Charles University

Faculty of Social Sciences
Institute of Economic Studies

MASTER'S THESIS

Rational Inattention in DSGE Model

Author: **Bc. David Vostřák**

Supervisor: **PhDr. Simona Malovaná**

Academic Year: **2017/2018**

Declaration of Authorship

The author hereby declares that he compiled this thesis independently; using only the listed resources and literature, and the thesis has not been used to obtain a different or the same degree.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, May 10, 2018

Signature

Acknowledgments

The author is grateful especially to PhDr.Simona Malovaná for her valuable comments and patience. The usual caveat applies.

Abstract

A great amount of available information over the internet makes it impossible for anyone to process it all. In this thesis, we use the rational inattention theory to see how the perceived signals about the exogenous variables would change under different levels of information capacity. Those signals are then applied in the New Keynesian model and corresponding impulse responses are compared with the case of unlimited attention. We found that for some autoregressive processes the differences from the perfect attention case are not very profound while for others the results vary considerably.

JEL Classification	E1, E2, , E3, E7
Keywords	rational inattention, DSGE, New Keynesian, shocks

Author's e-mail	vostrakd@seznam.cz
Supervisor's e-mail	simona.malovana@gmail.com

Abstrakt

Velké množství dostupných informací na internetu neumožňuje nikomu je všechny zpracovat. V této práci používáme teorii racionální nepozornosti, abychom zjistili, jak se liší vnímané signály ohledně exogenních šoků pro různé úrovně informační kapacity. Tyto signály jsou poté aplikovány v modelu nové keynesiánské makroekonomie a příslušné odezvy jsou porovnány s případem neomezené pozornosti. Zjistili jsme, že pro některé autoregresivní procesy není rozdíl s případem neomezené pozornosti příliš velký, zatímco pro jiné procesy jsou rozdíly značné.

Klasifikace	E1, E2, , E3, E7
Klíčová slova	racionální nepozornost, DSGE, nová keynesiánská makroekonomie, šoky

E-mail autora	vostrakd@seznam.cz
E-mail vedoucího práce	simona.malovana@gmail.com

Contents

List of Figures

Master's Thesis Proposal

Institute of Economic Studies
Faculty of Social Sciences
Charles University in Prague

Author:	**Bc. David Vostřák**	Supervisor:	PhDr. Simona Malovaná
E-mail:	vostrakd@seznam.cz	E-mail:	simona.malovana@cnb.cz
Phone:	732228645	Phone:	224 414 328
Specialization:	ET	Defense Planned:	June 2017

Notes: The proposal should be 2-3 pages long. Save it as "yoursurname_proposal.doc" and send it to mejstrik@fsv.cuni.cz, tomas.havranek@fsv.cuni.cz, and zuzana.havrankova@fsv.cuni.cz. Subject of the e-mail must be: "JEM001 Proposal (Yoursurname)".

Proposed Topic:

Rational inattention in DSGE model

Motivation:

In the world of unprecedented information available to general public, it seems unavoidable that decision makers have to carefully choose on which information to allocate their bounded time. Even the firms have only limited power regarding how much attention they are able to devote obtaining and assessing all the information available to them. As Wiederholt (2010) in his work about rational inattention puts it, economists should focus on modeling decision maker problem of optimizing attention as one of many other scare resources.

In the theory of rational inattention, the decision makers are not aware about true states of some variables, but instead, see them as random ones. From this perspective, for a decision maker, this represents huge uncertainty, which she is willing to decrease by receiving some signals about the true state of these variables. The signals nevertheless have some inherent noise and moreover, the decision maker can receive only limited amount of them, so the optimization technique are needed to solve the problem. I will add this new paradigm into the benchmark New Keynesian dynamic stochastic general equilibrium (NK DSGE) model, where my mainly focus will be on the monetary policy implications.

The literature on rational inattention in DSGE models is still modest but expanding rapidly. The first studies made by Sims (2003), Sims (2006) and Luo (2008) were focused on consumption-saving problem in rational inattention framework with an assumption about constant real interest rates. Mackowiak and Wiederholt (2009) developed price setting equilibrium model with rational inattention by firms, followed by a study of Paciello (2010), where general equilibrium model with rationally inattentive firms was solved. The full DSGE model with rational inattention was then developed by Maćkowiak and Wiederholt (2015).

Hypotheses:

1. Hypothesis #1: DSGE models with rational inattention by households and firms can better explain inflation inertia than standard DSGE models with Calvo price setting used for policy analysis.
2. Hypothesis #2: Some of the stylized facts about output–inflation dynamics can be explained by introducing the rational inattention into DSGE models
3. Hypothesis #3: The way how signals are made by policy makers is important for short-run dynamics of key economic variables

Methodology:

In the thesis, I examine the optimization problem a decision maker must face when she is limited by resource constraint about how much signals she can receive. This feature then will be introduced into the DSGE modeling framework, where the signals will be about the true state of the economy and the decision makers will be households and firms. The constructed DSGE model will be for a small open economy and calibrated for the Czech Republic. The hypothesis will be examined via impulse responses for various shocks and compared with the benchmark NK model. The data and calibrated values will be collected from the Czech National Bank.

Expected Contribution:

Rational inattention modeling is a new and sharply evolving area of research with some very promising results, so introducing the rational inattention into DSGE model could explain some of the stylized facts about output-inflation dynamics.

Outline:

1. Research Summary
2. Rational Inattention Optimization Problem
3. Benchmark NK model
4. Introducing Rational Inattention into the Benchmark Model
5. Comparison of both models
6. Conclusion

Core Bibliography:

Sims, Christopher A. (2003). "Implications of Rational Inattention." *Journal of Monetary Economics* 50 (3): 665–90.

Mackowiak, Bartosz, and Mirko Wiederholt. (2009). "Optimal Sticky Prices under Rational Inattention. *American Economic Review* 99 (3):769–803.

Sims, Christopher A. (2010). Rational inattention and monetary economics. *Handbook of Monetary Economics*, *3*, 155-181.

Maćkowiak B, Wiederholt M. (2015). Business Cycle Dynamics under Rational Inattention*. The Review of Economic Studies.

Hellwig, Christian, and Laura Veldkamp (2009). "Knowing what others know: Coordination motives in information acquisition." *The Review of Economic Studies* 76.1, 223-251.

Matějka, F. (2015). Rigid pricing and rationally inattentive consumer. *Journal of Economic Theory*, *158*, 656-678.

Woodford, Michael. (2009) "Information-constrained state-dependent pricing." *Journal of Monetary Economics* 56: S100-S124.

Caplin, Andrew, and Mark Dean. (2016). Revealed preference, rational inattention, and costly information acquisition. *The American Economic Review ,* 105(7): 2183-2203.

Paciello, Luigi (2010). "Monetary Policy Activism and Price Responsiveness to Aggregate Shocks under Rational Inattention." Discussion paper, Einaudi Institute for Economics and Finance.

Author **Supervisor**

1 Introduction

In today´s world, a lot of information can be freely accessed over the internet and still almost everyone reacts to them rather sporadically. It is hard to imagine a world where everyone would be aware of for example daily movements of food prices and reacts to it accordingly when doing for example daily shopping. It seems more natural to assume that a lot of people have only rough awareness about the economic state of the world because not everyone is paying close attention to all economic charts and figures one can find over the internet on the daily basis. It is therefore natural to expect that people would not react to some new economic figures immediately. Because we have only limited amount of cognitive power, one has to be very picky about where to put his eye on. As Sims (2010) puts it, people have often more important things to think about than to look at some economic variables. Nevertheless, it is not only the people who are limited in their attention, the same logic of course applies also to the firms. Even if our processing capacity is scarce, we can still to large extent decide how uninformed we want to be about some variables of interest. We could see the problem as a task where one has the power of limited variance reduction but has to decide on which noises she wants to put in. This feature is known in the information theory as so-called "water-filling", named after the example where we have a bucket with a limited amount of water and are filling several empty bins representing variances of some noises. To make some real-world example, let say that we are interested in oil prices. Then there would be a difference if we would be watching them daily, monthly or at all. Even the large financial houses with a vast amount of human labour are unable to pay attention to all information relevant to them, so they have to make a trade-off on where to put their attention. One would naturally expect that a bank would assign more people to work on a transaction more important to the bank than on the less important one.

It is empirically well-documented fact that prices show some stickiness which is pervasive in the data. A lot of prices are changing seldomly even within a year, so the models relying on adjustment costs should have these costs practically everywhere. The sluggishness of the many macroeconomic variables is hard to be explained just by adjustment costs, so Sims (1998) suggested to replace them by the information processing constraints. Unlike in the sticky prices models, where adjustment costs can be seen as an external source of stickiness, in the rational inattention models, the agents are sticky in the information they use rather than in the prices they are facing. The rational inattention does not make the sticky prices obsolete since in practice we see the prices change rather sporadically over the year, but rather endogenous them as can be seen in Matějka and Sims (2009). Rational inattention models try to enlight possible sources of persistent inertia found in the empirical literature on macroeconomic variables. In this model environment, the agents are continually "learning" the state of the world, so they cannot react immediately. As we will see in the next chapter, there were some attempts to replace the sticky prices models by the sticky information ones, for example by introducing some costs on the acquisition of relevant information, but they make it hard to explain why it should be the driving force in decision-making problems. The sticky information models see the information as one of the many scarce resources, so there are according to the theory some frictions in obtaining them. The story behind the rational inattention models is exactly opposite, there is so much information available at no cost that it is impossible to grasp it all.

As Sims (2010) argues, it is hard to defend dynamically optimizing rational agent models, when an agent does not utilize all free information available to him. We, therefore, need to introduce into models some natural friction regarding the information flow. It is what is made in rational inattention models, where an agent has only a limited amount of attention, so he is facing some "capacity constraints". Only if there would be some surprising new data, then we could expect some swift reaction. As an example made by Sims (2003), imagine a large headliner in newspapers announcing some new shocking data, then a lot of people would definitely pay more attention toward it. Rational inattention models are capable to some extent explain this paradoxically sluggish and at the same time erratic behaviour. We can take another real-life example from the textbook written by Bodie et al. (2014). Imagine a situation where two friends are arguing about the future exchange rate of the euro against the Czech crown 1-year ahead. One of the friends strongly believes that the exchange rate will be higher than 25 Czech crowns for 1 euro one year ahead and the other friend disagrees and says that the exchange rate will surely be lower. They make a bet, which is a reasonable choice in this case since the expected value for both is positive as probability p for the exchange rate being higher than 25 is >0.5 for one and <0.5 for the other. This situation should not be surprising since even the experts in the field have sometimes different opinions about the future rates. What is striking is that this situation could practically happen even if the bet would be about the current exchange rate. If the last time they both check the exchange rates were 1 month ago and their opinions about the development afterward differ, then we would have the same situation as before and therefore it would be reasonable for both to make a bet. (naturally under the condition that neither one is allowed to look it up on the internet or in the newspapers before the bet would take place). Using the terminology we will further develop in the next chapters, the signal about the current exchange rate was different since they both used only limited attention toward it (in this highly hypothetical case they used forecast one month ahead as their signals). If this bet would be theoretically recreated many times and the winning rate would be largely disproportionate, then we could say the friend who would be winning the bet more often is either more informed or has more luck. We will focus in this work on the first case.

The work is structured as follows. The first introductory chapter served as a basic motivation of usefulness of rational inattention for the purpose of this work. It is followed by a chapter devoted to a literature review, where more profound motivation for rational inattention will be provided. We begin the chapter from Lucas (1976) and his critique of policy evaluation to argue later that rational inattention approach can eliminate much of this critique. The following text is then devoted to the heated debate between economists at the time about °atheoretical° school of economics versus the rational expectation one. Next, we swiftly move to a comparison of the three prevalent approaches in macroeconomic models, mentioned briefly earlier in this chapter, to deepen the discussion about their strengths and weaknesses. It will serve as an important step toward motivation for use of rational inattention later in this work. We then cover the application of inattentive agent in various fields in economics used in the literature. In the next chapter in this work, as the rational inattention is closely related to the information theory, some basic notions known in the information theory such as the definition of entropy, data transmission maximum, data comprehension limit, mutual information and some fundamental relations between them will be depicted. It will be useful in the next chapter, where the rational inattention filter, as was developed by Maćkowiak et al. (2017), will be presented. It will be accompanied by some of the main findings from the same paper since they will be very useful for the interpretation of the results in this work. Following the chapter about rational inattention filter, we develop the New Keynesian DSGE model with sticky prices, as it will be our baseline model of the economy in which we compare different shocks using both the perfect attention case and the rational inattention one. The model will be calibrated using values mainly from the model used by the Ministry of Finance of the Czech Republic. Shocks used in this work will have autoregressive structure and we will use two different AR(2) processes and one ARMA(2,1) for two different values of information constraints and compare it to the perfect attention case. The work is ended with concluding remarks.

2 Literature Review

The literature on rational inattention is still rather scarce, but the number of economic papers written about the topic is increasing each year. In order to do a proper literature review, we will divide this section into two main blocks, particularly we make such division that the first part of the section will be devoted to the evolution of the idea of rational inattention and its development and the second part will be rather more about present-day state of the literature about the topic. Such division seems to be a reasonable one given a long-lasting discussion about a proper modeling structure for the economic policy analysis.

