Notowidigdo. M. and Ostling. R.** 2017. "The Effect of Wealth on Individual and Household Labor Supply: Evidence from Swedish Lotteries", mimeo.
- [14] **Challe. E. , Matheron. J. , Ragot. X. and Rubio Ramirez. J.** 2017. "Precautionary saving and aggregate demand", *Quantitative Economics*, 8, 435-478.
- [15] **Christiano. L.J. , Eichenbaum. M. and Evans. C.L.** 1997. "Sticky Price and Limited Participation Models of Money: A Comparison", *European Economic Review*, 41, 1201-1249.
- [16] **Christiano. L.J. , Trabandt. M. and Walentin. K.** 2010. "Involuntary Unemployment and the Business Cycles", NBER working papers.
- [17] **Christiano. L.J. , Eichenbaum. M. and Evans. C.L.** 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy", *Journal of Political Economy*, 113(1), 1-45.
- [18] **Del Negro. M., Giannoni. M., and Patterson, C.** 2013. "The Forward Guidance Puzzle", Federal Reserve Bank of New York Staff Report 574.

-
- [19] **Fahri. E. and Werning. I.** 2017. "Monetary Policy, Bounded Rationality and Incomplete Markets", mimeo.
- [20] **Frisch. R.** 1959. "A Complete Scheme for Computing All Direct and Cross Demand Elasticities in a Model with Many Sectors", *Econometrica*, 177-196.
- [21] **Gabaix. X.** 2014. "A Sparsity-Based Model of Bounded Rationality", *Quarterly Journal of Economics*, 1661-1710.
- [22] **Gabaix. X.** 2017. "A Behavioral New Keynesian Model", mimeo.
- [23] **Gali. J.** 2008. *Monetary Policy, Inflation, and the Business Cycle - An Introduction to the New Keynesian Framework*, Princeton University Press.
- [24] **Gali. J.** 2011. *Unemployment Fluctuations and Stabilization Policies*, MIT Press.
- [25] **Gali. J.** 2015. "Hysteresis and the European Unemployment Revisited", NBER Working Papers.
- [26] **Gali. J., Smets. F., Wouters. R.** 2011. "Unemployment in an Estimated New Keynesian Model", NBER Working Papers.
- [27] **Gertler. M. and Karadi. P.** 2015. "Monetary Policy Surprises, Credit Costs, and Economic Activity", *American Economic Journal: Macroeconomics*, 7(1), 44-76.
- [28] **Gourinchas. P.O. and Parker. J.** 2002. "Consumption over the Life Cycle", *Econometrica*, 70(1), 47-89.
- [29] **Greenwood. M., Hercowitz. S. and Huffman. G.** 1994. "Investment, capacity utilization, and the real business cycle", *American Economic Review*, 78, 402-417.
- [30] **Hall. R.** 1988. "Intertemporal Substitutability in Consumption", *Journal of Political Economy*, 96, 339-357.

-
- [31] **Kaplan. G., Moll. B. and Violante. G** 2018. "Monetary Policy According to HANK", *American Economic Review*, 108(3), 697-743.
- [32] **Kimball. M. and Shapiro. M.** 2008. "Labor Supply: Are the Income and Substitution Effects Both Large or Both Small?", NBER Working Papers.
- [33] **Krishnamurthy. A. and Vissing Jorgensen. A.** 2012. "The Aggregate Demand for Treasury Debt", *Journal of Political Economy*, 120(2), 233-267.
- [34] **Kumhof. M. , Ranciere. R. and Winant. P.** 2015. "Inequality, leverage and crisis", *American Economic Review*, 105(3), 1217-1245.
- [35] **Lusardi. A.** 2008. "Household Saving Behavior: The Role Of Financial Literacy, Information, And Financial Education Programs", NBER working papers, 13824.
- [36] **Michaillat. P. and Saez. E.** 2014. "An Economical Business-Cycle Model", NBER working papers.
- [37] **Michaillat. P. and Saez. E.** 2018. "Wealth in the Utility function and the New Keynesian model", mimeo.
- [38] **Michau. J.B.** 2017. "Secular Stagnation: Theory and Remedies", mimeo.
- [39] **McKay. A., Nakamura. E., and Steinsson. J.** 2016. "The Power of Forward Guidance Revisited", *American Economic Review*, 106(10), 3133–3158.
- [40] **McKay. A., Nakamura. E., and Steinsson. J.** 2016. "The Discounted Euler Equation: A note", mimeo.
- [41] **Normandin. M.** 2006. "The Effects of Monetary-Policy Shocks on Real Wages: A Multi-Country Investigation", mimeo.
- [42] **Peersman. G. and Smets. F.** 2001. "The monetary transmission mechanism in the euro area : more evidence from VAR analysis", ECB working papers.

-
- [43] **Romer. C.D., and Romer. D.H.** 1989. " Does Monetary Policy Matter? A New Test in the Spirit of Friedman and Schwartz," in Olivier J. Blanchard and Stanley Fischer, eds., NBER macroeconomics annual 1989. Cambridge, MA: MIT Press, 121–70.
- [44] **Romer. C.D., and Romer. D.H.** 2004. "A New Measure of Monetary Shocks: Derivation and Implications", *American Economic Review*, 94(4), 1055-1084.
- [45] **Saez. E. and Stantcheva. S.** 2016. "Simpler Theory of Optimal Capital Taxation", NBER working papers.
- [46] **Sims. C. and Zha. T.** 1998. "Does Monetary Policy Generate Recessions?", FRB Atlanta Working Papers.
- [47] **Smets. F., Wouters. R.** 2007. "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach", *American Economic Review*, 97(3), 586-606.
- [48] **Uhlig. H.** 2005. "What are the effects of monetary policy on output? Results from an agnostic identification procedure", *Journal of Monetary Economics*, 52, 381-419.
- [49] **Woodford. M.** 2005. *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University press.
- [50] **Woodford. M.** 2010. "Simple Analytics of the Government Expenditure Multiplier", NBER Working Papers.
- [51] **Yaari. M.** 1964. "On the consumer's lifetime allocation process", *International Economic Review*, 304-317.
- [52] **Yogo. M.** 2004. "Estimating the Elasticity of Intertemporal Substitution when instruments are weak", *The Review of Economics and Statistics*, 797-810.

Chapter 3

Corporate Investment and Adverse Selection, A Reappraisal

3.1 Introduction

The literature integrating financial frictions on the corporate sector into macroeconomic models has been successful at explaining some stylized facts, like the substantial rise in corporate spreads during the financial crisis. However, some doubts may still be cast about the relevant microfoundations. The more popular friction in macroeconomic model is the costly state verification model introduced by Townsend (1979). In this model, lenders should pay an auditing cost in order to observe the outcome of the firm. The borrower may declare himself bankrupt, avoid the repayment of the debt and run away with the profit and the assets of the firm unless the lender pay the auditing cost and verify the outcome. Such ability to divert assets seems implausible. It could be relevant for small firms but it seems hard to imagine shareholders of publicly listed firms announcing large losses before running away with the buildings, patents or machines tools of the corporation. Despite this issue, medium scale macroeconomic models like Bernanke, Gertler and Gilchrist(1999) or Christiano, Motto and Rostagno (2014) assume that the friction is relevant for all firms and identify borrowers with shareholders. Other popular friction like limited commitment to repay or moral hazard

also assume wide diversion possibilities for borrowers.

Given these limitations, it seems interesting to explore macroeconomic implications of alternative frictions. This paper contributes to the adverse selection literature by providing a tractable framework, easy to integrate into a macroeconomic model. Adverse selection on capital markets was introduced in the seminal paper of Stiglitz and Weiss (1981). In Stiglitz and Weiss, debtors have private information about the riskiness of investment project. They can signal their type and pay a lower interest rate if they accept a lower probability to get a loan. This offers an explanation for credit rationing. My model modifies and extends the Stiglitz and Weiss framework. Borrowers do not signal themselves by accepting a lower probability to get a loan but by constraining the amount they borrow to be a fraction of their retained earnings. I solve the problem in infinite horizon whereas the Stiglitz and Weiss framework is purely static. Surprisingly, it allows me to obtain a simple analytic solution for the incentive compatibility constraint. I embed the friction in a dynamic model of corporate investment whose solution is straightforward and derive some implications for the response of investment to changes in borrowing and opportunity cost.

Another paper close to mine is Kurlat(2013). In Kurlat's paper, borrowers run several projects which differ by their quality and have private information about the type of each project. Thus, the borrower has an incentive to resell "bad" projects to uninformed investors whereas keeping "good" projects in his portfolio. This lemon problem provides a microfoundation for the resaleability constraint introduced by Kiyotaki and Moore (2008). By contrast, in my model, borrowers run a single project that can be either bad or good. "Bad" borrowers have an incentive to imitate "good" borrowers in order to pay a lower interest rate. "Good" borrowers send a signal to avoid being imitated by bad borrowers and thus to avoid paying a higher interest rate.

The paper is organized as follows. I specify and solve the adverse selection problem in the first section. I develop the model of corporate investment in the second section

and derive some macroeconomic implications in the third section.

3.2 The adverse selection problem

In this section, I outline a partial equilibrium model of capital markets in which firms try to get funds from lenders in order to finance their investment. Firms hold private information about their riskiness. Firms try to signal their riskiness by constraining the amount they borrow to be a fraction of their retained earnings.

3.2.1 Firms and lenders

Production function The economy is populated by a continuum of firms. A firm j has a production function

$$Y^j = \pi K^j \tag{3.1}$$

.

Two type of firms There are two type of firms in the economy: Bad firms denoted with superscripts B and good firms denoted by the superscript G . Firms can observe their type but lenders cannot. Bad and good firms do not differ by their productivity but by their riskiness. At each period, a good firm has a probability κ to become a bad firm. With probability $1 - \kappa$, it remains a good firm. A bad firm still produces πK^j but has a probability $1 - \lambda$ to exit at the next period. When it exits, a firm produces nothing and its capital stock is worthless. Neither the lender nor the borrower recover anything. I summarize the timing by the tree [3.1](#).

The firm problem Firms accumulate capital by borrowing or using internal funds. They split their income between interest repayments, dividends and retained earnings.

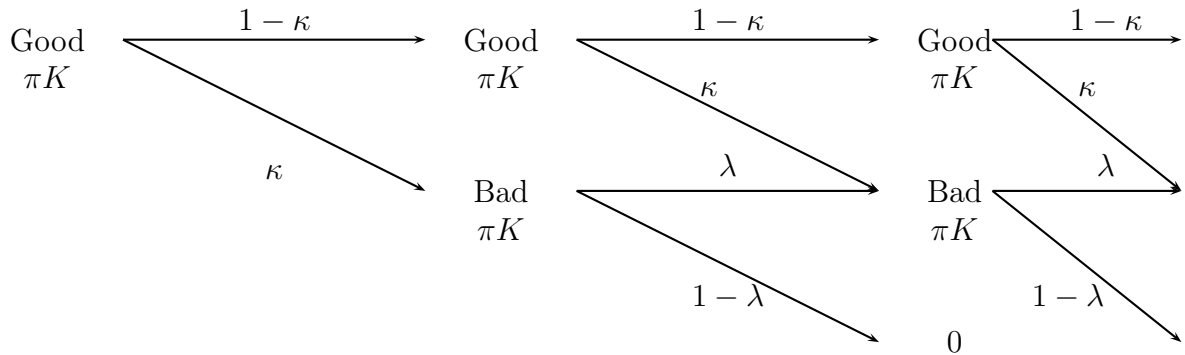


FIGURE 3.1 – State Tree

I assume that loans have an infinite maturity. Interest rates on past loans are fixed. Only interest rate on new loans may vary. Thus for an amount E_0 borrowed at period 0, the firm should pay the lender $r_0^e E_0$ at each period. At a given period t , the total repayment b_t of the firm is the sum

$$B_t = r_{t-1}^e E_{t-1} + r_{t-2}^e E_{t-2} + r_{t-3}^e E_{t-3} + \dots + r_{t-n}^e E_{t-n} + \dots$$

where E_{t-n} is the amount of money borrowed at period $t - n$ and r_{t-n}^e is the interest rate at period $t - n$ which may include a firm specific risk premium. The interest rate on good firms is denoted r^G and the interest rate on bad firms is denoted r^B

This assumption allows for a more tractable model. Introducing shorter maturities is interesting but creates complex issues about optimal maturity design that are not the core of this paper.

Firms maximize the discounted sum of expected dividends. Unlike lenders, firms' shareholders may invest outside the firm in a high return asset producing a yield γ . Thus, dividends are discounted using this yield γ . It is important to note that this discount rate is different from lenders' discount rate which is equal to the safe interest rate r . Distinguishing between the borrowing cost r and the opportunity cost for borrowers γ is useful because they play different roles in the model. However, the

model may be solved with $\gamma = r$. The value function of a firm is

$$V(K, B) = \max_{K', B', I, S, d} d + \frac{1}{1 + \gamma} EV'(K', B') \quad (3.2a)$$

$$w.r.t \quad K' = K + I \quad (3.2b)$$

$$B' = B + r^e(I - S) \quad (3.2c)$$

$$\pi K - B = d + S \quad (3.2d)$$

$$\pi K - B \geq d \geq (1 - s)(\pi K - B) \quad (3.2e)$$

$$I \geq 0 \quad (3.2f)$$

The first constraint is the law of motion for capital stock. Next period capital stock is equal to previous capital stock plus investment. There is no depreciation. The second constraint is the law of motion for *interest repayments*. Next period interest repayments B' are equal to current one plus the new borrowing $I - S$ times the specific interest rate charged on the firm r^e . The equation (3.2d) splits the income of the firm $\pi K - B$, equal to the production minus interests repayments, between dividends, denoted d , and retained earnings, denoted S . Dividends are constrained to have a minimal and a maximal value. They cannot be higher than the income of the firm and should be bigger than $(1 - s)(\pi K - B)$ where s is an exogenous parameter between 0 and 1. The last equation states that disinvestment is not possible.

An alternative interpretation of this program is that firms want to maximize the discounted value of shareholders' utility streams assuming they consume distributed dividends and that utility is linear with respect to consumption. The parameter γ would become the specific discount factor of firms' shareholders. This interpretation is close to the analysis developed in Kiyotaki and Moore (1997). Another assumption common with Kiyotaki and Moore (1997) is that a minimal fraction of earnings should be distributed as dividends, this minimal fraction $1 - s$ being set exogenously. Kiyotaki

and Moore (1997) interprets it as a nontradable output of the production process.

Given there are three equality constraints, firms have only to choose two variables to solve their optimization problem. Choosing investment I and retained earnings S automatically gives next period capital stock, next period debt repayments and dividends. Alternatively, firms may choose S and the investment over saving ratio $\frac{I}{S}$.

Lenders and safe interest rate There is a continuum of lenders behaving in a competitive way. They have the choice between lending to firms and buying a one period bond that generates a safe return r . They can observe current capital stock K , debt repayments B , earnings πK , dividends D , investment I and savings S of a firm but they do not observe the true type of the firm. Firms will try to reveal their type by sending an appropriate signal to lenders. Lenders are able to perfectly diversify their claims across good and bad firms. Thus, the price of a security issued by a given firm can be priced as if lenders are risk neutral.

I perform a partial equilibrium analysis. the safe interest rate is assumed to be exogenous. In this section, it is constant equal to r .

3.2.2 Equilibrium Definition under private information

Solving this problem under infinite horizon requires to define value functions of bad and good firms. However, value functions in period t depends on the interest rate paid by the firm hence on the belief of lenders regarding the type of the firm in period t but also from beliefs of lenders after the period t . To make the problem tractable, my approach is to assume that (i) good firms may signal their type. The signal consist in limiting their investment over retained earnings ratio to be below a certain threshold (ii) Good and bad firms compute the value of sending correct and incorrect signals in period t by assuming that they will send the correct one after period t (iii) value functions of bad and good firms are well defined. I characterize the equilibrium under these assumptions

and then verify that, at equilibrium, (i) The incentive compatibility constraint takes the form of a constraint on the ratio of investment over retained earnings (ii) Always sending the correct signal is superior to always sending the incorrect one for bad firms (iii) Value functions are well defined.

Definition of the signal Firms may signal their type by respecting an incentive compatibility constraint (ICC thereafter). The signal takes the form of a vector containing retained earnings denoted by S and the investment over retained earnings ratio. I call this ratio the leverage. I denote it by $\psi \equiv \frac{I}{S}$. The ICC implies that the leverage of good firms should be inferior to a certain value. A formal definition of the signal may be given

Definition 3.1. *The signal is a vector (S^G, ψ^G) . For a given capital stock and a given debt, good firms signal their true type to lenders by retaining some earnings $S^G > 0$ and by constraining their leverage ψ^G to be below a certain threshold $0 < \frac{I^G}{S^G} = \psi^G \leq \bar{\psi}$. Bad firms signal their true type either by having a leverage superior to $\bar{\psi}$ or by distributing all their income as dividends (e.g $S = 0$)*

I derive the ICC in proposition 3.7 and shows it actually takes the form of a constraint on the investment over savings ratio.

The intuition behind this signaling device is the following. Consider a firm investing one unit of good in period t and financing this investment through debt. At each following periods, the investment generates a return π and a debt repayment r^B if it is considered as a bad firm and r^G if it is considered as a good firm. The return is lower than the cost of borrowing for bad firms but higher than the cost of borrowing for good firms. Thus, if they can emulate good firms, bad firms have an incentive to invest. However, the return of investment in bad firms is also lower than the return of outside investment opportunities for bad firms' shareholders. Bad firms' shareholders have no incentive to invest their own funds in the firm. Thus, shareholders of good firms may

deter shareholders of bad firms to emulate them by financing a large fraction of their investment with internal funds.

Definition of the value functions The value of a given firm depends on its true type and on the signal send to lenders. If the firm sends the signal which allows lenders to identify it as a bad firm (the bad signal thereafter), it has to pay the interest rate charged on bad firms r^B . If it sends the signal of good firms (the good signal thereafter), it pays the interest rate r^G , but investment and savings are constrained by the threshold ratio $\bar{\psi}$.

I denote $V^{G,G}$ and $V^{G,B}$ the values for a good firm to send respectively a correct and a incorrect signal. $V^{B,G}$ and $V^{B,B}$ are the values for a bad firm to send respectively the signal of good firms (the incorrect one) and the correct signal. I denote (ψ^G, S^G) the signal send by a good firm. The policy rule of good firms (I^G, S^G) can be immediately deduced from the signal. I denote the policy rule of bad firms by (I^B, S^B) . Both values, signals and policy rules are *a priori* function of state variables, e.g capital stock K and debt B . However, the following analysis shows that they are actually independent from them.

The value of sending the incorrect signal is the value derived from an unconstrained optimization program under the assumption that the interest rate charged is the bad firm interest rate. The value of sending the correct signal is the value derived from the optimization program with the incentive compatibility constraint.

Computing the respective value associated to these two strategies raises some issue. In finite horizon, it is possible to start from the last period and to compute what is the best strategy in the last period. Then, you can compute the value of the two strategies for the previous period and repeat the same process until the current period. Such a solution is not available in infinite horizon. I use the following trick. I start by assuming that a the maximal value of a bad (resp. a good) firm sending the incorrect

signal in the current period is obtained if the firm sends the correct signal after the current period. I derive the ICC and solve for the equilibrium under this assumption. After deriving the ICC and solving for the equilibrium, I show that the value of always sending the correct signal is superior to the value of always sending the incorrect signal. Hence, the value sending the incorrect signal in current period and the correct one after is above the value of always sending the incorrect signal.

Hypothesis 3.1. *In period t , a firm computes its value using the assumption it will send the correct signal after period t .*

With this assumption, I can formally define values of good and bad firms for good and bad signals.

The value of sending the good signal for a good firm is

$$V^{G,G} = \max_{I,S,d,K',B'} d + (1 - \kappa) \frac{1}{1 + \gamma} V'^{G,G} + \kappa \frac{1}{1 + \gamma} V'^{B,B} \quad (3.3a)$$

$$w.r.t \quad K' = K + I \quad (3.3b)$$

$$B' = B + r^G(I - S) \quad (3.3c)$$

$$\pi K - B = d + S \quad (3.3d)$$

$$(\pi K - B) \geq d \geq (1 - s)(\pi K - B) \quad (3.3e)$$

$$I \geq 0 \quad (3.3f)$$

$$\frac{I}{S} \leq \bar{\psi} \quad (3.3g)$$

The value of sending the "bad" signal in period t is the value of the unconstrained

program in period t .

$$V^{G,B} = \max_{I,S,d,K',B'} d + (1 - \kappa) \frac{1}{1 + \gamma} V'^{G,G} + \kappa \frac{1}{1 + \gamma} V'^{B,B} \quad (3.4a)$$

$$w.r.t \quad K' = K + I \quad (3.4b)$$

$$B' = B + r^B(I - S) \quad (3.4c)$$

$$\pi K - B = d + S \quad (3.4d)$$

$$\pi K - B \geq d \geq (1 - s)(\pi K - B) \quad (3.4e)$$

$$I \geq 0 \quad (3.4f)$$

For bad firms, the two programs are similar

$$V^{B,G} = \max_{I,S,d,K',B'} d + \lambda \frac{1}{1 + \gamma} V'^{B,B} \quad (3.5a)$$

$$w.r.t \quad K' = K + I \quad (3.5b)$$

$$B' = B + r^G(I - S) \quad (3.5c)$$

$$\pi K - B = d + S \quad (3.5d)$$

$$\pi K - B \geq d \geq (1 - s)(\pi K - B) \quad (3.5e)$$

$$I \geq 0 \quad (3.5f)$$

$$\frac{I}{S} \leq \bar{\psi} \quad (3.5g)$$

$$V^{B,B} = \max_{I,S,d,K',B'} d + \lambda \frac{1}{1+\gamma} V'^{B,B} \quad (3.6a)$$

$$w.r.t \quad K' = K + I \quad (3.6b)$$

$$B' = B + r^B(I - S) \quad (3.6c)$$

$$\pi K - B = d + S \quad (3.6d)$$

$$\pi K - B \geq d \geq (1 - s)(\pi K - B) \quad (3.6e)$$

$$I \geq 0 \quad (3.6f)$$

$$(3.6g)$$

For the moment, I assume that these value functions exist and I verify *ex post* this is actually the case in proposition 3.10 and 3.12.

