Behavioral New Keynesian Macroeconomic Models
A Literature Review

Benedikt J. F. Höcherl

Tutored by Professor Rosemarie Nagel

Code: EWC04

Universitat Pompeu Fabra • End-of-Degree Paper in Economics • Academic Year 2019-2020

Abstract

The present paper tries to fill the gap of a missing up-to-date literature review of the most vanguard Behavioral New Keynesian (BNK) macroeconomic models. BNK models differ from traditional New Keynesian (NK) models insofar, as they challenge the notion of fully rational expectations of economic agents and introduce some kind of bounded rationality in the model. The models reviewed include *Adaptive Learning* by De Grauwe (2012), *Level K* by Mauersberger et al. (2020) and *Sparse Dynamics* by Gabaix (2019). The paper especially highlights different implications for monetary policy by critically comparing the heterodox BNK models to a standard version of the mainstream NK model. In doing so, it detects so far unidentified links and gaps in the current literature.

Keywords: Behavioral Macroeconomics, New Keynesian Models, Level K Reasoning, Sparse Dynamics, Adaptive Learning, Bounded Rationality

Contents

1 Introduction and Motivation

When conducting monetary policy, central banks use models to predict and estimate the impact their policy will eventually have on the economy and financial markets and thus, ultimately the impact it will have over individual lives. As all models, central banks' models are simplified versions of the world and rely on certain assumptions and specifications. Should those assumptions be flawed, the potential scope for erroneous decision making might increase and one could improve the ultimate outcome determined by decisions based on improved assumptions of the model. However, as expressed in Milton Friedman's classic *Essays in Positive Economics* (1953) the primary question when considering a model's assumption should not be so much whether the assumptions of the model are most realistic, but whether the model most accurately predicts the outcome of the research question at hand. This very methodological approach can also be applied to one of the crucial debates in modern macroeconomics, namely how to accurately incorporate (aggregate) behavior in the applied models.

With the advent of modern macroeconomics under Keynes's *General Theory* (1936) after the Second World War, human behavior was considered to obey to tendencies of *Animal Spirits* - herd like behavior of the masses when it comes to aggregate decision making in the markets. This view was in line with the back then contemporaneous psychological opinion of limited free will of the masses represented by the likes of Sigmund Freud (1917) and Edward Bernays (1947). However Keynes's notion came under severe attack in the 1970's and 1980's after Keynesian demand stimuli had failed to deliver the anticipated results during the infamous stagflation period. Under John Muth the rational revolution began and finally materialized in a new - rational expectations - approach led by Lucas and his disciples (De Vroey and Malgrange 2011). This approach under model consistent expectations implies the assumptions that "agents know the model and the probability distributions for the stochastic variables [and] agents are able to calculate the equilibrium of the model, and belief that all other agents will choose equilibrium strategies, in essence implying a coordination of expectations" (Driscoll and Holden 2014). By the 1990's a newly, micro-founded, class of macroeconomic models had been developed that today are known as New Keynesian (NK) DSGE models as distilled for example in Galí (2008) and Woodford (2003). This class of models can be considered as part of the main modern macroeconomic workhorses and finds wide use in central banks around the developed world.

However, over the last two decades economists have used an ever increasing number of results from behavioral and experimental economics in order to improve the mainstream versions of macroeconomic models. Driscoll and Holden (2014) identify that - in line with Friedman's notion as stated in the beginning of the introduction - these modifications of orthodox models are driven by two main motivations: models based on optimizing behavior under rational expectations might systematically account wrongly for real world observations and behavioral macroeconomics models might predict empirical data better. Thus, it might methodologically make sense to improve upon traditional NK models accounting for insights from behavioral as well as experimental economics - creating *Behavioral New Keynesian* (BNK) models. More so, insights from behavioral economics have experienced ever more scientific support across the last two decades by cognitive psychologists (see for example Tversky and Kahneman 1981, Thaler 1994 and Damasio 2003) and experimental economists such as for example Nagel and Coricelli (2012), as well as Duffy and Arifovic (2018) who find large differences in individuals' ability to use higher levels of reasoning in a laboratory neural-experiment using functional magnetic resonance imaging. On the basis of such new behavioral insights, researchers try to improve the microeconomic foundations underlying the macroeconomic models in order to obtain better predictions, which in turn will potentially improve policy realizations.

In his fundamental survey on experimental macroeconomics *Macroeconomics: A Survey of Laboratory Research* John Duffy (2016) states that "the field of macroeconomics is among the final frontiers in the continuing transformation of economics into an experimental science [where] another revolution in macroeconomic methodology may well be at hand" and George Akerlof claimed already almost 20 years ago in his by now classic paper *Behavioral Macroeconomics and Macroeconomic Behavior* (2001) that "if there is any subject in economics which should be behavioral, it is macroeconomics".

Besides the purely academic debate, however, it was in particular the eruption of the financial crisis in 2007 that questioned the hitherto well functioning NK macroeconomics. Until 2007 many macroeconomists attributed the success of the *Great Moderation* partly to "scientific insights provided by modern macroeconomic theory" (De Grauwe 2012). Aforementioned NK models are characterized by rational agents - aware of the assumptions and design of the very model they inhabit - whose behavioral implications lead to stable economic aggregates, if not disturbed by an exogenous shock every so often. However, with the 2008 *Great Recession* the view of a self-stabilizing economy inhabited by rational individuals naturally came under heavy criticism and confirmed theorists who were suspicious of the model's assumptions already before the crisis, fostering research on bounded rationality in macroeconomics.

As it becomes obvious from the above mentioned, macroeconomic modeling matters a great deal not only for academic discussion but because it shapes the debate policy makers lead when thinking of any kind of monetary or fiscal policy in order to stabilize or revitalize the economy during a crisis. It also becomes clear that new behavioral versions of widely used macroeconomic models potentially might lead to either more robust or even fundamentally different conclusions than their mainstream counterparts.

Especially New Keynesian models are widely understood and are being used by policymakers (for a list of central banks using NK DSGE models see Chung, Herbst and Kiley, 2014). Therefore, the aim of this paper is to explore the implications of the introduction of insights from behavioral and experimental economics into the New Keynesian framework for monetary policy, by the means of a literature review.

For this purpose I will critically review three distinct types of vanguard Behavioral New Keynesian models and examine what their advantages and disadvantages are compared to the mainstream NK benchmark. In doing so the paper identifies how these models help to explain more aspects of macroeconomic phenomena, as observed in the real world. Finally, the review detects so far unidentified links and gaps in the current literature.

In the present paper I will address: firstly, Paul De Grauwe's Adaptive Learning approach as presented in his book *Lectures on Behavioral Macroeconomics* (De Grauwe 2020). Secondly, fundamental insights for BNK models of Felix Mauersberger, Rosemarie Nagel and Christoph Bühren from their survey *Bounded rationality in Keynesian beauty contests: a lesson for central bankers?*. And lastly Xavier Gabaix's sparse dynamics model as presented in his paper *A Behavioral New Keynesian Model* (Gabaix 2019). All BNK models are being compared, analyzed and criticised with respect to the NK benchmark.

The literature review is structured as follows: section 2 presents the terminology and methodology used to chose the three models examined amongst the wide literature on behavioral macroeconomics. Section 3.1 brushes up a common version of the canonical NK benchmark model. Sections

3.2 through 3.4 introduce De Grauwe's Adaptive Learning Model (2012), Nagel's Level K-Thinking Model as in Mauersberger et al. (2020) and Gabaix's Sparse Dynamics Model (2019), respectively. Section 7 provides a brief discussion of the macroeconomic policy implications. Section 8 concludes.