We begin this paragraph by Robert Lucas's (1976) famous critique of econometric policy evaluation, where he concerned proper ways of interpreting the policy evaluations using a vector autoregression. When Lucas wrote his seminal work on this topic, there was increasing usage of econometric techniques for economic policy recommendations at that time and Lucas expresses his worries about the usefulness of such approach for long-term policy evaluations. Although he recognized a huge success of forecasting performance of econometric models for policy evaluations, he argued that such techniques are useful in case of short-term forecasting only and that in fact long-term forecasting is ill-modeled by lacking a rational expectation part on the side of economic agents. His main concerns were about a difference between the true state of an estimated structure before some policy change takes place, and the true structure afterward, where he saw a potential difference between these two. To state formally his thoughts, we can think of an economy motion described by a difference equation of the form $y_{t+1} = f(y_t, x_t, \varepsilon_t)$, where one can estimates its vector of parameters θ from regressing $f(y, x, \varepsilon) = F(y, x, \theta, \varepsilon)$. As he argued, when estimating the parametres θ, we are treating them as constant over time, which genarally does not have to be true for all values of possible policy actions x_t. Instead of this approach, we should according to Lucas (1976) treated vector of parameters θ as having its own dynamic structure, for example by letting it follow a random walk process of the form $\theta_{t+1} = \theta_t + \vartheta_t$. Possible way how we should make an inference about the system would be to regress the system of equations in the form $x_t = G(y_t, \lambda, \eta_t)$ and $y_{t+1} = F(y_t, x_t, \theta(\lambda), \varepsilon_t)$ and seeing the econometric problem as the one of having to estimate the function $\theta(\lambda)$. According to Lucas, the agents are aware of the policy changes which are set by policy makers upon them and behave correspondingly. That line of thought gave rise to the school of economics known as rational expectation models.

Lucas's critique had made a huge impact on the modeling practices at the time and raised heated debate over usage of econometric technique as a proper tool for policy evaluation. As many papers were published defending Lucas's critique, the same was true for advocates of the conventional use of simultaneous equations. Sims (1982) argued that rather then rational expectation should make econometric analysis obsolete, we should view it more as a cautionary footnote. According to Sims (1982), policy shifts in the economic system are very often short-lived and in many cases not known in advance, and moreover, many policy announcements are not likely to be believed by the general public, which creates some natural obstacles in the theory of rational expectations. As Sims (1982) noted, if we are up to make a quantitative analysis of policy actions, we must begin with a model which give us outcome for each possible policy action. Policymaker then chooses such action which maximizes some objective function, usually in the form of conditional distribution. A typical case in macroeconomics is to choose c_t in order to maximize y_{t+1} , or in other cases, to find the best conditional probality distribution of y_{t+1}. In control theory problems such as this y_{t+1} frequently depends not only on current the state of variables, but on the future ones as well. Therefore a policy maker has to decide on c_t bearing in mind the contingency plan for all future c_{t+s} as well as function of the information set available at that time. We can see in some sense this control theory problem to be a static one given that once we set c_t, we know exactly how to proceed in future for all possible outcomes. That is as Sims (1982) pointed out would not be the case in practise, since policy makers often make decisions just a few quarters ahead and often reconsider their plans frequently. It seems unlikely that policy makers are in favour of optimal policy as a deterministic function of actions given information set available in that particular time. Even for an agent having the same information set as a policy maker, there is still a large uncertainty for an agent about which policy would take place. This uncertainty creates huge demands on information processing on the part of an agent, since in a sense policy setting is partially random from the perspective of an agent.

It also worth noted a commentary on the issue made by Sargent (1984) and later by Cooley and Leroy (1985). The first mentioned trying to narrow the gap between rational expectation econometrists endorsed by Lucas (1976) and the other one he called °atheoretical° or uninterpreted approach supported by Sims (1982). What rational expectation theory suggested was that once different policy action is set, a purposeful agent dynamically resets its expectations so that a former dynamic structure in an economy no longer exists. In practice, it means that when policy maker is evaluating different policies using vector autoregressive equations from the past, he cannot simply extrapolate the resulting equations, since they do not have to be the same, or put it differently, unaffected, by the altered policy being in place. This represented a huge problem from a control theory perspective because the main goal for each policy maker is ultimately quantifying set of different policy actions and its forecasted implications in the years beyond the sample period in order to set the optimal rules of the regime by the government. From this point of view, the government would be guided by different principles during the sample period and during the future period, which would meant that while during the sample period there exists an asymmetry between a government and an agent, this would be removed in the future period when we expect both government and agents to act purposely. This was viewed by Sims (1982) to be a major problem with Lucas's theory. According to Sims (1982) and later formulated by Sargent (1984), we should see a model economy to be rather a dynamic game where different players represented by institutions and agents take place and where there is no clear certainty about how will each player behave. In this environment, there is always present some level of uncertainty and this uncertainty is presumably very similar in both sample period as it is in the future period. These ideas then can be found in Sims (1985), where the more formal critique of Lucas's (1976) seminal work is provided.

A huge step forward toward the rational inattention models came with the academic paper called Stickiness made by Sims (1998). All macroeconomists at the time were aware about a sluggish and pervasive nature of real and nominal inertia in responses to monetary policy, although the underlying story behind them was subject to heated debates among academists. The notion that prices and wages are not changing continuously is well-established fact, but the source of such slow and irregular changes remains unclear up till today. We can track down the realization of sluggish price and wage movements back to Keynes, which was his major point in defending a strong and long-lasting effect of fiscal and monetary policy on real economic activity, although he did not provide a microeconomic foundation for this fact. Sims (1998) outlined the idea of stickiness as a consequence of limited information processing capacity. In markets, there are a great variety of products and assessment of average price level and their relevance to each individual consumer and producer is a quite demanding task. According to Sims (1998), people have only limited information capacity and moreover, they are not willing to pay all their intellectual resources on data-gathering and analyzing for their economic decision making. It makes them pay only limited attention to the economic variables of interest, which in turn explains rather sporadical changes in prices and wages. Sims (1998) reasoning was quite simple, people have just better things to do than constantly updating price levels relevant to their economic activity. This explanation is quite compelling since it can clarify why would agents do not make use of information, which is generally available and for free, such as for example monthly updates of price levels provided by state institutions or by newspapers and over the internet. They simply choose not pay attention to it, which is a different story than one proposed by Lucas (1973), where he saw the stickiness as part of signal-extraction problem induced by technical barriers on the acquisition of relevant information.

The notion of rational inattention was first used in Sims (2003), where Sims proposed a theoretical framework for stickiness found in many empirical studies. At the time when Sims (2003) wrote his paper, there was increasing number of academic papers challenging the postulate about "homo economicus", mainly coming from field of behavioural economics literature (see for example textbook on the topic written by Ryan and Deci (1985), or other similar studies made by Laibson (1997), Thaler (2000), Carrillo and Mariotti (2000) and Brocas and Carrillo (2003)). With an exception of Laibson (1997) and Carrillo and Mariotti (2000), who provided a great formal description on the information acquisition and strategic ignorance from the microeconomic theory standpoint, the problem with the behavioral economics literature was that despite the great insight on economic behaviour from the psychological perspective, there was no theory which could be tested against already existing models used by policy makers in analysing outcomes for different policy actions. Among those theories, which attempted to explain the stickiness using formalized microeconomic theory, we should name already mentioned Lucas (1973), who tried to clarify the sluggish responses of price setters using information delay. Sims (2003) argued that such approach is hardly defendable given easy access to relevant information with almost no delay, which is even more true nowadays thanks to the incredible boom in information technology seen in every corner of our lives. Another branch of academic papers saw a visible stickiness as part of adjustment costs imposed on price setters, and those costs function as an incentive to not to reset prices or wages continuosly, because it would exceed the benefits of optimal price adjustment (see for example Sheshinski and Weiss (1977), Caplin and Spulber (1987), or more recent studies made by Golosov and Lucas (2007), Wolman (2007) or article made by Levy and Smets (2010), briefly reviewing other empirical studies on the topic made in the EU area). Although there are apparently some adjustment costs in play for decision making of price setters, there are doubts that these costs are dominant forces behind sluggish responses of price adjustment processes. According to Sims (2003), the underlying process is not because of information delay as suggested by Lucas (1973), but due to the information-processing constraint. By this theory, people have only limited channel capacity they are willing to devote to economic realities, which in turn creates less frequent responses in behaviour on relative changes in price levels. In order to model information processing constraints in a rigorous way, Sims (2003) uses theoretical framework known from information theory, which is field usually seen as not so closely related to economic theory. Although the concept and notions which Sims (2003) uses in his seminal work are very well known for decades among information theorists (in fact some of the notions can be tracked down back to invention of telegraph in the 1830s), there is still huge novelty in his paper by incorporating them as crucial part of macroeconomic models.

The competiting idea for the rational inattention models, which is using the information constraints in macroeconomic models as well, is the idea of sticky information, promoted by Mankiw and Reis (2002). The reasoning behind the sticky information can be seen as very similar to the rational inattention, nevertheless there are some major differences worth mentioning. In sticky prices model, the producers know everything about the model and are aware about the optimal price for them in each period, but unfortunately, some of them are unable to reset their prices in the current period. Contrary to that, in sticky information models, all producers can reset their prices, however, some of them are unaware about their optimal prices because the information did not reach them. Mankiw and Reis (2002) argue that cost of acquiring information or cost of reoptimizing could be the reasons. In practice, it means that portion of producers update the information using the current state of the economy and the rest rely on the outdated information. Rational inattention advocates argue that everyone is constrained in the information acquisition, simply because there is too much information needed to assess optimal prices perfectly. What differs is the level of willingness and information channel capacity among population and institutions. We should, therefore, see less involvement on the side of the consumers and much more on the side of the producers, not to mention institutions like for example the central banks. Contrary to the sticky prices, the rational inattention models are capable to explain inflation inertia even in a situation when the cost of information or cost of reoptimizing is zero. Although there are some appealing features in rational inattention models, the discussion about sticky prices, sticky information and rational inattention is still far from over as is suggested in recent paper written by Solorzano (2017), where they empirically evaluated these three approaches by examining each microfoundation of price-setting firm behaviour and its ability to explain monetary transmission mechanism, which turned out poorly for the rational inattention model used in Mackowiak and Wiederholt (2009) compared to the other two.

Papers by Mankiw and Reis (2002) and Sims (2003) spurred many other researchers to dive into the area of information constraint economic models and to examine the extent of their applications. One of the first followers was Reis (2006a), who studied the inattentive consumers updating their consumption plans only sporadically. In this model, the consumers set their consumption plans according to the newest information, but once they are done, they then remain inattentive for several following periods and do not update their consumption plan at all. Because each consumer has a different period in which he or she is updating the consumption scheme, we should witness sluggish and relatively smooth consumer reaction to the news. According to Reis (2006a), the same reasoning applies even to the situation when consumer receive permanent income shock, since only a small portion of the consumers react instantly, while the rest remain inattentive about such shocks. If some event could be easily predicted, then most of the consumers would change their consumption plans accordingly. Reis (2006a) predicted that there might be some events, which would capture everyone's attention and therefore ignite sharp and quick reaction, a feature used by many other researchers as well. What is interesting about Reis (2006a) is that the model predicts that around one-third of the U.S. citizens will choose to never be attentive and are living from hand-to-mouth, resulting in savings near zero for those people. We should mention that similar study was made by the same author (Reis (2006b)), but this time with a focus on the inattentive producers. In this study the firms decide if they would like to make plans about the prices they will charge or about quantities they would like to sell, but at the same time deciding about their news updating scheme, which in other words means how many periods they would choose to be inattentive, allowing for the richer structure of the model.

There are other shortcomings arising when using models of sticky information as in Mankiw and Reis (2002) and Reis (2006a, 2006b). These models provide no clear explanation why would some of the agents choose not pay attention at all and some would pay infinite attention to the state variables. What is missing in the models is some sort of noise involved in the state variables, which would make acquisition of the precise state variable costly. Strictly speaking, from the information theory point of view, when an agent is perfectly aware about everything at the time when updating of information occurs, then he has unlimited information capacity and the cost of acquiring a new information is zero, resulting in updating a new information in every period. A different approach to this problem was developed by Moscarini (2004). In his work, he also worked with infrequent updating scheme for a new information, but since he introduced noise factor to the state variables, then the frequency when updating occurs was directly dependent on the information capacity constraint and on the volatility of the noise, which is the feature used later in this work. To put it simply, when the volatility of our variable of interest increases, then in order to preserve the same frequency of the updating scheme, we have to increase our information channel in a way which would balance the increased volatility. Such reasoning seems natural, we would expect for example a broker to take extra time in work in a situation when stock markets “go crazy” and their volatility is hugely increased. This phenomena was later developed by Peng and Xiong (2006), wherein the work they focused on the behaviour of representative investor under limited information capacity. Since there is fundamental uncertainty in the world, an investor has to decide where to put her limited capacity in order to minimize her wealth uncertainty. The paper showed that the investor will invest more capacity toward more volatile assets, a feature very often occurring in the limited information literature. Slightly different approach regarding portfolio selection problem in the rational inattention framework was introduced by Huang and Liu (2007). In their model, the representative investor receives both continuous and discrete economic news relevant to her investment decision and decide about frequency and accuracy of that news. The main difference between Huang and Liu (2007) and Moscarini (2004) together with Peng and Xiong (2005) and Sims (2003) is that in case of the later named the agent is limited in information processing, while in the case of the former mentioned the agent is unlimited, but the news is the bearer of the noise and agent decide about their precision and frequency.