As I said previously, I conjecture that $V^{B,G}$ is superior to the value of always sending the incorrect signal for a bad firm defined as

$$U^{B,G} = \max_{I,S,d,K',B'} d + \lambda \frac{1}{1+\gamma} U'^{B,G} \quad (3.7a)$$

$$w.r.t \quad K' = K + I \quad (3.7b)$$

$$B' = B + r^G(I - S) \quad (3.7c)$$

$$\pi K - B = d + S \quad (3.7d)$$

$$\pi K - B \geq d \geq (1 - s)(\pi K - B) \quad (3.7e)$$

$$I \geq 0 \quad (3.7f)$$

$$\frac{I}{S} \leq \bar{\psi} \quad (3.7g)$$

Conjecture 2. *At the separating equilibrium, we have*

$$V^{B,G} > U^{B,G}$$

Definition of the separating equilibrium I can now define the separating equilibrium of this economy

Definition 3.2. *A separating equilibrium is a vector (r^B, r^G) , a vector of policy function for bad firms (I^B, S^B) , and a signal sent by good firms (ψ^G, S^G) such that*

- *The pair (r^B, r^G) respects the participation constraint of lenders*
- *(I^B, S^B) are the optimal policy function for the program (6).*
- *The signal (ψ^G, S^G) sent by good firms maximizes the program (3).*
- $\forall K, B, V^{G,G}(K, B) \geq V^{G,B}(K, B)$
- $\forall K, B, V^{B,B}(K, B) \geq V^{B,G}(K, B)$

The first condition states that lenders should prefer lending to firms rather than investing in the safe asset or should be indifferent between the two options. The second and the third imply that policy rules of good and bad firms maximize their respective optimization program. The fourth states that good firms have an incentive to send the good signal. The fifth is the incentive compatibility constraint for bad firms.

3.2.3 Characterization of the separating equilibrium

Interest rates

First, I compute interest rate for good and bad firms. Lenders have to be indifferent between lending to bad or good firms and buying the safe asset.

Proposition 3.1. *Interest rate for good and bad firms are given respectively by*

$$r^B \lambda Q'^B = 1 + r \quad (3.8a)$$

$$r^G [(1 - \kappa)Q'^G + \kappa Q'^B] = 1 + r \quad (3.8b)$$

Where

$$Q^B = 1 + \frac{\lambda}{1 + r} Q'^B \quad (3.9a)$$

$$Q^G = 1 + \frac{1}{1 + r} [(1 - \kappa)Q'^G + \kappa Q'^B] \quad (3.9b)$$

Equations (3.8a) and (3.8b) are non-arbitrage equations. On the left side is the return of the safe bond whereas on the right side lies the return of lending to a bad firm (resp. a good firms). Q^B can be interpreted as the value for a risk neutral lender of an income stream delivering one unit of consumption good starting from period t and ending when a given bad firm exits. Q^G has the same interpretation for a given good firm.

Similarly, I define $\gamma^B, \gamma^G, \Gamma^G, \Gamma^B$. They are the counterparts of r^B, r^G, Q^G, Q^B when the discount rate is given by γ instead of r

Definition 3.3.

$$\gamma^B \lambda \Gamma'^B = 1 + \gamma \quad (3.10a)$$

$$\gamma^G [(1 - \kappa)\Gamma'^G + \kappa\Gamma'^B] = 1 + r \quad (3.10b)$$

Where

$$\Gamma^B = 1 + \frac{\lambda}{1 + \gamma} \Gamma'^B \quad (3.11a)$$

$$\Gamma^G = 1 + \frac{1}{1 + \gamma} [(1 - \kappa) \Gamma'^G + \kappa \Gamma'^B] \quad (3.11b)$$

If γ and r are constant, all these values can be computed explicitly. In particular

Proposition 3.2. *Values of r^B , r^G , γ^G and γ^B are*

$$\begin{aligned} r^B &= \frac{1 + r - \lambda}{\lambda} \\ r^G &= \frac{(r + \kappa)(1 + r - \lambda)}{(1 - \kappa)(1 + r - \lambda + \kappa) + \kappa(r + \kappa)} \\ \gamma^B &= \frac{1 + \gamma - \lambda}{\lambda} \\ \gamma^G &= \frac{(\gamma + \kappa)(1 + \gamma - \lambda)}{(1 - \kappa)(1 + \gamma - \lambda + \kappa) + \kappa(\gamma + \kappa)} \end{aligned}$$

Proof Results follow from standard results on geometric sums. Computations for r^B and r^G are given in appendix C.1.1 □

The behavior of lenders is now easy to characterize.

Proposition 3.3. *The interest rate charged on bad firms is $r^B = \frac{1+r-\lambda}{\lambda}$ and the rate charged on good firms is $r^G = \frac{(r+\kappa)(1+r-\lambda)}{(1-\kappa)(1+r-\lambda+\kappa)+\kappa(r+\kappa)}$*

Proof The proof follows immediately from proposition 3.2 □

Policy rules of bad and good firms

To compute the optimal behavior of borrowers, I need to make two parametric assumptions.

Hypothesis 3.2. Parameters λ, π, γ and the real interest rate r are such that

$$r^B = \frac{1 + r - \lambda}{\lambda} > \pi$$

and

$$\gamma^B = \frac{1 + \gamma - \lambda}{\lambda} > \pi$$

The first condition states that the borrowing cost should be superior to the investment return in bad firms. The second condition is equivalent to assume that the return of outside opportunity for shareholders is superior to the return of investment in bad firms.

Hypothesis 3.3. Parameters $\lambda, \pi, \gamma, \kappa$ and the real interest rate r are such that

$$\pi > \frac{(\gamma + \kappa)(1 + \gamma - \lambda)}{(1 - \kappa)(1 + \gamma - \lambda + \kappa) + \kappa(\gamma + \kappa)} = \gamma^G$$

and

$$\pi > \frac{(r + \kappa)(1 + r - \lambda)}{(1 - \kappa)(1 + r - \lambda + \kappa) + \kappa(r + \kappa)} = r^G$$

The two conditions states that the return of investment in good firms should be superior to the borrowing cost of good firms and the return of outside opportunities.

Those restrictions raise some questions about the relevance of the model. I investigate their quantitative importance in a dedicated subsection.

With this set of assumptions, I am able to determine and characterize the best policy of the two type of firms when their true type is revealed to lenders.

Proposition 3.4. *If hypothesis 2 holds, under separating equilibrium, bad firms do not save, invest and borrow. $S^B = 0, I^B = 0, d^B = \pi K - B$*

Proof An investment financed by debt in a bad firm generates a benefit π and a cost r^B at each period. $r^B = \frac{1+r-\lambda}{\lambda} > \pi$. Thus, bad firms do not want to borrow.

The case of an investment financed by equity raises more issue. To demonstrate that a bad firm do not want to invest its earnings if $\frac{1+\gamma-\lambda}{\lambda} > \pi$, I assume that the firm invest zero and saves zero and I consider a small deviation from this path. There is an increase in investment by one unit only in period t . The investment implies a lower dividend in period t but generates a constant income stream after. The variation of the value of the firm is

$$\Delta V^{B,B} = -1 + \sum_{k=1}^{+\infty} \left(\frac{\lambda}{1+\gamma} \right)^k \pi$$

I show that (see detailed proof in appendix C.1.2) $\Delta V^{B,B} < 0$ if $\pi < \frac{1+\gamma-\lambda}{\lambda}$. Thus an investment financed by equity will generate a net loss for the bad firm. Note that if the firm reinvests the earnings generated by the investment in period t , its loss increases. Thus, a bad firm do not want to invest under hypothesis 2. It distributes all its income as dividends. \square

Proposition 3.5. *If hypothesis 3 holds, under separating equilibrium, good firms wants to invest and borrows as much as possible. They also distribute the minimum level of dividends $\frac{I^G}{S^G} = \bar{\psi}$, $d^G = (1-s)(\pi K - B)$*

Proof The proof follows the same strategy than the previous one. I show that both an investment financed by debt and an investment financed by retained earnings are profitable for the good firm.

Hypothesis 3 implies that $\pi > r^G$. Thus, investment financed by debt is profitable.

I denote $\Delta V^{G,G}$ the variation of value following a small rise in investment financed by retained earnings in period t I show (see detailed proof in appendix C.1.3) that $\Delta V^{G,G} > 0$ if $\pi > \frac{(\gamma+\kappa)(1+\gamma-\lambda)}{(1-\kappa)(1+\gamma-\lambda+\kappa)+\kappa(\gamma+\kappa)}$.

Under hypothesis 3, equity financed investment is profitable. Thus, good firms want

to invest and to borrow as much as they can. They distribute the minimal level of dividends and invest until the ratio of investment over saving reaches the threshold ratio $\bar{\psi}$ □

Proposition 3.6. *If hypothesis 2 holds, good firms prefer revealing their true type.*

$$V^{G,G}(K, B) \geq V^{G,B}(K, B)$$

Proof If a good firm borrows above its leverage threshold, it has to pay the interest rate charged on bad firm r^B . The additional investment generates an annual income π and an annual debt repayment r^B . Under hypothesis 2, $r^B > \pi$. Thus, this strategy carries no benefit for the good firm. □

The incentive compatibility constraint

I now explore the incentive compatibility constraint and determine the leverage threshold $\bar{\psi}$. For the bad firm, the value of distributing all its earnings as dividend and not investing whereas paying a high interest rate on its debt should be superior to the value of paying a lower interest rate whereas investing the same fraction of its earnings as good firms. The incentive compatibility constraint can be expressed in a simple way

Proposition 3.7. *We have $V^{B,B} \geq V^{B,G}$ if and only if*

$$I^G \leq \frac{\gamma^B - r^G}{\pi - r^G} S^G = \psi S^G$$

Remarkably, the signal send by good firms does not depend on state variables. Whereas different generations of firms coexist in the model, aggregation is straightforward. The incentive compatibility constraint allows me to fully characterize the behavior of good firms.

Proposition 3.8. *Under assumption (3) Good firms want to invest and to borrow as*

much as possible, thus the incentive compatibility constraint is binding

$$I^G = \frac{\gamma^B - r^G}{\pi - r^G} S^G = \psi S^G \quad (3.12)$$

Proof It is consequence of propositions 3.7 and 3.5 □

To better understand the underlying mechanism, it is useful to consider all the steps of the proof of proposition 3.7. The incentive compatibility constraint implies that the value for a bad firm of sending the "bad" signal should be superior to the value of sending the "good" signal.

$$V^{B,B} \geq V^{B,G} \quad (3.13)$$

I need to compute these value functions. The trick is that, under separating equilibrium, bad firms do not invest nor borrow. Their capital stock and their debt remain equal to those of previous period. Thus, the value $V^{B,B}$ can be explicitly computed.

$$V^{B,B}(K, B) = \pi K - B + \lambda \frac{1}{1 + \gamma} V'^{B,B}(K, B) \quad (3.14)$$

Iterating forward the right side of the expression and factorizing by $\pi K - B$ leads to

$$V^{B,B} = (\pi K - B) \Gamma^B \quad (3.15)$$

By a similar reasoning, I obtain $V'^{B,B} = (\pi K - B) \Gamma'^B$. The value of function $V^{B,G}$ may be deduced from $V^{B,B}$ thanks to assumption 3.1. In the current period, the bad firm does not distribute all its earnings as dividends but must retain as much earnings as a good firm. The bad firm also has to invest and borrow, leading to higher capital and debt stock at the next period. But, at the next period, the bad firm is supposed to send the correct signal and the relevant value function is $V'^{B,B}$, hence

$$V^{B,G}(K, B) = \pi K - B - S^G + \frac{\lambda}{1+\gamma} V'^{B,B}(K + I^G, B + r^G(I^G - S^G)) \quad (3.16)$$

I now rewrite the equation (3.13). I replace $V_{B,B}$ and $V_{B,G}$ by their expressions from (3.14) and (3.16).

$$\pi K - B + \lambda \frac{1}{1+\gamma} V'^{B,B}(K, B) \geq \pi K - B - S^G + \frac{\lambda}{1+\gamma} V'^{B,B}(K + I^G, B + r^G(S^G - I^G))$$

Using the expression of $V'^{B,B}$

$$\pi K - B + \frac{\lambda}{1+\gamma} (\pi K - B) \Gamma'^B \geq \pi K - B - S^G + \frac{\lambda}{1+\gamma} (\pi(K + I^G) - (B + r^G(I^G - S^G))) \Gamma'^B$$

By simplifying, I get

$$\frac{1+\gamma}{\lambda} S^G \geq \Gamma'^B \pi I^G - \Gamma'^B r^G (I^G - S^G)$$

By equation (3.10a)

$$\gamma^B \Gamma'^B = \frac{1+\gamma}{\lambda}$$

It is easy to see that the incentive compatibility constraint can be rewritten in the following way

$$I^G \leq \frac{\gamma^B - r^G}{\pi - r^G} S^G \quad (3.17)$$

The incentive compatibility constraint takes the form of a linear relation between investment and cash flows or retained earnings. This constraint is about flows and not stocks unlike most of popular financial frictions in macroeconomics.

Summary of the equilibrium

I now summarize equilibrium of the capital market.

Proposition 3.9. *Under hypothesis (1)-(3), the equilibrium is characterized by the following policy rules and the following price vector (here the vector of interest rates).*

1. *For a given capital stock K and a given level of debt repayment B , a good firm retains the maximum level of earnings $d^G = (1-s)(\pi K - B)$ and $S^G = s(\pi K - B)$ and borrows as much as possible whereas respecting the incentive compatibility constraint $I^G = \frac{\gamma^B - r^G}{\pi - r^G} S^G$.*
2. *For a given capital stock K and a given level of debt repayment B , a bad firm does not retain any earning $d^B = \pi K - B$ and $S^B = 0$ and does not invest nor borrow $I^B = 0$*
3. *Interest rate charged on bad firms are $r^B = \frac{1+r-\lambda}{\lambda}$ and the rate charged on good firms is $r^G = \frac{(r+\kappa)(1+r-\lambda)}{(1-\kappa)(1+r-\lambda)+\kappa(r+\kappa)}$*

Proof The proof of propositions 3.2 shows that the first condition of the equilibrium, the participation constraint of lenders, is verified. The proof of proposition 3.4 shows that the policy rule of bad firms is optimal, hence the second condition is also verified. Proof of proposition 3.5 shows that the good firms' signal is also the optimal policy rule, hence the third condition holds. Proof of proposition 3.6 shows that the fourth condition, the incentive compatibility constraint for good firms, also holds. Proposition 3.7 characterizes the incentive compatibility constraint for bad firms hence the fifth condition is verified. □

Demonstration of conjectures

In addition, I verify that value functions for bad and good firms exist.

Proposition 3.10. *Value functions are well defined for bad firms under separating equilibrium. $V^{B,B}$ and $V^{B,G}$ exist.*

Proof Under separating equilibrium, bad firms do not invest if they reveal their type. The growth rate of their capital stock is equal to zero. Their profits are stationary and the discounted sum of their profits converges. Thus, $V^{B,B}$ exists. $V^{B,G}$ is given by $V^{B,B}$ and exists too. \square

To verify that value functions are well defined for good firms, I have to compute the growth rate of their profits

Proposition 3.11. *At the separating equilibrium, profits grow at the rate $s\gamma^B$.*

Proof Under the separating equilibrium, at each period, $S^G = s(\pi K - B)$ and $I^G = \psi S^G$, where ψ is the leverage of good firms. When computing $\pi K' - B'$, I obtain

$$\begin{aligned}\pi K' - B' &= \pi(K + \psi s(\pi K - B)) - (B + r^G(\psi - 1)s(\pi K - B)) \\ &= (\pi K - B)(1 + \psi s\pi - r^G(\psi - 1)s)\end{aligned}$$

The growth rate is $s(\psi\pi - r^G(\psi - 1))$. If the incentive compatibility constraint binds, $\psi = \frac{\gamma^B - r^G}{\pi - r^G} S^G$ and the growth rate of profit is $s\gamma^B$ \square

I deduce an existence condition for $V^{G,G}$.

Proposition 3.12. *Value functions $V^{G,G}$ and $V^{G,B}$ are well defined if*

$$s\gamma^B < \gamma + \kappa$$

Proof If the condition is verified, The growth rate of good firms' profits is smaller than the discount rate plus the rate of transition to bad firms. The value function converges. $V^{G,B}$ is given by $V^{G,G}$ and $V^{G,G}$ and thus is also defined \square

The difference between γ and γ^B is not large, thus the condition is automatically fulfilled for most values of s . Note however, that the condition is not fulfilled in the extreme case where $s = 1$ (firms retain all earnings).

I am also able to verify that always sending the correct signal is for bad firms is superior

to always sending the incorrect signal.

Proposition 3.13. *Under the separating equilibrium defined in proposition 3.9, the value of always sending the correct signal is superior to the value of always sending the incorrect signal*

$$V^{B,B} > U^{B,G}$$

As a consequence the maximal value of sending the incorrect signal in the current period is $V^{B,G}$

$$V^{B,G} > U^{B,G}$$

Proof A bad firm always sending the correct signal neither invests nor borrows, hence the value $V^{B,B}$ is

$$V^{B,B}(K, B) = (\pi K - B) \frac{1 + \gamma}{1 + \gamma - \lambda}$$

Under the separating equilibrium, at each period, $S^G = s(\pi K - B)$ and $I^G = \psi S^G$ where ψ is the leverage of good firms. Profits of a bad firm always sending the incorrect signal grow at a rate $s(\psi\pi - r^G(\psi - 1))$. Using standard results on geometric sum

$$U^{B,G}(K, B) = (1 - s)(\pi K - B) \frac{1 + \gamma}{1 + \gamma - \lambda(1 + s(\psi\pi - r^G(\psi - 1)))}$$

I demonstrate in appendix C.1.4 that

$$\begin{aligned} (\pi K - B) \frac{1 + \gamma}{1 + \gamma - \lambda} &\geq (1 - s)(\pi K - B) \frac{1 + \gamma}{1 + \gamma - \lambda(1 + s(\psi\pi - r^G(\psi - 1)))} \\ \Leftrightarrow \psi &\leq \frac{\gamma^B - r^G}{\pi - r^G} S^G \end{aligned}$$

Because $V^{B,B} \geq U^{B,G}$, it follows immediately that $\pi K - B - S^G + \frac{\lambda}{1 + \gamma} V^{B,B} \geq \pi K - B - S^G + \frac{\lambda}{1 + \gamma} U^{B,G}$, hence, $V^{B,G} \geq U^{B,G}$. \square

3.2.4 The scope of the equilibrium

The separating equilibrium holds under quite restrictive assumptions. In this section, I make a quantitative exploration to determine whether these restrictions are plausible. Hypothesis (2) and (3) provides an upper and a lower bound for the rate of profit π . π should be between values of r^G and r^B and between values of γ^B and γ^G . An issue is that empirical counterparts of r^G and r^B are not easy to identify. There is a wide range of interest rate for corporate bonds and not a clear separation between "bad" and "good firms". A better way to look at the problem is to compute values of the safe interest rate r that allow π to be between r^B and r^G .

To perform this quantitative exercise, I use the calibration displayed in table 3.1. The return on capital π is set at 0.08, targeting the average value of the ratio between the net operating surplus of corporate firms and their nonfinancial assets between 1953 and 2018 in the United States. The outside investment yield γ is set at 0.05 and aims at replicating the average return on equity. Transition probabilities from good to bad firms and from bad firms to exit are set at 0.02 and 0.05. The firm saving rate is set at 0.5 targeting a ratio of net corporate investment over corporate value added around 0.06. All parameters are calibrated in annual data.

A simple computation shows that, for values of π , λ and κ displayed in table 3.1, the safe real interest rate and the discount rate should be above 2.6 percent but below 7.3 percent. The lower bound allows for π to be lower than r^B and γ^B whereas the upper bound makes π larger than r^G and γ^G . The scope of the equilibrium is relatively large. A safe real interest rate and a discount rate of firms' shareholders below 7.3 percent seems plausible. The lower bound raises more issues. 2.6 percent is a plausible value for the private sector discount rate but seems high for the safe real rate.

However, this lower bound is very sensitive to the value of λ . For an annual exit probability of seven percent instead of five percent, the safe real rate should only be

above 0.5 percent. Such larger exit probabilities are not incompatible with observed data on corporate bonds. For example, for a safe real rate of 1.5 percent, the interest rate charged on bad firms is around 9 percent in the model. The nominal yield for CCC corporate bonds is around 10 percent in good times and in low inflation period for much shorter maturity. For a eight percent annual exit probability, the lower bound is minus 0.5 percent.