2 Methodology and Terminology

This section presents the terminology of the literature review and the methodology used in order to select the papers to be surveyed.

2.1 Methodology

Given the widespread impact of behavioral and experimental economics on macroeconomics it has been necessary to narrow the material somewhat selecting the relevant publications by following a coherent methodological approach.

Amongst papers on behavioral macroeconomics it is worthwhile mentioning the salient strand of work on models creating sunspot equilibria, which are able to follow Keynes's (1936) awareness of animal spirits by producing multiple rational expectation equilibria under extraneous variables - the "sunspots" (De Grauwe 2012). Amongst influential papers are the likes of Shell (1977), Azariadis (1981) as well as Azariadis and Guesnerie (1986) and - providing a modified version of the NK framework - Clarida et al. (1999).

Another branch of publications in behavioral macroeconomics is investigating global indeterminacies - Howitt and McAfee (1992), Evans et al. (1998) and Evans and Honkapohja (2001) are worthwhile examples of this kind of multiple equilibrium modeling.

However, literature in this review has been selected and separated from the aforementioned by several criteria. Firstly, I only consider works published after 2012 (explicit timeframe from 2012 to 2020). Secondly, I focus on papers that apply insights from behavioral or experimental economics explicitly within the New Keynesian framework providing the necessary tools for analyzing general equilibria - compared to a large strand of literature that does so in partial equilibrium analysis[1].

The paper uses both, forward snowballing (finding citations to a paper), as well as backward snowballing (finding citations in a paper). However, all meeting the aforementioned criteria, the specific papers of De Grauwe (2012), Mauersberger et al. (2020) and Gabaix (2019) have been chosen be-

[1] For an extensive review of this kind of literature see Duffy (2016) and Hommes et al. (2014).

cause of their contentwise diverse and mathematically rigorous contributions to macroeconomics. It is worth mentioning that Branch and Evans (2007) introduced a similar regime-switching approach as De Grauwe (2012), however, besides slight theoretical differences, the paper did not meet the methodological time frame criteria of this review.

The critiques of the models reviewed will follow the methodology of *positive economics* as set out by Friedman (1953), where ”economics as a positive science is a body of tentatively accepted generalizations about economic phenomena that can be used to predict the consequences of changes in circumstances“ and a ”theory is to be judged by its predictive power for the class of phenomena which it is intended to ”explain““. Thus, here a model will be positively evaluated when it can be said to be simpler, when ”the less the initial knowledge needed to make a prediction within a given field of phenomena“ and more fruitful, that is ”the more precise the resulting prediction, the wider the area waiting which a theory yields predictions, and the more additional lines for further research it suggests.”

A potential weakness in the above described methodology could be the leaving out of earlier relevant literature, since the considered time frame is relatively short, nonetheless the reader is also referred extensively to older works that can be accessed via the reference section at the end of the paper. Another point is that the positivist critique of economic models can be dismissed itself for example by the classic criticism of positive economics provided by Lewis E. Hill (1968) who states that ”although positive economics is capable of making reliable predictions, it is completely incapable of explaining why the predictions are fulfilled“. However, as this paper assumes mostly positivist criticism of heterodox models from mainstream theorists it is reasonable to directly anticipate it in the review. The strength in the paper's methodological approach can be identified in that it focuses on the most recent literature in BNK modeling and therefore generates practical value. Also the clear delimitation from other strands of literature as mentioned above facilitates compact particulars to the reader.

2.2 Terminology

Since for this review paper, a wide range of literature is drawn from different authors across the economic profession, there might in principle be different terminologies used. In order to bring readers of this review on the same page and to guarantee a clear understanding of terminology, I will outline in the following some of the key technical terms according to the definitions of well established economists which are used throughout the text.

-*Behavioral Economics*: the study of "the combination of psychology and economics that investigates what happens in markets in which some of the agents display human limitations and complications." (Thaler and Mullainathan 2001)
- *Behavioral Macroeconomics*: the study of the effects of psychological and cognitive limitations of individuals and their decisions on aggregate economic phenomena and "how those predicted phenomena vary from those implied by classical economic theory." (Lin 2012)
-*Experimental Economics*: "the study of individual decision making in which an isolated individual chooses many alternatives that have a monetary (or commodity) value defined quantitatively by the experiment." (Smith 1990)
-*Experimental Macroeconomics*: "a sub-field of experimental economics that makes use of controlled laboratory methods to understand aggregate economic phenomena and to test the specific assumptions and predictions of macroeconomic models." (Duffy 2006)
-*(Classical) Economic Rationality*: "postulate[s] that people behave in rational ways and consider options and decisions within logical structures of thought, as opposed to involving emotional, moral, or psychological elements." (Michalos 2014)
-*Bounded Rationality*: "willingness of agents to maximize their economic performance subject to their model specific cognitive limitations, usually using some kind of simple heuristic." (De Grauwe 2012)
-*Rational Expectations*: "equality between agents' subjective probabilities and the probabilities emerging from the economic model containing those agents" (Savage 1954). This "implies that agents are assumed to know the model they inhabit and the probability distributions for the stochastic variables." (Driscoll and Holden 2014)
-*Bounded Rational Expectation*: "expectations about future values of variables formed under cognitive limitations of the decision maker — limitations of both knowledge and computational capacity." (Simon 1957)

3 Behavioral New Keynesian Macroeconomic Models

3.1 The Benchmark New Keynesian Rational Expectations DSGE Model

In this section I briefly brush up the textbook version of a rational expectation New Keynesian DSGE model [2] as distilled by Galí (2008), which will be used as a benchmark for comparisons to the other models introduced throughout the paper. BNK models mostly direct their criticism towards this very benchmark.

The canonical NK model consists of the following three equations:

$$\tilde{y}_t = E_t y^n_{t+1} - \frac{1}{\sigma}(i_t - r^n_t - E_t \pi_{t+1}) + \epsilon^{IS}_t \tag{1}$$

$$\pi_t = \kappa \tilde{y}_t + \beta E_t \pi_{t+1} + \epsilon^{PC}_t \tag{2}$$

$$i_t = \rho + \phi_\pi \pi_t + \phi_y \tilde{y}_t + \epsilon^{TY}_t \tag{3}$$

where (1) is the so called Dynamic IS equation representing aggregate demand, (2) represents aggregate supply in form of the NK Phillips curve and (3) is central bank behavior as expressed by a Taylor Rule (Taylor 1993). Within the dynamic IS curve (1), $\tilde{y}_t$ represents the output gap (difference between natural and actual output), y^n_t is the natural output, i_t is the nominal interest, r^n_t the natural rate of interest and π_t inflation. E serves as the rational expectations operator and the subscripts t for the different time periods. The IS curve is derived from the NK-Euler equation, where fully model-informed forward looking rational agents act in a utility-maximizing manner. In that sense lagged output enters the equation due to habit formation of the agents in order to introduce certain inertia to the model (Galí 2008, chapter 8). When future aggregate income increases (corresponding to an increase in future output gap), agents spend more on consumption.
The NK Phillipys curve (2) (Galí 2008, chapter 3) is derived from the assumption of profit-maximizing firms under Calvo price-setting (Calvo 1983), where only a fraction of firms changes their prices in any period t. In that sense, κ and β are functions of structural parameters indicating the level of price stickiness due to Calvo pricing (Hornstein 2008). Together, (1) and (2) solve for

[2] DSGE stands for *dynamic stochastic general equilibrium* and is a widely used approach in modeling macroeconomic relationships, where expectations are the main channel through which policy affects economy (for a detailed explanation see Romer 2012, Chapter 7).

the equilibrium values of output gap and inflation.