The idea of rational inattention proved to be vital even in the area of international finance, namely as one of the possible explanations behind one of the greatest international finance puzzles, the uncovered interest rate parity and why it does not hold the way as theory suggests. This question is of great importance since many international macroeconomics models lie upon the idea of uncovered interest rate parity. According to the theory, the regression coefficient of monthly change in the exchange rate and one-month forward premium should be one, but as was shown in Fama (1984) and by many other academists, this equality empirically does not hold true and is in fact negative. Bacchetta and Wincoop (2005) suggested the rational inattention to be the main driving force behind this important paradox. Their approach is similar to those of Moscarini (2004), the significant costs associated with collecting and processing information generate motives for investors to update their portfolio only infrequently, which is therefore responsible for predictable expectational errors in forward discounts. Another great problem in economics which was tried to be explained using the rational inattention approach is the equity premium puzzle. This problem was suggested by Mehra and Prescott (1985) using the general equilibrium model, where they showed that covariance of consumption growth and stock returns is too low for explaining the overall equity premium and has to lead to implausible high risk aversion. One of the possible solutions to this puzzle was limited stock market participation hypothesis suggested by Vissing-Jørgensen (2002). In this approach, only limited portion of the population participate in stock markets, leading to a lower covariance between the consumption growth and the stock returns as seen in the data. The problem with the limited stock market participation hypothesis is that it provides no clear explanation why some part of the population decided not to participate. In the paper written by Luo (2006), he tried to endogenize the problem using the rational inattention and applied it to the consumption-based capital asset pricing model. It was shown in the paper that incorporating rational inattention in the model can lead to a smooth consumption process and low covariance between consumption growth and asset prices while leaving the coefficient of the relative aversion of the representative agent unchanged. Another example of usage of the rational inattention in the economic literature can be found in Luo (2008). In this paper, the rational inattention is proposed as an alternative hypothesis to other hypotheses in the permanent income model such as for example the habit formation.

Very important area of economics is large macroeconomics models. These remain still a huge challenge for the future research, as many complexities arise by introducing the rational inattention and a lot still has to be done. In Máckowiak and Wiederholt (2009), they developed an equilibrium model of price setting using the rational inattention. In the model, there is a continuum of firms and they decide about their prices in each time period in order to maximize their discounted future real profits. Although in the model, the profits are a function of firm´s own price level, the overall price level, aggregate demand and firm-specific demand conditions, the firm cannot see the later three and receive only a signal about them instead. So each firm has to decide its own price using the current signal as well as all the previously received signals, which may or may not be a good approximation of the underlying state variables. The main finding of the paper (Máckowiak and Wiederholt (2009)) is that firms react more swiftly and greatly to their own idiosyncratic shocks and more slowly and by a little amount to aggregate shocks. The even more ambitious model was developed by the same authors two years later (Máckowiak and Wiederholt (2015)). It was the first rational inattention dynamic stochastic general equilibrium model up to date, featuring both households and firms optimization problem. The motivation behind the paper was clear, to provide an alternative explanation for a large inertia found in empirical macroeconomic studies. The authors summarized their findings into four points. Firstly, the households choose to pay a little attention toward changes in the real interest rates and therefore the consumption responds very slowly to these movements. Since the interest rates are the main tool of the central banks, this result has consequences for monetary policy recommendations. Secondly, in the line with Máckowiak and Wiederholt (2009), the model predicts large adjustments to market specific or firm-specific shocks, smaller to the technology shocks and the smallest for monetary policy shocks. Thirdly, the optimal allocation of the firm's attention is largely influenced by the optimal allocation of attention made by the households, so if the households choose not to pay much attention toward some variable of interest, so does the firms. The model sometimes predicts very different behaviour than in the standard DSGE model and thus very different policy recommendations.

Not all of the papers we mentioned earlier were accepted without any criticism. In Sims (2005) and later in Sims (2006), he argues that the papers often assume Gaussian prior distribution as well as Gaussian posterior distribution, after the information flow takes place, which is not always defendable. There can be cases, when even if the prior distribution of the state variables would be independent and identically distributed, the same would not have to be true for the posterior distributions (Sims (2005), Sims (2006)). It is an agent, who can choose his posterior distribution according to his information capacity restraints. In practice, many time series in economics show some kind of autocorrelation structure as well as a lot of interdependencies between the variables. Moreover, as Sims (2005) shows, even the standard forms of utility functions can create non-Gaussian posterior distributions. He uses a simple two-period saving problem to illustrate this point and even if an agent is faced with a hard budget condition, the optimization problem is still tractable in the face of limited information flows. The problem which arises in saving optimization model from the rational inattention point of view is that an agent´s net worth, or put it differently, an agent´s lifetime savings, are itself variables that an agent is not perfectly aware of. There is some inherent noise in these variables as an agent changes his occupation or retirement plan for example but still has to make a decision about his savings versus consumption in each period. Sims (2005) main finding is that numerical solution to this simple problem implies a discrete distribution of actions, even when the underlying uncertainties are continuously distributed. It resulted as a consequence of numerical calculation, nevertheless, this idea was later developed by Matějka and Sims (2009), where they provide an analytical solution to a broad number of such problems, i.e. to have continuous objective function and continuous uncertainty. Their result is that an agent´s optimal actions can under these conditions resulted in a discrete behaviour, which is important result given that in real life we see practically everyone taking discrete actions. The example they use in the paper is a grocery store, changing its prices rather sporadically and from a limited range of values, which is something we can observe in practice.

We can see in the most recent development in the field application of the rational inattention to the even broader spectrum of economic theories. Matějka and McKay (2012) for example consider market equilibria with consumers under the information constraints, while having perfectly informed firms. In the model, the consumers are uncertain about firms´ offers and can evaluate them with any precision under the cost of information acquisition. The authors came with some interesting results. Firstly, they found out that increasing cost of processing information results in increasing prices set by the firms. Next, the prior knowledge of consumers increase the prices of those firms which are favorable to them, which is appealing result in the branding culture we can see nowadays. Interestingly, when in the model consumers have the heterogeneous cost of information acquisition, the low-quality firms tend to set higher prices. The authors are explaining this finding by the fact that if firms see consumers to be inattentive to a low quality of their products, then consumers are supposed to be inattentive to higher prices as well. Another study was made by Matějka (2016a) and what is interesting about this work is that the rational inattention on the consumer's side provide a feasible explanation about the rigidity of prices as it is an optimal price setting strategy for a monopolistic seller. In another study made by Matějka (2016b), the author looked at the problem from a perspective of the rationally inattentive seller. We can sometimes see in the market very large price rigidities and discreteness, i.e. that some prices go up and down only from two or three distinct points, which is very hardly defendable behaviour in a lot of theories.

Matějka (2016b) showed that in the rational inattention setting such behaviour can occur and can even be the explaining force while having no other adjustment costs.

3 Information Theory

Nowadays it impossible to imagine the world without the information theory. Everyone comes into contact with practical usage of this theory on a daily basis, whether playing DVD, using a mobile phone or surfing over the internet. The central concepts in the information theory are the notions of the ultimate data compression (defined as the entropy H) and the optimal transmission rate of communication (defined as the channel capacity C). For this reason, the information theory can be understood as a subset of communication theory from a purely electrical engineering point of view, but its many aplications in other fields such as computer science, statistics, biology or economics prove that this limitiation would be very shortsighted. To be more concrete, we present some examples of usefullness of the information theory in some particular fields. In electrical engineering, the main focus of interest lies on the information flow between some systems (e.g. communication between some electronical devices) and what is the volume of such flow, in other words, what is the channel capacity of those systems. The notion of entropy is very well known in physics, as one of the most fundamental law in nature, the second law of thermodynamics, is directly related to the notion of entropy, which is in fact always increasing in time in a given system. In statistics, we are usually interested in reliability of the information flow, i.e. how much are some signals received errors ridden. It was Shannon (1948) who showed that probability of error can be asymptotically reduced toward zero for the communication rates below the channel capacity. Unfortunatelly, there are many practical obstaclesin attaining those theoretical goals, nevertheless, the advancements in integrated cicuits together with faster communication channels and effective coding/decoding mechanisms make theoretical limits foreshadowed by Shannon more and more achievable.

For the purpose of this work, some concepts and notions from the information theory outlined here will be crucial in our rational inattention model described later. The main question underlying the information theory is how to measure uncertainty reduction as a function of the information flow. To do so in some manageable fashion, it is needed to reduce our focus only on the extreme points of the communication theory, namely the data compression limit $I(X;\hat{X})$ and the data transmission maximum $I(X;Y)$. There is of course obvious conection with the economic theory, where for an agent to be acknowleged about the true state of the world from relevant signals usually makes an agent better off. Many of the random processes have inherent complexicity, which cannot be further reduced. We can take as an example a coin flip or a dice toss, where beforehand there is a lot of uncertainty about how the result turn out to be. Such understanding is very well known to the field of econometrics, where every model contains an error term, or some other irreducible part of the uncertainty. This is in the information theory literature called the Shannon entropy, named after the father of modern information theory. The concept of entropy was first introduced in the seminal work of Shannon (1948) and in order to formalize this intuitive meaning, now we will give a formal definition of the concept.

For a random variable X with probability mass function $p(x)$, we say that the entropy H of the variable X given $p(x)$ is equal to $H(X) = -\sum_x p(x)\log_2 p(x)$. The entropy is a measure of an average uncertainty contained in some variable and it is always nonnegative as it is clear from the definition since $p(X) \geq 0$. We defined entropy so far only for the discrete random variables, but one can naturally extend the definition for the continuous case as well. Then the entropy would take the form $H(X) = E_p \log_2 \frac{1}{p(X)}$. As a convention, it is usually assumed that 0*log0 = 0, which is justified by the theory of limits as $\lim_{x\to 0} x \log x = 0$. We should point out that the use of the logarithm to the base 2 is rather practical, one can use the logarithm to any base b and defined its entropy as $H_b(X) = -\sum_x p(x)\log_b p(x)$, but then the units of measurement would be different since for a given random variable X and $b \neq 2$, the entropy $H(X)$ and the entropy $H_b(X)$ are never the same. The logarithm to the base 2 is nevertheless a stardard in the information theory literature, since the entropy with the logarithm to the base 2 is equal to 1 for a distrete random variable with a uniform distribution over 2 outcomes. The unit of information in which the entropy is measured is in the literature called one "bit". What it practically means is that we would need information channel at least equal to 1 bit to fully overcome our uncertainty about fair coin toss or decision between 0 and 1. Another special case is the entropy measured in "nats", then we would use the natural logarithm in our definition.

In order to really grasp the definition of entropy, one should pay attention to the fact that the entropy is an average uncertainty of some variable. To illustrate this, take for example the entropy contained in the uniform distribution over the eight outcomes, equal by the definition to $H(X) = -\sum_8 \frac{1}{8} \log_2 \frac{1}{8} = 3$ bits, and compare it with discrete distribution over eight outcomes with corresponding probabilities being $(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \frac{1}{64}, \frac{1}{64}, \frac{1}{64}, \frac{1}{64})$. Then the entropy in this case is equal to $H(Y) = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{4}\log_2\frac{1}{4} - \frac{1}{8}\log_2\frac{1}{8} - \frac{1}{16}\log_2\frac{1}{16} - \frac{4}{64}\log_2\frac{1}{64} = 2$ bits. The reason for that is that there is less uncertainty in the later random variable, since we already have some prior knowledge about the outcome. In this particular case, we a priori know that the first outcome is much more likely than the other outcomes. What it means for us is that it would suffice enough to have channel transmitting only 2 bites on average to be left with no uncertainty at all. If we would like to send a message about the outcome of this variable, we could make use of inequalities in the probability distribution by a proper coding, in our case for example by assigning the first outcome as 0, the second one as 10, the third one as 110, the fourth one as 1110 and the last four as 111110, 111101, 111100 and 111111, respectively. The bottom line is that we can see the entropy as the lower bound of the average bits needed to be sent to decribe some outcome of a random variable.

The mutual information $I(X;Y)$ about random variables X and Y with the joint probability distribution $p(X,Y)$ is equal to $I(X;Y) = E_{p(X.Y)} \log \frac{p(X,Y)}{p(X)p(Y)}$. The meaning of the mutual information can be intuitively seen as the answer to the question: How much uncertainty can be reduced in one variable thanks to the knowledge of the second variable? The mutual information is symetric, so the reduction of the uncertainty is the same regardless if we are conditioning X on Y or other way around. Another way how to look at it is that if we would like to get some information about X given some sequences of observations on Y and on Z, there would be no difference at all if we would see the observations on Y first and the observations on Z second or reversely. As one would expect, if X and Y are mutualy independent, then their mutual information equals 0 since $p(X,Y) = p(X)p(Y)$ for an independent variables.

Closely related to the entropy and the mutual information is the notion of the conditional entropy $H(X|Y)$, which is the entropy of a random variable X, given the knowledge of a variable Y. It can be shown that in fact $I(X;Y) = H(X) - H(X|Y)$. In the information theory, the mutual information is often interchanged with the term channel capacity. It is because if we take $\max I(X;Y)$ over the corresponding probability functions, then we say that our information flow reached the channel capacity, so there is no uncertainty left from our source. To see how $H(X)$, $H(X|Y)$ and $I(X;Y)$ are related, we use an example from the Shannon (1948). Let say we need some message from a source and this message is long 1000 zeros and ones with equal probabilities for both. Thus we have the entropy equal 1000 bits per second, so if we are transmitting the information flow equal to 1000 bits per second, then we are at the channel capacity of the system. Now imagine that there is some noise in the trasmission, so our source have probability of generating an error, i.e. sending 1 when 0 is true and otherwise, being equal to 1%. Then we know that there would be 10 mistakes in the message on average, but since we do not know the position of these error, we need to take into consideration this source of uncertainty as well. After some calculation we would get that $H(X|Y) = -0.01 \log_2 0.01 - 0.99 \log_2 0.99 = 0.081$ per bit send, so our system is transmitting at rate 919 bit per second. From the example above it should be clear that $H(X|Y)$ is in fact information lost by using Y as an output instead of X, while we are in fact interested in X. Thus, Y can be seen in some sense as a part of natural or technical barriers between the true states and the signals received.