Thus, whereas hypothesis (2) and (3) are restrictive, they do not seem unrealistic.

Variable	parameter value	Variable definition
π	0.08	Average Return on capital
s	0.5	Saving rate of firms
γ	0.05	Return of outside investment
κ	0.02	Transition probability from good to bad
λ	0.95	survival rate

TABLE 3.1 – Calibration for the Investment Model

3.3 The macroeconomic model

From the viewpoint of the economist, the investment equation (3.12) of the previous section has interesting features.

First, this provides a simple theoretical justification for a linear relation between cash flows and investment. Such relation is empirically plausible. The sensitivity of investment to cash flows is a regular feature of empirical studies (unlike Tobin's Q or user cost of capital). The formula for investment is tractable which is quite surprising for an adverse selection problem in infinite horizon. Then, it is a microfounded relation. The elasticity of investment to cash flows is governed by structural parameters and in particular by the real interest rate.

I embed the framework of the previous section in a dynamic macroeconomic model.

This model is composed of three blocks: Prices equations, dynamic equations for good firms and dynamic equations for bad firms. The model developed in this section remains a partial equilibrium model. Real interest rate is still assumed exogenous but is time varying. I get a system of recursive equations. As I deal with transitional dynamics, I add the subscript t to all variables.

3.3.1 Extension of the equilibrium for time varying interest and discount rate

The equilibrium concept developed in the previous section can be extended to environment with time varying borrowing cost and discount rate.

First, I define interest rate for good and bad firms for a given sequence of real interest rate $(r_t)_{t \in \mathcal{N}}$

$$r_t^B \lambda Q_{t+1}^B = 1 + r_t \quad (3.18a)$$

$$r_t^G [(1 - \kappa)Q_{t+1}^G + \kappa Q_{t+1}^B] = 1 + r_t \quad (3.18b)$$

Where

$$Q_t^B = 1 + \frac{\lambda}{1 + r} Q_{t+1}^B \quad (3.19a)$$

$$Q_t^G = 1 + \frac{1}{1 + r} [(1 - \kappa)Q_t^G + \kappa Q_{t+1}^B] \quad (3.19b)$$

Similarly, I define γ_t^B and γ_t^G for a given sequence of discount rate $(\gamma_t)_{t \in \mathbb{N}}$

$$\gamma_t^B \lambda \Gamma_{t+1}^B = 1 + \gamma_t \quad (3.20a)$$

$$\gamma_t^G [(1 - \kappa) \Gamma_{t+1}^G + \kappa \Gamma_{t+1}^B] = 1 + r_t \quad (3.20b)$$

Where

$$\Gamma_t^B = 1 + \frac{\lambda}{1+r} \Gamma_{t+1}^B \quad (3.21a)$$

$$\Gamma_t^G = 1 + \frac{1}{1+r} [(1 - \kappa) \Gamma_t^G + \kappa \Gamma_{t+1}^B] \quad (3.21b)$$

Γ_t^B is the counterpart of Q_t^B for a risk neutral lender.

Having these definitions, I can reformulate hypothesis (2) and (3).

Hypothesis 3.4. For all $t \in \mathbb{N}$

$$r_t^B > \pi$$

and

$$\gamma_t^B > \pi$$

Hypothesis 3.5. For all $t \in \mathbb{N}$

$$\pi > r_t^G$$

and

$$\pi > \gamma_t^G$$

It is easy to see that if hypothesis 4 is verified, bad firms have no incentive to finance an investment through borrowing or through equity. Hypothesis 5 implies that good firms want to invest and to borrow as much as possible. Because of hypothesis 4, they also have no incentives to pay the interest rate r^B and thus to deviate.

Strikingly, the incentive compatibility constraint remains the same

Proposition 3.14. *Bad firms do not send the same signal as good firms if*

$$I_t^G \leq \frac{\gamma_t^B - r_t^G}{\pi - r_t^G} S_t^G = \psi_t S_t^G$$

Proof The proof actually follows the same steps as in the previous section. Detailed computations can be found in appendix [C.2.1](#) □

Thus, the leverage equation is

$$\psi_t = \frac{\gamma_t^B - r_t^G}{\pi - r_t^G} \quad (3.22)$$

3.3.2 Dynamic equations for good and bad firms

Equations (3.18a), (3.18b), (3.19a), (3.19b) and (3.22) form the first block of the model. Those are price equations. They only depend on exogenous variables, including real interest rate and are not impacted by capital stock, debt levels or investment.

Dynamic equations for good firms are

$$K_{t+1}^G = (1 - \kappa)(K_t^G + I_t^G) \quad (3.23a)$$

$$B_{t+1}^G = (1 - \kappa)(B_t^G + r_t^G(I_t^G - S_t^G)) \quad (3.23b)$$

$$\pi K_t^G - B_t^G = S_t^G + d_t^G \quad (3.23c)$$

$$d_t^G = (1 - s)(\pi K_t^G - B_t^G) \quad (3.23d)$$

$$I_t^G = \psi_t S_t^G \quad (3.23e)$$

Equation (3.23a) gives the capital accumulation of good firms at the aggregate level. Aggregate capital stock of good firms in period $t + 1$ is the capital stock at period t

plus investment at period t , times the "survival" rate of good firms. Equation (3.23b) gives the debt accumulation equation for good firms, debt being measured by interests repayments. Equation (3.23c) is an accounting equation dividing firm income between internal savings and dividends. All variables are in period t , so the survival rate is not relevant here. The fourth equation gives aggregate dividend as a linear function of firm income. The last equation is the investment equation.

Similar equations can be given for bad firms

$$K_{t+1}^B = \kappa K_t^G + \lambda K_t^B \quad (3.24a)$$

$$B_{t+1}^B = \lambda B_t^B + \kappa [B_t^G + r_t^G (I_t^G - S_t^G)] \quad (3.24b)$$

$$d_t^B = \pi_t K_t^B - b_t^B \quad (3.24c)$$

Bad firms do not invest but a fraction κ of good firms becomes bad at period t and "feed" capital stock and debt of the bad firm sector. All the income of bad firms is distributed as dividends.

Proposition 3.15. *The system of price equations (3.18a)-(3.22) can be solved independently from the system of recursive equations for good firms (3.23a)-(3.23e) and from the system of recursive equations for bad firms (3.24a)-(3.24c)*

Indeed, none of the eight variables of the two systems appear in the five equations of the system (8).

Proposition 3.16. *The system of recursive equations for good firms (3.23a)-(3.23e) can be solved independently from the system of recursive equation for bad firms (3.24a)-(3.24c)*

The justification of this proposition is also straightforward. Bad firms always remain bad and do not invest. Thus, they do not affect interest rates, capital stock, and have no feedback effect on good firms.

3.3.3 Investment Dynamics

I now show that the system (1.14) can be transformed into one recursive equation for debt from which I infer the dynamic of investment and growth.

The system has to be stationary. Indeed, the AK production function allows for endogenous growth. The solution is to divide all variables by K_t^G and to replace K_{t+1}^G by $1 + g_{t+1}$ where g is the growth rate. Stationary variables are denoted in lowercase. I drop the superscript G to make notations easier. For example, $i_t \equiv \frac{I_t^G}{K_t^G}$.

The five stationary equations can be simplified in a system of three equations whose variables are debt b , investment i and growth

$$(1 + g_{t+1})b_{t+1} = (1 - \kappa)(b_t + r_t^G(\psi_t - 1)s(\pi - b_t)) \quad (3.25a)$$

$$1 + g_{t+1} = (1 - \kappa)(1 + i_t) \quad (3.25b)$$

$$i_t = \psi_t s(\pi - b_t) \quad (3.25c)$$

The first two equations are the two recursive equations for capital stock and debt repayments. They are stationary versions of equations (3.23a) and (3.23b). The last equation is the investment equation and has been given by equation (3.23e) combined to (3.23c) and (3.23d). (See appendix C.2 for details).

Replacing (3.25c) in (3.25b) and (3.25b) in (3.25a) lead to a recursive equation for b

$$b_{t+1} = \frac{b_t + r_t^G(\psi_t - 1)s(\pi - b_t)}{1 + \psi_t s(\pi - b_t)} \quad (3.26)$$

The equation is nonlinear. To give some intuition, I plot a typical graph associated to the recursive equation in figure 3.2. The red dotted line is the 45 degrees line whereas the blue line represents equation (3.26). There are two steady states, the upper one

being unstable whereas the lower one is stable.

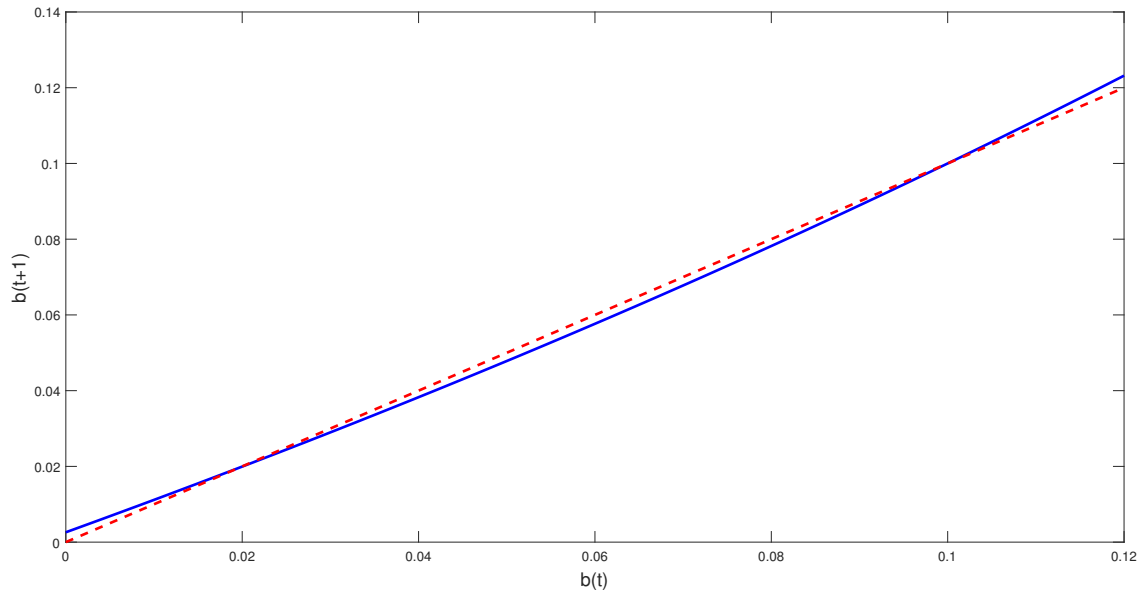


FIGURE 3.2 – Representation of the recursive equation

3.3.4 Steady state

I solve analytically for the steady state. The two stationary debt levels is given by the roots of the recursive equation (3.26)

Proposition 3.17. *Equation (3.26) have two roots b^* and b^{**}*

$$b^* = \pi$$

$$b^{**} = \frac{r^G(\psi - 1)}{\psi}$$

It is easy to see that one of the roots is equal to π . in that case, all the firms income is distributed as dividends, there is no growth and no investment, so interest repayments over capital stock remain at a fixed level. This equilibrium is quite degenerate because the economy will grow at a negative rate $-\kappa$. The root given by $r^G \frac{\psi-1}{\psi}$ is also intuitive.

When a firm invest ψ , it borrows $\psi - 1$ and thus will repay $r^G(\psi - 1)$ forever. The figure (1) suggests that one root is stable and one unstable. Fortunately, the stable root is the non degenerate one.

Proposition 3.18. *b^* is an unstable root and b^{**} is the stable root.*

Proof I denote f the function $f(b_t) = \frac{b_t + r^G(\psi - 1)s(\pi - b_t)}{1 + \psi s(\pi - b_t)}$

The derivative $f'(\pi) = 1 + s(\psi\pi - r^G(\psi - 1))$. Because by assumption 1 $\pi > r^G$ it is clear that the derivative of f is always superior to one when $b_t = b^*$. thus b^* is an unstable root.

The derivative $f'(r^G \frac{\psi - 1}{\psi}) = \frac{1 + s(\psi\pi - r^G(\psi - 1))}{[1 + s(\psi\pi - r^G(\psi - 1))]^2} < 1$. So b^{**} is a stable root. \square

I now focus on the stable root and compute the corresponding investment and growth rate.

Proposition 3.19. *For the stable steady state, growth and investment rate given by the system of equations (3.25a)-(3.25c) are*

$$i^{**} = s\psi\pi - sr^G(\psi - 1) \quad (3.27a)$$

$$g^{**} \simeq s\psi\pi - sr^G(\psi - 1) - \kappa \quad (3.27b)$$

An even simpler result may be obtained when I replace the leverage by its value with respect to r^G, r^B and π given by the incentive compatibility constraint

Proposition 3.20. *When replacing the leverage ψ by its value given by equation (3.22)*

$$i^{**} = s\gamma^B \quad (3.28a)$$

$$g^{**} \simeq s\gamma^B - \kappa \quad (3.28b)$$

3.4 Implications: Response to changes in borrowing cost

In this section, I explore some implications of the model. Beside inducing a relation between retained earnings and investment, the model has interesting consequences for the response of investment to a change in interest rate. The effects of the borrowing cost, determined by r and the effects of the opportunity cost for firms' shareholders, determined by γ are quite different and hence are analyzed separately. Interestingly, both a rise in borrowing cost through r and a rise in the shareholders' discount rate γ cause a *rise* in investment on impact.

3.4.1 The determinants of corporate investment

First, I analyze the response of investment to a change in the borrowing cost. Consider again the equation of corporate investment

$$i_t = \psi_t s(\pi - b_t)$$

Investment is sensitive to the leverage ψ and the debt level b_t . Both react to a change in the interest rate r but not necessarily in the same direction and with a different timing. Only the leverage reacts to a change in the discount rate γ .

A fall in r lowers r^G , pushing downward b_{t+1} . The income of good firm is unaffected in period t but increases in period $t + 1$ allowing an increase in retained earnings and in investment. This income distribution channel is very intuitive and appears in every model with financial friction (see Bernanke Gertler 1989 and Kiyotaki and Moore 1995). However, interest repayments b_t are a state variable. A fall in r does not

affect interest repayments on impact, hence investment remains initially unchanged.

What is specific to the model is the response of the leverage ψ_t . It affects investment directly and indirectly by altering the debt dynamic.

3.4.2 Response of leverage to interest rate

The leverage is given by

$$\psi = \frac{\gamma^B - r^G}{\pi - r^G} \quad (3.29)$$

It is not independent from the borrowing cost, because of the two terms r^G . A fall in r^G increases both numerator and denominator and thus is ambiguous.

I compute the derivative of the leverage with respect to r

$$\frac{\partial \psi}{\partial r} = \frac{-\frac{\partial r^G}{\partial r}(\pi - r^G) + (\gamma^B - r^G)\frac{\partial r^G}{\partial r}}{(\pi - r^G)^2}$$

Rearranging, I obtain

$$\frac{\partial \psi}{\partial r} = \frac{\partial r^G}{\partial r} \frac{\gamma^B - \pi}{(\pi - r^G)^2} > 0$$

r^G is an increasing function of r and its derivative is positive. Moreover, the hypothesis (2) implies that $\gamma^B > \pi$. Thus, $\frac{\partial \psi}{\partial r} > 0$. This result implies that good firms will tight their investment policy for a given level of cash flows when real interest rate falls. The intuition behind this result is the following. When the borrowing cost is lower, the profit generated by a debt financed investment increases for bad firms. Sending the signal of good firms and paying the low interest rate becomes more attractive for them. In response, good firms have to tight their investment policy in order to continue to deter bad firms.

How large is this effect ? To investigate the issue, I compute the elasticity of leverage with respect to real interest rate r

$$\frac{1}{\psi} \frac{\partial \psi}{\partial r} = \frac{\partial r^G}{\partial r} \frac{\gamma^B - \pi}{(\pi - r^G)(\gamma^B - r^G)}$$

Intuitively, a large return on firm specific capital π , close to the value of γ^B and far from the value of r^G , provides a low elasticity. But, it also implies a low leverage. Return of bad firms is high compared to their outside investment opportunity. Thus, good firms have to tight their investment policy to keep signaling to lenders they are good firms. This relation between low leverage and low elasticity is intuitive. The leverage cannot be below one, thus, closer it is to one, less elastic it is with respect to a change in borrowing cost.

3.4.3 Long run effects of changes in borrowing costs

The leverage effect and the debt repayment effect go in opposite direction. A fall in interest rate lowers investment through the first channel but increases it through the second. Thus, the total effect is indeterminate. However, a very powerful result is that the two effects cancel each other in the long run.

Proposition 3.21. *In the long run, corporate investment is not affected by the borrowing cost r*

Indeed, the steady state value of investment i^{**} is equal to $s\gamma^B$. Both parameters s and γ^B are invariant with respect to r .

3.4.4 Short run effects

Intuitively, the leverage channel dominates the debt repayment channel in the short run. Indeed, leverage adjusts immediately whereas debt repayments evolve very slowly because of the infinite maturity assumption. Investment should fall on impact after an interest rate cut and increases progressively thereafter.

To confirm this intuition, I explore the transition dynamics following a permanent change in interest rate in figure 3.3 and the effects of a temporary change in figure 3.4. I consider a permanent then temporary drop of the real rate from 3 percent to 1 percent. For the temporary shock, the fall lasts five periods. Then, r goes back to 3 percent. I follow the calibration displayed in table 3.1, except for the survival probability of bad firms which is set at 0.93.

Following the permanent shock, investment slightly falls and recovers progressively whereas debt adjusts slowly to its new steady state value. Following the temporary shock, leverage and corporate investment falls on impact. They recover after five periods. After these five periods, interests repayments are lower than their steady state value. As a consequence, investment has a peak and is slightly higher than its steady state value. Then, debt and investment converge to their long run values. For both the temporary and the permanent shock, the large shock on interest rate has nearly no effects on investment, debt and leverage.

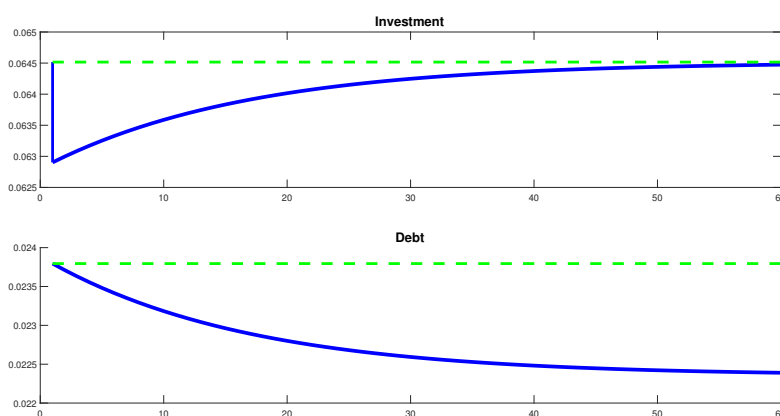


FIGURE 3.3 – Response to a permanent change in borrowing cost

3.4.5 Response to changes in opportunity cost

I also consider the effects of a change in the return of outside opportunities γ . I reproduce the same exercise as in the previous paragraph. γ is supposed to fall from 5

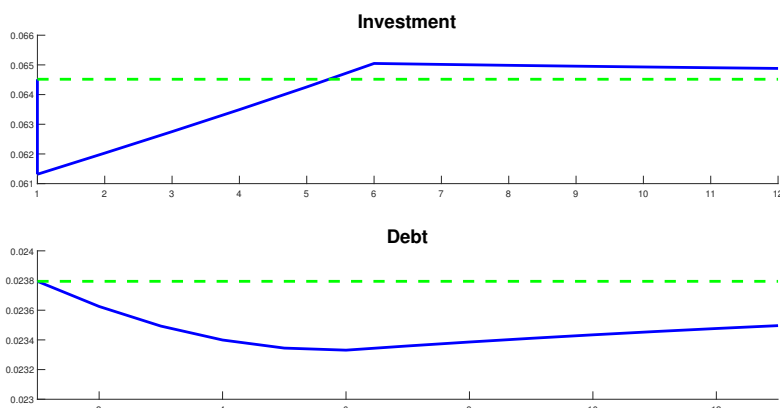


FIGURE 3.4 – Response to a temporary change in borrowing cost

percent to 3 percent. Figure 3.5 represents the response to a permanent shock whereas figure 3.6 shows the response to a temporary shock. The calibration is the same as in the previous section. The safe real rate is set at 3 percent.

