



Talitha Faustino Speranza

**A Proposal for Setting Central Banks' Interest
Rate Using Neural Networks and Genetic
Algorithms**

Dissertação de Mestrado

Dissertation presented to the Programa de Pós-graduação em Engenharia Elétrica of PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Engenharia Elétrica.

Advisor : Prof. Ricardo Tanscheit
Co-advisor: Prof. Marley M.B.R. Vellasco

Rio de Janeiro
September 2019



Talitha Faustino Speranza

**A Proposal for Setting Central Banks' Interest
Rate Using Neural Networks and Genetic
Algorithms**

Dissertation presented to the Programa de Pós-graduação em Engenharia Elétrica of PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Engenharia Elétrica. Approved by the Examination Committee.

Prof. Ricardo Tanscheit

Advisor

Departamento de Engenharia Elétrica – PUC-Rio

Prof. Marley M.B.R. Vellasco

Co-advisor

Departamento de Engenharia Elétrica – PUC-Rio

Prof. Pedro Guilherme Costa Ferreira

Fundação Getulio Vargas – FGV

Prof. Harold Dias de Mello Junior

Universidade do Estado do Rio de Janeiro – UERJ

Rio de Janeiro, September the 12th, 2019

All rights reserved.

Talitha Faustino Speranza

Holds a Bachelor in Computer Engineering from PUC-Rio and a Master's Degree in Economics from Barcelona Graduate School of Economics and Pompeu Fabra University (Barcelona, Spain), with focus in economic growth and development. Developed advanced statistical software and conducted studies in macroeconometric forecasting for the Brazilian Institute of Economics at Getulio Vargas Foundation.

Bibliographic data

Faustino Speranza, Talitha

A Proposal for Setting Central Banks' Interest Rate Using Neural Networks and Genetic Algorithms / Talitha Faustino Speranza; advisor: Ricardo Tanscheit; co-advisor: Marley M.B.R. Vellasco. – Rio de Janeiro: PUC-Rio, Electrical Engineering Department, 2019.

v., 62 f: il. color. ; 30 cm

Master's Thesis - Pontifical Catholic University of Rio de Janeiro, Electrical Engineering Department.

Inclui bibliografia

1. Engenharia Elétrica – Teses. 2. Economia – Teses. 3. Algoritmos Genéticos;. 4. Política Monetária;. 5. Modelagem DSGE;. 6. Redes Neurais;. 7. Macroeconomia Brasileira;. I. Tanscheit, Ricardo. II. Vellasco, M.B.R. Marley. III. Pontifical Catholic University of Rio de Janeiro. Electrical Engineering Department. IV. Título.

CDD: 620.11

To all the people that crossed my path
in these intense last years

Acknowledgments

Throughout this journey, many people contributed, directly or indirectly, to the completion of this work. Unfortunately, there is no space here to thank everyone, but that does not mean those that will not be mentioned here are less important.

First of all, I would like to thank my advisors, Ricardo and Marley, whose advice, criticism, suggestions and patience, of course, were fundamental. The success of this dissertation, its merits and qualities, are as much theirs as mine. Its weaknesses and problems are solely my responsibility.

I would also like to thank my colleagues and the entire FGV/IBRE team. Being by their side was one of the most pleasant and intellectually compelling moments in my life. They are the greatest proof that a great institution is measured by the greatness of its members.

In particular, I would like to thank my boss, Pedro, a friend and great inspiration, for his help in this work and for having kindly accepted to compose the examining board. I am also thankful to Harold, another member of the examining board, for accepting the invitation and for making valuable comments that were indispensable to tie up loose ends.

I thank my many friends for making my life lighter even in those hardest moments. My special thanks to Magno, for his help and support in the most critical points of this dissertation. I also thank all my colleagues at the Pompeu Fabra University, whose joy and intelligence gave me strength and confidence to move forward in the difficult final stages of this dissertation.

I thank my parents, Mauro and Glória, and my sister, Giulia, for always having been by my side, even when they often did not understand my decisions or when I could not be so close to them.

I am deeply grateful to Alexandre, the biggest victim of this dissertation, who consciously, and in a sometimes reckless way, moved me in all stages of this master's degree and in many other things.

Finally, I would like to thank CAPES and PUC-RIO. I could never have completed this master's degree without their trust and financial aid.