4 Rational Inattention Filter

In the following pages, we will follow the work by Maćkowiak et al. (2017). Our particular focus will be on modeling behaviour of the representative agent in a dynamic environment setting.

Many of the processes found in macroeconomics can be well described by ARMA(p,q) in the form

$$X_t = \gamma_1 X_{t-1} + \cdots + \gamma_p X_{t-p} + \theta_0 \varepsilon_t + \cdots + \theta_q \varepsilon_{t-q},$$

where X_t is our variable of interest in period t and ε_t stands for innovation in period t. Further, in the text, we will assume X_t to be a stationary process. Our representative agent has only limited attention available to him, so he has to decide how his scare resource will be allocated. In each period of time, the agent receive a signal about his variable of interest in the form

$$S_t = a_0 X_t + \cdots + a_{p-1} X_{t-(p-1)} + b_0 \varepsilon_t + \cdots + b_{q-1} \varepsilon_{t-(q-1)} + \psi_t,$$

where ψ_t is i.i.d. noise of the signal. Since the agent has limited attention, he is forced to have trade-off between the increased variance of the variable ψ_t and the particular signal weights $a_0, \ldots, a_{p-1}, b_0, \ldots, b_{q-1}$ on the variables. Variable of interest for our agent in time t is X_t, so it would be reasonable to expect that optimal choice takes the form $S_t = X_t + \psi_t$, which means to learn about X_t as much as possible today and be left only with the variance of the signal noise σ_ψ^2. As was shown in Maćkowiak et al. (2017), that is generally not the case. Because we have a dynamic environment, the agent wants to be well informed today as well as in the future and this creates tension between these two goals. As it turns out, the agent in fact puts unzero weights only on those variables, which are affecting expected value of our variable of interest in the next period, i.e. only on those who are present in $E[X_{t+1}|I_t]$. It means for our ARMA(p,q) process up to period p-1 and q-1. There is no usage for an agent to put nonzero weights on let say a_p or b_q, since they are not helping him improving the signal about his variable of interest neither at time t or at time t+1. . This result proved by Maćkowiak et al. (2017) is of great importance, since it easify the usage of the Kalman filter in much more computationaly tractable way.To see why it is so in rather intuitive way, let take as an example ARMA(1,1) process of the form $X_t = \gamma_1 X_{t-1} + \theta_1 \varepsilon_{t-1}$, $\gamma_1 \neq 0, \theta_1 \neq 0$. Then the expectation about X_{t+1} given the information set in time t is $E[X_{t+1}|I_t] = \gamma_1 X_t + \theta_1 \varepsilon_t$, so the agent can utilize in next period devoting some attention at time t to innovation process ε_t. Analytical results in Maćkowiak et al. (2017) show that an agent chooses signals in the form $S_t = X_t + \psi_t$only if the underlying process driving the X_t is AR(1), so there is no tension between the current and the future period. Agent has information set available to him at time t≥1 equal to

$$I_t = I_0 \cup \{S_1, \dots, S_t\},$$

where I_0 is the initial information set.

In our dynamic rational inattention decision problem, we will follow Maćkowiak et al. (2017). The very similar problem was solved by Sims (2003), where he solved the problem using brute force numerical optimization technique, while in Maćkowiak et al. (2017) they derived the analytical solution of the problem for the particular class of ARMA(p,q) processes. They showed that dynamic rational inattention problem can actually be conveniently solved by using the Kalman filter. Our agent solves Gaussian linear-quadratic optimization problem, specifically, he aims to minimize the mean squared error, subject to the capacity constraint in the form

$$\min_{K,A,B,\Sigma_\psi} E[(X_t - E[X_t|I_t])^2]$$

subject to

$$X_t = \gamma_1 X_{t-1} + \dots + \gamma_p X_{t-p} + \theta_0 \varepsilon_t + \dots + \theta_q \varepsilon_{t-q},$$

$$S_t^K = A X_t^M + B \varepsilon_t^N + \Psi_t^K,$$

$$I_t = I_0 \cup \{S_1, \dots, S_t\},$$

$$\lim_{T\to\infty} \frac{1}{T} I(Z_0, \varepsilon_1, \dots \dots, \varepsilon_T, S_1^K, \dots \dots, S_T^K) \leq \kappa,$$

where S_t^K is K-dimensional signal vector received by the agent, $K \geq 1$ is an integer, $A \epsilon R^{KxM}$ and $B \epsilon R^{KxN}$ are matrices of weights set by the agent, M≥1 and N≥0 are positive integers, Ψ_t^K is K-dimensional vector of signal noises, Z_0 is the vector of initial conditions such that $Z_0 = \left(X_0, \ldots\ldots, X_{1-p}, \varepsilon_0, \ldots\ldots, \varepsilon_{1-q}\right)'$ and κ is a positive real parameter. Here $I(Z_0, \varepsilon_1, \ldots\ldots, \varepsilon_T, S_1^K, \ldots\ldots, S_T^K)$ denotes the mutual information between the variables in brackets, defined in the same way as in the previous chapter. As we can see, here we have dynamic optimization problem and the main distinction from the static case is that our agent is forward looking, so he is aware of the benefits which he could get by optimizing not only the current gains by setting $S_t = X_t + \psi_t$, but also the future ones by paying attention to the variables which improves his predictions over the next periods. This feature is present because the agent is limited in his attention, since as κ is approaching the infinite, then in fact the signal choosed by the agent has the form $S_t = X_t + \psi_t$ in the limit, as was proved by Maćkowiak et al. (2017). In other words, if the information flow is very large (i.e. the κ is large) and agent is paying large amount of attention to the variable of interest, then the variance in σ_ψ^2 is decreasing and cease to be the key variable in determining the optimal weights in the signal. Intuitively, if monitoring of some variable becomes crucial for the agent, then one would expect from him to watch very carefully this particular variable and increase his efforts and therefore information capacity devoted to it.

In this part, we will present some of the main findings in the work made by Maćkowiak et al. (2017) since they will enable us to solve the dynamic optimization rational inattention problem in analytically tractable fashion using the Kalman filter. Now let ξ_t be the vector defined as

$$\xi_t = \begin{cases} \left(X_t, \ldots\ldots, X_{t-\max\{M,N+p\}+1}\right)' & if\ q = 0 \\ \left(X_t, \ldots\ldots, X_{t-\max\{M,N+p\}+1}, \varepsilon_t, \ldots\ldots, \varepsilon_{t-q+1}\right)' & if\ q > 0 \end{cases}$$

and let $S^{K,t}$ be the set of all signals recevied up to time t, so $S^{K,t} = \{S_1^K, \ldots\ldots, S_t^K\}$. Then Maćkowiak et al. (2017) proved that the information constraint $\lim_{T\to\infty} \frac{1}{T} I(Z_0, \varepsilon_1, \ldots\ldots, \varepsilon_T, S_1^K, \ldots\ldots, S_T^K) \leq \kappa$ from the decision problem can be equivalently described as

$$\lim_{T\to\infty}[H(\xi_t|S^{K,T-1}) - H(\xi_t|S^{K,T})] \leq \kappa$$

as long as the vector ξ_t is such that vectors X_t^M and ε_t^N can be computed from it and ξ_t does not contain any redundant components. $H(\xi_t|S^{K,T-1})$ is here a prior entropy before the signal in time T is received and $H(\xi_t|S^{K,T})$ is posterior entropy after receiving the signal. This result is important, since it gives us some insight about the information constraint. Basically, it tells us the restriction on how much information can be contained in the signal S_t^K, so that our entropy would not decrease below the limit given by the parameter κ. This equivalency was used by Maćkowiak et al. (2017) in proving that an agent can in fact reach the optimum by receiving only one-dimensional signal vector.

We can now move to our rational inattention filter, as was presented in Maćkowiak et al. (2017). Let have state-space representation in the usual form

$$\xi_{t+1} = F\xi_t + \upsilon_{t+1},$$

$$S_t = h'\xi_t + \psi_t,$$

where ξ_t and S_t are the same vectors as defined earlier, F is a square matrix with the length corresponding to the length of the vector ξ_t, υ_{t+1} is noise vector of the same length as ξ_t, h is a vector of signal weights with the length max{1, p}+q. We assume that ψ_t is signal noise following a Gaussian white noise process.

Let $\Sigma_{t|t}$ and $\Sigma_{t|t-1}$ to be conditional variance matrices of ξ_t given the information set in time t and t-1, respectively, and Q to be variance matrix of the vector υ_{t+1}. Then the Kalman filter equations are

$$\Sigma_{t+1|t} = F\Sigma_{t|t}F' + Q,$$

$$\Sigma_{t|t} = \Sigma_{t|t-1} - \Sigma_{t|t-1}h\left(h'\Sigma_{t|t-1}h + \sigma_\psi^2\right)^{-1}h'\Sigma_{t|t-1},$$

Now let $\Sigma_1 = \lim_{t\to\infty}\Sigma_{t|t-1}$ and $\Sigma_0 = \lim_{t\to\infty}\Sigma_{t|t}$. Existence of those limits is guaranteed by the assumption about X_t being a stationary process and they are equal to

$$\Sigma_1 = F\Sigma_0F' + Q,$$

$$\Sigma_0 = \Sigma_1 - \Sigma_1 h\left(h'\Sigma_1 h + \sigma_\psi^2\right)^{-1} h'\Sigma_1,$$

As was shown in Maćkowiak et al. (2017), given the Kalman filter equations, we can rewrite our information flow constraint to so-called "signal-to-noise" ratio in the form

$$\frac{1}{2}\log_2\left(\frac{h'\Sigma_1 h}{\sigma_\psi^2} + 1\right) \le \kappa.$$

Since the information constraint is always binding, we can change inequality into equality and by slightly rearranging the terms, we get that

$$\sigma_\psi^2 = \frac{h'\Sigma_1 h}{2^{2\kappa} - 1}.$$

Now if we substitute the variance of the signal into our Kalman equations, then we can rewrite our decision problem from the beginning as

$$\min_{h \epsilon R^{\max\{1,p\}+q}} (1 \quad 0 \quad \dots \quad 0)\Sigma_0 \begin{pmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}$$

where matrices Σ_0 and $\Sigma_1$are given by the Kalman filter equations as

$$\Sigma_1 = F\Sigma_0 F' + Q,$$

$$\Sigma_0 = \Sigma_1 - \left(\frac{1 - 2^{-2\kappa}}{h'\Sigma_1 h}\right)^{-1} \Sigma_1 h h'\Sigma_1.$$

We have used the information flow constraint to remove the signal noise variance from the objective of our decision problem, so now the agent must decide only about the optimal signal weights h^*. In our Kalman filter equations, the prior variance matrix $\Sigma_1$of the state vector and the posterior variance matrix Σ_0 of the state vector together create implicit functions of matrices F and Q, the vector of signal weights h, , and the parameter κ. Solving this yields the optimal signal weights h^* and subsequently the optimal variance σ_ψ^*. To find a solution for the system is not a trivial task and usualy requires extensive use of computation power. Some of the results which easify the computation can be found in Maćkowiak et al. (2017).

5 The Basic New Keynesian Model

In this section, we will derive the basic New Keynesian model in a standard setting with utility-maximizing households and profit-maximizing firms. This model will later be extended to account for rational inattention, here we just sketch the basic results for a better comparison with the rational inattention one. Many of the steps in deriving the model will therefore be skipped, as they can be found in many macroeconomic textbooks. The New Keynesian models have become standard for many institutions around the world and the basic New Keynesian can be very often found in the literature as a baseline model for comparison with models including further extensions. For an interested reader, we encourage to look for example into Gali (2008) for the more detailed derivation of the model presented below. The choice of the DSGE model is motivated by the fact that the model is dynamic, so an agent is well aware that his today´s actions are affecting his future gains. It has a direct connection with the results depicted in the previous chapter, where we have a dynamic environment for an agent´s decision problem as well. The New Keynesian models have some appealing features, which might be the reason why they become so popular in the literature. Firstly, the firms produce different goods, which gives them a monopolistic power over the market. It means that they are their own price setters, contrary to the Walrasian auctioneer, so they can freely determine what will be the price for a supplied output. Secondly, the firms are facing the nominal rigidities, in the case of the basic New Keynesian the pricing rigidities. It creates frictions over the markets, which leads to variations in the real interest rates and thus inducing short-run non-neutrality of monetary policy.

5.1 Households

We will first present the household´s optimization problem and later the firm´s optimization problem. In our setting, we have identical infinitely-lived households normalized to measure 1. The representative household is facing the optimization problem of maximizing its lifetime utility

$$E_0 \sum_{t=0}^{\infty} \beta^t U(C_t, N_t, \frac{M_t}{P_t})$$

where C_t is consumption in time t, N_t is labour in time t,$\frac{M_t}{P_t}$ are real money holdings in time t and $\beta\epsilon(0,1)$. The utility function is increasing and concave in C_t , decreasing and concave in N_t and increasing and concave in $\frac{M_t}{P_t}$.

We assume that there is a continuum of goods over the interval (0,1) and the consumption is the overall sum of those goods given by

$$C_t = \left(\int_0^1 C_{it}^{\frac{\varepsilon-1}{\varepsilon}} di \right)^{\frac{\varepsilon}{\varepsilon-1}}$$

The household is facing the budget constraint

$$\int_0^1 P_{it} C_{it} di + M_t + Q_t B_t \leq M_{t-1} + B_{t-1} + W_t N_t + T_t,$$

where $\int_0^1 P_{it} C_{it} di$ are total expenditures, M_t are nominal money, $Q_t = \frac{1}{1+i_t}$ is the price per bond, B_t is the number of bonds bought, W_t is the wage per unit of labour and finally T_t is the tax imposed on the household in time t. The budget constraint is sometimes called "No-Ponzi Game" constraint, since it rules out the possibility of having strategy that would keep rolling an explosive debt indefinitively.