Within the Taylor rule (3), ϕ_π and ϕ_y describe the intensity of the central bank's reaction to the economy's deviations from the target inflation rate and output gap, respectively. The intercept ρ makes the rules consistent with a zero inflation steady state (Galí 2008, page 64). In that sense, under strict inflation targeting, the central banks sets $\phi_y = 0$. In general the so called *Taylor principle* - implying $\phi_\pi > 1$ - needs to hold for model stability. Thus, should measured inflation surpass its target, the central bank reacts by increasing the nominal interest rate by intensity ϕ_π. The same holds for the output-gap under lenient inflation targeting.

Equations (1), (2) and (3) come with with stochastic random shocks ϵ_t^{IS},ϵ_t^{PC} and ϵ_t^{TY}, respectively. These shocks all have mean zero, a constant standard deviation, are i.i.d. distributed and can be interpreted as demand shocks (eg. preference shocks), supply shocks (eg. post push shocks) and interest rate shocks (eg. change in policy environment).

In the model it is price stickiness that makes money *non-neutral* and allows monetary policy to affect real variables (see Galí 2010). Furthermore, *Ricardian Equivalence* [3] holds. A common choice of calibrated parameters for numerical solution to all equations seen above are those suggested in Galí (2008). For a more detailed digest of the NK DSGE Model also see Driscoll and Holden (2014).

3.1.1 Analysis and criticism of the New Keynesian Rational Expectations DSGE

As already touched on in the introduction, a widely stated criticism of the NK DSGE model is limited realism of its assumptions - especially the notion of the fully informed model optimizing agent. In his extensive criticism of the NK DSGE model, De Grauwe (2012) states under the model's specification of rationality there cannot exist endogenous dynamics of the economy. Business cycle movement consequently only can arise because of exogenous shocks (ϵ_t^{IS},ϵ_t^{PC} and ϵ_t^{TY}) - if there are no shocks, there are no ripple effects causing business cycles. The problem here is that mainstream theory delivers no explanation for where these random shocks come from. Thus, there is no possibility for policies to identify and directly attack the origins of such shock.

Following a positivist critique, all this would not impose a problem if the economy behaved *as if* the model's assumptions were true and predictions were to be delivered with utmost accuracy. However, using data from US Department of Commerce and Congressional Budget Office (1960-2009),

[3] the concept of *Ricardian Equivalence* is illustrated in greater detail in section 3.4

De Grauwe (2012) finds that the rational NK model fails to predict non-normality in the output gap and underestimates its autocorrelation (applying calibrated parameters as in Galí 2008). Using a normal distribution for predictions on non-normal data might underestimate the probabilities for significant booms and busts in the economy. Therefore, one could argue that the NK model fails to deliver sufficiently accurate empirical predictions and thus is to be challenged as a trustworthy macroeconomic forecasting tool.

3.2 De Grauwe's Adaptive Learning Model

In this section I present Paul De Grauwe's approach to behavioral macroeconomic modeling as developed in his book *Lectures on Behavioral Macroeconomics* (2012). The author motivated his work by noting that the mainstream NK model fails to predict sufficient amounts of auto-correlation and non-normality in the output gap for US data from 1960 to 2009, whereas the data indicates the opposite (rejecting normality with a *Jaque-Bera test* giving a p-value=0.027). Therefore, the author calls for a behavioral enrichment in order to improve current orthodox macroeconomic modeling and the resulting monetary policy implications. De Grauwe's main alternative proposals are outlined in the following.

De Grauwe specifies the overall economic model dynamics as:

$$y_t = a_1 \bar{E}_t y_{t+1} + (1 - a_1) y_{t-1} + c_2 (r_t - \bar{E} \pi_{t+1}) + \epsilon_t^{IS} \tag{4}$$

$$\pi_t = b_1 \bar{E}_t \pi_{t+1} + (1 - b_1) \pi_{t-1} + b_2 y_t + \epsilon_t^{PC} \tag{5}$$

$$r_t = c_t (\pi_t - \pi^*) + c_2 y_t + c_3 r_{t-1} + \epsilon_t^{TY} \tag{6}$$

where in contrast to the benchmark model, the expectations operator $\bar{E}_t$ stands for bounded rational expectations. The nominal interest rate is given by r_t and lagged output is added to the IS curve. Also lagged interest rate is added to the Taylor rule describing *interest rate smoothing* by the central bank. Furthermore, the model contrasts in that heterogeneous agents choose - via heuristic switching - one out of two simplified rules that best predicted past inflation and output gap. It represents thus an *adaptive* learning version of the switching model as introduced by Brock and Hommes (1997) implemented in a NK framework. Learning in that sense is rational, as agents try to optimize their forecasting results acknowledging their own cognitive limitations.

The model is inhabited by two types of agents: firstly, the "fundamentalists", who predict the output gap to be the steady state value of the output gap of the economy, thus they use $\bar{E}_t^f y_{t+1} = 0$ (steady state output gap normalized to 0). Secondly, the "extrapolists", who assume future output gap to correspond to last period's output gap, thus using $\bar{E}_t^e y_{t+1} = y_t - 1$. Finally, the overall market forecast for future output gap is described by a weighted combination of the two forecasting heuristics: $\bar{E}_t y_{t+1} = \alpha_{f,t} E_t^f y_{t+1} + \alpha_{e,t} E_t^e y_{t+1}$. One interprets $\alpha_{f,t}$ and $\alpha_{e,t}$ as the probability of being either a fundamentalist or extrapolist at any given time point t - the probability being determined by the forecasting performance of either rule observed so far and measured by its mean squared error:

$$U_{f,t} = -\sum_{k=0}^{\infty} \omega_k [y_{t-k-1} - \bar{E}_{t-k-2}^f y_{t-k-1}]^2 \tag{7}$$

$$U_{e,t} = -\sum_{k=0}^{\infty} \omega_k [y_{t-k-1} - \bar{E}_{t-k-2}^e y_{t-k-1}]^2 \tag{8}$$

where ω stands for increasing weighted "forgetfulness" with respect to past forecasting errors. However, the model does not finish here. Instead of letting his agents decide exclusively on the basis of higher forecasting utility (what purely rational agents would do), De Grauwe introduces a mechanism that only lets a part of the model's inhabitants switch forecasting regime and in doing so simulates more boundedness to rational behavior. The probability being part of the fundamentalist group is expressed in discrete choice specification as:

$$\alpha_{f,t} = \frac{\exp(\gamma U_{f,t})}{\exp(\gamma U_{f,t}) + \exp(\gamma U_{e,t})} \tag{9}$$

where the parameter γ stands for agents' willingness to learn; $\gamma = 0$ means there is no learning at all and $\gamma \to \infty$ means agents choose randomly between rules. Accordingly, the probability of being an extrapolist is $\alpha_{e,t} = 1 - \alpha_{f,t}$.

The same mechanism as described above also applies to the expectation formation on inflation.

3.2.1 Analysis and criticism of De Grauwe's Adaptive Learning Model

In *Lectures on Behavioral Macroeconomics* (2012) Paul De Grauwe provides a comprehensive approach to macroeconomic modeling under bounded rationality within a New Keynesian framework.