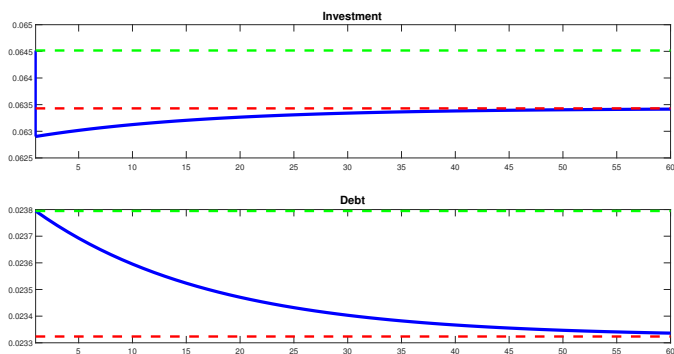


FIGURE 3.5 – Response to a permanent change in opportunity cost

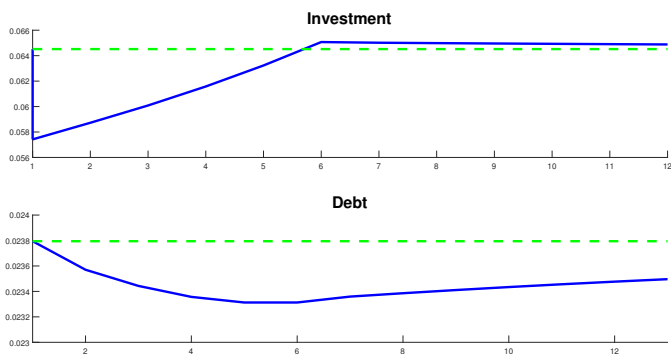


FIGURE 3.6 – Response to a temporary change in opportunity cost

A permanent fall in the opportunity cost of firms' shareholders generates a fall in steady state leverage, investment and debt. Interestingly, investment slightly overshoots on impact. Indeed, the lower borrowing cost leads to lower debt repayments, increasing firm earnings, leading to a small and progressive rise in investment after the initial drop. A temporary fall in opportunity cost also lowers investment and debt. The intuition for this response is similar to the intuition for borrowing costs. When the return of outside opportunity is lower, an investment in their own firm is more profitable for shareholders of bad firms. They have more incentives to send the false signal, hence good firms have to reduce their investment over saving ratio.

3.5 Conclusion

I have provided a tractable model of corporate investment with adverse selection. A first implication of the model is that aggregate investment depends linearly on retained earnings. This constraint directly relates investment and cash flows and is not derived from a relation between capital stock and debt. A second implication is that a lower borrowing cost or a lower opportunity cost may *reduce* investment on impact. The impact of real interest rate on corporate investment remains a subject of controversy. VAR evidence from Bernanke and Gertler (1995) shows that the response of corporate investment to monetary policy shocks is small and delayed compared to the response of residential investment. Estimations of investment with aggregate or firm level data give mixed results (see Sharpe and Suarez 2014 for a survey). However, a depressing effect on impact seems counterfactual. In my opinion, it does not mean that the model is irrelevant but it should be refined before being properly confronted to data. For example, the constraint may apply to a subset of firms whereas other firms are unconstrained. Another possibility is that a lower interest rate provide an incentive to retain more earnings, counterbalancing the fall in leverage.

Bibliography

- [1] **Albuquerque. R. and Hopenhayn. H.** 2004. "Optimal Lending Contracts and Firm Dynamics" *Review of economic studies*, 71, 285-315.
- [2] **Azariadis. C. and Smith. B.** 1998. "Financial Intermediation and Regime Switching in Business Cycles", *American Economic Review*, 88, 516-536.
- [3] **Bernanke. B., and Gertler. M.** 1989. "Agency Costs, Net Worth, and Business Fluctuations", *American Economic Review*, 79(1), 14-31.
- [4] **Bernanke. B., and Gertler. M.** 1995. "Inside the black box: The Credit Channel of Monetary Policy Transmission", *NBER working papers*, 5146.
- [5] **Bernanke. B., Gertler. M., and Gilchrist S.** 1999. "The Financial Accelerator in a Quantitative Business Cycle Framework", *in Handbook of Macroeconomics*, North-Holland, 1341-1393.
- [6] **Blanchard. O.** 1986. "Comments and Discussion on Investment, Output and The cost of Capital", *Brookings Paper on Economic Activity*.
- [7] **Carlstrom. C. and Fuerst. T.** 1997. "Agency Costs, Net Worth, and Business Fluctuations: A Computable General Equilibrium Analysis", *American Economic Review*, 87, 893-910.
- [8] **Clementi. G. and Hopenhayn. H.** 2006. "A theory of Financing Constraints and Firm Dynamics", *Quarterly Journal of Economics*, 121(1), 229-265.
- [9] **Cooley. T., Marimon. R. and Quadrini.V.** 2003. "Aggregate Consequence of Limited Contract Enforceability", *NBER workin paper 10132*.

-
- [10] **Chirinko. R.** 1993. "Business Fixed Investment Spending: Modelling Strategies, Empirical Results, and Policy Implications", *Journal of Economic Literature*, 31(4), 1875-1911.
- [11] **Christiano. L., Motto. R. and Rostagno. M.** 2014. "Risk Shocks", *American Economic Review*, 104, 27-65.
- [12] **Cummins. J., Hasset. K., and Hubbard. R.** 1994. "A Reconsideration of Investment Behavior using Tax Reforms as Natural Experiments", *Brookings Paper on Economic Activity*.
- [13] **Dressler. S. and Li. V.** 2009. "Inside Money, Credit, and Investment", *Journal of Economic Dynamics and Control*, 33, 970-984.
- [14] **Fazzari. S., Hubbard. G. and Petersen. B.** 1988. "Financing Constraints and Corporate Investment", *Brookings Papers on Economic Activity*, 141-195.
- [15] **Gertler. M. and Gilchrist. S.** 1994. "Monetary Policy, Business Cycles, and the Behavior of Small Manufacturing Firms", *The Quarterly Journal of Economics*, 109, 309-340.
- [16] **Gilchrist. S. and Zakrajsek. E.** 2012. "Credit Spread and Business Cycles Fluctuations", *American Economic Review*, 102, 1692-1720.
- [17] **Hart. O. and Moore. J.** 1994. "A Theory of Debt Based on the Inalienability of Human Capital", *The Quarterly Journal of Economics*, 109, 841-879.
- [18] **Holmstrom. B. and Tirole. J.** 2011. *Inside and Outside Liquidity*, MIT press.
- [19] **Iacoviello. M.** 2005. "House Prices, Borrowing Constraints, and Monetary Policy in the Business Cycle", *American Economic Review*, 95, 739-764.
- [20] **Kaplan. S. and Zingales. L.** 1997. "Do Investment-Cash Flow Sensitivities Provide Useful Measures of Financing Constraints?", *The Quarterly Journal of Economics*, 112, 169-215.

-
- [21] **Khan. A. and Thomas. J.** 2013. "Credit Shocks and Aggregate Fluctuations in an Economy with Production Heterogeneity", *Journal of Political Economy*, 121, 1055-1107.
- [22] **Kiyotaki. N. and Moore. J.** 1997. "Credit Cycles", *Journal of Political Economy*, 105, 211-248.
- [23] **Kiyotaki. N. and Moore. J.** 2008. "Liquidity, Business Cycles and Monetary Policy", *Mimeo*, Princeton University.
- [24] **Kocherlakota. N.** 2000. "Creating Business Cycles Through Credit Constraints", *Federal Reserve Bank of Minneapolis Quarterly Review*, 24, 2-10.
- [25] **Kothari. S., Lewellen. J., and Warner. J.** 2015. "The Behavior of Aggregate Corporate Investment", Working paper.
- [26] **Kurlat. P.** 2013. "Lemons Market and the Transmission of Aggregate Shocks", *American Economic Review*, 103(4), 1463-1489.
- [27] **Kurlat. P.** 2016. "Asset Markets with Heterogenous Information", *Econometrica*, 84, 33-85.
- [28] **Myers. S. and Maljuf. N.** 1984. "Corporate Financing and Investment Decisions when Firms have Information Investors do not have", *Journal of Financial Economics*, 13, 187-21.
- [29] **Saint Paul. G.** 2005. "Fiscal Policy and Economic Growth: The Role of Financial Intermediation", *Review of International Economics*, 13, 612-629.
- [30] **Shapiro. A.** 1986. "Investment, Output and The Cost of Capital", *Brookings Paper on Economic Activity*.
- [31] **Sharpe. S. and Suarez. G.** 2014. "The Insensitivity of Investment to Interest Rates: Evidence from a Survey of CFOs", *Finance and Economics Discussion*

Series Divisions of Research and Statistics and Monetary Affairs Federal Reserve Board, Washington, D.C..

- [32] **Stiglitz. J. and Weiss. A.** 1981. "Credit Rationning in Market with Imperfect Information", *American Economic Review*, 71, 393-410.
- [33] **Tirole. J.** 2005. *Theory of Corporate Finance*, Princeton University press.
- [34] **Townsend. R.** 1979. "Optimal contracts and competitive markets with costly state verification", *Journal of Economic Theory*, 21, 265-293.

Appendix A

Appendix of chapter 1

A.1 Proofs

A.1.1 Proof of proposition 1.1

π_f is the outcome of an univariate linear regression : $\pi_f = \frac{Cov(x_t, y_t)}{V(y_t)}$.

$$Cov(x_t, y_t) = cov(\alpha + \lambda\varphi_f + (\beta + \lambda\pi_f)y_t + u_t, y_t)$$

y_t is neither correlated with residual u_t nor with the constant, thus

$$Cov(x_t, y_t) = cov((\beta + \lambda\pi_f)y_t, y_t)$$

$$Cov(x_t, y_t) = (\beta + \lambda\pi_f)V(y_t)$$

π_f is the root of the equation

$$\pi_f = \beta + \lambda\pi_f$$

Thus, $\pi_f = \frac{\beta}{1-\lambda}$.

The constant of the linear regression φ_f is given by $E(x_t) - \frac{\beta}{1-\lambda}E(y_t)$. The process for y_t is given by equation (1.2) and implies that $E(y_t) = 0$.

The expected value of x_t is given by

$$E(x_t) = \alpha + \lambda\varphi_f + (\beta + \lambda\pi_f)E(y_t)$$

$$E(x_t) = \alpha + \lambda\varphi_f$$

Thus, I have $\varphi_f = \alpha + \lambda\varphi_f$.

φ_f is the root of this equation leading to $\varphi_f = \frac{\alpha}{1-\lambda}$ □

A.1.2 Proof of proposition 1.4

π_o is also the outcome of an univariate linear regression : $\pi_o = \frac{Cov(x_t, x_{t-1})}{V(x_{t-1})}$.

First, I compute the covariance

$$Cov(x_t, x_{t-1}) = cov(\alpha + \lambda\varphi_f + (\beta + \lambda\pi_f)y_t + u_t, \alpha + \lambda\varphi_f + (\beta + \lambda\pi_f)y_{t-1} + u_{t-1})$$

u and y are given by two independent processes. They are uncorrelated. Moreover, u is the white noise, thus $cov(u_t, u_{t-1}) = 0$. y is an autoregressive process whose persistence parameter is θ . It leads to

$$Cov(x_t, x_{t-1}) = \left(\frac{\beta}{1-\lambda}\right)^2 cov(y_t, y_{t-1})$$

$$Cov(x_t, x_{t-1}) = \left(\frac{\beta}{1-\lambda}\right)^2 \theta V(y_t)$$

Then, I compute the unconditional variance of x_t

$$V(x_{t-1}) = V(x_t) = \left(\frac{\beta}{1-\lambda}\right)^2 V(y_t) + \sigma_u^2$$

Moreover, using standard results on autoregressive process, I have $V(y_t) = \frac{\sigma_y^2}{1-\theta^2}$. Thus

$$\pi_o = \frac{\left(\frac{\beta}{1-\lambda}\right)^2 \theta \frac{\sigma_y^2}{1-\theta^2}}{\left(\frac{\beta}{1-\lambda}\right)^2 \frac{\sigma_y^2}{1-\theta^2} + \sigma_u^2}$$

On another hand

$$\begin{aligned} \theta - \frac{\sigma_u^2 \theta}{\left(\frac{\beta}{1-\lambda}\right)^2 \frac{\sigma_y^2}{1-\theta^2} + \sigma_u^2} &= \frac{\theta \left(\left(\frac{\beta}{1-\lambda}\right)^2 \frac{\sigma_y^2}{1-\theta^2} + \sigma_u^2 \right) - \sigma_u^2 \theta}{\left(\frac{\beta}{1-\lambda}\right)^2 \frac{\sigma_y^2}{1-\theta^2} + \sigma_u^2} \\ &= \frac{\left(\frac{\beta}{1-\lambda}\right)^2 \theta \frac{\sigma_y^2}{1-\theta^2}}{\left(\frac{\beta}{1-\lambda}\right)^2 \frac{\sigma_y^2}{1-\theta^2} + \sigma_u^2} \end{aligned}$$

QED

The constant of the linear regression φ_o is given by $E(x_t) - \pi_o E(x_{t-1})$. As $E(x_t) = \frac{\alpha}{1-\lambda}$, I have $\varphi_o = \frac{\alpha}{1-\lambda}(1 - \pi_o)$. \square

A.1.3 Proof of proposition 1.5

Before computing the average forecast error, I highlight an important intermediate result

$$\varphi_o + \pi_o x_{t-1} = \frac{\alpha}{1-\lambda}(1 - \pi_o) + \pi_o \frac{\alpha}{1-\lambda} + \pi_o \frac{\beta}{1-\lambda} y_{t-1} + \pi_o u_{t-1} = \frac{\alpha}{1-\lambda} + \pi_o \frac{\beta}{1-\lambda} y_{t-1} + \pi_o u_{t-1}$$

The previous result allows me to eliminate the constant. The average forecast error is

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=0}^T [(x_k - \varphi_o - \pi_o x_{k-1})^2] = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=0}^T \left[\frac{\beta}{1-\lambda} y_k + u_k - \pi_o \frac{\beta}{1-\lambda} y_{k-1} - \pi_o u_{k-1} \right]^2$$

All components of the error term have a zero expected value, thus

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=0}^T \left[\frac{\beta}{1-\lambda} y_k + u_k - \pi_o \frac{\beta}{1-\lambda} y_{k-1} - \pi_o u_{k-1} \right]^2 = V \left(\frac{\beta}{1-\lambda} y_t + u_t - \pi_o \frac{\beta}{1-\lambda} y_{t-1} - \pi_o u_{t-1} \right)$$

Variables y and u are uncorrelated, thus the variance is

$$V \left(\frac{\beta}{1-\lambda} y_t + u_t - \pi_o \frac{\beta}{1-\lambda} y_{t-1} - \pi_o u_{t-1} \right) = (1 + \pi_o^2) \sigma_u^2 + (1 + \pi_o^2) \left(\frac{\beta}{1-\lambda} \right)^2 V(y_t) - 2\pi_o \left(\frac{\beta}{1-\lambda} \right)^2 \text{cov}(y_t, y_{t-1})$$

Previous results for variance and covariance of y lead to

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=0}^T [(x_k - \varphi_o - \pi_o x_{k-1})^2] = (1 + \pi_o^2) \sigma_u^2 + (1 + \pi_o^2) \left(\frac{\beta}{1-\lambda} \right)^2 \frac{\sigma_y^2}{1 - \theta^2} - 2\pi_o \left(\frac{\beta}{1-\lambda} \right)^2 \theta \frac{\sigma_y^2}{1 - \theta^2}$$

□

A.1.4 Proof of proposition 1.6

The proof follows from proposition 1.2 and 1.5. Let me remind that $\theta < 1$. It follows

$(1 + \pi_o^2 - 2\pi_o\theta) > (1 + \pi_o^2 - 2\pi_o) = (1 - \pi_o)^2 > 0$. Then

$$(1 + \pi_o^2 - 2\pi_o\theta) \left(\frac{\beta}{1-\lambda} \right)^2 \frac{\sigma_y^2}{1 - \theta^2} > 0$$

As a consequence

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=0}^T [(x_k - \varphi_o - \pi_o x_{k-1})^2] > (1 + \pi_o^2) \sigma_u^2 > \sigma_u^2$$

□

A.1.5 Proof for proposition 1.9

In this proof, I show that the right hand side of the equation (1.10) is the asymptotic least square estimator.

π_o is the outcome of an univariate linear regression : $\pi_o = \frac{Cov(x_t, x_{t-1})}{V(x_{t-1})}$.

Before computing the covariance and the variance, I need to compute $cov(y_t, x_{t-1})$. To make computation easier , I replace π_o by π . I also denote $V(y_t)$ by $V(y)$

$$\begin{aligned} cov(y_t, x_{t-1}) &= \theta cov(y_{t-1}, x_{t-1}) = \theta cov(y_{t-1}, \beta y_{t-1} + \lambda \pi x_{t-2} + u_{t-1}) \\ &= \beta \theta V(y_t) + \lambda \pi \theta cov(y_{t-1}, x_{t-2}) \end{aligned}$$

I have eliminated the constant from x_{t-1} . Moreover, residuals u_t are governed by a white noise process and are uncorrelated with both y_{t-1} and x_{t-2}

Now, using $cov(y_t, x_{t-1}) = cov(y_{t-1}, x_{t-2})$, I find

$$cov(y_t, x_{t-1}) = \frac{\beta \theta}{1 - \lambda \pi \theta} V(y_t)$$

I need another intermediate result. I compute the variance of x_t

$$\begin{aligned} V(x_t) &= V(\beta y_t + \lambda \pi x_{t-1} + u_t) \\ V(x_t) &= (\lambda \pi)^2 V(x_{t-1}) + \beta^2 V(y_t) + \sigma_u^2 + 2\beta \lambda \pi cov(y_t, x_{t-1}) \end{aligned}$$

Using $V(x_t) = V(x_{t-1})$, I get

$$[1 - (\lambda \pi)^2] V(x_t) = \beta^2 V(y_t) + \sigma_u^2 + 2\beta \lambda \pi \frac{\beta \theta}{1 - \lambda \pi \theta} V(y_t)$$

After eliminating the denominator term

$$(1 - \lambda\pi\theta)[1 - (\lambda\pi)^2]V(x_t) = (1 - \lambda\pi\theta)\beta^2V(y_t) + (1 - \lambda\pi\theta)\sigma_u^2 + 2\beta^2\lambda\pi\theta V(y_t)$$

$$(1 - \lambda\pi\theta)[1 - (\lambda\pi)^2]V(x_t) = (1 + \lambda\pi\theta)\beta^2V(y_t) + (1 - \lambda\pi\theta)\sigma_u^2$$

I now compute the covariance $Cov(x_t, x_{t-1})$

$$Cov(x_t, x_{t-1}) = Cov(\beta y_t + \lambda\pi x_{t-1} + u_t, x_{t-1})$$

$$Cov(x_t, x_{t-1}) = \lambda\pi V(x_{t-1}) + \beta Cov(y_t, x_{t-1})$$

$$Cov(x_t, x_{t-1}) = \lambda\pi V(x_{t-1}) + \frac{\beta^2\theta}{1 - \lambda\pi\theta}V(y_t)$$

$$\Rightarrow (1 - \lambda\pi\theta)Cov(x_t, x_{t-1}) = (1 - \lambda\pi\theta)\lambda\pi V(x_{t-1}) + \beta^2\theta V(y_t)$$

$$\Rightarrow (1 - \lambda\pi\theta)[1 - (\lambda\pi)^2]Cov(x_t, x_{t-1}) = (1 - \lambda\pi\theta)[1 - (\lambda\pi)^2]\lambda\pi V(x_{t-1}) + \beta^2\theta[1 - (\lambda\pi)^2]V(y_t)$$

$(1 - \lambda\pi\theta)[1 - (\lambda\pi)^2]V(x_{t-1})$ is equal to $(1 - \lambda\pi\theta)[1 - (\lambda\pi)^2]V(x_t)$ whose expression has been found above

$$(1 - \lambda\pi\theta)[1 - (\lambda\pi)^2]Cov(x_t, x_{t-1}) = \lambda\pi \left((1 + \lambda\pi\theta)\beta^2V(y_t) + (1 - \lambda\pi\theta)\sigma_u^2 \right) + \beta^2\theta[1 - (\lambda\pi)^2]V(y_t)$$

I now develop all terms in $V(y)$ and simplify

$$(1 - \lambda\pi\theta)[1 - (\lambda\pi)^2]Cov(x_t, x_{t-1}) = \lambda\pi(1 - \lambda\pi\theta)\sigma_u^2 + V(y_t) [\lambda\pi\beta^2 + (\lambda\pi)^2\theta\beta^2 + \beta^2\theta - \beta^2\theta(\lambda\pi)^2]$$

$$(1 - \lambda\pi\theta)[1 - (\lambda\pi)^2]Cov(x_t, x_{t-1}) = \lambda\pi(1 - \lambda\pi\theta)\sigma_u^2 + \beta^2V(y_t)(\lambda\pi + \theta)$$

The estimator may be now computed

$$\frac{Cov(x_t, x_{t-1})}{V(x_{t-1})} = \frac{\beta^2(\lambda\pi + \theta)V(y_t) + \lambda\pi(1 - \lambda\pi\theta)\sigma_u^2}{\beta^2(1 + \lambda\pi\theta)V(y_t) + (1 - \lambda\pi\theta)\sigma_u^2}$$

□

A.1.6 Proof for proposition 1.11

I transform equation (1.10) into a quadratic equation.

$$\begin{aligned} \pi &= \frac{\beta^2 V(y_t)(\theta + \lambda\pi) + \lambda\pi(1 - \lambda\pi\theta)\sigma_u^2}{\beta^2 V(y_t)(1 + \lambda\pi\theta) + (1 - \lambda\pi\theta)\sigma_u^2} \\ \Rightarrow \pi (\beta^2 V(y_t)(1 + \lambda\pi\theta) + (1 - \lambda\pi\theta)\sigma_u^2) &= \beta^2 V(y_t)(\theta + \lambda\pi) + \lambda\pi(1 - \lambda\pi\theta)\sigma_u^2 \\ \Rightarrow \lambda\theta[\beta^2 V(y_t) - \sigma_u^2(1 - \lambda)]\pi^2 + (\beta^2 V(y_t) + \sigma_u^2)(1 - \lambda)\pi - \beta^2\theta V(y_t) &= 0 \end{aligned}$$