In order to solve household´s decision problem, we need to proceed in two stages. First, we have to find the optimal composition of C_t for any level of consumption and second, we need to find the combination of consumption, labour and money, which maximaze the household´s utility. The first task can be solved by finding the consumption vector which maximize a total consumption for a given level of expenditures, i.e.

$$\max_{C_{it}} \left(\int_0^1 C_{it}^{\frac{\varepsilon-1}{\varepsilon}} di \right)^{\frac{\varepsilon}{\varepsilon-1}}$$

subject to

$$\int_0^1 P_{it} C_{it} di \leq Z_t,$$

where Z_t are the consumption expenditures. After solving the problem, we get the equation for an aggregate price index

$$P_t = \left(\int_0^1 P_{it}^{1-\varepsilon} di \right)^{\frac{1}{1-\varepsilon}},$$

and the demand function for good i

$$C_{it} = \left(\frac{P_{it}}{P_t} \right)^{-\varepsilon} C_t$$

Here ε function as the elasticity of substitution, so with increasing ε we are getting closer to perfect competition. Now we have to find levels of consumption, labour and money, which maximize the households utility, i.e.

$$\max_{C_t, N_t, \frac{M_t}{P_t}, B_t} E_0 \sum_{t=0}^{\infty} \beta^t U(C_t, N_t, \frac{M_t}{P_t})$$

s.t.

$$P_t C_t + M_t + Q_t B_t \leq M_{t-1} + B_{t-1} + W_t N_t + T_t,$$

where we used$P_t C_t = \int_0^1 P_{it} C_{it} di$. Before we move next, we will assume that household´s utility takes the form

$$U\left(C_t, N_t, \frac{M_t}{P_t}\right) \equiv \frac{C_t^{1-\sigma}}{1-\sigma} - \frac{N_t^{1+\varphi}}{1+\varphi} + \frac{\left(\frac{M_t}{P_t}\right)^{1-\nu}}{1-\nu}$$

After solving the optimization problem, we get the Euler equation

$$1 = \beta Q_t^{-1} E_t \left\{ \left(\frac{C_{t+1}}{C_t} \right)^{-\sigma} \frac{P_t}{P_{t+1}} \right\},$$

the labour-leisure choice equation

$$C_t^{\sigma} N_t^{\varphi} = \frac{W_t}{P_t},$$

and the money demand equation

$$\frac{\left(\frac{M_t}{P_t}\right)^{-\nu}}{C_t^{-\sigma}} = \frac{i_t}{1 + i_t}$$

These equations are rather impractical to work with as they are obviously non-linear. In order to linearize them, we use log-linearization together with the Taylor expansion around the steady state. From now on, we will denote the logarithm of some variable by using a small letter. Furthermore, we will define the following variables

$$\rho = -ln\beta$$

$$i_t = -lnQ_t$$

$$\Delta c_{t+1} = c_{t+1} - c_t = lnC_{t+1} - lnC_t$$

$$\pi_{t+1} = p_{t+1} - p_t = lnP_{t+1} - lnP_t$$

Using this, the Euler equation can be log-linearized into

$$1 = E_t\left(e^{-\rho + i_t - \sigma\Delta c_{t+1} - \pi_{t+1}}\right)$$

Further, taking the first-order Taylor expansion around steady state gives us

$$c_t \approx E_t c_{t+1} - \frac{1}{\sigma}(i_t - \rho - E_t \pi_{t+1})$$

Next, the log-linearized version of the labour-leisure choice equation is equal to

$$w_t - p_t = \sigma c_t + \varphi n_t$$

Now only the log-linearization of the money demand remains to be done. It is equal to

$$m_t - p_t = \frac{\sigma}{\nu} c_t + \frac{1}{\nu} \ln\left(\frac{1 + i_t}{i_t}\right)$$

After taking Taylor expansion around the steady state we are left with

$$m_t - p_t \approx \frac{\sigma}{\nu} c_t - \frac{1}{\nu}\frac{1}{(1+i)i} i_t + \frac{1}{\nu}\left[\ln\left(\frac{1+i}{i}\right) + \frac{1}{1+i}\right]$$

Now if we remove the constant term on the right side of the equation and assume $\sigma = \nu$ (i.e. assuming income elasticity being one), we will finally get

$$m_t - p_t = c_t - \eta i_t$$

where we defined $\eta = \frac{1}{\nu}\frac{1}{(1+i)i}$. This finishes the households optimization problem and we can now move to firms behavior, which has much richer structure than the households.

5.2 Firms

Production of the firm-specific output is our model made via labour hired by the firm and technology level available across all firms at given specific time. One can see the omission of the capital from the production function, which in our model plays no role and capital stock is treated as equal to fixed level with investments being zero in short run. Our production function has the form

$$Y_{it} = A_t N_{it}^{1-\alpha}$$

The key ingredience in the New Keynesian models lies in some form of rigidities involved, in our case, it is the firms which are unable to always set their prices optimally in each period of time. To be more specific, in our model there is a portion of firms which cannot change their prices in the current period and are left with the prices set in the previous period. The probability for the firm to be stuck with the last period prices is equal to parameter θ. Which of the firms would face such limitation depends purely on luck and there is no way firms could possibly know in which period it would apply for them. The current aggregate price level is therefore a composition of optimal prices set by the firms able to reset their prices to the new optimal level and prices from firms stuck with the prices from the previous period.

$$P_t = \left[\int_{S(t)}^{1} P_{it-1}^{1-\varepsilon} di + (1-\theta) P_t^{*1-\varepsilon}\right]^{\frac{1}{1-\varepsilon}}, S(t) \in [0;1]$$

We define the aggregate gross inflation $\Pi_t = \frac{P_t}{P_{t-1}}$. After little arrangements with the above equation we can express the inflation as

$$\Pi_t^{1-\varepsilon} = \theta + (1-\theta)\left(\frac{P_t^*}{P_{t-1}}\right)^{1-\varepsilon}$$

In the steady state there is zero inflation, implying $\Pi = 1$. Linearizing around sthe teady state the following equation of the aggregate inflation, where logarithms of variables are denoted the same way by small letters as in the chapter devoted to the households.

$$\pi_t = (1-\theta)(p_t^* - p_{t-1})$$

The equation shows that inflation is present only if there are firms willing to change their prices and are moreover given the opportunity to do so. This is important for the model since there would be no point in limiting some firms to not be able to reset the prices if there is no need for them to do so. Now we will derive the decision of the firms over the optimal price level and why it may differ from the previous period. The firms are profit maximizers and decide over their production level and corresponding prices. This can be written as following optimization problem

$$\max_{P_t^*}\left\{\sum_{k=0}^{\infty} \theta^k E_t\left[Q_{t,t+k}(P_t^* Y_{it+k|t} - TC_{it+k|t}^n(Y_{it+k|t})\right]\right\}$$

$$s.t. Y_{it+k|t} = \left(\frac{P_t^*}{P_{t+k}}\right)^{-\varepsilon} C_{t+k}$$

The representative firm wants to maximize its future stream of profits by choosing over the current price level. In the maximization problem, we denote by $Y_{it+k|t}$ the output in period $t+k$ by firm i, which reset its price last time in period t. How likely this is for the particular firm corresponds with the parameter θ^k. We define $Q_{t,t+k} = \beta^k \left(\frac{C_{t+k}}{C_t}\right)^{-\sigma} \frac{P_t}{P_{t+k}}$. This has much to do with the Euler equation from the last chapter, as the firms are naturally discounting the future profit streams the similar way as the households do. The profit in each period is calculated as the output times price minus total nominal costs, which are increasing function of the output. Lastly, the firm is subject to demand constraint and the market clearing conditions.

In order to solve the optimization problem, we write down the Lagrange function and solve for the optimal price

$$\mathcal{L} = \sum_{k=0}^{\infty} \theta^k E_t \left[\beta^k \left(\frac{C_{t+k}}{C_t}\right)^{-\sigma} \frac{P_t}{P_{t+k}} \left(P_t^* \left(\frac{P_t^*}{P_{t+k}}\right)^{-\varepsilon} C_{t+k} - TC_{it+k|t}^n \left(\left(\frac{P_t^*}{P_{t+k}}\right)^{-\varepsilon} C_{t+k} \right) \right) \right]$$

The first order condition with respect to optimal price P_t^* is given by

$$\sum_{k=0}^{\infty} \theta^k E_t \left[\beta^k \left(\frac{C_{t+k}}{C_t}\right)^{-\sigma} \frac{P_t}{P_{t+k}} \left((1-\varepsilon) \left(\frac{P_t^*}{P_{t+k}}\right)^{-\varepsilon} C_{t+k} + MC_{t+k|t}^n \varepsilon \left(\frac{P_t^*}{P_{t+k}}\right)^{-\varepsilon-1} C_{t+k} \frac{1}{P_{t+k}} \right) \right] = 0$$

After some rearrangements, we get finally the equation for the optimal price in the current period. Here the real marginal costs for the firm are achieved by dividing the nominal marginal costs by appropriate price level, i.e. $MC_{t+k|t}^r = \frac{MC_{t+k|t}^n}{P_{t+k}}$.

$$P_t^* = \frac{\varepsilon}{\varepsilon - 1} \frac{E_t \sum_{k=0}^{\infty} \theta^k \beta^k C_{t+k}^{1-\sigma} P_{t+k}^{\varepsilon} MC_{t+k|t}^r}{E_t \sum_{k=0}^{\infty} \theta^k \beta^k C_{t+k}^{1-\sigma} P_{t+k}^{\varepsilon-1}}$$

From the equation, we can see that the optimal price in the current period is an actually weighted average of future real marginal costs, where the firms are also accounting for the possibility of being unable to reset the price for the future periods. It should be noted that in the case of any price rigidities involved (parameter θ being zero) our equation collapses into $P_t^* = \frac{\varepsilon}{\varepsilon-1} MC_{t|t}^n$, so the optimal price would be driven by the nominal costs only.

For practical purposes, it is convenient to have the equation for the optimal price in the linearized form, which can be done by log-linearization of the left and right-hand side of the equation. Both sides will be taken as depicted below

$$P_t^* E_t \sum_{k=0}^{\infty} \theta^k \beta^k C_{t+k}^{1-\sigma} P_{t+k}^{\varepsilon-1} = \frac{\varepsilon}{\varepsilon - 1} E_t \sum_{k=0}^{\infty} \theta^k \beta^k C_{t+k}^{1-\sigma} P_{t+k}^{\varepsilon} MC_{t+k|t}^r$$

In the steady state, we have zero inflation and stable prices as well as consumption and output levels, which means that $\Pi_t = 1, Y_{t+k|t} = Y_{t|t}, Q_{t,t+k} = \beta^k, MC_{t+k|t}^r = MC^r = \frac{\varepsilon-1}{\varepsilon}$.

We begin by log-linearization of the left-hand side of the equation, which gives us

$$LHS: C^{1-\sigma} P^{\varepsilon-1} E_t \sum_{k=0}^{\infty} \theta^k \beta^k [1 + (p_t^* - p) + (\varepsilon - 1)(p_{t+k} - p) + (1 - \sigma)(c_{t+k} - c)]$$

Followed by log-linearization of the right-hand side

$$RHS: \frac{\varepsilon}{\varepsilon - 1} C^{1-\sigma} P^{\varepsilon-1} MC^r E_t \sum_{k=0}^{\infty} \theta^k \beta^k \left[1 + \varepsilon(p_{t+k} - p) + (1 - \sigma)(c_{t+k} - c) + \left(mc_{t+k|t}^r - mc^r\right)\right]$$

Finally, by equating both sides, we get the equation for the optimal price in the desired form

$$p_t^* = \mu + (1 - \theta\beta) E_t \sum_{k=0}^{\infty} \theta^k \beta^k \left[mc_{t+k|t}^r + p_{t+k}\right]$$

where $\mu = -mc^r$, which is desired markup for the firms given their market power.