The author is able to account for waves of optimism and pessimism in his model and derives particular insights for monetary policy. Firstly, under perfectly credible inflation targeting *animal spirits* are weak and the *divine coincidence* holds: there is no trade-off between inflation and output gap (in a well defined region). However, at the same time, strict inflation targeting is unable to provide full inflation credibility. This finding contrasts its benchmark NK counterpart, where in principle the traditional interpretation of the Phillips curve holds and a central bank only can stabilize output by paying with an increase in inflation. Secondly, *self-fulfilling* booms and busts in the economy are created by the dominance of extrapolists (expecting a positive output gap) or of fundamentalists (pessimistic when current output gap is above its steady state level), respectively. Thereby the model is able to generate an endogenously driven business cycle. An important consequence hereby is that monetary policy can be rendered ineffective hinging on the level of optimism or pessimism in the economy. Thirdly, considering shocks to the economy, their consequences are qualitatively similar to those in the NK model, however, quantitatively the BNK version yields wider short run fluctuations to the impact of productivity and monetary policy shocks. Thus, these shocks - enhanced by animal spirits - are less predictable than their NK counterpart and stem from behavioral and not parameter uncertainty. A very important insight here is that shocks' impacts are *time dependent* in the sense that it matters in which mood the economy currently is. For monetary policy this also means that its effectiveness is less predictable than in the mainstream world when the central bank has no profound knowledge on the markets' moods. The same holds for *Ricardian Equivalence* in de Grauwe's model: it might exist when the economy is riding a wave of optimism and thus increasing the fiscal multiplier, however, fiscal stimulus might as well be totally ineffective in other scenarios. In general the author supplies arguments for a central bank targeting both, inflation *and* output gap - in contrast to the *single mandate* doctrine.
My criticism to the technical argumentation of the model is that ultimately an exogenous shock is still needed in order to first push animal spirits in one direction or another and to "start" the endogenously contained business cycle - especially supply shocks are important in creating "wild" animal spirits (see De Grauwe 2012, chapter 7). This version of a BNK model thus cannot be said to be exclusively endogenously driven or *truly Keynesian.* Furthermore, animal spirits only work under price rigidity (see De Grauwe 2012, page 77) - either way one of the main pillars of traditional NK models. Finally, even though De Grauwe provides an extension of his model with more than

two types of forecasters, it is doubtful to what extent the assumption of different homogeneous forecasting rules should really be realistic - especially as his results are neither being tested in laboratory nor in field experiments. I therefore do not see improvements in the underlying assumptions that clearly outperform the traditional version of the model in terms of realism. Nevertheless, from a positivist point of view, the model is to be appreciated since it appears to yield a great improvement in terms of predictive power on the data used by the author. Even if De Grauwe's "trial and error" approach does not account for completely realistic assumptions, as long as agents act *as if* they did learn this way, policy decisions can be improved through improved predictability. Also working with statistical non-normality allows for novel interpretations of monetary policy with consequences of interest.

3.3 Nagel's Level K Thinking Model

The following is devoted to presenting Rosemarie Nagel's *Level K thinking* model in the context of New Keynesian macroeconomics, especially as introduced in the paper *Bounded rationality in Keynesian beauty contests: a lesson for central bankers?* (2020) by Mauersberger, Nagel and Bühren. Level K thinking is a term from literature on game theory and experimental economics first coined by Nagel (1995). It describes a certain decision heuristic in situations where individuals are supposed to take into account the reasoning of other players as well, but frequently fail to choose a level of reasoning high enough to collectively reach the rational expectations equilibrium (REE). This might be due to increased cognitive costs to applying higher levels of reasoning or simply due to cognitive limitations of individuals (Alaoui and Penta 2016). Level K thinking is the underlying mechanism of a particularly well known game, namely the classic *Keynesian Beauty Contest Game* (BC) - established as a term by Nagel and Duffy (1997) - as introduced by Keynes in his *General Theory* (1936, Chapter 12), where he specifies the following game:

"*Professional investment may be likened to those newspaper competitions in which the competitors have to pick out the six prettiest faces from a hundred photographs, the prize being awarded to the competitor whose choice most nearly corresponds to the average preferences of the competitors as a whole; [...] It is not a case of choosing those [faces] which, to the best of one's judgment, are really the prettiest, nor even those which average opinion genuinely thinks the prettiest. We* *have reached the third degree where we devote our intelligences to anticipating what average opinion expects the average opinion to be. And there are some, I believe, who practice the fourth, fifth and higher degrees.*"

After all the Keynesian BC reflects the belief that investment decisions are significantly influenced by expectations on other investors' credence - and not solely on the expected value of future discounted cash flows. This might result in markets under so called *irrational exuberance* (Le Roy 2004) creating *asset bubbles* and observed deviations of eg. stock prices from their fundamental value. Keynesian animal spirits thus can be thought of in terms of BC games, since agents base their decisions on the aggregated decisions of all other agents making investment prone to volatility and the proper understanding of mass psychology a crucial tool in the Keynesian world.
Mauersberger et al. motivate their approach especially by a growing literature on experimental findings that confirm Level K thinking heuristic on the individual and group (mass psychological) level (see Hommes et al. 2014). In that sense, NK models can be interpreted as games in which individuals are being asked to forecast the mean future inflation and output gap, which in turn depend on the forecasts of all the other agents.
In the following I will briefly introduce the reader to the technical aspects of the *step-level-reasoning* approach as presented by Mauersberger et. al. (2020), as well as its implications for monetary policy of the application in a BNK framework.
The easiest way to present the technical concept of Level K reasoning to the reader is by thinking through the classic version of the model as introduced by Nagel (1995):

"*Imagine every participant in a group of people is asked to choose a number from 0 to 100. The person whose number is closest to two thirds times the average wins a prize. If all players are rational, the optimal choice in this game is zero, which is the unique Nash equilibrium. This is because if everyone else chooses zero, the choice you need to make is zero. If everyone chooses a different number from zero, a player could deviate and win the game by choosing a lower number. The only number for which that is impossible is the lowest admissible number, zero. Consider the naive player who chooses a random number from 0 to 100 with equal probability. The expected value of this player's choice is 50. Such behavior is referred to as "level 0." A slightly more sophisticated* *player anticipates such behavior and best-responds to level 0 by choosing* $\frac{2}{3} \cdot 50$. *This player type is referred to as "level 1." Such behavior can also be anticipated, and some players may best respond to level 1 and choose* $(\frac{2}{3})^2 \cdot 50$. *Such a player type is called "level 2." Similarly, one can define k such thinking steps by "level k". A level k player chooses* $(\frac{2}{3})^k \cdot 50$. *The higher k is, the closer the behavior corresponds to the Nash-equilibrium of zero.* "

By slowly going through the above example, the reader should easily grasp the intuition behind BC games. This form of the game generally does not result in the Nash equilibrium of 0 when tried out with humans in experimental settings, however, equilibrium might be reached over time when individuals change their thinking steps by learning. Importantly, in the above setting *Level K reasoning* accounts for the fact that there are players that might assume that some of their opponents do to fully understand the game creating heterogeneity amongst players that ultimately prevents the realization of the Nash equilibrium. From here linking experimental Level K thinking to the micro-foundations of macroeconomics in order to create a BNK model is relatively intuitive: Woodford (2013) and Angeletos and Lian (2018) make use of the notion that "all individual expectations depend on all others' average expectations" (Mauersberger et al. 2020) and specify a model where agents' and companies' pay-offs depend on their own consumption and pricing decisions as well as on average, aggregated consumption and pricing decisions of all other agents and companies, respectively. In the same line, Mauersberger et al. establish their BNK model as a game where the target of forecasting are average values for inflation and output gap, where the best response of agent i at time point t becomes:

$$y_t^i = \hat{E}_t^i\{c + b \cdot f(y^1,_t, y_t^2, ..., y_t^N) + d \cdot g(y_{t+1}^1, y_{t+1}^2, ..., y_{t+1}^N) + \epsilon_t^i\} \tag{10}$$