□

A.1.7 Proof for proposition 1.14

I defined the function $f : \pi \rightarrow A\pi^2 + B\pi + C$ where A , B and C are defined in proposition 1.11

$$f(\pi) = \lambda\theta[\beta^2 V(y_t) - \sigma_u^2(1 - \lambda)]\pi^2 + (\beta^2 V(y_t) + \sigma_u^2)(1 - \lambda)\pi - \beta^2\theta V(y_t)$$

I compute $f(1)$

$$\begin{aligned} f(1) &= \lambda\theta[\beta^2V(y_t) - \sigma_u^2(1 - \lambda)] + (\beta^2V(y_t) + \sigma_u^2)(1 - \lambda) - \beta^2\theta V(y_t) \\ &= \beta^2V(y_t)[\lambda\theta + (1 - \lambda) - \theta] + \sigma_u^2(1 - \lambda)(1 - \lambda\theta) \\ &= \beta^2V(y_t)(1 - \lambda)(1 - \theta) + \sigma_u^2(1 - \lambda)(1 - \lambda\theta) \end{aligned}$$

Hypothesis 1.1 implies $f(1) > 0$

Now I compute the derivative of f

$$f'(\pi) = 2\lambda\theta[\beta^2V(y_t) - \sigma_u^2(1 - \lambda)]\pi + (\beta^2V(y_t) + \sigma_u^2)(1 - \lambda)$$

For $\pi > 0$ and a fortiori for $\pi \geq 1$, $f'(\pi) > 0$

Thus, $\forall \pi > 1, f(\pi) > 0$. Thus, roots of the equation $f(\pi) = 0$ are all inferior to 1. \square

A.1.8 Proof of proposition 1.16

I use the intermediate result found in proposition 1.5 to simplify the expression of the average error

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=0}^T [x_k - \varphi_o - \pi_o x_{k-1}]^2 = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=0}^T [\beta y_k + \lambda \pi_o (x_{k-1} - \frac{\alpha}{1 - \lambda}) + u_k - \pi_o (x_{k-1} - \frac{\alpha}{1 - \lambda})]^2$$

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=0}^T [x_k - \varphi_o - \pi_o x_{k-1}]^2 = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=0}^T [\beta y_k + u_k - (1 - \lambda) \pi_o (x_{k-1} - \frac{\alpha}{1 - \lambda})]^2$$

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=0}^T [x_k - \varphi_o - \pi_o x_{k-1}]^2 = V(\beta y_t + u_t - (1 - \lambda) \pi_o x_{t-1})$$

I now use standard results on variance to find the formula highlighted in the proposition

$$V(\beta y_t + u_t - (1 - \lambda)\pi_o x_{t-1}) = \beta^2 V(y_t) + \sigma_u^2 + (1 - \lambda)^2 \pi_o^2 V(x_t) - 2\beta(1 - \lambda)\pi_o \theta \text{Cov}(y_t, x_t)$$

□

A.1.9 Proof of proposition 1.18

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=0}^T [x_k - \varphi_f - \pi_f y_k]^2 = \lim_{T \rightarrow \infty} \sum_{k=0}^T \frac{1}{T} [\beta y_k + \lambda \pi_o (x_{k-1} - \frac{\alpha}{1 - \lambda}) + u_k - \pi_f y_k]^2$$

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{k=0}^T [x_k - \varphi_f - \pi_f y_k]^2 = V((\beta - \pi_f)y_t + \lambda \pi_o x_{t-1} + u_t)$$

And by standard properties of variance and covariance

$$V((\beta - \pi_f)y_t + \lambda \pi_o x_{t-1} + u_t) = (\pi_f - \beta)^2 V(y_t) + (\lambda \pi_o)^2 V(x_t) + \sigma_u^2 - 2\lambda \pi_o (\pi_f - \beta) \theta \text{cov}(x_t, y_t)$$

□

Appendix B

Appendix of chapter 2

B.1 Proofs

B.1.1 Proof of proposition 2.1

Consider the lagrangean

$$U(C_t, L_t, A_{t+1}) + \Lambda_t (W_t + A_t - \Pi_t - W_t L_t - C_t - Q_t A_{t+1})$$

First order condition are

$$U_C(C_t, L_t, A_{t+1}) - \Lambda_t = 0 \tag{B.1a}$$

$$U_L(C_t, L_t, A_{t+1}) - W_t \Lambda_t = 0 \tag{B.1b}$$

$$U_A(C_t, L_t, A_{t+1}) - Q_t \Lambda_t = 0 \tag{B.1c}$$

B.1.2 Proof of proposition 2.2

A first order Taylor expansion around the steady state of $U_C(C_t, L_t, A_{t+1})$ gives

$$U_c(C_t, L_t, A_{t+1}) = U_c + U_{cc}(C_t - C) + U_{ca}(A_{t+1} - A) + U_{cl}(L_t - L) + \epsilon$$

$$\frac{U_c(C_t, L_t, A_{t+1}) - U_c}{U_C} = \frac{U_{cc}C}{U_c} \frac{(C_t - C)}{C} + \frac{U_{ca}A}{U_c} \frac{(A_{t+1} - A)}{A} + \frac{U_{cl}L}{U_c} \frac{(L_t - L)}{L}$$

$$u_{c,t} = \frac{U_{CC}C}{U_C} c_t + \frac{U_{CAA}}{U_C} a_{t+1} + \frac{U_{CL}L}{U_C} l_t$$

Where lowercase letters denotes log linear values. For example $c_t = \frac{C_t - C}{C}$ and Letters without t index denotes steady state value.

The linearization of the two other equations use the same method.

B.1.3 Proof of proposition 2.3

The proposition rewrites the system of linear equations (2.8) with notations introduced in the paper and eliminate the Lagrange multiplier λ_t .

B.1.4 Proof of proposition 2.4

The proposition rewrites the system of linear equations from proposition 2.2 with assumption 2.3 and 3.2

B.1.5 Proof of corollary 2

The proposition rewrites the system of linear equations from proposition 2.4 with notations introduced in the paper and eliminate the Lagrange multiplier λ_t .

B.1.6 Proof of proposition 2.5

The hessian matrix of the utility function

$$\begin{pmatrix} U_{ll} & U_{cl} & U_{al} \\ U_{cl} & U_{cc} & U_{ca} \\ U_{al} & U_{ca} & U_{aa} \end{pmatrix}$$

Using assumption 2.3, it becomes

$$\begin{pmatrix} U_{ll} & 0 & 0 \\ 0 & U_{cc} & U_{ca} \\ 0 & U_{ca} & U_{aa} \end{pmatrix}$$

The utility function is concave if the hessian matrix is semidefinite negative, thus if all eigenvalues of the hessian are negative. The eigenvalues are given by U_{ll} and eigenvalues of the matrix

$$\begin{pmatrix} U_{cc} & U_{ca} \\ U_{ca} & U_{aa} \end{pmatrix}$$

Eigenvalues of this matrix are negative if the trace is negative and the determinant positive. Sufficient conditions are

$$U_{cc} \leq 0$$

$$U_{aa} \leq 0$$

$$U_{aa}U_{cc} - U_{ca}^2 \geq 0$$

B.1.7 Proof of proposition 2.6

The recursive formulation for the optimization problem is

$$V(A_t) = \text{Max}_{A_{t+1}, L_t, C_t} U(C_t, L_t, A_{t+1}) + \beta E_t V(A_{t+1})$$

The lagrangean is

$$U(C_t, L_t, A_{t+1}) + E_t V(A_{t+1}) + \Lambda_t (W_t + A_t - \Pi_t - W_t L_t - C_t - Q_t A_{t+1})$$

First order condition are

$$\begin{aligned} U_C(C_t, L_t, A_{t+1}) - \Lambda_t &= 0 \\ U_L(C_t, L_t, A_{t+1}) - W_t \Lambda_t &= 0 \\ U_A(C_t, L_t, A_{t+1}) + \beta E_t \frac{\partial V}{\partial A_{t+1}} - Q_t \Lambda_t &= 0 \end{aligned}$$

Using envelope theorem, the derivative of the value function is

$$\frac{\partial V}{\partial A_{t+1}} = \Lambda_{t+1} = U_C(C_{t+1}, L_{t+1}, A_{t+2})$$

Leading to equations highlighted in the proposition.

B.1.8 Proof of proposition 2.7

Consider the modified euler condition of the optimization problem

$$U_A(C_t, L_t, A_{t+1}) + \beta U_C(C_{t+1}, L_{t+1}, A_{t+2}) - Q_t U_C(C_t, L_t, A_{t+1}) = 0$$

Consider now this equation at the steady state

$$U_A + \beta U_C - Q U_C = 0 \Rightarrow \beta = Q - \frac{U_A}{U_C}$$

B.1.9 Proof of proposition 2.8

The linearization method is the same as in the proof of the proposition 2.2. The log linear equation for leisure is the same.

I now compute the log linear Euler equation. I denote the log linear approximation of x as $\mathcal{M}(x)$, except for variables A , C , Q and L whose log linear approximations are denoted in lowercase.

$$\mathcal{M}(U_A(C_t, L_t, A_{t+1}) + \beta E_t U_C(C_{t+1}, L_{t+1}, A_{t+2})) = \mathcal{M}(Q_t U_C(C_t, L_t, A_{t+1}))$$

By standard properties of log linear first order approximations, I have

$$\frac{U_A}{U_A + \beta U_C} \mathcal{M}(U_A(C_t, L_t, A_t)) + \frac{\beta U_C}{U_A + \beta U_C} \mathcal{M}(E_t U_C(C_{t+1}, L_{t+1}, A_{t+2})) = q_t + \mathcal{M}(U_C(C_t, L_t, A_{t+1})) \quad (\text{B.2})$$

I use the steady state condition $U_A + \beta U_C = Q U_C$ to obtain

$$\left(1 - \frac{\beta}{Q}\right) \mathcal{M}(U_A(C_t, L_t, A_t)) + \frac{\beta}{Q} \mathcal{M}(E_t U_C(C_{t+1}, L_{t+1}, A_{t+2})) = q_t + \mathcal{M}(U_C(C_t, L_t, A_{t+1}))$$

Using standard properties for log linear approximations of expectations, first order

taylor expansion of U_c and U_A and the hypothesis $a_{t+1} = 0$, I obtain

$$\left(1 - \frac{\beta}{Q}\right) \frac{U_{CAC}}{U_A} c_t + \frac{\beta}{Q} \frac{U_{CC}C}{U_C} E_t c_{t+1} = q_t + \frac{U_{CC}C}{U_C} c_t$$

B.1.10 Proof of proposition 2.9

The proposition rewrites the system of equations from proposition 2.8 with notations introduced in the paper.

B.1.11 Proof of proposition 2.10

Combination of equations (2.14) and (2.15) gives

$$\begin{aligned} \theta \eta n_t &= w_t + \frac{\sigma}{\sigma + \kappa \nu} r r_t + \frac{(1 - \kappa) \sigma}{\sigma + \kappa \nu} (\theta \eta n_{t+1} - w_{t+1}) \\ \theta \eta \left(n_t - \frac{(1 - \kappa) \sigma}{\sigma + \kappa \nu} n_{t+1} \right) &= \left(w_t - \frac{(1 - \kappa) \sigma}{\sigma + \kappa \nu} w_{t+1} \right) + \frac{\sigma}{\sigma + \kappa \nu} r r_t \end{aligned}$$

Consider a non persistent and exogenous change in real interest rate. Forward looking terms are no longer relevant. The equation becomes

$$\theta \eta n_t = w_t + \frac{\sigma}{\sigma + \kappa \nu} r r_t$$

If $U_{CA} > 0$, then $\nu > 0$ and $\frac{\sigma}{\sigma + \kappa \nu} < 1$. The elasticity of hours with respect to real interest rate is lower than the elasticity with respect to real wages.

B.1.12 Proof of proposition 2.11

The new keynesian labor demand becomes

$$(\sigma + \kappa\nu)\alpha n_t = -rr_t + (1 - \kappa)\sigma\alpha n_{t+1} \quad (\text{B.3})$$

From it, I deduce

$$\left(n_t - \frac{(1 - \kappa)\sigma}{\sigma + \kappa\nu}n_{t+1}\right) = -\frac{1}{\alpha(\sigma + \kappa\nu)}rr_t \quad (\text{B.4})$$

I use the equation giving labor supply as a function of real wages and real interest rate

$$\theta\eta\left(n_t - \frac{(1 - \kappa)\sigma}{\sigma + \kappa\nu}n_{t+1}\right) = \left(w_t - \frac{(1 - \kappa)\sigma}{\sigma + \kappa\nu}w_{t+1}\right) + \frac{\sigma}{\sigma + \kappa\nu}rr_t$$

And replace hours by real interest rate using the relation derived from the labor demand

$$-\frac{\theta\eta}{\alpha(\sigma + \kappa\nu)}rr_t = \left(w_t - \frac{(1 - \kappa)\sigma}{\sigma + \kappa\nu}w_{t+1}\right) + \frac{\sigma}{\sigma + \kappa\nu}rr_t \quad (\text{B.5})$$

rearranging gives

$$w_t = -\left(\underbrace{\frac{\sigma}{\sigma + \kappa\nu}}_{\text{Supply effect}} + \underbrace{\frac{\theta\eta}{(\sigma + \kappa\nu)\alpha}}_{\text{Demand effect}}\right)rr_t + \frac{(1 - \kappa)\sigma}{\sigma + \kappa\nu}w_{t+1} \quad (\text{B.6})$$

B.2 Wealth in the utility function and indivisible labor

In this section, I show how the framework developed by Gali(2011) may be adapted to accommodate wealth in the utility function and nonseparability between wealth and consumption.

The household is composed of a continuum of members indexed by j . Each member may work either one or zero unit of time. The labor is indivisible. Members also differ by the disutility associated to their work time. The j th household suffer a disutility

equal to j^θ . There is perfect risk sharing across household's members for consumption and assets. The head of the household chooses household's consumption assets and working household's members. He maximizes

$$\begin{aligned} U(C_t, A_{t+1}) + \int_0^{N_t} j^\theta dj \\ \Rightarrow U(C_t, A_{t+1}) + \frac{N_t^{1+\theta}}{1+\theta} \end{aligned}$$

Utility is separable between consumption and labor force participation but nonseparable between assets and consumption like in the intertemporal complementarity model.

B.3 Bayesian estimation

First, I display the equations of the estimated model

Equation	Interpretation
$y_t = \alpha n_t + a_t$	Production Function
$y_t = \varphi c_t + d_t$	Market clearing on good market
$r_t - \pi_{t+1} = r r_t$	Accounting equation
$\mu_t = w_t + n_t - y_t$	Marginal cost equation
$\pi_t = \frac{\beta_\pi}{1+\beta_\pi\tau_p} E_t \pi_{t+1} + \frac{\tau_p}{1+\beta_\pi\tau_p} \pi_{t-1} + \gamma_p \mu_t + e p_t$	Philips Curve
$r g_t = \phi_\pi \pi_t + \phi_y (y_t - \tilde{y}_t) + e_t$	Monetary policy rule
$r_t = \lambda r_{t-1} + (1 - \lambda) r g_t$	Effective nominal interest rate
$\theta \eta n_t^d = w_t - \sigma c_t$	Labor supply
$u_t = n_t^d - n_t$	Unemployment
$\pi_{w,t} = \frac{\beta_w}{1+\beta_w\tau_w} E_t \pi_{w,t+1} + \frac{\tau_w}{1+\beta_w\tau_w} \pi_{w,t-1} + \gamma_w u_t + e w_t$	Wage Philips curve
$w_t = w_{t-1} + \pi_{w,t} - \pi_t$	Real wage equation
$(\sigma + \kappa\nu) c_t = -r r_t + (1 - \kappa)\sigma E_t c_{t+1} + e c_t$	Consumption equation

TABLE B.1 – Estimated model

I now display posteriors for the different estimation

Variable	distribution	mean	std
κ	beta	0.5	0.2
ν	normal	0	1.5
σ	normal	1	1.5
θ	gamma	2	0.25
γ_p	normal	0.25	0.05
γ_w	normal	0.25	0.05
ϕ_π	normal	1.5	0.25
ϕ_y	normal	0.12	0.25
h	beta	0.5	0.15
τ_p	beta	0.5	0.15
τ_w	beta	0.5	0.15

TABLE B.2 – Priors for parameters

Variable	distribution	mean	std
ρ_d	beta	0.5	0.2
ρ_e	beta	0.5	0.2
ρ_a	beta	0.5	0.2
ρ_{ec}	beta	0.5	0.2
ρ_{ep}	beta	0.5	0.2
ρ_{ew}	beta	0.5	0.2
σ_d	inv gamma	0.4	0.5
σ_e	inv gamma	0.4	0.5
σ_a	inv gamma	0.4	0.5
σ_{ec}	inv gamma	0.4	0.5
σ_{ep}	inv gamma	0.4	0.5
σ_{ew}	inv gamma	0.4	0.5

TABLE B.3 – Priors for standard deviation and persistence

Variable	mean	mode	inf	sup
σ	2.77	2.63	2.16	3.37
θ	2.58	2.51	2.14	2.98
γ_p	0.0042	0.0019	-0.0015	0.0109
γ_w	0.206	0.247	0.0513	0.318
ϕ_y	1.24	1.27	0.952	1.57
ϕ_π	1.75	1.66	1.23	2.3
τ_p	0.374	0.356	0.153	0.592
τ_w	0.138	0.104	0.0471	0.225

TABLE B.4 – Posteriors for the standard model estimation

Variable	mean	mode	inf	sup
κ	0.797	0.839	0.644	0.961
ν	2.13	2.01	1.03	3.19
σ	1.17	1.11	0.627	1.69
θ	2.31	2.26	1.9	2.7
γ_p	0.0004	0.0005	-0.0006	0.0013
γ_w	0.264	0.284	0.178	0.374
ϕ_y	0.463	0.44	0.113	0.788
ϕ_π	1.47	1.46	1.07	1.86
τ_p	0.287	0.315	0.131	0.435
τ_w	0.126	0.0956	0.0378	0.21
ρ_a	0.925	0.924	0.884	0.969
ρ_d	0.849	0.853	0.782	0.917
ρ_{ep}	0.101	0.0663	0.014	0.181
ρ_e	0.65	0.651	0.521	0.778
ρ_{en}	0.958	0.969	0.93	0.986
ρ_{ec}	0.94	0.953	0.907	0.973
ρ_{ew}	0.752	0.811	0.644	0.913
σ_d	0.516	0.506	0.454	0.577
σ_a	0.202	0.199	0.177	0.226
σ_e	0.12	0.118	0.105	0.134
σ_{ep}	0.0696	0.0678	0.059	0.0789
σ_{ec}	0.209	0.207	0.179	0.238
σ_{en}	0.288	0.281	0.248	0.325
σ_{ew}	0.221	0.206	0.176	0.267

TABLE B.5 – Posteriors for the baseline estimation

B.4 Generalized nonseparable preferences in the model

In this section, I consider the case where all form of nonseparabilities are possible.

Reconsider the model with all forms of nonseparabilities.

$$\begin{aligned}\frac{U_{CC}C}{U_C}c_t + \frac{U_{CAA}}{U_C}a_{t+1} + \frac{U_{CL}L}{U_C}l_t &= \lambda_t \\ \frac{U_{CL}C}{U_L}c_t + \frac{U_{LAA}}{U_L}a_{t+1} + \frac{U_{LL}L}{U_L}l_t &= w_t + \lambda_t \\ \frac{U_{CA}C}{U_A}c_t + \frac{U_{AAA}}{U_A}a_{t+1} + \frac{U_{LAL}}{U_A}l_t &= q_t + \lambda_t\end{aligned}$$

I adopt more friendly notations. All elasticities are denoted by δ plus a subscript corresponding to the second derivative. Thus, I have $\frac{U_{CC}C}{U_C} = \delta_{cc}$, $\frac{U_{AAA}}{U_A} = \delta_{aa}$ etc. For cross derivative, I have $\delta_{la} = \frac{U_{LAL}}{U_A}$, $\delta_{ca} = \frac{U_{CA}C}{U_A}$ and $\delta_{cl} = \frac{U_{CL}C}{U_L}$. Moreover, I denote $m_{cl} = \frac{WL}{C}$, $m_{ac} = \frac{QA}{C}$ and $m_{al} = \frac{QA}{WL}$.