5.3 Equilibrium

Once we have solved the household optimization problem as well as the optimal prices for the firms, we can combine the results to bring the equilibrium of the overall economy and close the model. The market clearing condition implies that level of output produced by the specific firm has to equal the amount demanded by the households, which can be formally written as the market clearing condition in the goods market

$$Y_{it} = C_{it}$$

In the following, we define the overall output as a function of specific output using the constant elasticity of substitution (CES) function of the form

$$Y_t = \left(\int_0^1 Y_{it}^{\frac{\varepsilon-1}{\varepsilon}} \, di \right)^{\frac{\varepsilon}{\varepsilon-1}}$$

Such structure allows us to state the aggregate market clearing condition. It can be shown that such aggregate output function combined with the specific product market clearing conditions implies the aggregate market clearing conditions in the goods market as well. In other words, the total level of output is equal to total consumption level in given period, thus

$$Y_t = C_t$$

Using the Euler equation from the chapter on the households and taking into account the aggregate market clearing conditions, we can write the function of output the same way as their consumption counterparts

$$y_t = E_t y_{t+1} - \frac{1}{\sigma}(i_t - \rho - E_t \pi_{t+1})$$

As a next step, we state the aggregate labour function, which is given by labour market clearing conditions. The labour force in this model is perfectly substitutable, therefore the form of the aggregate labour function

$$N_t = \int_0^1 N_{it} di$$

By combining the production function of the good i from the beginning of the chapter on firms with the labour market clearing conditions and goods market clearing conditions, as well as with the consumption demand equation, we get after taking logarithms the following equation

$$y_t = a_t + (1-\alpha)n_t - (1-\alpha)ln\left[\int_0^1 \left(\frac{P_{it}}{P_t}\right)^{-\frac{\varepsilon}{1-\alpha}} di\right]$$

It can be shown that last term in the equation is up to second order error equal to zero, which easifies our equation into

$$y_t = a_t + (1-\alpha)n_t$$

In our model without the capital investments, the nominal marginal cost is equal to the wage given to the labour W_t. In equilibrium, this cost has to equal the marginal gains from acquiring an extra labour, which corresponds to the nominal marginal product of labour times the current price. Thus, the real marginal cost equals the nominal cost divided by the nominal gains, i.e.

$$MC_t^r = \frac{W_t}{P_t MPN_t^n}$$

Which after simply taking the logarithms written as

$$mc_t^r = w_t - p_t - mpn_t^n$$

The marginal product of labour can be easily obtained from the production function by taking its derivative with respect to N_t, which gives us

$$mpn_t^n = a_t + ln(1-\alpha) - \alpha n_t$$

Finally, if we substitute the marginal product of labour into our equation of the real marginal cost, we obtain

$$mc_t^r = w_t - p_t - a_t - l\,n(1-\alpha) + \alpha n_t = w_t - p_t - \frac{a_t - \alpha y_t}{1-\alpha} - ln(1-\alpha)$$

In very similar manner, we can derive the real marginal cost in time t+k for a firm which lastly set its price in time t

$$mc_{it+k|t}^r = w_{t+k} - p_{t+k} - \frac{a_{t+k} - \alpha y_{it+k|t}}{1-\alpha} - ln(1-\alpha)$$

This equation is of great importance since it tells us the real cost of producing more output for a specific firm. The wage, price level and current technology are the same across all firms, so the only way how to influence the real marginal cost lies in determining the actual output, which is dependent on the marginal product of labour. In case when $\alpha = 0$ and thus we have constant returns to scale, the real marginal costs are constant and shared with all firms, but when $\alpha > 0$, a firm has to match its increased production with increasing marginal cost due to diminishing marginal product of labour, while paying the same wage for an extra labour.

As pointed out earlier, for equilibrium determinacy a firm has to meet the market clearing conditions, which together with the demand function from the household optimization problem gives that a firm output is equal to

$$Y_{it+k|t} = \left(\frac{P_{t+k|t}}{P_{t+k}}\right)^{-\varepsilon} Y_{t+k}$$

After taking logarithms, we get

$$y_{it+k|t} = -\varepsilon\left(p_{t+k|t} - p_{t+k}\right) + y_{t+k}$$

Now we will compute the difference of the real marginal costs between a firm which could not reset its price since period t and a firm which reset its price in current period t+k, i.e. the difference

$$mc^r_{it+k|t} - mc^r_{t+k} = \frac{\alpha}{1-\alpha}(y_{it+k|t} - y_{t+k})$$

Note that the difference in the real marginal costs is all about differences in output levels, which is the feature we discussed earlier. If we plug demand function into this equation and notice that $P_{t+k|t} = P_t^*$, we can rewrite the above equation as

$$mc^r_{it+k|t} = mc^r_{t+k} - \frac{\alpha\varepsilon}{1-\alpha}(p_t^* - p_{t+k})$$

Now we have done all the necessary steps to derive an equation for the inflation, which is a cornerstone in the New Keynesian models. For that, we will use the equation of the optimal price from the previous chapter

$$p_t^* - p_{t-1} = (1-\theta\beta)E_t\sum_{k=0}^{\infty}\theta^k\beta^k\left(mc^r_{t+k|t} - mc^r + p_{t+k} - p_{t-1}\right)$$

After some tedious algebra and defining $\Theta \equiv \frac{1-\alpha}{1-\alpha+\alpha\varepsilon} < 1$ and $\widehat{mc}_{t+k}^{r} \equiv mc_{t+k}^{r} - mc^{r}$, we can rewrite the above equation as

$$p_t^* - p_{t-1} = (1-\theta\beta)\Theta E_t \sum_{k=0}^{\infty} \theta^k \beta^k \widehat{mc}_{t+k}^{r} + E_t \sum_{k=0}^{\infty} \theta^k \beta^k \pi_{t+k}$$

From this, we can see the difference in the new and the old prices is due to future expected differences of the real marginal costs from the optimal markup $-mc^r$ as well as future inflation. The above equation can be written even more compactly as a difference equation of infation. In the previous chapter, we showed that inflation is equal to $\pi_t = (1-\theta)(p_t^* - p_{t-1})$. Using this together with $\lambda \equiv \frac{(1-\theta)(1-\theta\beta)}{\theta}\Theta$ to ease the notation, we finally get the equation of inflation expressed as a difference equation

$$\pi_t = \beta E_t \pi_{t+1} + \lambda \widehat{mc}_t^{r}$$

We can iterate this difference equation forward to actually see that inflation is in fact pushed by the real marginal costs only

$$\pi_t = \beta\left[\beta(\beta E_t \pi_{t+3} + \lambda E_t \widehat{mc}_{t+2}^{r}) + \lambda \widehat{E_t mc}_{t+1}^{r}\right] + \lambda \widehat{mc}_t^{r} = \lambda \sum_{k=0}^{\infty} \beta^k E_t\, \widehat{mc}_{t+k}^{r}$$

At this point it useful to remind that the real marginal cost in steady state is given by $MC^r = \frac{\varepsilon-1}{\varepsilon}$, which in log means $mc^r = ln\left(\frac{\varepsilon-1}{\varepsilon}\right) = -\mu$. If the average markup in the future is expected to be below the optimal markup μ, the firm will choose the price above the current level in order to push markup closer to the desired level and thus enhancing the inflation to rise. From above it follows that inflation is created solely by the decision of firms. In next section, we would like to show inflation as function of economic activity, i.e. as function of output rather than marginal cost. In order to do so, we will use some identities derived earlier, for example the equation of the real marginal cost such as

$$mc_t^r = w_t - p_t - mpn_t^n = \sigma c_t + \varphi n_t - [y_t - n_t + l\,n(1-\alpha)],$$

where we used labour supply equation $w_t - p_t = \sigma c_t + \varphi n_t$. Further, by using identity $\frac{y_t - a_t}{(1-\alpha)} = n_t$, we can rewrite the equation as function of output and technology only, i.e.

$$mc_t^r = \frac{\sigma(1-\alpha)+\varphi+\alpha}{(1-\alpha)} y_t - \frac{1+\varphi}{1-\alpha} a_t - ln(1-\alpha)$$

For further on, we will define very important notion of natural output level, which is crucial in the New Keynesian models. Let the natural output y_t^n be the output under full price flexibility, which is the situation where an average markup is equal to the optimal one and an economy is relaxed from any inflationary pressures. Quantitatively speaking it means the equation of the form

$$mc^r = -\mu = \frac{\sigma(1-\alpha)+\varphi+\alpha}{(1-\alpha)} y_t^n - \frac{1+\varphi}{1-\alpha} a_t - ln(1-\alpha)$$

Since the optimal markup is a constant in our model, we can express the natural output as the function of technology only

$$y_t^n = \frac{1+\varphi}{\sigma(1-\alpha)+\varphi+\alpha} a_t - \frac{(1-\alpha)[\mu - ln(1-\alpha)]}{\sigma(1-\alpha)+\varphi+\alpha}$$

This can be more compactly written after defining $\psi_{ya}^n = \frac{1+\varphi}{\sigma(1-\alpha)+\varphi+\alpha}$ and $\vartheta_y^n = -\frac{(1-\alpha)[\mu - ln(1-\alpha)]}{\sigma(1-\alpha)+\varphi+\alpha}$ as

$$y_t^n = \psi_{ya}^n a_t + \vartheta_y^n$$

What it means is that once the parameters of the model are set, the natural output is determined by an exogenous variable, in our case by a technology level, and nor households or firms have any means to influence it. From the above, we can easily compute the alternative representation of the deviation of the real marginal cost from the optimal one as

$$\widehat{mc}_t^r = \frac{\sigma(1-\alpha)+\varphi+\alpha}{1-\alpha} \tilde{y}_t,$$

where we define the deviation of an output $\tilde{y}_t$ (in the literature reffered as the output gap) as a difference between the actual output and the natural output $\tilde{y}_t = y_t - y_t^n$. Now we can give a different representation of an inflation dynamics

$$\pi_t = \beta E_t \pi_{t+1} + \lambda \widehat{mc}_t^r = \beta E_t \pi_{t+1} + \lambda \frac{\sigma(1-\alpha)+\varphi+\alpha}{1-\alpha} \tilde{y}_t,$$

which we can be easified further by setting $\kappa = \lambda \frac{\sigma(1-\alpha)+\varphi+\alpha}{1-\alpha}$ as

$$\pi_t = \beta E_t \pi_{t+1} + \kappa \tilde{y}_t$$

This is the equation of the New Keynesian Phillips curve (NKPC) and represents one of the key equations in the model. Now we turn our attention to dynamic IS equation, which together with the NKPC constitutes the foundation of all New Keynesian models. The real interest rate consists of the nominal interest rate decreased by inflation, in our case by expectations about future inflation, and is given by famous expectations augmented Fisher equation $r_t \equiv i_t - E_t \pi_{t+1}$. Using this, we can recreate the Euler equation for output as

$$y_t = E_t y_{t+1} - \frac{1}{\sigma}(r_t - \rho),$$

or similarly as the Euler equation for natural output

$$y_t^n = E_t y_{t+1}^n - \frac{1}{\sigma}(r_t^n - \rho)$$

By combining those results, we will get the dynamic IS equation (DIS) as an equation of the output gap in the form

$$\tilde{y}_t = E_t \tilde{y}_{t+1} - \frac{1}{\sigma}(i_t - E_t \pi_{t+1} - r_t^n)$$

The DIS with NKPC represents a system of two equations which builds the whole model to function. Once we know rule how the nominal interest rate is determined by a central bank and we know how the natural real interest rate reacts to exogenous variables, then the model is complete. For an stability of the system, we need to assure that there are convergence forces which make the system to return to its steady state after some unexpected shock. In order to do that, it is important that an agent knows each deviation of the output gap is only temporary, so that $\lim_{k\to\infty} E_t\tilde{y}_{t+k} = 0$ (transversality condition). By this assumption, we can iterate the DIS equation further to see that DIS can be equally written as

$$\tilde{y}_t = -\frac{1}{\sigma}E_t\sum_{k=0}^{\infty}(r_{t+k} - r_{t+k}^n)$$

In this light, similar to the inflation dynamics, the output gap is the sum of anticipated deviations of the real interest rates from its natural counterparts. What we need to do next is to find an equation of the natural real interest rate as being determined by exogenous variables only. From the Euler equation we have $E_t\Delta y_{t+1} = \frac{1}{\sigma}(i_t - E_t\pi_{t+1} - \rho)$ and we defined before the the natural output as an function of technology progress, therefore $E_t\Delta y_t^n = \psi_{ya}^n E_t\Delta a_{t+1}$. If we take an difference of those two we get

$$r_t^n = i_t - E_t\pi_{t+1} - \sigma(E_t\tilde{y}_{t+1} - \tilde{y}_t)$$

After few adjustments, we finally get an equation for the natural interest rate

$$r_t^n = \rho + \sigma\psi_{ya}^n E_t\Delta a_{t+1}$$

This equation gives a natural representation of the natural real interest rate. Remember from earlier that ρ is the dicsount rate set by households, or put it differently, how much the households value the present consumption over the future one. If there would be no technological progress, our natural real interest rate would be simply equal to this discount rate. However, in presence of technological advancements, the agents would like to benefit from those advancements which push the natural interest rate to rise.

Next thing we need to resolve in order to complete our model is to determine a monetary policy rule. This rule is of crucial importance for the central banks since in the New Keynesian models the monetary policy is not neutral in the short run, which gives some power over the real variables to the central banks. I will use here the Taylor rule, used in seminal work of Taylor (1993) and in many subsequent papers. It takes a simple form

$$i_t = \rho + \phi_\pi \pi_t + \phi_y \tilde{y}_t + v_t,$$

where v_t is an exogenous variable. The equation above can take many forms, for example by putting the inflation several periods ahead instead of in the current period, which seems more realistic given the delay when the monetary action takes place, but for the purpose of this work it will sufice to use the Taylor rule in the form suggested above. This concludes our model, since exogenous variables of technology and monetary shocks and their determinancy represents main part of the work and will be given a seperate chapter.

5.4 Calibration

After completing the model, we now have to specify the parameters used in equations above. Such task is always difficult and one has to be very careful in specifying the model. As the point of the work is to create a model which fits data for the Czech Republic, we will use as a primary source for our calibration the values used by the Ministry of Finance of the Czech Republic. They use a New Keynesian DSGE model for an open economy called HUBERT for the purposes of the macroeconomic policy analysis. If not stated otherwise, all coefficient presented below will be taken from Štork et al. (2014), the paper where the macroeconomic model for the Ministry of Finance is presented (HUBERT3). It should be noted that the model takes some of the coefficients from the previous version of the model made by Štork et al. (2009).

As our discount factor will take value $\beta = 0.99$, which corresponds to steady state interest rate of 4% given by the relationship $\rho = -ln\beta$. We assume risk aversion in consumption utility to be equal to 2 and risk aversion of labour, or put it differently, inverse of substitution of labour supply, to be equal to 5. Next, we take the share of capital as our production function parameter to be $\alpha = 0.5$ and the elasticity of substitution equal to $\varepsilon = 1.5$. Another important parameter in our model is a price stickiness. We assume an average duration of a price set to be 2 periods, so our coefficient is $\theta = 0.5$. Regarding monetary policy rule set by the central bank, we take the sensitivity of the central bank with respect to inflation equal to 1.5 and sensitivity with respect to the output gap to be 0.25. All the other parameters used in the model can be computed directly from the parameters presented above, so they will be omitted here.