where a, b, d are coefficients, $g(.)$ stands for the aggregated mean of forecasts and ϵ_t^i describes an idiosyncratic shock. The behavioral version of the NK model is then specified following an experimental approach where participants are supposed to forecast inflation and output gap in a learning-to-forecast-setting (see Hommes 2014). Players are subsequently paid according to the distance function $U_{t+1}^i = A - Q(\hat{E}_t^i \pi_{t+1} - \pi_{t+1})^2$ measuring the distance between individual forecasts and actual solutions, which are given by the IS (11) and Philips curve (12) of the following model:

$$y_t = \bar{y}^e_{t+1} - \sigma(i_t - \bar{\pi}^e_{t+1} - \rho) \tag{11}$$

$$\pi_t = \kappa y_t + \beta \bar{\pi}^e_{t+1} \tag{12}$$

$$i_t = \rho + \phi_\pi(\pi_t - \pi) \tag{13}$$

This model as first seen in Woodford (2003) departs from the representative agents assumption of the NK model seen in section 3 by introducing heterogeneous expectations, were $\bar{\pi}^e_{t+1}$ and $\bar{y}^e_{t+1}$ describe the means of these expectations across every agent. In the model, it is this specific notion of expectations that makes the interaction between individual and aggregate decisions possible. Furthermore, using the *ad hoc* assumption that "expectations of output gap equals its long run steady state value" Mauersberger et al. derive inflation to be forecasted by $\pi_t = c + d\bar{\pi}^e_{t+1}$, where π_t stands for the REE value of inflation and c is a constant [4]. The authors furthermore present a closer look at the micro-foundations of macroeconomics by adapting a BNK model as introduced by Woodford (2013) and Garcia-Schmidt and Woodford (2019) in which households and firms behave in a non-myopic way since optimal consumption and pricing decisions depend on realizations in the far future:

$$c^i_t = \lambda \hat{E}^i_t c_t - \sigma(i_t - \hat{E}^i_t \pi_t) + \beta \hat{E}^i_t c^i_{t+1} \tag{14}$$

$$\pi^j_t = \phi \hat{E}^j_t c_t + \xi \hat{E}^i_t \pi_t + \alpha\beta \hat{E}^j_t \pi^j_{t+1} \tag{15}$$

introducing *boundedly expectations* operators $\hat{E}^i_t$ and $\hat{E}^j_t$, where optimizing consumption and prices depend on a Level-K-like thinking process described by individual beliefs about the others' behavior on expected values of consumption or prices. Ultimately Mauersberger et al. present the aforementioned specifications in matrix form, as followed:

$$\begin{bmatrix} c^i_t \\ \pi^j_t \end{bmatrix} = \begin{bmatrix} -\sigma i_t \\ 0 \end{bmatrix} + \begin{bmatrix} \lambda & \sigma \\ \phi & \xi \end{bmatrix} \begin{bmatrix} \hat{E}^i_t c_t \\ \hat{E}^i_t \pi_t \end{bmatrix} + \epsilon^i_t \tag{16}$$

where ϵ^i_t is assumed to be a idiosyncratic shock. Mauersberger et al. hereby take into account the important intuition derived by Woodford (2013) who interprets $\hat{E}^i_t c_t$ and $\hat{E}^i_t \pi_t$ as individual forecast choices, which opens the door to a model with endogenous dynamics not depending on exogenous shocks - similar to the one seen in section 4 by De Grauwe. Here, endogenous dynamics are provoked as agents try to infer the beliefs of shocks of the other players.

[4] where $c \equiv \frac{(1-\beta)+\sigma\kappa\phi_\pi}{1+\sigma\kappa\phi_\pi}\pi$

3.3.1 Analysis and criticism of Nagel's Level K Thinking Model

The BNK model presented in *Bounded rationality in Keynesian beauty contests: a lesson for central bankers?* (2020) by Mauersberger, Nagel and Bühren captures an intuitive aspect in an elegant way, namely that humans in their daily economic lives base their decisions not only on their own, but also on the others' actions - however, with individually distinct levels of sophistication. What is especially attractive about the approach is that it enhances the micro-foundation of macroeconomics by experimentally backed evidence.

The main conclusions for monetary policy from the paper are to be identified as: firstly, the existence of a unique rational expectations equilibrium may depend on the variation of the BC game that is being played (for an extensive list of variations of the BC games in the literature see Nagel, Mauersberger and Bühren, 2020, page 6). For example games introducing an idiosyncratic signal [5] (similar to a random shock) to the actual target might reach REE almost instantaneously. Secondly, the degree of the central bank's reaction to deviations of inflation from its target can transform the BNK model into a game of either pure strategic complementarity or a mix of *strategic complementarity and substitutability* [6] and thereby determining the existence of possible REEs. In that sense, when the *Taylor Principle* [7] holds, the BNK model is characterized by a mix of strategic complementarity and substitutability. Thirdly, due to "low level thinkers" there is increased inertia in adjustments of the economy after a shock. Since level 0 (the *naive level*) is regarded to be the previous period's average of a forecasted value, a shock to such a value makes the learning-to-forecast process start anew. Consequently, positive stimuli as well as inflationary effects of long run low nominal interest rates are less noticeable compared to the classic NK model. This is underlined by a so called *horizon effect*, where innovations far in the future only have a small effect on the present - the further away the innovation, the more levels of reasoning would have to be considered

[5] Within a reduced game form assuming stationary output gap.

[6] strategic complements and substitutes here stand for the incentive to adjust decisions to the average, or for the incentive to bet against the average in order to maximize pay-offs, respectively. Experimental evidence shows that in games with strategic substitutes convergence is much faster and therefore different from behavior in games with strategic complements in which behavior typically oscillates for a long time around the equilibrium (see Heemeijer et al., 2009).

[7] under a Taylor rule of $i_t = \bar{i} + \phi_\pi(\pi_t - \pi)$, where ϕ_π stands for the severity the central bank reacts to deviations off inflation from its target and $\phi_\pi > 1$.

by agents. Lastly, the Level K BNK model challenges the *Neo-Fisherian paradox* [8] because of the same reason.

Nevertheless, I criticise that the present model may come to significantly different results whenever the underlying BC game varies its specification (Mauersberger et al. 2020, p. 6). It appears that that this BNK model - without applying extra specifications to the BC game - enhances inertia in the adjustment process after shocks and in a sense repeats the notion of price stickiness by another route. Also, the real world counterpart to the above mentioned notion of "variations in idiosyncratic signals" remains a bit vague.

Another criticism - claimed and countered in Mauersberger et al. (2020) - to the model is that Level K implicitly assumes that a given player believes that all other players adopt a level lower than himself. However, the authors refer the reader to a survey by Crawford et al. (2013) that reviews successful attempts to solving this problem.

From a positivist point of view the question is what exactly the BNK model predicts better, since the mere notion of increased realism on the assumptions of human decision making would not justify its application. What stands out here is that, when introducing variations of the BC game, the model yields exceptional conclusions different from the ones obtained in the mainstream model such as the notion of strategic *substitutability* and *complementarity.* These conclusions potentially might inspire further fruitful research.

3.4 Gabaix's Spare Dynamics Model

In the last twenty five years the fields of behavioral and experimental economics as well as psychology have found compelling empirical evidence for the shortsightedness of human behavior in economic decision making (see eg. Thaler and Benartzi 1995 and Zhang and Yu 2013). However, what appears to be intuitive - that real human beings do not discount each future period in a perfectly exponential manner - is being systematically discarded by the rational expectations assumption of mainstream NK models. Xavier Gabaix picks upon exactly these findings on human myopic behavior in his paper *Behavioral New Keynesian Model* (2019) providing an application of myopic agents paying relatively less attention to changes in future interest rate, consumption

[8] The *Neo-Fisherian paradox* states that a negative interest rate gap is not a prerequisite for increasing inflation. More so, a negative interest rate gap is "said to lower inflation" (see Gerke and Hauzenberger 2017)

and shocks to the economy, compared to the fully rational NK counterpart. Technically, the paper offers a novel approach introduced by Gabaix via the so called *sparsity operator*, which constitutes a "less than fully attentive and rational version of the traditional *max operator*" (Gabaix 2014).