The system becomes

$$\begin{aligned}\delta_{cc}c_t + \delta_{ca}m_{ca}a_{t+1} + \delta_{cl}m_{cl}l_t &= \lambda_t \\ \delta_{cl}c_t + \delta_{la}m_{la}a_{t+1} + \delta_{ll}l_t &= w_t + \lambda_t \\ \delta_{ca}c_t + \delta_{aa}a_{t+1} + \delta_{al}l_t &= q_t + \lambda_t\end{aligned}$$

If $a_{t+1} = 0$, we get

$$\begin{aligned}\delta_{cc}c_t + \delta_{cl}m_{cl}l_t &= \lambda_t \\ \delta_{cl}c_t + \delta_{ll}l_t &= w_t + \lambda_t \\ \delta_{ca}c_t + \delta_{al}l_t &= q_t + \lambda_t\end{aligned}$$

eliminating λ_t

$$\begin{aligned} -(\delta_{cl}m_{cl} - \delta_{ll})l_t &= w_t - (\delta_{cl} - \delta_{cc})c_t \\ (\delta_{ca} - \delta_{cc})c_t + (\delta_{al} - \delta_{cl}m_{cl})l_t &= q_t \end{aligned}$$

The consumption equation can be written with respect to wages and interest factors

$$\left(\delta_{ca} - \delta_{cc} + \frac{(\delta_{al} - \delta_{cl}m_{cl})(\delta_{cl} - \delta_{cc})}{\delta_{cl}m_{cl} - \delta_{ll}} \right) c_t = q_t + \frac{\delta_{al} - \delta_{cl}m_{cl}}{\delta_{cl}m_{cl} - \delta_{ll}} w_t$$

B.5 Adjustment cost for consumption in an overlapping generations model

Consider an household living two periods, receiving wages only in the first period. The utility at the second period features habit consumption

$$\begin{aligned} \max_{c_y, c_o} u(c_y) + u(c_o - hc_y) \\ \text{w.r.t } w &= c_y + a \\ \text{w.r.t } a &= c_o \end{aligned}$$

Where a are assets accumulated at the first period. c_y and c_o are consumption levels the agent is young and old. h is between 0 and 1. $u'(\cdot) > 0$, $u''(\cdot) < 0$

The objective function can also be written under the form

$$u(c^y) + v(a, c^y)$$

where $v(a, c^y) = u(a - hc_y)$

I compute first order and cross derivative of v

$$\frac{dv}{da} = u'(a - hc_y)$$

The cross derivative with respect to consumption of the young is

$$\frac{d^2v}{dadc_y} = -hu''(a - hc_y) > 0 \quad (\text{B.7})$$

The cross derivative is positive

B.6 Forward guidance computations

B.6.1 Forward guidance in the standard model

I consider a standard linear New Keynesian model.

$$\sigma(c_{t+1} - c_t) = r_t - \pi_{t+1}$$

$$-\theta l_t = w_t - \sigma c_t$$

$$l_t = -\eta n_t$$

$$\mu_t = w_t + n_t - y_t$$

$$y_t = \alpha n_t$$

$$\pi_t = \psi \mu_t + \beta_\pi \pi_{t+1}$$

$$r_t = \phi_\pi \pi_t + \epsilon_t$$

$$y_t = c_t$$

Where c is consumption, r is nominal interest rate, π is inflation, l is leisure, n is hours

worked, y is output, w is real wage, μ is the marginal cost (the inverse of the markup), and ϵ is a monetary policy shock.

First, I compute multipliers for a contemporaneous monetary policy shock. They are denoted $\psi_{(\cdot)}$. the subscript denotes the variable of interest.

$$\begin{aligned}\Psi_w &= \frac{\theta\eta + \sigma\alpha}{\alpha}\psi_y \\ \Psi_\mu &= \frac{\theta\eta + \sigma\alpha + 1 - \alpha}{\alpha}\psi_y \\ \Psi_\pi &= \frac{\psi}{1 - \beta_\pi\rho} \frac{\theta\eta + \sigma\alpha + 1 - \alpha}{\alpha}\Psi_y \\ \Psi_y &= \frac{(1 - \beta_\pi\rho)\alpha}{(1 - \beta_\pi\rho)\alpha\sigma(\rho - 1) - (\phi_\pi - \rho)\psi(\theta\eta + \sigma\alpha + 1 - \alpha)}\end{aligned}$$

The first multiplier is obtained by combining labor supply equation, production function and market clearing condition. The second comes from the combination of the markup equation, the previous result for wages and the production function. The markup multiplier immediately gives the multiplier for inflation.

I now compute multipliers for a shock occurring in $t + 1$ under the assumption that nominal interest rate in t is fixed. I denote these multipliers $\mathcal{M}_{(\cdot)}$. I derive the system relating current multiplier for output and inflation to future ones.

$$\begin{aligned}\sigma\mathcal{M}_y &= \sigma\Psi_y + \Psi_\pi \\ \mathcal{M}_\pi &= \psi \frac{\theta\eta + \sigma\alpha + 1 - \alpha}{\alpha}\mathcal{M}_y + \beta_\pi\Psi_\pi\end{aligned}$$

Solving the system leads to

$$\mathcal{M}_y = \left(1 + \frac{\psi(\theta\eta + \sigma\alpha + 1 - \alpha)}{\sigma(1 - \beta_\pi\rho)\alpha}\right)\Psi_y \geq \Psi_y \quad (\text{B.8a})$$

$$\mathcal{M}_\pi = \Psi_\pi \left(\beta_\pi + \frac{\psi}{\alpha\sigma}(\theta\eta + \sigma\alpha + 1 - \alpha)\right) \quad (\text{B.8b})$$

B.6.2 Forward guidance with wealth in the utility

The equation for consumption is now

$$(\sigma + \kappa\nu)c_t = -r_t + \pi_{t+1} + (1 - \kappa)\sigma c_{t+1} \quad (\text{B.9})$$

Ψ_w , Ψ_μ and Ψ_π remains unchanged.

The output multiplier associated with a contemporaneous monetary policy shock is

$$\Psi_y = \frac{(1 - \beta_\pi\rho)\alpha}{(1 - \beta_\pi\rho)\alpha(\sigma + \kappa\nu - (1 - \kappa)\sigma\rho) - (\phi_\pi - \rho)\psi(\theta\eta + \sigma\alpha + 1 - \alpha)}$$

Computing multiplier for a shock occurring in $t + 1$ gives

$$\begin{aligned} (\sigma + \kappa\nu)\mathcal{M}_y &= (1 - \kappa)\sigma\Psi_y + \Psi_\pi \\ \mathcal{M}_\pi &= \psi\frac{\theta\eta + \sigma\alpha + 1 - \alpha}{\alpha}\mathcal{M}_y + \beta_\pi\Psi_\pi \end{aligned}$$

For the output multiplier, the solution is now

$$\mathcal{M}_y = \left(\frac{(1 - \kappa)\sigma}{\sigma + \kappa\nu} + \frac{\psi(\theta\eta + \sigma\alpha + 1 - \alpha)}{(\sigma + \kappa\nu)(1 - \beta_\pi\rho)\alpha} \right) \Psi_y \quad (\text{B.10})$$

B.7 Model with creditors and debtors

Compare to the standard model, the labor supply and the consumption equation are replaced by first order conditions for both agents, and aggregation equation . In their linear version, it gives

$$\begin{aligned}
-\sigma c_t^d &= -\gamma^d a_{t+1}^d - q_t \\
-\theta_d l_t^d &= w_t - q_t - \gamma a_{t+1}^d \\
-\sigma c_t^c &= -\gamma^c a_{t+1}^c - q_t \\
-\theta_c l_t^c &= w_t - q_t - \gamma a_{t+1}^c \\
\lambda_c c_t^c + (1 - \lambda_c) c_t^d &= c_t \\
\lambda_l l_t^c + (1 - \lambda_c) l_t^d &= l_t \\
a_{t+1}^d &= -a_{t+1}^c
\end{aligned}$$

Where λ_c and λ_l are the share of debtors for respectively consumption and leisure at steady state. Parameters γ , θ and σ have the same meaning as in the paper but are specific to each type of household and are indexed by subscripts d or c . To solve the problem, I need the budget constraint of one of the two agents. The budget constraint of debtors is

$$W_t + A_t^d = W_t L_t + C_t + Q_t A_{t+1}^d$$

I denote $m = \frac{C}{W+A+\Pi}$ the average propensity to consume at steady state and $v = \frac{WL}{W+A}$ its equivalent for leisure.

The linearized budget constraint gives

$$v w_t + v l_t + m c_t^d + (1 - m - v)(q_t + a_{t+1}^d) = \frac{1 - m - v}{q} a_t^d + \frac{m + v}{q} w_t \quad (\text{B.11})$$

I can now derive an equation for debtor asset law of motion by combining the budget

constraint with first order conditions for debtors

$$\left(1 - m - v + \frac{v\gamma_d}{\theta_d}\right) a_{t+1}^d = w_t \left(\frac{m+v}{q} - v + \frac{v}{\theta_d}\right) + a_t^d \left(\frac{1-m-v}{q}\right) - q_t \left(1 - m - v + \frac{v}{\theta_d} + \frac{m}{\sigma}\right) \quad (\text{B.12})$$

Equations for consumption and aggregate leisure are

$$\begin{aligned} -\sigma c_t &= -q_t + [\lambda_c \gamma_c - \gamma_d(1 - \lambda_c)] a_{t+1}^d \\ -(\theta_d(1 - \lambda_l) + \theta_c \lambda_l) l_t &= w_t - q_t + (\lambda_l \gamma_c - \gamma_d(1 - \lambda_l)) a_{t+1}^d \end{aligned}$$

Appendix C

Appendix of chapter 3

C.1 The Adverse Selection Problem

C.1.1 Proof for proposition 3.2

When an investor lends to bad firms, it will receive r^B until the firm dies. The non arbitrage condition is

$$\begin{aligned}
 1 &= \sum_{k=1}^{+\infty} \left(\frac{\lambda}{1+r} \right)^k r^B \\
 \Rightarrow \frac{1+r}{\lambda} &= r^B \sum_{k=0}^{+\infty} \left(\frac{\lambda}{1+r} \right)^k \\
 r^B &= \frac{1+r-\lambda}{\lambda}
 \end{aligned}$$

Computing the interest rate charged on good firms is a little bit more complex, because a good firm first becomes a bad firm before dying.

I denote $Q^B \equiv \sum_{k=0}^{+\infty} \left(\frac{\lambda}{1+r} \right)^k$. Q^B is the discounted value of an income stream generating one unit of good at each period as long as a given bad firm survives.

Q^B is also given by the recursive equation

$$Q^B = 1 + \frac{\lambda}{1+r} Q^B$$

$$Q^B = \frac{1}{1 - \frac{\lambda}{1+r}} = \frac{1+r}{1+r-\lambda}$$

I also define Q^G as the discounted value of an income stream generating one unit of good at each period as long as the good firm survive.

$$Q^G = 1 + \frac{1-\kappa}{1+r} Q^G + \frac{\kappa}{1+r} Q^B$$

$$Q^G \left(\frac{r+\kappa}{1+r} \right) = \frac{1+r-\lambda}{1+r-\lambda} + \frac{\kappa}{1+r-\lambda}$$

$$Q^G = \frac{(1+r-\lambda+\kappa)(1+r)}{(r+\kappa)(1+r-\lambda)}$$

$$1 = \left(\frac{1-\kappa}{1+r} \right) Q^G r^G + \frac{\kappa}{1+r} Q^B r^G$$

$$\Rightarrow \frac{1}{r^G} = \frac{1-\kappa}{1+r} \frac{(1+r-\lambda+\kappa)(1+r)}{(r+\kappa)(1+r-\lambda)} + \frac{\kappa}{1+r} \frac{1+r}{1+r-\lambda}$$

$$\Rightarrow \frac{1}{r^G} = \frac{(1+r-\lambda+\kappa)(1-\kappa) + \kappa(r+\kappa)}{(r+\kappa)(1+r-\lambda)}$$

$$r^G = \frac{(r+\kappa)(1+r-\lambda)}{(1-\kappa)(1+r-\lambda+\kappa) + \kappa(r+\kappa)}$$

□

The proofs for γ^B and γ^G are symmetrical from the previous ones.

C.1.2 Proof of proposition 3.4

An investment financed by debt in a bad firm generate a benefit π and a cost r_b at each period. $r_B = \frac{1+r-\lambda}{\lambda} > \pi$. Thus, bad firms do not want to borrow.

To demonstrate that a bad firm do not want to invest its earnings if $\frac{1+\gamma-\lambda}{\lambda} > \pi$, I assume that the firm invest zero and saves zero and I consider a small deviation from this path. There is an increase in investment by one unit only in period t . The investment implies a lower dividend in period t but generates a constant income stream after. The variation of the value of the firm is

$$\Delta V^{B,B} = -1 + \sum_{k=1}^{+\infty} \left(\frac{\lambda}{1+\gamma} \right)^k \pi$$

The variation is negative if

$$\begin{aligned} \Delta V^{B,B} &< 0 \\ \Rightarrow \sum_{k=1}^{+\infty} \left(\frac{\lambda}{1+\gamma} \right)^k \pi &< 1 \\ \Rightarrow \pi \frac{\lambda}{1+\gamma} \frac{1}{1 - \frac{\lambda}{1+\gamma}} &< 1 \\ \Rightarrow \pi \frac{\lambda}{1+\gamma-\lambda} &< 1 \\ \Rightarrow \pi &< \frac{1+\gamma-\lambda}{\lambda} \end{aligned}$$

Thus an investment financed by equity will generate a net loss for the bad firm. Note that if the firm reinvest the earnings generated by the investment in period t , it will increase its loss. Thus, a bad firm do not want to invest under hypothesis 2. It will distribute all its income as dividends. \square

C.1.3 Proof of proposition 3.5

The proof follows the same strategy than the previous one. I show that both an investment financed by debt and an investment financed by retained earnings is profitable

for the good firm.

Hypothesis 3 implies that $\pi > r^G$. Thus, investment financed by debt is profitable.

The variation of value following a small rise in investment financed by retained earnings in period t gives

$$\Delta V^{G,G} = -1 + \frac{1-\kappa}{1+\gamma} \Delta V'_{G,G} + \frac{\kappa}{1+\gamma} \Delta V'_{B,B}$$

and

$$\Delta V'_{G,G} = \pi + \frac{1-\kappa}{1+\gamma} \Delta V'_{G,G} + \frac{\kappa}{1+\gamma} \Delta V'_{B,B}$$

The previous proof have shown that

$$\Delta V'_{B,B} = \frac{\lambda\pi}{1+\gamma-\lambda}$$

After computations, it gives

$$\begin{aligned} \Delta V^{G,G} &> 0 \\ \Rightarrow \pi &> \frac{(\gamma+\kappa)(1+\gamma-\lambda)}{(1-\kappa)(1+\gamma-\lambda+\kappa) + \kappa(\gamma+\kappa)} \end{aligned}$$

Under hypothesis 3, equity financed investment is profitable. Thus, the good firm want to invest and borrow as much as they can. They distribute the minimal level of dividends and invest until the ratio of investment over saving reaches the threshold ratio $\bar{\psi}$ □

C.1.4 Proof of proposition 3.13

A bad firm always sending the correct signal neither invests nor borrows, hence the value $V^{B,B}$ is

$$V^{B,B}(K, B) = (\pi K - B) \frac{1 + \gamma}{1 + \gamma - \lambda}$$

Under the separating equilibrium, at each period, $S^G = s(\pi K - B)$ and $I^G = \psi S^G$ where ψ is the leverage of good firms. Profits of a bad firm always sending the incorrect signal grow at a rate $s(\psi\pi - r^G(\psi - 1))$. Using standard results on geometric sum

$$U^{B,G}(K, B) = (1 - s)(\pi K - B) \frac{1 + \gamma}{1 + \gamma - \lambda(1 + s(\psi\pi - r^G(\psi - 1)))}$$

I now demonstrate in appendix that

$$(\pi K - B) \frac{1 + \gamma}{1 + \gamma - \lambda} \geq (1 - s)(\pi K - B) \frac{1 + \gamma}{1 + \gamma - \lambda(1 + s(\psi\pi - r^G(\psi - 1)))} \quad (\text{C.1})$$

$$\Rightarrow \psi \leq \frac{\gamma^B - r^G}{\pi - r^G} S^G \quad (\text{C.2})$$

Indeed,

$$(\pi K - B) \frac{1 + \gamma}{1 + \gamma - \lambda} \geq (1 - s)(\pi K - B) \frac{1 + \gamma}{1 + \gamma - \lambda(1 + s(\psi\pi - r^G(\psi - 1)))} \quad (\text{C.3})$$

$$\Rightarrow 1 \geq (1 - s) \frac{1 + \gamma - \lambda}{1 + \gamma - \lambda(1 + s(\psi\pi - r^G(\psi - 1)))} \quad (\text{C.4})$$

$$\Rightarrow 1 + \gamma - \lambda(1 + s(\psi\pi - r^G(\psi - 1))) \geq (1 - s)(1 + \gamma - \lambda) \quad (\text{C.5})$$

$$\Rightarrow \lambda s(\psi\pi - r^G(\psi - 1)) \leq s(1 + \gamma - \lambda) \quad (\text{C.6})$$

$$\Rightarrow \frac{\lambda}{1 + \gamma - \lambda} (\psi\pi - r^G(\psi - 1)) \leq 1 \quad (\text{C.7})$$

$$(\text{C.8})$$

Proposition 3.2 gives $\gamma^B = \frac{1 + \gamma - \lambda}{\lambda}$, hence

$$\Rightarrow \psi\pi - r^G(\psi - 1) \leq \gamma^B \quad (\text{C.9})$$

$$\Rightarrow \psi \leq \frac{\gamma^B - r^G}{\pi - r^G} \quad (\text{C.10})$$

□

C.2 The Macroeconomic Model

C.2.1 Proof of proposition 3.14

The incentive compatibility constraint implies that the value for a bad firm to send the "bad" signal should be superior to the value of sending the "good" signal.

$$V_t^{B,B} \geq V_t^{B,G} \quad (\text{C.11})$$

Under separating equilibrium, bad firms never invests or borrows thus their capital stock and their debt remains at the same level. The value function $V^{B,B}$ can be computed quite simply

$$V_t^{B,B}(K_t, B_t) = \pi K_t - B_t + \lambda \frac{1}{1 + \gamma_t} V_{t+1}^{B,B}(K_t, B_t)$$

I factorize the right side expression by K_t and B_t and get

$$V_t^{B,B}(K_t, B_t) = (\pi K_t - B_t) \Gamma_t^B$$

The value function $V_t^{B,G}$ follows

$$V_t^{B,G} = \pi K_t - B_t - S_t^G + \lambda \frac{1}{1 + \gamma_t} V_{t+1}^{B,B}(K_t + I_t^G, B_t + r_t^G(I_t^G - S_t^G))$$

The ICC becomes

$$\pi K_t - B_t + \lambda \frac{1}{1 + \gamma_t} V_{t+1}^{B,B}(K_t, B_t) \geq \pi K_t - B_t - S_t^G + \lambda \frac{1}{1 + \gamma_t} V_{t+1}^{B,B}(K + I^G, B + r^G(S^G - I^G))$$

Using the expression found for $V^{B,B}$

$$\begin{aligned} & \pi K_t - B_t + \lambda \frac{1}{1 + \gamma_t} (\pi K_t - B_t) \Gamma_{t+1}^B \\ & \geq \pi K_t - B_t - S_t^G + \lambda \frac{1}{1 + \gamma_t} (\pi(K_t + I_t^G) - B_t - r_t^G(I_t^G - S_t^G)) \Gamma_{t+1}^B \end{aligned}$$

By simplifying, I get

$$\frac{1 + \gamma_t}{\lambda} S_t^G \geq + [\Gamma_{t+1}^B \pi I_t^G - \Gamma_{t+1}^B r_t^G (I_t^G - S_t^G)]$$

By equation (3.20a)

$$\gamma_t^B \Gamma_{t+1}^B = \frac{1 + \gamma_t}{\lambda}$$

then

$$I_t^G \leq \frac{\gamma_t^B - r_t^G}{\pi - r_t^G} S_t^G \quad (\text{C.12})$$

□

C.2.2 Model derivation

The system of recursive equations for good firms can be transformed into one recursive equation.

The original system is

$$\begin{aligned}
K_{t+1}^G &= (1 - \kappa)(K_t^G + I_t^G) \\
B_{t+1}^G &= (1 - \kappa)(B_t^G + r_t^G(I_t^G - S_t^G)) \\
\pi K_t^G - B_t^G &= S_t^G + d_t^G \\
d_t^G &= (1 - s)(\pi K_t^G - B_t^G) \\
I_t^G &= \psi_t S_t^G
\end{aligned}$$

I divide all variables by K_t^G and label $\frac{K_{t+1}^G}{K_t^G} \equiv 1 + g_{t+1}$. Stationnarized variables are denoted by lowercase without subscript : $\frac{I_t^G}{K_t^G} \equiv i_t$ etc. \tilde{d}_t is the stationary value of d_t^G . It gives

$$\begin{aligned}
(1 + g_{t+1}) &= (1 - \kappa)(1 + i_t) \\
b_{t+1} &= (1 - \kappa)(b_t + r_t^G(i_t - s_t)) \\
\pi - b_t &= s_t + \tilde{d}_t \\
\tilde{d}_t &= (1 - s)(\pi - b_t) \\
i_t &= \psi_t s_t
\end{aligned}$$

Combining the last three equations and eliminating \tilde{d}_t and s_t gives

$$\begin{aligned}
(1 + g_{t+1})b_{t+1} &= (1 - \kappa)(b_t + r_t^G(\psi_t - 1)s(\pi - b_t)) \\
1 + g_{t+1} &= (1 - \kappa)(1 + i_t) \\
i_t &= \psi_t s(\pi - b_t)
\end{aligned}$$

I reintroduce the last two equations in the first one and get the recursive equation for

debt

$$b_{t+1} = \frac{b_t + r_t^G(\psi_t - 1)s(\pi - b_t)}{1 + \psi_t s(\pi - b_t)}$$

Calibration Tables

Variable	parameter value	Variable definition
π	0.08	Average Return on capital
s	0.5	Saving rate of firms
γ	0.05	Return of outside investment
κ	0.02	Transition probability from good to bad
λ	0.95	survival rate

TABLE C.1 – Calibration for the Investment Model

Proof of proposition 14 I denote f the function $f(b_t) = \frac{b_t + r^G(\psi - 1)s(\pi - b_t)}{1 + \psi s(\pi - b_t)}$

Let's replace b_t by x . The equation of the mapping f is

$$f(x) = \frac{x + r^G(\psi - 1)s(\pi - x)}{1 + \psi s(\pi - x)}$$

Computing the derivative

$$\begin{aligned} f'(x) &= \frac{(1 - r^G(\psi - 1)s)(1 + \psi s(\pi - x)) + \psi s(x + r^G(\psi - 1)s(\pi - x))}{(1 + \psi s(\pi - x))^2} \\ \Rightarrow f'(x) &= \frac{1 - r^G(\psi - 1)s + \psi s x}{(1 + \psi s(\pi - x))^2} \end{aligned}$$

The derivative $f'(\pi) = 1 + s(\psi\pi - r^G(\psi - 1))$. Because by assumption 1 $\pi > r^G$ it is clear that the derivative of f is always superior to one when $b_t = b^*$. thus b^* is an unstable root.