5.5 Shocks

In this section, we will calculate the impulse responses to different shocks. For each shock, we will calculate its rational inattention counterpart and plot impulse responses for both the perfect attention problem as well as for the rational inattention one. The choice of autoregression functions is rather arbitrary and serve as an initial point from which the signals under limited attention are constructed. We know from earlier that in case of ARMA(p,q) processes in the form

$$X_t = \gamma_1 X_{t-1} + \cdots + \gamma_p X_{t-p} + \theta_0 \varepsilon_t + \cdots + \theta_q \varepsilon_{t-q},$$

which is the form usually expected when assuming technology and monetary shocks dynamics. Our agent receives a signal from such process in time t corresponding to

$$S_t = a_0 X_t + \cdots + a_{p-1} X_{t-(p-1)} + b_0 \varepsilon_t + \cdots + b_{q-1} \varepsilon_{t-(q-1)} + \psi_t,$$

Remember that choice of the signal weights relies entirely on our agent and his willingness to spend attention k towards this signal. In the case of unlimited attention devoted to a variable of interest X_t, i.e. to have our k approaching infinity, we would then in fact have $S_t = X_t$, the situation which we will from now on call the perfect attention case. In this particular scenario the rational inattention theory would coincide with prevalent theory of perfect attention and thus there would be no need to model signals received by agents separetely. We also omit the case when the underlying process is the random walk or autoregressive of order 1, even though these processes are not of any interest, since the receive signals are under limited information still spoiled by a noise represented by ψ_t, whose variance is largely affected by k as we discussed earlier. Because we model only the cases where a signal is changed due to a shock in the underlying process X_t and not by ψ_t, we will not consider these case here and therefore we start with an autoregressive process of order 2 with very common structure used in time series literature, particularly

$$Y_t = 0.8Y_{t-1} + 0.09Y_{t-2} + \varepsilon_t$$

Such process is a stable one, without any humps and with geometrically decreasing autocorrelation function, so it should, in theory, does not require too much attention compared to other more complex processes to receive signal with some given precision. For calculating the structure of signal that agent receives about the process, we use the Rational inattention filter used by Maćkowiak et al. (2017) described in the previous chapter. We set κ for this case to be equal to 1. The calculation was made using Matlab software. The resulting signal has the form

$$S_t = Y_t + 0.0314Y_{t-1} + \psi_t$$

$$k = 1, \sigma_\psi^2 = 0.4061, \Sigma_1 = \begin{bmatrix} 1.2027 & 0.2457 \\ 0.2457 & 0.3008 \end{bmatrix}$$

As we can see, an agent does not pay attention only toward Y_t in time period t, but due to his limited attention also toward variables in the previous time period, which is in this case Y_{t-1}. The reasonig behind this was already suggested in the chapter devoted to Rational inattentiion filter, because of the fact that agent is limited in his attention, he cannot just simple receive signal in the form $S_t = Y_t + \psi_t$, since extrapolating the information from the past improves his ability to predict the signals in future periods.

It is interesting to see what happens to the signal with different values of information processing constraint given by parameter κ. According to the theory, we should see with increasing κ the received signals to be closer to the perfect attention case, so that the signal processing error should diminish as well as the attention devoted toward other variables than Y_t. In next we calculated the signal for the same underlying process as in the previous case, but this time with parameter κ equal to 2, which corresponds to the channel capacity two times bigger than before. The resulting signal is

$$S_t = Y_t + 0.0075Y_{t-1} + \psi_t$$

$$k = 2, \sigma_\psi^2 = 0.0696, \Sigma_1 = \begin{bmatrix} 1.0426 & 0.0524 \\ 0.0524 & 0.0652 \end{bmatrix}$$

Now, in this case, we see the signal to be in the line with theory predictions made above, so the weight given to Y_{t-1} decreased rapidly as well as the variance of the signal error σ_{ψ}^2, which in turn means much higher precision of our signal about the underlying process.

In the picture below we depicted the difference in impulse responses from the technology shock of 1 standard deviation in the New Keynesian model described above. The calibrated values for parameters in the model are the same as described in the chapter about calibration. The solid line is representing the rational inattention case and the dashed line perfect attention case as defined before. From the picture, we can see that differences between the two cases are not very profound, which can be given by the fact that the underlying process is not too complex for an agent to make quite an accurate signal extraction. In later cases, where we take richer structures of autoregressive processes, the differences became more noticeable. The impulse responses for the rational inattention case are made for κ=1.

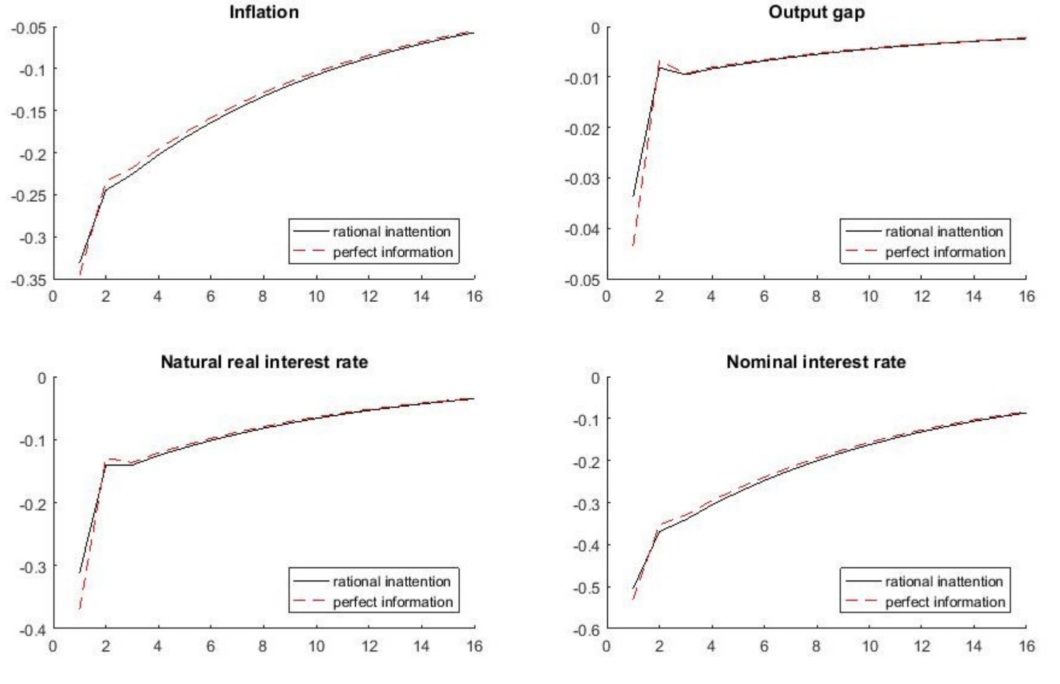

Figure 1, Impulse responses, k=1, own calculations

Next case we will consider is an autoregressive process of order 2 in the form

$$Y_t = 1.5Y_{t-1} - 0.8Y_{t-2} + \varepsilon_t$$

This process is stationary as was the case of the previous one, but this time the process allows for more curved behaviour, which in turns increases the demands for an agent to his attention in order to get the desired precision of the signal given the underlying process. With the same calculations made as before, an agent with channel capacity κ=1 receives the signal as

$$S_t = Y_t - 0.4179Y_{t-1} + \psi_t$$

$$k = 1, \sigma_\psi^2 = 0.4676, \Sigma_1 = \begin{bmatrix} 1.7601 & 0.5310 \\ 0.5310 & 0.4951 \end{bmatrix}$$

Now it is apparent that the weight devoted to Y_{t-1} is much larger than in the previous case, which means that for given parameter κ, we should see more profound differences between the signal and the underlying process. The same is also true for the the variance of signal processing error, which is now larger compared to the previous case for the same level of information processing capacity κ. In the next we do the same thing we did before, we increase the channel capacity of an agent two times and see what happens to the received signal. Now it has the form

$$S_t = Y_t - 0.1883Y_{t-1} + \psi_t$$

$$k = 2, \sigma_\psi^2 = 0.0757, \Sigma_1 = \begin{bmatrix} 1.1702 & 0.0985 \\ 0.0985 & 0.0754 \end{bmatrix}$$

As was the case before, the signal with increased channel capacity moved closer to the true process and the variance of the signal error tends to zero. Nevertheless, even as the precision of the signal improved considerably with increased κ, there are still significant differences as almost one-fifth of the Y_{t-1} takes place in the signal. This could be of course improved further by allowing for higher capacity channel κ.

In the picture below, we can see considerable differences regarding impulse responses in our New Keynesian model. Due to a rationally inattentive agent, the impulse responses got flattened.

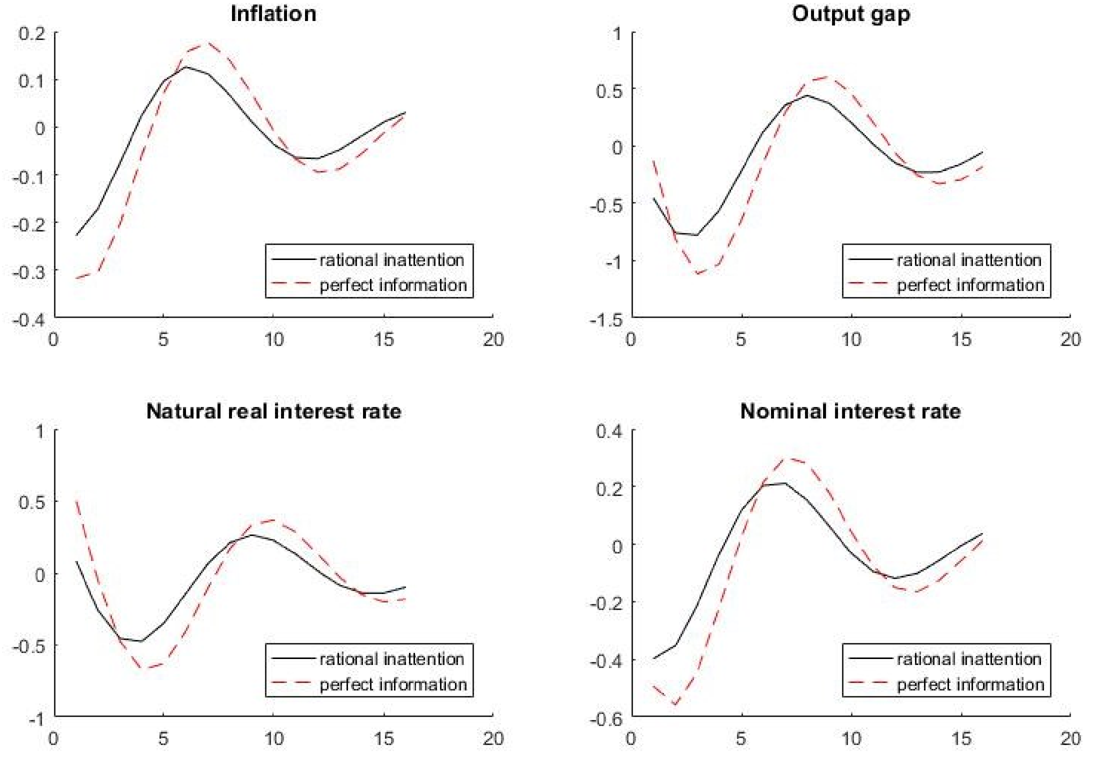

Figure 2, Impulse responses, k=1, own calculations

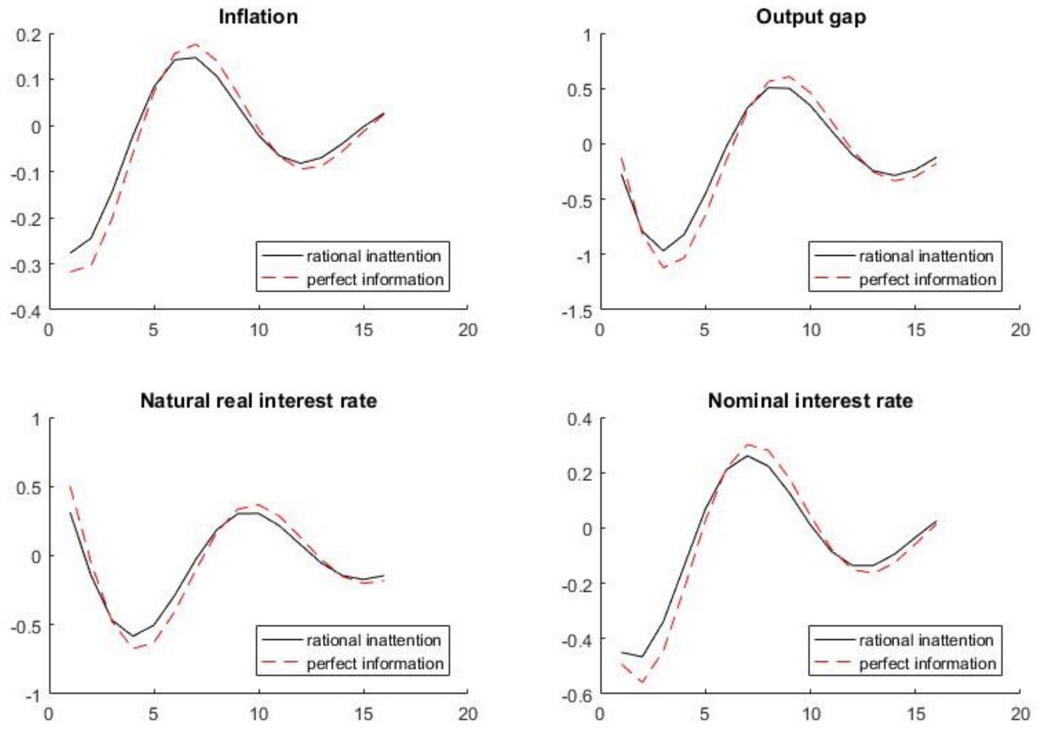

Figure 3, Impulse responses, k=2, own calculations

Now the last case we will consider will be the autoregressive moving average process of order (2,1). This process is even more complex as innovations in earlier periods are directly present in the equation, which according to the theory means that they should be present in some form in the signal. We take ARMA(2,1) to have the form

$$Y_t = 0.9Y_{t-1} - 0.3Y_{t-2} + \varepsilon_t + 0.2\varepsilon_{t-1}$$

The corresponding signal for κ=1 will be

$$S_t = Y_t - 0.1071Y_{t-1} + 0.2665\varepsilon_t + \psi_t$$

$$k = 1, \sigma_\psi^2 = 0.6282, \Sigma_1 = \begin{bmatrix} 1.344 & 0.3181 & 1.0005 \\ 0.3181 & 0.3543 & -0.0004 \\ 1.0005 & -0.0004 & 0.9994 \end{bmatrix}$$