Since the paper is highly technical it is useful for reading ease to induce the model backwardly. The author specifies his BNK model as followed:

$$x_t = ME_t[x_{t+1}] - \sigma(i_t - E_t\pi_{t+1} - r_t^n) \tag{17}$$

$$\pi_t = \beta M^f E_t[\pi_{t+1}] + \kappa x_t \tag{18}$$

$$i_t = \phi_\pi \pi_t + \phi_x x_t + j_t \tag{19}$$

where, in comparison to the benchmark, x_t denotes the output gap, j_t is a constant and $M, M^f \in [0,1]$ are the aggregated levels of attention of consumers and firms (note that by setting $M = M^f = 0$, one obtains the classic NK curves); a table of calibrated parameter estimates is given by the author (see Gabaix 2020; page 14, 15). These aggregated levels stem from the introduction of an individual attention parameter for agents and the further presence of shortsighted firms. This parameter enters the model explicitly via the idea that agents see every variable more "dimly" the further in the future it is being realized, leading to:

$$E_t^{BR}[z(X_{t+k})] = \bar{m}^k E_t[z(X_{t+k})] \tag{20}$$

where $\bar{m} \in (0,1)$ is measuring attention to the future [9] depend, z is any variable to be predicted by agents or firms in the economy and E_t^{BR} is subjective expectation by a behavioral agent using a misspecified law of motion. It is worth mentioning that Gabaix specifies E_t^{BR} as "individual beliefs under a subjective model", implying the option of flexibly adapting distinct characteristics for different groups of agents.

However, implicitly the parameter enters the BNK model via a rewritten behavioral Euler equation, specified as:

$$\hat{c}_t = ME_t[\hat{c}_{t+1}] - \sigma\hat{r}_t \tag{21}$$

[9] Note that the aggregating $\bar{m}$ yields $M = \bar{m}$ and $M^f = \bar{m}(\theta + \frac{1-\beta\theta}{1-\beta\theta\bar{m}}(1-\theta))$.

yielding:

$$x_t = -\sigma \sum_{k \geq 0} M^k E_t[\hat{r}_{t+k} - \hat{r}^n_{t+k}] \tag{22}$$

where r^n is the natural rate of interest.

On the side of the firms, imperfect attention rules amongst producers with respect to marginal costs and inflation. In order to maximize profits, firms want to choose the optimal price, however, their decisions also are influenced by the misspecified law of motion, so that:

$$\max_{q_{it}} E_t^{BR} \sum_{\tau=t}^{\infty} (\beta\theta)^{\tau-t} \frac{c(X_\tau)^{-\gamma}}{c(X_t)^{-\gamma}} \upsilon(q_{it}, X_\tau) \tag{23}$$

What's more, as in the traditional model introduced in section 3, firms underlie a Calvo pricing regime with only a fraction of firms changing prices in any given period t.

In general, with the incapacity of accurately discounting far ahead shocks, the author introduces a form of hyperbolic discounting (as described by Laibson 1997), where agents discount in a time inconsistent manner. However, conveniently, if all of the above parameters turn out to be 1, the model defaults on the fully rational NK version.

3.4.1 Analysis and criticism of Gabaix's Spare Dynamics Model

In *Behavioral New Keynesian Model* (2019) Gabaix provides a compelling approach to behavioral macroeconomic modeling with the explicit goal of showing how bounded rationality affects monetary and fiscal policy. The paper is technically rigorous and introduces a new way of describing shortsightedness of economic agents' behavior as observed by a wide field of behavioral sciences. The main advantage this version of a BNK model yields is its very high degree of flexibility - especially due to its focus on various elements within the consumption function: hyperbolic discounting in multiple areas, where agents assign attention differently to the individual components of their consumption function. Agents are potentially shortsighted towards future adjustments of the nominal interest rate, income and shocks to the economy - however with varying degrees of shortsightedness to the respective variables. It is easy to see that, when adjusting all inattention parameters (M and M^f) to 1, the models defaults on the traditional NK framework with all of its assumptions. Therefore, Gabaix's proposition on formalizing bounded rationality in macroeconomic models provides an easy way for comparing heterodox and orthodox results. Furthermore the models facilitates experimentation with different experimental results simply by adjusting parameters and thus allowing policy makers to look at the consequences of different degrees and types of shortsightedness in an easy way.

The heterodox insights relevant for monetary (and fiscal) policy are numerous and are summarized subsequently. A lot of these insights can be derived after having proved that the Gabaix's model yields a unique steady state equilibrium solution. Firstly, when monetary policy is passive ($\phi_\pi = \phi_x = 0$ eg. at the zero lower bound) there is a unique equilibrium if agents are boundedly rational enough“ (see Gabaix, 2019, page 16), because agents react less strongly to the future. This implies that the impact of an increase in the nominal interest rate in $T(T > t)$ periods, today's inflation π_t only suffers a small effect. This stands in stark contrast to the model's rational NK counterpart, where the same interest rate shock provides much higher variation in inflation today. Following the same logic forward guidance by the central bank is less effective the more behavioral the agents are since distant policy adjustments are perceived less clearly. Secondly, optimal monetary policy is possible in the NK environment when the zero lower bound is not binding [10]. Thirdly, the central bank can react less to future cost-push shocks, the more shortsighted firms are. Fourthly, since consumers as well as companies are not able to anticipate tax increases after a fiscal expansion in a fully rational manner, the *Ricardian Equivalence* does not hold. Ricardian Equivalence states that rationally forward looking individuals would anticipate future tax increases after fiscal stimuli, thus leaving aggregate demand and output unchanged (see Barro, 1974). However, should this equivalence not hold this implies that during recessions aggregate demand can be stimulated by government spending - a well known technique first applied under President Hoover's *New Deal* during the *Great Depression* (see Gould, 1989) - in order to stabilize output. As the idea of using countercyclical government spending to alleviate economies during periods of *animal spirits driven busts* an essential notion of Keynes's *General Theory* (1936), Gabaix's BNK model can be claimed to be truly Keynesian. In mainstream NK models, however, this crucial Keynesian notion is not met, thus leaving fiscal stimuli ineffective. Fifthly, due to the relatively weakened power of monetary policy under the sparse model, there exists a perfect substitutability between monetary and fiscal policy. Sixthly, so called *helicopter drops of money* (direct government cash transfers to agents) are an optimal measure when optimal nominal interest rate is at the zero lower

[10](where optimal monetary policy is defined as perfectly replicating the flexible price equilibrium)

bound and inflation is low. Because of the non-existence of Ricardian Equivalence and an increased government spending multiplier, the best response is a nominal rate of zero and deficient spending to stimulate demand - a policy currently proposed in order to dampen the economic effects of the Coronavirus crisis (see Galí 2020). Lastly, Gabaix shows that with myopic agents and his sparsity specification of the model, *Fisher Neutrality holds.* In the aforementioned BNK model a permanent increase in nominal interest translates into a corresponding increase in inflation overtime, leaving real interest constant. This is also the result of Fisher's famous proposition (see Tobin, 2005) that yields: $r = i + \phi_t$.