The derivative $f'(r^G \frac{\psi - 1}{\psi}) = \frac{1}{[1 + s(\psi\pi - r^G(\psi - 1))]^2} < 1$. So b^{**} is a stable root. \square

List of Figures

1.1	Scope of multiple equilibrium	35
1.2	Convergence for several calibrations	44
1.3	Spread between rational and consistent errors for several calibrations	45
1.4	Convergence with respect to λ and θ	46
1.5	Error difference with respect to λ and θ	47
1.6	Sensitivity to σ_y and σ_u	48
1.7	Sensitivity to γ_0 and ν	48
1.8	Sensitivity to λ and θ when initialization parameters are random	49
1.9	Convergence for different initialization	50
1.10	Error spread for different initialization	50
1.11	Sensitivity to λ and θ : average mass of consistent learners after multiple simulations	52
1.12	Sensitivity of the response to λ and ν with structural breaks	54
2.1	Real wages following a rise in real rate	82
2.2	Unemployment following a rise in real rate	84
2.3	Responses to monetary policy shocks	86
2.4	Responses for several calibration	86
2.5	Labor supply and demand after a monetary policy shock	87
2.6	Response of Real wages to demand shocks	88
2.7	Response of output gap and output to demand shocks	90
2.8	Posterior and prior distribution for κ and ν	92
2.9	Predicted and actual labor supply	94

2.10	Impulse responses to Monetary Policy shocks	97
2.11	Responses of labor supply and demand to Monetary Policy Shocks . . .	97
2.12	Response of output gap to demand shocks	98
3.1	State Tree	118
3.2	Representation of the recursive equation	144
3.3	Response to a permanent change in borrowing cost	149
3.4	Response to a temporary change in borrowing cost	150
3.5	Response to a permanent change in opportunity cost	150
3.6	Response to a temporary change in opportunity cost	150

List of Tables

1.1	Calibration Table for figure 1.1	35
1.2	Calibration Table for the Simulated Economy	42
2.1	Simple Model	76
2.2	Medium scale Model	77
2.3	calibration	77
2.4	Priors for parameters	91
2.5	Posteriors for the baseline estimation	92
2.6	Posteriors for the standard model estimation	93
2.7	Posterior mean of key parameters for different variants	95
3.1	Calibration for the Investment Model	138

Résumé français

Cette thèse aborde plusieurs problématiques posées par les cycles économiques. Elle se compose de trois chapitres, chacun pouvant être lu indépendamment. Cependant, chaque chapitre a potentiellement des implications en matière de politique économique, en particulier pour la transmission de la politique monétaire.

Chapitre 1

Le contexte théorique Le premier chapitre est une contribution théorique à la littérature sur les anticipations. La modélisation des anticipations est un sujet difficile mais important à la fois pour la théorie macroéconomique et l'analyse macroéconométrique.

Du point de vue d'un agent économique, beaucoup de décisions impliquent de prévoir les valeurs que sont susceptibles de prendre les grandes variables macroéconomiques dans l'avenir. Par exemple, une firme désirant savoir si construire une nouvelle usine sera profitable ou non, doit prévoir les futurs profits et les futurs coûts engendrés par cet investissement. Un investisseur choisissant entre deux actions doit anticiper la croissance future des profits des deux entreprises. Les anticipations sont également importantes pour la consommation, l'offre de travail, ou encore la fixation des prix et salaires. Au niveau agrégé, la plupart des variables macroéconomiques sont affectées par les prévisions des agents.

La question de la formation des anticipations est donc cruciale. Pour prévoir la valeur

future d'une variable x , la possibilité la plus immédiate est d'utiliser la valeur actuelle de x . Le problème de cette approche est que la valeur présente de x peut délivrer un signal très imparfait de sa valeur future. Une solution est d'obtenir davantage d'informations, par exemple en utilisant une moyenne mobile. Il s'agit du principe des anticipations adaptatives.

L'inconvénient de cette approche est que les agents n'utilisent pas toute l'information disponible, et en particulier, les nouvelles informations. Une crise financière, une banque centrale plus conservatrice, un nouveau gouvernement souhaitant accroître le déficit public, affecteront vraisemblablement les futures valeurs du revenu national. Un agent purement adaptatif ne réagira pas immédiatement à ces changements parce qu'ils n'ont pas encore modifié le revenu national ou les autres variables macroéconomiques. L'hypothèse des anticipations rationnelles, introduite par Muth(1961) et popularisée par Lucas, implique que les agents utilisent toute l'information à leur disposition et réagissent donc à ce type de changement. L'anticipation de la variable x est égale à son espérance mathématique, conditionnelle à l'information dont l'agent dispose. Formellement, on note l'anticipation x^e , l'ensemble d'information I . Utilisant la notation habituelle pour l'opérateur espérance, l'anticipation est donnée par

$$x^e = E(X | I)$$

En pratique, dans la plupart des modèles utilisant cette hypothèse, l'agent dispose du même niveau de connaissance que le modélisateur. Par exemple, dans un modèle de croissance optimale à agent représentatif vivant indéfiniment, chaque agent calcule les trajectoires possibles de la productivité et les probabilités associées, résout le modèle macroéconomique, et en déduit les futures trajectoires du stock de capital, du salaire réel et du taux d'intérêt réel. Une fois obtenues les trajectoires du salaire et du taux d'intérêt, il résout son problème d'optimisation intertemporelle.

Il est évident que les anticipations rationnelles implique des capacités cognitives phénoménales de la part des agents, et ce même dans des modèles assez simples, ainsi qu'une connaissance exceptionnelle de leur environnement. Ils connaissent le vrai modèle de l'économie et les vraies valeurs des paramètres du modèle. Ils savent aussi comment les autres agents forment leurs anticipations. Ces hypothèses sont très fortes et les anticipations rationnelles ont été critiquées pour cela (voir par exemple Guesnerie 2001, 2005). Plusieurs littératures ont exploré des théories alternatives des anticipations. L'une d'entre elles est la littérature sur l' "apprentissage". Les agents ne connaissent pas les vraies valeurs de ces paramètres mais il les obtiennent en estimant la solution du modèle à anticipations rationnelles via des techniques économétriques. Les anticipations cohérentes, définies par Hommes et Sorger (1998), représentent une autre alternative. Elles sont proches des anticipations adaptatives. L'anticipation de x dépend de ses valeurs passées. La différence est que la relation entre l'anticipation et les valeurs passées doit être cohérentes avec les valeurs observées des auto-corrélations de la variable x .

Le papier Le premier chapitre s'inspire de ces deux littératures. J'étudie une économie très simple. Une variable macroéconomique x est donnée par l'équation

$$x_t = \alpha + \beta y_t + \lambda x_t^E + u_t$$

α est une constante. β et λ sont des paramètres. u_t est un bruit blanc dont l'écart type est σ_u . y est une variable exogène suivant un processus AR(1) $y_t = \theta y_{t-1} + \epsilon_t$. θ est un paramètre. ϵ_t est un bruit blanc d'écart type σ_y . x_t^E est l'anticipation de la valeur x en t .

Dans cette économie, les agents utilisent différents modèles de prévisions. L'un de ces modèles est le modèle "rationnel". La variable endogène est fonction des variables exogènes de l'économie. Les agents rationnels estiment à chaque période le modèle

$x_k = \pi_f y_k + \varphi_f + u_{f,k}$. Une fois les paramètres π_f, φ_f estimés, le modèle est utilisé pour former la prévision des agents rationnels à la période t .

Le second modèle utilise le concept des anticipations cohérentes d'Hommes et Sorger (1998). Les agents cohérents utilisent le modèle $x_k = \pi_o x_{k-1} + \varphi_o + u_{o,t}$. Comme pour les agents rationnels, le modèle est ré-estimé à chaque période et les agents cohérents utilisent la nouvelle estimation pour former leur prévision en t .

Les agents adoptent un modèle en fonction de sa performance passée en terme de prévisions. Je définis un équilibre de long terme de cette économie comme une situation dans laquelle (i) les agents ont une estimation stable des deux modèles (ii) Il y a un modèle dominant (i.e tout les agents utilisent un seul des deux modèles) (iii) les erreurs de prévision du modèle dominant sont inférieures aux erreurs de prévision de l'autre modèle. Deux équilibres peuvent exister, l'équilibre rationnel dans lequel les agents utilisent le modèle rationnel (cet équilibre correspond à l'équilibre à anticipations rationnelles) et l'équilibre "cohérent" dans lequel le modèle d'anticipations cohérentes domine. Ma première contribution est de montrer que celui-ci peut exister pour un ensemble de paramètres assez large. Je montre en effet que l'équilibre cohérent existe si

$$\begin{aligned} & (\pi_f - \beta)^2 V(y_t) + (\lambda \pi_o)^2 V(x_t) + \sigma_u^2 - 2\lambda \pi_o (\pi_f - \beta) \theta \text{Cov}(x_t, y_t) \\ & \geq \beta^2 V(y_t) + (1 - \lambda)^2 \pi_o^2 V(x_t) + \sigma_u^2 - 2(1 - \lambda) \pi_o \beta \theta \text{Cov}(y_t, x_t) \end{aligned}$$

Le coté gauche de l'expression représente l'erreur de prévision moyenne du modèle rationnel si tout les agents sont cohérents. L'expression de droite représente l'erreur de prévision du modèle cohérent quand tout les agents sont cohérents. Cette inégalité implique donc que lorsque tout les agents sont cohérents, ils forment de meilleures prévisions en utilisant le modèle cohérent plutôt que le modèle rationnel. Pour rendre cette expression plus lisible, j'essaie de représenter l'ensemble des paramètres pour

lequel l'équilibre existe. Je représente, dans la figure ci-après, la différence entre les erreurs de prévision du modèle rationnel et celles du modèle cohérent en fonction de θ et de λ . La figure représentée est verte quand la différence est supérieure à 0 et donc quand l'équilibre existe.

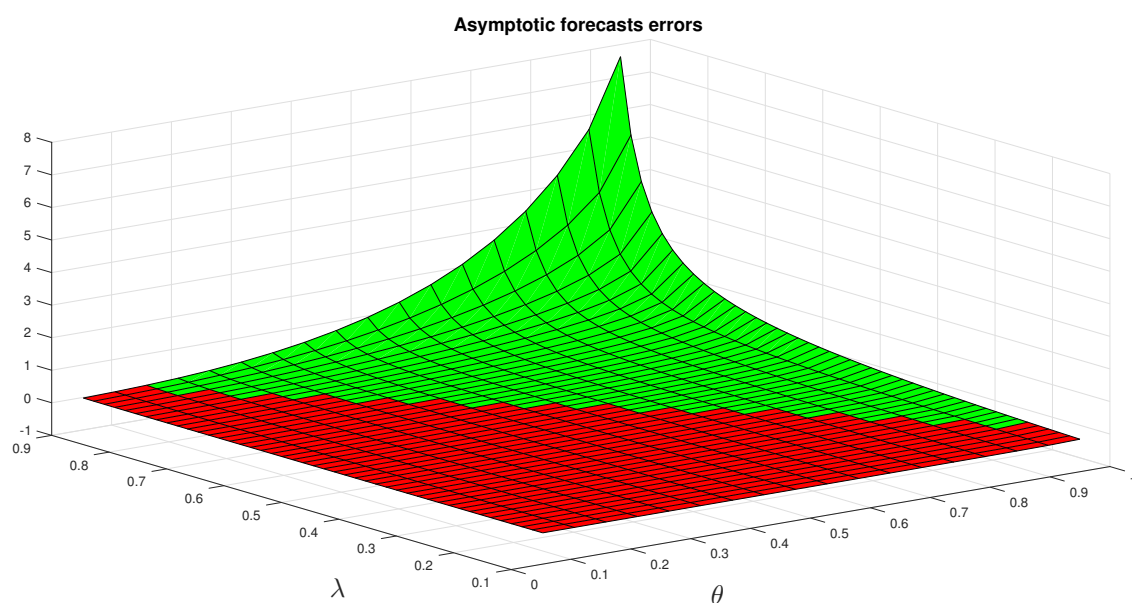


FIGURE 1 – Représentation graphique de la condition d'existence de l'équilibre

Ma seconde contribution est de simuler l'économie et de montrer qu'elle peut converger vers cette équilibre cohérent. Cette partie suit la littérature sur l'apprentissage des agents (voir notamment Evans et Honkapoja 2004). Les agents apprennent leurs modèles respectifs en utilisant les algorithmes des moindres carrés récursifs. A chaque période, les agents comparent les performances de leurs modèles de prévisions. Une fraction des agents utilisant le modèle le moins performant change de modèle à chaque période. Je simule l'économie sur un grand nombre de périodes. Ces simulations montrent que l'économie a une forte probabilité de converger vers l'équilibre cohérent si celui-ci existe.

Chapitre 2

Le contexte Le second chapitre explore une spécification alternative du problème d'optimisation intertemporelle des ménages. Dans la plupart des modèles de cycles économiques, comme le modèle du cycle réel ou le modèle Nouveau Keynesien, les ménages choisissent leur consommation et leur offre de travail en résolvant un problème de maximisation en horizon infini

$$\max_{C_t, L_t, A_{t+1}} \sum_{t=0}^{+\infty} \beta^t U(C_t, L_t)$$

$$w.r.t \quad \forall t \quad (1 + r_t)A_t + W_t = W_t L_t + C_t + A_{t+1}$$

Où C_t est la consommation en t , L_t est le loisir, W_t est le salaire et A_t est un actif, acquis en $t-1$ et donnant droit à une unité de bien ¹ à la période t . U est la fonction objectif du consommateur (dite d'"utilité"). Cette fonction objectif est souvent supposée séparable. Par exemple, une forme fonctionnelle fréquemment utilisée est

$$U(C_t, L_t) = \frac{C_t^{1-\sigma}}{1-\sigma} + \frac{L_t^{1-\theta}}{1-\theta}$$

Résoudre ce programme donne deux conditions du premier ordre.

$$C_t^{-\sigma} = \beta(1 + r_t)C_{t+1}^{-\sigma}$$

$$L_t^{-\theta} = W_t C_t^{-\sigma}$$

La première est l'équation d'Euler et la seconde est la condition du premier ordre pour le loisir. Ces deux équations forment le coeur du modèle de croissance optimale,

¹comme d'habitude dans ce type de modèle, il n'y a qu'un seul bien qui sert à la fois de bien de consommation et de production

l'un des principaux modèles de la macroéconomie moderne. Si l'hypothèse de l'agent représentatif est retenue, i.e tout les agents ont la même fonction objectif et la même dotation initiale, ces deux équations sont pertinentes pour la consommation et le loisir agrégé.

Je considère une légère modification du modèle. A chaque période, l'utilité est fournie par la consommation, le loisir et la richesse, assimilée à la variable A . Le programme devient

$$\max_{C_t, L_t} \sum_{t=0}^{+\infty} \beta^t U(C_t, L_t, A_{t+1})$$

$$w.r.t \quad \forall t \quad (1 + r_t)A_t + W_t = W_t L_t + C_t + A_{t+1}$$

Introduire la richesse dans la fonction d'utilité semble contre-intuitif. *A priori*, les actifs A ne fournissent pas d' "utilité" en eux mêmes, mais seulement à travers les biens qu'ils permettent d'acheter dans le futur. Il y a toutefois plusieurs raisons d'étudier cette spécification. Premièrement, elle est susceptible de capturer plusieurs motifs d'épargne absents du modèle standard, comme l'épargne accumulée pour la retraite ou pour faire face au risque de baisse des revenus, ou celle accumulée dans le but de transmettre un héritage. Ces aspects sont quantitativement importants pour expliquer les comportements d'épargne (voir par exemple Gourinchas et Parker 2002). Les modéliser explicitement serait préférable mais il est compliqué de les intégrer *tous* dans un modèle macroéconomique d'équilibre général présentant de nombreuses frictions. Une spécification incorrecte mais capturant de nombreux aspects est de mon point de vue préférable à une spécification correcte mais ne capturant qu'un seul aspect. Ensuite, une littérature importante (Krishnamurthy et Vissing-Jorgensen 2012, Saez et Stantcheva 2016, Michaillat et Saez 2018, Kumhof-Rancièrè et Winant 2015) a utilisé cette spécification pour répliquer des faits stylisés que le modèle standard n'est pas

pas capable de reproduire, par exemple la réponse de la consommation à la "forward guidance" des banques centrales ou certains aspects de la distribution de la richesse, en particulier sa concentration extrême au sommet de la distribution. D'autres explications sont possibles et pourraient être intellectuellement plus satisfaisantes, mais la richesse dans la fonction d'utilité est probablement la manière la plus simple de résoudre ces problèmes, et l'une des plus faciles à intégrer dans un modèle Nouveau Keynesien comportant de nombreuses frictions.

Le papier Ma contribution est d'étudier le cas de non-séparabilité entre la consommation et la richesse dans la fonction d'utilité. A la période t , l'utilité instantanée d'un agent est fonction de la consommation, du loisir et de la richesse assimilée aux obligations détenues par l'agent notées A .

$$U(C_t, L_t, A_{t+1})$$

Je n'effectue aucune hypothèse sur la forme fonctionnelle de U . J'effectue néanmoins l'hypothèse suivante sur les dérivées secondes croisées.

$$U_{CL} = 0$$

$$U_{AL} = 0$$

$$U_{CA} \neq 0$$

La dérivée croisée par rapport à la consommation et à la richesse est différente de 0. Cela signifie que la fonction d'utilité est non-séparable. Si la dérivée croisée est positive, consommation et richesse sont compléments au sens d'Edgeworth, si elle est négative, ils sont substitués. Je suppose que les agents maximisent la somme actualisée de leurs

flux d'utilité

$$\sum_{T=t}^{+\infty} \beta^{T-t} E_t U(C_T, L_T, A_{T+1})$$

Je dérive les conditions du premier ordre du programme. Utilisant l'hypothèse de l'agent représentatif et des anticipations rationnelles, j'agrège ces conditions du premier ordre. Je peux ensuite linéariser ces équations autour des valeurs d'état stationnaire de ces variables agrégés (via un développement de Taylor d'ordre 1). J'obtiens ainsi les équation suivantes pour la consommation et le l'offre de travail

$$\begin{aligned} \theta \eta n_t &= w_t - \sigma c_t \\ (\sigma + \kappa \nu) c_t &= -r r_t + (1 - \kappa) \sigma E_t c_{t+1} \end{aligned}$$

L'équation d'offre de travail n'est pas modifiée par rapport au modèle standard. Le changement intervient dans l'équation de la consommation. Deux nouveau paramètres sont introduits. Le paramètre κ est un facteur d'escompte. C'est la conséquence de la présence de la richesse dans la fonction d'utilité. Le paramètre ν est introduit par la non-séparabilité. Il est en effet égal à $\frac{U_{CA}C}{U_A}$.

Analyser le cas non-séparable (et plus précisément le cas de la complémentarité entre les deux) est un moyen de séparer l'élasticité de substitution intertemporelle, déterminée par $\sigma + \kappa \nu$, et l'effet revenu sur l'offre de travail donné par σ . Dans le modèle standard, les deux sont déterminés par σ et l'un est l'inverse de l'autre. Cette restriction imposée par le modèle standard implique que la quantité de travail désirée d'un agent représentatif a la même élasticité par rapport au salaire réel et par rapport au taux d'intérêt réel. Si on suit l'identification de l'offre de travail proposée par Galí(2011), cela implique que la réponse du taux de chômage à un choc de politique

monétaire est d'une part très large, et d'autre part provient autant de la variation de l'offre de travail que de la variation de la demande de travail. Cela implique aussi que la réponse du taux d'intérêt naturel ² à un choc de demande³ est faible, proche de la réponse effective du taux d'intérêt réel. Par conséquent, il y a une faible différence entre la réponse de la production nationale et la réponse de la production naturelle (i.e la production dans un modèle à prix flexibles). Autrement dit, la réponse de l'output gap est faible. La richesse dans la fonction d'utilité et la non-séparabilité permettent un effet revenu sur l'offre de travail plus faible, tout en maintenant un faible effet de substitution intertemporelle. Cela réduit la réponse du chômage à un choc de politique monétaire et accroît la réponse de l'output gap à un choc de demande.