As the complexity of the underlying process increased, the same is true for the variance σ_ψ^2, which is now considerable higher at the same level of κ compared to the previous cases. What is different now is that current innovations are directly present in the signal and not indirectly through the underlying process, which means we have another variable ruffling the signal.

$$S_t = Y_t - 0.0328Y_{t-1} + 0.2897\varepsilon_t + \psi_t$$

$$k = 2, \sigma_\psi^2 = 0.116, \Sigma_1 = \begin{bmatrix} 1.0808 & 0.0689 & 1.0003 \\ 0.0689 & 0.0715 & -0.0003 \\ 1.0003 & -0.0003 & 0.9996 \end{bmatrix}$$

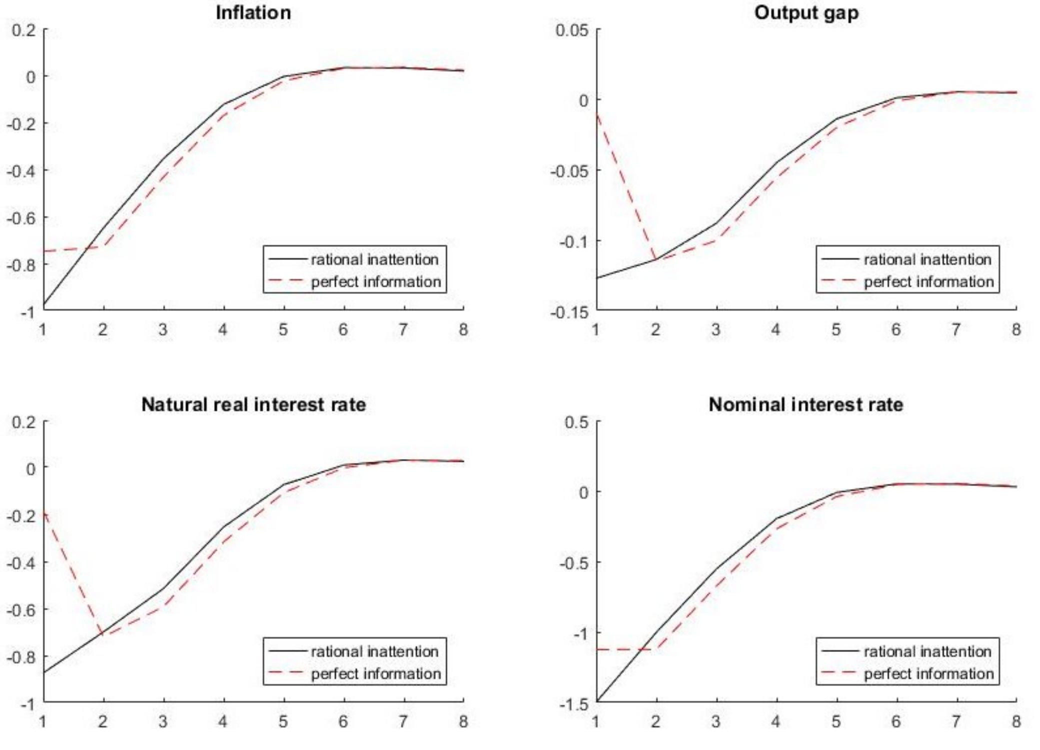

Figure 4, Impulse responses, k=1, own calculations

6 Conclusion

In this thesis, we used rational inattention filter as was suggested in Maćkowiak et al. (2017) to derive optimal signals received by a representative agent about exogenous variables of interest. We then applied those signals to the New Keynesian DSGE model calibrated for the Czech economy to see if there would be some distinctions from the perfect attention case. We found out that for the case of AR (2) process with a little value put on the second variable, the differences from the perfect attention are not very profound. Nevertheless, there is still a motivation for an agent to increase his information capacity k due to decreased noise in the signal and thus making better predictions about the underlying process. This was shown as we increased the information capacity by two times, the resulting variance σ_{ψ}^{2} was less than a fifth of the former value. When the more richer structure of the AR (2) process was allowed, the difference from the perfect attention case gets bigger as well as a demand for increased information capacity. In ARMA (2,1) case, as is suggested by the theory, the received signal is spoiled by innovations directly in the function and not only indirectly through the underlying process. This means a greater distiction from the perfect attention case and more jumps in the impulse responses in the model.

What policy recommendations can be depicted from the above? The first are policies related to the increasing value of k for representative agents, which directly increases their precision of the received signals about the world around them. This is of course the huge question, because we would have to have complete utility function regarding all activities for each agent in order to say that he or she would be better of if devoting more attention toward economic variables and not for example to raising of their offsprings. Nevertheless, we can still say that overally increased information capacity make one better of since that one can then make more informed choices while left with extra information capacity which can then be used elsewere. This make ongoing research aimed to the increasing human cognitive capacities very valuable area to pursue for the mankind. The question is also largely related to the state government providing the basic education, whose enormous importance does not have to be mention. We could nevertheless speculate that for example, state-backed programmes targeted to fincancial literacy should lead to the incresed capacity k, altought it is not clear by which amount and for which costs. The endogenousness of the k was completely ommited in this thesis, but represents definitely a vital area for future research. What can rational inattention say about the topic is that the overall level of k can be endogenously decided, once we are aware about the structure of noises and have fully developed optimization problem, then the capacity k can be part of the cost function as anything else and have its value as shadow price in terms of money.

Another important topic relevant to the work is the variable responsible for the noise in the signal ψ_t. In our setup, the variable is independently identicaly distributed. Only part in the variable which could our representative agent influenced was a variance of the process σ_ψ^2. There are no doubts that the future research finds that the signals have richer structure than one suggested in this work, part of which could be covered by ψ_t. One of the suggestion made in this area is one from the Sims (2005) that an agent can partially ifluence other features of the signal such as its distribution or how it is skewed etc..The perceived signals about reality around us is one of the biggest question in economics as well as in other areas ranging from philosophy to biology and it is clear that a lot of interesting findinds from the future research is ahead. In the literature review section, we believe that we provided a strong motivation for rational inattention approach in various economic problems.

Bibliography

Bacchetta, P., Wincoop, E., 2005. “Rational inattention: A solution to the forward discount and predictability puzzles”, NBER Working Papers No. 11633, National Bureau of Economic Research, Inc.

Bodie, Z., Kane, A., Marcus, A. J., 2014. Investments (10th Ed.), McGraw-Hill

Brocas, I., Carrillo, J., 2003. “Information and Self-Control”. The Psychology of Economic Decisions, Oxford University Press, vol.1.

Caplin, A. S., Spulber, D.F., 1987. “Menu Costs and the Neutrality of Money”. Quarterly Journal of Economics, vol. 102, pages 703-726.

Carrillo, J., Mariotti, T., 2000. “Strategic Ignorance as a Self—Disciplining Device,”. Review of Economic Studies, vol. 67, pages 529—544.

Cover, T. M., Thomas, J. A., 1991. Elements of Information Theory. Wiley-Interscience, New York.

Deci, E., Ryan, R., 1985. Intrinsic Motivation and Self-Determination in Human Behavior, New York: Plenum Press.

Fama, E., 1984. “Forward and Spot Exchange Rates”. Journal of Monetary Economics, vol.14, pages 319–338.

Galí, J.,2008. "Monetary Policy Design in the Basic New Keynesian Model," Introductory Chapters,in: Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework, Princeton University Press.

Golosov, M., Lucas, R.E., 2007. "Menu Costs and Phillips Curves". Journal of Political Economy, University of Chicago Press, vol. 115, pages 171-199.

Huang, L., Liu, H., 2007. “Rational Inattention and Portfolio Selection”. The Journal of Finance, vol. 62, Issue 4, pages 1999-2040.

Laibson, D. I., 1997. “Golden Eggs and Hyperbolic Discounting”, Quarterly Journal of Economics, vol. 112, pages 443-477.

Levy, D., Smets, F., 2010. “Price Setting and Price Adjustment in Some European Union Countries: Introduction to the Special Issue”. Managerial and Decision Economics, vol. 31, pages 63-66.

Lucas, R. E., 1973. “Some International Evidence on Output-Inflation Tradeoffs". American Economic Review, 63(3), pages 326-334.

Lucas, R. E., 1976. “Econometric Policy Evaluation: A Critique”. The Phillips Curve and Labour Markets, vol. 1, pages 19-46.

Luo, Y., 2008. “Consumption Dynamics under Information Processing Constraints”. Review of Economic Dynamics, vol. 11, pages 366-385.

Maćkowiak, B., Wiederholt, M., 2009. “Optimal Sticky Prices under Rational Inattention,”American Economic Review, vol. 99(3), pages 769–803.

Maćkowiak, B., Wiederholt, M., 2015. “ Business cycle dynamics under rational inattention“. The Review of Economic Studies, vol. 82(4), pages 1502-1532.

Maćkowiak, B., Matějka, F., Wiederholt, M., 2017. “The Rational Inattention Filter“. Working Paper Series No.2007, January 2017, European Central Bank.

Mankiw, N. G., Reis, R., 2002. “Sticky information versus sticky prices: A proposal to replace the new Keynesian Phillips curve”. Quarterly Journal of Economics, vol. 117(4), pages 1295-1328.

Matějka, F., Sims, Ch. A., 2009. “Discrete actions in information-constrained tracking problems,” Discussion paper, Princeton University.

Matějka, F., McKay A., 2012. "Simple Market Equilibria with Rationally Inattentive Consumers," American Economic Review, American Economic Association, vol. 102(3), pages 24-29.

Matějka, F., 2015. "Rigid pricing and rationally inattentive consumer," Journal of Economic Theory, vol. 158(PB), pages 656-678.

Matějka, F., 2016. "Rationally Inattentive Seller: Sales and Discrete Pricing," Review of Economic Studies, Oxford University Press, vol. 83(3), pages 1125-1155.

Mehra, R., Prescott, E., 1985. "The Equity Premium: A Puzzle". Journal of Monetary Economics, vol. 15, pages 145-161.

Moscarini, G., 2004. "Limited Information Capacity as a Source of Inertia". Journal of Economic Dynamics and Control, vol.28, pages 2003-2035.

Peng, L., Xiong, W., 2006. "Investor Attention, Overconfidence And Category Learning," Journal of Financial Economics, vol. 80, pages 563-602.

Reis, R., 2006. "Inattentive Producers". Review of Economic Studies, vol. 73, pages 793-821.

Reis, R., 2006. "Inattentive consumers". Journal of Monetary Economics, vol. 53(8), pages 1761-1800.

Sargent, T. J., 1984. "Autoregressions, Expectations and Advice". The American Economic Review, vol. 74, issue 2, pages 408-415.

Shannon, C., 1948. "A Mathematical Theory of Communication". The Bell System Technical Journal, vol. 27, pages 623-656.

Sheshinski, E., Weiss, Y., 1977. "Inflation and Costs of Price Adjustment". The Review of Economic Studies, Volume 44, Issue 2, Pages 287–303.

Sims, C. A., 1992. "Policy Analysis with Econometric Models". Brookings Papers on Economic Activity, vol. 1, pages 64-107.

Sims, C. A., 1998. "Stickiness". Carnegie-rochester Conference Series On Public Policy, vol. 49(1), pages 317–356.

Sims, C. A., 2003. "Implications of Rational Inattention". Journal of Monetary Economics, vol. 50(3), pages 665–690.

Sims, C., A., 2005. "Rational inattention: A research agenda". Working Paper, Princeton University.

Sims, Ch. A., 2006. “Rational Inattention: Beyond the Linear-Quadratic Case.” American Economic Review Papers and Proceedings, vol. 96, pages 158-163.

Solorzano, J. D., 2017. “Sticky Prices, Sticky information or Rational Inattention? Evaluating Microfoundations”. Working Paper, University of Warvick

Thaler, R. H., 2000. “From Homo Economicus to Homo Sapiens,”. Journal of Economic Perspectives, vol. 14(1), pages 114—142.

Vissing-Jørgensen, A., 2002. “Limited Asset Market Participation and the Elasticity of Intertemporal Substitution”. Journal of Political Economy, vol. 110, No. 4, pages 825-853.

Wolman, A., 2007. “The frequency and costs of individual price adjustment”. Managerial and Decision Economics, vol. 28, pages 531–552.