However, what remains slightly odd is the double specification of myopic behavior with respect to firms in the Phillips curve. Not only does there exist the parameter M^f indicating shortsighted behavior of the companies, but also Calvo pricing - itself a version of shortsightedness of firms towards prices. Not only might this formulation lead to overly sticky prices in the economy when not updating parameters accordingly, it is simply expressing the same phenomena twice within the same equation making one of the specifications potentially obsolete.

Overall, the paper captivates with technical rigor and a variety of policy relevant results leaving little to be desired.

4 Discussion

Section 3 introduced heterodox behavioral macroeconomic models applying different approaches derived from insights from behavioral and experimental economics to the New Keynesian framework. These models contribute considerably to the current literature, however, they also face some critical limitations.

The major monetary policy results of the models analyzed in this paper are summarized in the following - always as compared to the mainstream NK framework of section 3:

- *De Grauwe's Adaptive Learning Model* enriches the literature by exogenously driven business cycles that explain why timing and market mood matter for the amplification of exogenous shocks. Monetary policy can be rendered ineffective in extreme states of optimism or pessimism in agents' beliefs - with heterogeneous agents learning adaptively from the past. The central bank may face a *convex* trade-off between inflation and output gap variance. However, after all, this model as well as the mainstream model relies on exogenous shocks that initially trigger optimism or pessimism among agents.

- *Nagel et al.'s Level K Thinking Model* enriches the literature by an enhanced micro-foundation of the NK world - heterogeneity by agents is introduced by different levels of reasoning (Level K). At the same time the model is also able to challenge the Neo-Fisherian paradox and to introduce more endogenous inertia in expectation adaption, rendering interest rate effects less pronounced. However, implications for monetary policy do not differ overly extremely from benchmark model.

- *Gabaix's Spare Dynamics Model* enriches the literature by a truly Keynesian unique equilibrium approach - Ricardian equivalence does not hold making fiscal policy an effective tool. Fisher-Neutrality does hold under the introduction of a sparsity parameter. Interest rate adjustments in T periods only have a small effect on today's inflation, weakening the scope of monetary policy and forward guidance. Helicopter drops of money can be a first best response in order to stimulate the economy. Overall, the approach is highly flexible and can easily be readjusted to the benchmark NK model. However, the model contains unnecessary double specification of myopic behavior of firms.

In order to select the most favourable approach amongst several, Friedman (1953) introduced a trade-off between the accuracy and the cost-effectiveness of alternative models, when equally acceptable on other grounds. Following this, Gabaix's flexible sparsity approach seems to perform very well on the cost dimension as it is easy to come back to the mainstream model and so to effortlessly use solutions from both. Nonetheless, Nagel's Level K approach might be more acceptable on microeconomic grounds as well as on experimental evidence. De Grauwe's adaptive-learning clearly performs best on the prediction dimensions, especially on the data used in his book (De Grauwe 2012). However, increased accuracy here comes at the cost of less micro-founded rigor.

It can be said that amongst the three models reviewed in this paper, Gabaix's approach currently appears to be the most advanced one - both in terms of technical rigor as well as on general predictability and policy implications. Nevertheless, the other two models contribute with very specific insight from behavioral and experimental economics that cannot be overestimated enough, as they serve as door openers for possibly highly fruitful advancements in future versions of the New Keynesian framework.

To the best of my knowledge, a so far unidentified link in literature is the possibility of endogenously generated economic dynamics in both, the adaptive-learning and Level K models. This is especially of interest as the mainstream NK models depends on exogenous shocks to drive business cycle in new directions. The technical fact that in the model world it is rather unclear where these shocks ultimately originate from, led to severe criticism of mainstream NK models after the *Great Recession* in 2008. The models are said to not having been able to provide any convincing class of exogenous shocks that could have triggered the crisis. It is now often assumed that a sudden exogenous shock changed agents' risk aversion and triggered the crisis (see eg. Schumacher and Żochowski 2017). However, this intuition appears rather far-fetched and therefore, a new class of endogenously driven behavioral macroeconomic models could help to explain such recessions better. Following the above mentioned, one can also identify a gap in literature regarding De Grauwe and Nagel, namely an integration of the two approaches into one single model. García-Schmidt and Woodford (2019), as well as Evans and McCough (2018) provide examples of how to use Level K and learning dynamics to draw conclusions about the Neo Fisherian paradox, however, I suggest to integrate the concepts in order to create endogenously driven business cycles similar to De Grauwe (2012). For example, one could consider a world populated by agents with different initial levels of reasoning, however, able to eventually switch level after observing undesired forecasting results and a favourable cost-benefit analysis of learning (thinking through more levels). Agents would be aware of the existence of higher levels and could observe their corresponding utility and costs, however, the probability of switching furthermore could be distorted by a logistically distributed error term when comparing individual utilities of different levels (De Grauwe 2012, p.8). Thus, dynamically depending on each point t in time, the world could be dominated by different level-thinkers and thereby potentially generate endogenous macroeconomic dynamics and policy implications yet to be explored in future research.

5 Concluding Remarks

Besides excellent, existing literature reviews on behavioral and experimental economics, the present paper provides a critical review of the most up to date applications of these fields in New Keynesian macroeconomic models. *Behavioral* New Keynesian macroeconomic models are important insofar as they constitute an enrichment of the mainstream NK models heavily applied in central banks across the industrialized world. Therefore, as these models, via the central banks, affect a huge number of humans and real economic activity, it is important to have a clear overview of possible macroeconomic policy alternatives in order to improve current shortcomings of actual applications. The paper furthermore addresses the positivist critique of heterodox economics by anticipating this critique and self-applying it to every model. In that sense De Grauwe and Gabaix's models clearly improve the predictability of the NK model, while Mauersberger et al. contribute an important step to a more micro-founded macroeconomics. The review detects an important link in literature between De Grauwe and Mauersberger et al., namely their special feature of generating dynamics endogenously within the economic model world. A gap in the literature also refers to Nagel and De Grauwe, insofar as future research could fruitfully synthesize their two approaches into an extended *adaptive Level K learning* version within the NK world. A mind sketch for a possible version of such a model has been provided at the end of section 4. However, regarding the criticism of the dangers of deviation from pure rationality, it becomes clear that academia is still far from any full consensus on how to model irrational behaviour best. Some of the reviewed models change certain aspects of the benchmark significantly leading to novel monetary policy interpretations. However, often a lot of features remain relatively unchanged leading critics to question if the same results could not be performed by the mainstream model, too. I conclude that it is not so, since all of the models - De Grauwe (2012), Nagel (2020) and Gabaix (2019) - contribute to the relevant literature by providing either a) better predictions, b) better causal inference due to more realistic assumptions or c) enriched models by endogenous dynamics - or all of it together. Further synthesis of certain aspects of the models appear to be fruitful. Especially further research in experimental economics in order to estimate model parameters in an experimental set up might yield enhanced behavioral New Keynesian models with optimal calibration.

6 Acknowledgements

I want to thank Rosemarie Nagel for supervising this thesis paper and for helpful directions. For useful conversation and advice I thank Silvana Tenreyro, Michael McLeay, Jordi Galí, Wouter Den Haan and Christian Fons-Rosen. Also I want to thank *University Pompeu Fabra* for institutional support.

7 References

Angeletos, G., and C. **Lian**. 2018. *Forward guidance without common knowledge.* American Economic Review.

Akerlof, G. 2001. *Behavioral Macroeconomics and Macroeconomic Behavior.* The Nobel Foundation.

Alaoui, L., and A. **Penta**. 2016 *Endogenous depth of reasoning.* The Review of Economic Studies.

Azariadis, C. 1981. *Self-fullfiling prophecies.* Journal of Economic Theory.