Ensuite, j'estime le modèle et en particulier les deux paramètres clés κ et ν . Le modèle estimé est un "medium scale model" dans lequel sont introduits des salaires et prix imparfaitement flexibles et partiellement indexés. La politique monétaire est donnée par une règle de Taylor. Les composantes de la demande agrégée autres que la consommation sont supposées exogènes et sont représentés par un "choc" de demande. En plus, du choc monétaire et du choc de demande, j'introduis un choc sur les prix, sur les salaires, sur le facteur d'escompte de l'utilité, sur la désutilité du travail et un choc de productivité. Le modèle est estimé à l'aide des méthodes bayésiennes. Les variables observables sont des données macroéconomiques américaines prises entre 1985 :2 et 2007 :4. Les séries utilisées sont la production nationale, le déflateur du PIB, les salaires réels, le taux d'intérêt effectif de la réserve fédérale, le taux de chômage, les heures travaillées et les dépenses réelles de consommation privée. Mon estimation fait apparaître des valeurs positives et élevées pour κ et ν . L'interprétation que je donne de ce résultat est que les données pointent vers une faible élasticité de substitution intertemporelle et un effet revenu modéré, expliquant la valeur élevée trouvée pour ν . La figure ci-après appuie cette interprétation. Sur cette figure, je représente l'offre de

²Le taux d'intérêt naturel correspond ici au taux d'intérêt d'équilibre d'un modèle à prix flexibles

³Par exemple un choc de consommation publique financée par impôt

travail prédite par le modèle et l'offre de travail effective (identifiée à l'aide des données sur les heures travaillées et le taux de chômage) pour mon modèle (figure de droite) et pour la spécification standard dans laquelle κ et ν fixés à 0 (figure de gauche). Avec la spécification standard, la valeur estimée de σ est élevée, induisant un effet revenu substantiel sur l'offre de travail. L'offre de travail prédite est faiblement corrélée avec l'offre de travail effective et très négativement corrélée avec le résidu (le choc sur la désutilité du travail). En d'autres termes, la spécification retenue induit des mouvements assez large de l'offre de travail prédite par le modèle, mouvements qui n'ont pas de contreparties dans la réalité. Pour compenser cet écart entre la prédiction du modèle et la réalité, il faut introduire un choc artificiellement large et peu vraisemblable sur la désutilité du travail. Avec la spécification retenue dans le papier, la valeur estimée de σ est proche de 1, induisant une offre de travail prédite beaucoup plus proche de l'offre de travail effective et une série beaucoup plus raisonnable pour le choc sur la désutilité.

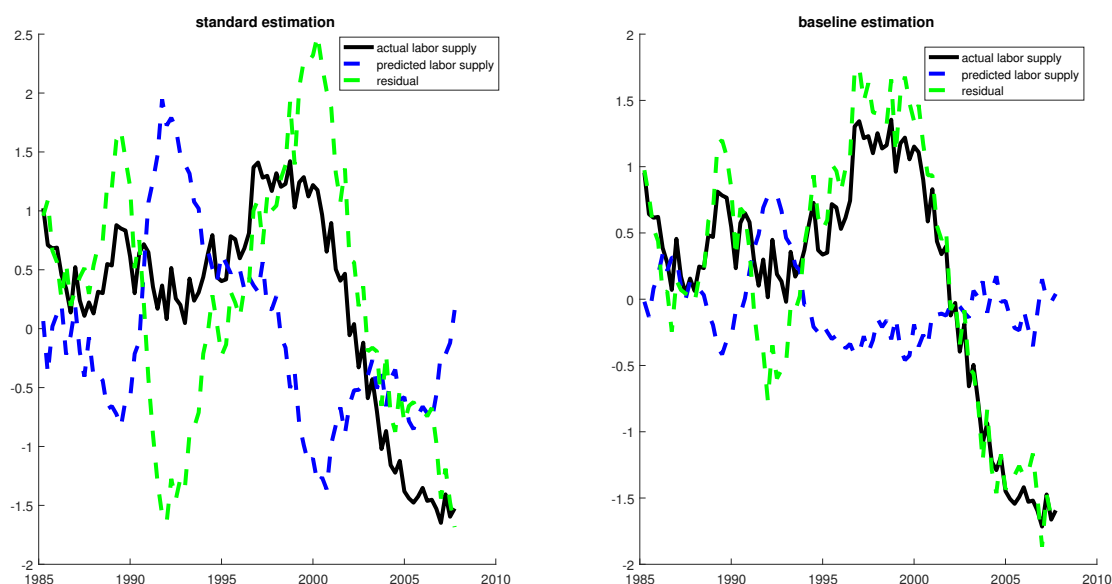


FIGURE 2 – Offre de travail prédite et effective

Les résultats sont robustes à plusieurs estimation alternatives, en particulier à une

prior moins favorable pour κ , à un échantillon de données beaucoup plus étendu dans le temps, allant de 1955 à 2008, à l'introduction d'habitudes de consommation, au remplacement des heures travaillées par la participation au marché du travail comme variable observable, à un facteur d'escompte plus important dans le bloc offre du modèle et à des valeurs calibrées pour les coefficients du coût marginal et du chômage dans les équations d'inflation et de salaires.

Chapitre 3

Le contexte Le troisième chapitre est consacré au comportement des entreprises et en particulier à leurs décisions d'investissement.

Selon la théorie néoclassique du capital, le stock de capital est fixé de manière à égaliser la productivité marginale du capital et le coût d'usage du capital dont une composante essentielle est le taux d'intérêt. Formellement, si les firmes maximisent leurs profits $f(K) - rK$ où f est la fonction de production, K le stock de capital et r le coût d'usage, la condition du premier ordre est $f'(k) = r$. Le stock de capital optimal est déduit de cette expression et l'investissement est la différence entre le stock de capital actuel et le stock optimal de la prochaine période. Ce modèle est très simple. Il est souvent complété par des hypothèses supplémentaires pour être plus proche des données. Par exemple, un coût d'ajustement est souvent introduit.

La théorie néoclassique de l'investissement suppose que les firmes ne sont soumises à aucune contrainte de financement. Un certain nombre d'évidences contredisent cette idée. Par exemple, une corrélation significative est observée dans les données d'entreprises entre l'investissement et les variations de trésorerie. L'interprétation de cette corrélation est difficile en raison d'une possible relation entre ces cash flows et les opportunités d'investissement, relation susceptible d'induire un problème d'endogénéité dans les estimations. Toutefois, les frictions financières forment une autre explication

à cette corrélation. De plus, la sensibilité de l'investissement au cash flow est plus forte dans les firmes connaissant des difficultés financières (Fazzari, Hubbard et Petersen 1988), ce qui peut accréditer l'idée que les contraintes financières jouent un rôle dans cette corrélation (ce point a toutefois été contesté par Kaplan et Zingales 1997). D'autres évidences ont été apportées par Gilchrist et Zakrajsek (2012). Ils montrent que certaines variations dans le prix des obligations émises par les entreprises ne peuvent s'expliquer ni par des variations de prix des obligations sans risques ni par des variations du risque de défaut lui-même. Ils suggèrent que des contraintes financières plus ou moins fortes pourraient expliquer ces variations. D'autre part, Étudiant les canaux de transmission de la politique monétaire, Bernanke et Gertler (1995) soutiennent que le canal du crédit et le canal du bilan des entreprises est plus important que le canal du coût d'usage du capital, suggérant ainsi que les frictions financières jouent un rôle important pour la politique monétaire.

Ces différents éléments ont conduit les économistes à développer des modèles de frictions financières et à les intégrer dans des modèles macroéconomiques plus larges. Le développement de cette littérature a également "bénéficié" de la crise financière de 2008. La friction la plus populaire dans les modèles macroéconomiques appliqués est le modèle de "costly state verification" introduit par Townsend (1979). Dans ce modèle, les prêteurs ne peuvent observer le résultat d'un projet d'investissement que si ils supportent un coût d'audit. Si ils ne paient pas ce coût d'audit, un emprunteur pourrait sous estimer la valeur de ses actifs, se déclarer en faillite et partir avec ces mêmes actifs. Pour éviter cela, les prêteurs paient le coût d'audit, créant ainsi un écart entre le rendement de l'investissement pour l'emprunteur et le rendement pour le prêteur. Bernanke, Gertler et Gilchrist (1999) introduisent cette friction dans un modèle Nouveau Keynesien de taille moyenne, calibré, et montrent que cela amplifie la réponse de l'investissement et de la production nationale à un choc monétaire. Christiano, Motto et Rostagno (2014) estiment un modèle similaire et suggèrent que l'accroissement de la volatilité de la valeur de l'actif des firmes pourrait être l'un des principaux facteurs

explicatifs du cycle économique.

Un aspect troublant de cette littérature est la capacité des emprunteurs à détourner pour leur bénéfice propre une fraction particulièrement importante des actifs des firmes. Si les emprunteurs sont identifiés comme les actionnaires, ce qui est le cas dans les modèles macroéconomiques appliqués, cela revient à soutenir l'idée que les actionnaires des entreprises cotées peuvent déclarer leurs entreprises en faillite et détourner ses actifs sans rencontrer la moindre difficulté pratique. Cette idée semble difficile à soutenir dans des économies développées qui se caractérisent par des institutions relativement fonctionnelles. Malgré le succès de ces modèles, par exemple pour expliquer l'accroissement de l'écart de taux entre les obligations d'entreprises et les obligations d'état au moment de la crise financière, il semble nécessaire de continuer à explorer des frictions alternatives et leurs conséquences.

Le papier Ma contribution relève de la littérature étudiant la sélection adverse sur le marché du capital. La sélection adverse sur le marché du crédit a été introduite par Stiglitz et Weiss (1981). Un papier récent par Kurlat (2015) montre que la sélection adverse fournit des micro-fondements pour la contrainte de revente introduite par Kiyotaki et Moore (2012). Je considère un problème de sélection adverse légèrement différent de celui de Stiglitz and Weiss en horizon infinie.

Ma contribution est de fournir une formule explicite et simple pour l'investissement des entreprises. L'investissement est une fonction linéaire des profits conservés par les firmes. J'intègre la relation dans un modèle dynamique et explore certaines implications, en particulier pour la sensibilité de l'investissement au coût des emprunts.

J'étudie l'équilibre partiel sur un marché du capital dans lequel les entreprises se concurrencent pour emprunter des fonds, et ainsi financer leurs investissements. Il y a de nombreux prêteurs en situation de concurrence sur ce marché. Les prêteurs sont capables de diversifier leurs placements et ne sont donc pas sensibles au risque

idiosyncratique supporté par les entreprises. De plus, ils ont accès à un actif sans risque produisant un rendement r . Les entreprises détiennent une information privée sur leur niveau de risque. Elles tentent de signaler ce niveau de risque en limitant leurs emprunts à une fraction des bénéfices mis en réserve. La firme J produit πK^j

Il y a deux types d'entreprises dans l'économie : les bonnes et les mauvaises entreprises qui diffèrent par leur niveau de risque. A chaque période, une bonne entreprise a une probabilité κ de devenir une mauvaise entreprise. Avec une probabilité $1 - \kappa$, elle reste bonne. Une mauvaise entreprise produit toujours πK^j mais a une probabilité $1 - \lambda$ de disparaître à la prochaine période. Quand elle disparaît, une entreprise ne produit plus rien et son stock de capital est sans valeur. Ce processus est résumé dans la figure suivante.

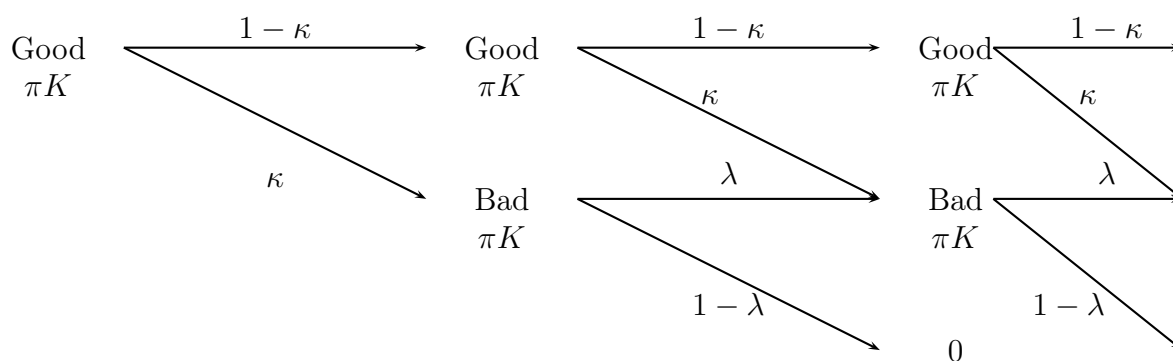


FIGURE 3 – Probabilités de transition pour les entreprises

Les entreprises maximisent la somme actualisée des dividendes versées. La fonction de valeur s'écrit

$$V(K, B) = \max_{K', B', I, S, d} d + \frac{1}{1 + \gamma} EV'(K', B')$$

w.r.t $K' = K + I$

$$B' = B + r^e(I - S)$$

$$\pi K - B = d + S$$

$$\pi K - B \geq d \geq (1 - s)(\pi K - B)$$

$$I \geq 0$$

La première contrainte décrit l'évolution du stock de capital. La seconde contrainte décrit l'évolution de la dette. La dette est supposée à maturité infinie. Seuls les intérêts sont payés à chaque période. L'équation suivante répartit le revenu de la firme entre les dividendes et les bénéfices mis en réserve. Les dividendes ne peuvent être supérieures au revenu et ne peuvent être inférieures à une certaine fraction du revenu $(1 - s)(\pi K - B)$, où s est un paramètre exogène compris entre 0 et 1. La dernière équation empêche la firme de désinvestir.

La résolution du problème de la firme en horizon infinie implique de définir les fonctions de valeur des bonnes et des mauvaises firmes. Le problème est que la fonction de valeur va dépendre du taux d'intérêt supporté par les firmes et donc des croyances présentes et *futures* des prêteurs sur le type auquel appartient l'entreprise. Pour résoudre le programme, j'effectue les trois conjectures suivantes (i) Les bonnes entreprises peuvent signaler leur type. Pour se signaler, elles contraignent leur ratio investissement/bénéfices mis en réserve à être inférieur à une certaine valeur. (ii) Bonnes et mauvaises entreprises calculent leur fonction de valeur à la période t sous l'hypothèse qu'elles enverront le signal correct à partir de la période suivante. (iii) Les fonctions de valeur existent⁴. Je caractérise l'équilibre sous ces trois hypothèses, et je montre a

⁴Le lecteur attentif notera que cette hypothèse est placée en troisième

posteriori qu'elles sont bien vérifiées.

Je caractérise l'équilibre de la manière suivante

1. Pour un niveau de capital donné K et un niveau donné de "dette" B , une bonne entreprise met en réserve ses bénéfices jusqu'au seuil minimal de dividendes $d^G = (1 - s)(\pi K - B)$ et $S^G = s(\pi K - B)$ et emprunte autant que possible en respectant la contrainte d'incitation $I^G = \frac{\gamma^B - r^G}{\pi - r^G} S^G$.
2. Pour un niveau donné de stock de capital K et un niveau donné de dette B , une mauvaise entreprise ne met aucun bénéfice en réserve, distribue tout son revenu comme dividende $d^B = \pi K - B$ et $S^B = 0$ et n'investit pas $I^B = 0$.
3. Le taux d'intérêt supporté par les mauvaises entreprises $r^B = \frac{1+r-\lambda}{\lambda}$ et le taux supporté sur les bonnes entreprises est $r^G = \frac{(r+\kappa)(1+r-\lambda)}{(1-\kappa)(1+r-\lambda)+\kappa(r+\kappa)}$.

A l'équilibre, seuls les bonnes entreprises investissent. L'équation de l'investissement est donc

$$I = I^G = \frac{\gamma^B - r^G}{\pi - r^G} S^G = \frac{\gamma^B - r^G}{\pi - r^G} s(\pi K^G - B^G)$$

où K^G et B^G représentent respectivement le stock de capital agrégé des mauvaises entreprises et en début de période et la "dette" agrégée en début de période.

J'intègre maintenant cet équilibre dans un modèle dynamique. J'ajoute un indice t à toutes les variables. Le système dynamique ainsi obtenu se caractérise par trois équations

$$(1 + g_{t+1})b_{t+1} = (1 - \kappa)(b_t + r_t^G(\psi_t - 1)s(\pi - b_t))$$

$$1 + g_{t+1} = (1 - \kappa)(1 + i_t)$$

$$i_t = \psi_t s(\pi - b_t)$$

dans lequel g_{t+1} est le taux de croissance du stock de capital des bonnes entreprises $\frac{K_{t+1}^G}{K_t^G}$, i_t est le ratio investissement sur capital des bonnes entreprises $i_t = \frac{I_t^G}{K_t^G}$, et b_t est le ratio dette sur capital des bonnes entreprises $\frac{B_t^G}{K_t^G}$.

Ce système peut être transformé assez facilement en une seule suite définie par l'équation

$$b_{t+1} = \frac{b_t + r^G(\psi_t - 1)s(\pi - b_t)}{1 + \psi_t s(\pi - b_t)}$$

Cette suite est définie par une équation non linéaire et admet deux points fixes, l'un stable, l'autre instable. Le ratio dette sur capital va converger vers la valeur

$$b^{**} = \frac{r^G(\psi - 1)}{\psi}$$

Il est facile d'en déduire les taux d'investissement et de croissance de long terme

$$i^{**} = s\psi\pi - sr^G(\psi - 1)$$

$$g^{**} \simeq s\psi\pi - sr^G(\psi - 1) - \kappa$$

Après une légère transformation, une expression encore plus simple est obtenue

$$i^{**} = s\gamma^B$$

$$g^{**} \simeq s\gamma^B - \kappa$$

J'effectue ensuite plusieurs exercices de dynamique comparative. En particulier, j'étudie la réponse de l'investissement à un choc sur le coût d'emprunt r . Un résultat intéressant est que l'investissement diminue en impact après une baisse du coût d'em-

prunt. En effet, une telle baisse rend plus intéressant pour les mauvaises entreprises d'imiter les bonnes entreprises car le rendement de l'investissement augmente. Pour continuer à se différencier des mauvaises entreprises, les bonnes entreprises doivent réduire leur ratio investissement sur bénéfices mis en réserve. Ce résultat est contre-factuel. Je ne pense cependant pas qu'il s'agisse d'un argument définitif en défaveur du modèle. Le modèle doit être approfondi avant d'être confronté aux données. Plusieurs pistes sont envisageables : explorer le modèle avec des maturités plus courtes pour la dette, combiner équilibres séparateurs et mélangeants (ce qui nécessiterait toutefois quelques hypothèses supplémentaires), ou encore endogénéiser le niveau de dividendes distribués.

Essays on Macroeconomic Theory

Abstract

This thesis is made of three independent chapters. The first chapter contributes to the literature on expectations. I argue that they may learn a misspecified model instead of learning the rational expectation model. I consider a simple economy with two types of agents. Rational learners learn the true model of the economy whereas consistent learners learn an autoregressive model. I show that a long run equilibrium exists in which consistent learners dominate. Simulations show that the economy may converge towards it. The second chapter deals with the intertemporal choice. I consider a model with wealth in the utility. I study the case of nonseparability. This disentangles between the income effect on labor supply and the intertemporal substitution effect. I derive several implications for economic policy. Then, I estimate the two new parameters introduced in the paper. I find large and positive values for both. The third chapter builds a model of corporate investment under adverse selection. My contribution is to provide a tractable model easy to embed into a macroeconomic model. Borrowers differs by the riskiness of their investment project like in Stiglitz and Weiss (1981). They have infinite horizon and signal their type by borrowing a fraction of their retained earnings. I get an analytic solution for the incentive constraint. I integrate the relation into a dynamic model and derive some implications.

Keywords : Expectations, Intertemporal Substitution, labor supply, corporate investment.

Essais sur la théorie macroéconomique

Résumé

Cette thèse est composée de trois chapitres indépendants. Le premier chapitre concerne la formation des anticipations. Je montre que les agents sont susceptibles d'utiliser un modèle mal spécifié plutôt que le "vrai" modèle de l'économie. Je considère une économie simple avec deux types d'agents. Les agents rationnels apprennent la solution à anticipation rationnelle tandis que les agents "cohérents" utilisent un modèle auto-régressif. Je montre qu'un équilibre de long terme dans lequel les agents cohérents sont dominants existe. Des simulations montrent que l'économie peut converger vers cet équilibre. Le deuxième chapitre concerne le choix intertemporel. Je considère un modèle dans lequel la richesse entre dans l'utilité. J'étudie le cas non-séparable, séparant ainsi l'effet revenu sur l'offre de travail de l'effet de substitution intertemporelle. Je déduis des implications pour la politique économique, puis, j'estime les deux paramètres introduits par cette spécification de l'utilité. Je trouve des valeurs positives et élevées pour les deux. Le troisième chapitre présente un modèle d'investissement en présence de sélection adverse. Ma contribution est de fournir une solution simple, facile à intégrer dans un modèle macroéconomique. Les emprunteurs diffèrent par le risque de leur projet d'investissement comme dans Stiglitz et Weiss (1981). Ils signalent le risque de leur projet en empruntant une fraction des bénéfices mis en réserve. J'obtiens une solution analytique pour la contrainte d'incitation. Je l'intègre dans un modèle dynamique et déduis certaines implications.

Keywords : Anticipations, Substitution Intertemporelle, Offre de travail, investissement.