Azariadis, C., and R. **Guesnerie**. 1986. *Sunspots and Cycles.* Review of Economics Studies.

Barro, R. 1974. *Are government bonds net wealth?.* Journal of political economy.

Bernays, E. 1947. *The engineering of consent.* The Annals of the American Academy of Political and Social Science

Branch, W., and G. **Evans**. 2007. *Model uncertainty and endogenous volatility.* Review of Economic Dynamics.

Brock, W., and C. **Hommes**. 1997. *A rational route to randomness.* Econometrica: Journal of the Econometric Society.

Calvo, G. 1983. *Staggered prices in a utility-maximizing framework.* Journal of monetary Economics.

Chung, H., **Herbst**, E., and M. **Kiley**. 2014. *Effective Monetary Policy Strategies in New Keynesian Models: A Reexamination.* NBER Macroeconomics Annual.

Clarida, R., **Galí**, J., and M. **Gertler**. 1999. *The science of monetary policy, a new Keynesian perspective.* Journal of Economic Literature.

Crawford, V., **Costa-Gomes**, M., and N. **Iriberri**. 2013. *Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications.* Journal of Economic Literature.

Damasio, A. 2003. *Looking for Spinoza, Joy, Sorrow and the Feeling Brain.* Orlando, FL: Harcourt.

De Grauwe, P. 2012. *Lectures on Behavioral Macroeconomics.* Princeton University Press.

De Vroey, M., and P. **Malgrange** 2011. *The History of Macroeconomics from Keynes's General Theory to the Present.* Insitute de Recherhces Économiques et Sociales de l'Université de Louvin.

Discroll, J., and S. **Holden**. 2014. *Behavioral economics and macroeconomic behavior.* Journal of Macroeconomics.

Duffy, J. 2016. *Macroeconomics: A Survey of Laboratory Research.* The Handbook of Experimental Economics, Volume 2.

Duffy, J. 2006. *Agent-based models and human subject experiments.* Handbook of computational economics.

Arifovic J., J Duffy and J. **Arifovic**. 2018. *Heterogeneous agent modeling: experimental evidence.* Handbook of Computational Economic.

Duffy, J., and R. **Nagel**. 1997. *On the robustness of behavior in experimental 'beauty contest' games.* The Economic Journal.

Evans G., and B. **McCough**. 2018. *Interest-Rate Pegs in New Keynesian Models.* Journal of Money, Credit and Banking.

Evans, G., and S. **Honkapohja**. 2001. *Learning and Expectations in Macroeconomics.* Princeton University Press.

Evans, G., **Romer**, P., and S. **Honkapohja**. 1998. *Growth Cycles.* American Economic Review.

Freud, S. 1917. *Eine Schwierigkeit der Psychoanalyse.* Fischer Verlag. Gesammelte Werke. Vol. 12. Frankfurt am Main (1917), p. 11.

Friedman, M. 1953. *Essays in Positive Economics.* University of Chicago Press.

Gabaix, X. 2019. *A Behavioral New Keynesian Model.* NBER Working Paper No. 22954.

Gabaix, X. 2014. *A sparsity-based model of bounded rationality.* The Quarterly Journal of Economics

Galí, J., 2020. *Helicopter money: The time is now.* Vox CEPR Portal. https://voxeu.org/article/helicopter-money-time-now

Galí, J., 2010. *Monetary policy and unemployment.* Handbook of Monetary Economics.

Galí, J., 2008. *Monetary Policy, Inflation and the Business Cycle.* Princeton University Press.

García-Schmidt, M., and M. **Woodford**. 2019. *Are low interest rates deflationary? A paradox of perfect-foresight analysis.* American Economic Review.

Gould, L. 1989. *The rise and fall of the New Deal order, 1930-1980.* Princeton University Press.

Gerke, R., and K.**Hauzenberger**. 2017. *The Fisher Paradox: A Primer.* Bundesbank Discussion Paper No. 20/2017.

Heemeijer, P.,**Hommes**, C., **Sonnemans**, J., and J. **Tuinstra**. 2009. *Price stability and volatility in markets with positive and negative expectations feedback: An experimental investigation.* Journal of Economic Dynamics and Control.

Hill, L. E. 1968. *A Critique of Positive Economics.* The American Journal of Economics and Sociology, Vol. 27, pp.259-266.

Hommes, C., **Assenza**, T., **Bao**, T., and D. **Massaro**. 2014. *Experiments on Expectations in Macroeconomics and Finance.* Emerald Group Publishing.

Hornstein, A. 2008. *Introduction to the New Keynesian Phillips curve.* Economic Quarterly.

Howitt, P., and R. **McAfee**. 1992. *Animal Spirits.* American Economic Review.

Keynes, J. M. *The General Theory of Employment, Interest and Money.* London: Macmillan.

Laibson, D., 1997. *Golden Eggs and Hyperbolic Discounting.* The Quarterly Journal of Economic.

Le Roy, S. 2004. *Rational exuberance.* Journal of Economic literature.

Lin, T. 2012. *A Behavioral Framework for Securities Risk.* Seattle University Law Review.

Mauersberger, F., and R. **Nagel**. 2018. *Levels of reasoning in Keynesian Beauty Contests: a generative framework.* Handbook of computational economics.

Mauersberger, F., **Nagel**, R., and C. **Bühren**. 2020. *Bounded rationality in Keynesian beauty contests: a lesson for central bankers?.* Economics Discussion Papers, No 2019-53, Kiel Institute for the World Economy.

Michalos, C. 2014. *Encyclopedia of Quality of Life and Well-Being Research.* Springer Science+Business Media Dordrecht.

Mullainathan, S., and R. **Thaler**. 2001. *Behavioral Economics.* Pergamon Press.

Nagel, R. 1995. *Unraveling in Guessing Games.* The American Economic Review.

Nagel, R., and G. **Coricelli**. 2012. *The Neural Basis of Bounded Rational Behavior.* Revista Internacional de Sociologia

Romer, D. 2012. *Advanced Macroeconomics.* McGraw-Hill Publishing.

Savage, L. 1954. *The Foundations of Statistics.* New York: John Wiley and Sons.

Schumacher. M., D. **Żochowski**. 2017. *The risk premium channel and long-term growth.* European Central Bank Working Paper Series.

Simon, H., 1957. *A behavioral model of rational choice.* Competition Policy International.

Shell, K., 1977. *Monnaie et allocation intertemporelle.* CNRS Séminaire Roy-Malinvaud, Paris.

Smith, V. 1990. *Experimental economics.* Edward Elgar Publishing.

Taylor, J. 1993. *Discretion versus policy rules in practice.* Carnegie-Rochester conference series on public policy.

Thaler, R. 1994. *Quasi Rational Economics.* New York: Russell Sage Foundation.

Thaler, R., and S. **Mullainathan**. 2001. *Behavioral economics.* International Encyclopedia of Social Sciences. Pergamon Press.

Thaler, R., and S. **Benartzi**. 1995. *Myopic Loss Aversion and the Equity Premium Puzzle.* The Quarterly Journal of Economics.

Tobin, J. 2005. Irving Fisher (1867–1947). American Journal of Economics and Sociology.

Tversky A., and D. **Kahneman**. 1981. *The framing of decisions and the psychology of choice.* Science 211:453-458.

Woodford, M. 2013. *Macroeconomic analysis without the rational expectations hypothesis.* Annual Review of Economics.

Woodford, M. 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy.* Princeton University Press.

Zhang S., and A. **Yu**. 2013. *Forgetful Bayes and myopic planning: Human learning and decision-making in a bandit setting.* Advances in Neural Information Processing Systems 26 (NIPS 2013).