To all the others who were not mentioned here, thank you very, very much.

Abstract

Faustino Speranza, Talitha; Tanscheit, Ricardo (Advisor); Vellasco, M.B.R. Marley (Co-Advisor). **A Proposal for Setting Central Banks' Interest Rate Using Neural Networks and Genetic Algorithms**. Rio de Janeiro, 2019. 62p. Master's Thesis – Electrical Engineering Department, Pontifical Catholic University of Rio de Janeiro.

Dynamic Stochastic General Equilibrium (DSGE) models are flawed, as became clear after the 2007-2008 financial crisis. Efforts to subdue the shortcomings have been insufficient: to this date, there is still a demand for building a new framework to study policy implications and make decisions. We propose a novel monetary policy strategy, in an attempt to provide an auxiliary tool to central banks, whose main predictive models are still from the DSGE family. We derive an objective function from three empirical relationships that have long been established in economic literature: Okun's Law, the Phillips Curve, and liquidity effects. Using data from Brazil, we seek to minimise the value of this function by choosing the interest rate via a genetic algorithm. Since the function is forward looking, we use a neural network to predict values of unemployment and inflation. Results suggest that had the Brazilian central bank applied our strategy, and all other economic conditions remained identical, inflation could have been lower for 62.48% of the time. Predicted unemployment, however, was lower only for 39.69% of covered periods, as it faces a trade-off with inflation. We discuss the applicability of the proposed strategy and argue for its theoretical soundness.

Keywords

Genetic Algorithms; Monetary Policy; DSGE Modelling; Neural Networks; Brazilian Macroeconomics;

Resumo

Faustino Speranza, Talitha; Tanscheit, Ricardo; Vellasco, M.B.R. Marley. **Uma Proposta para Determinar a Taxa de Juros de Bancos Centrais Usando Redes Neurais e Algoritmos Genéticos**. Rio de Janeiro, 2019. 62p. Dissertação (Mestrado) – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Os modelos Dinâmicos Estocásticos de Equilíbrio Geral (DSGE) contêm falhas diversas, como ficou claro após a crise financeira de 2007-2008. Esforços para mitigar as deficiências têm sido insuficientes: até hoje, ainda há uma demanda por construir uma nova estrutura para estudar as implicações de política econômica e tomar decisões. Propomos uma nova estratégia para resolver o problema do banco central, na tentativa de prover uma ferramenta auxiliar para os bancos centrais, cujos principais modelos ainda pertencem à família dos DSGEs. Derivamos uma função objetivo a partir de três relações empíricas estabelecidas há muito tempo na literatura econômica: a Lei de Okun, a Curva de Phillips e os efeitos de liquidez. Usando dados do Brasil, procuramos minimizar o valor dessa função, escolhendo a taxa de juros através de um algoritmo genético. Como a função é prospectiva, usamos uma rede neural para prever valores de desemprego e inflação. Os resultados sugerem que, se o banco central brasileiro houvesse aplicado nossa estratégia e todas as outras condições econômicas continuassem iguais, a inflação poderia ter sido mais baixa 62,48% do tempo. O desemprego previsto, contudo, foi mais baixo apenas 39,69% dos períodos cobertos, pois enfrenta um *trade-off* com a inflação. Discutimos a aplicabilidade da estratégia proposta e defendemos sua solidez teórica.

Palavras-chave

Algoritmos Genéticos; Política Monetária; Modelagem DSGE; Redes Neurais; Macroeconomia Brasileira;

Table of contents

1	Introduction	12
1.1	Motivation	13
1.2	Objective	15
1.3	Contribution	15
1.4	Organisation of the Thesis	16
2	Literature Review	17
2.1	Failure of DSGE Models and Proposed Alternatives	17
2.2	Macroeconomic Relationships	20
2.2.1	Okun's law	21
2.2.2	Phillips Curve	22
2.2.3	Liquidity effects	23
2.3	Machine Learning Applied to Monetary Policy	24
3	Developing the Framework	27
3.1	Overview	28
3.2	Central Bank's Objective Function	30
3.3	Genetic Algorithm	33
3.4	Neural Network Regression	34
4	Simulations	37
4.1	Data	37
4.2	Results	38
4.2.1	Neural Network's Predictions	38
4.2.2	Genetic Algorithm Evolution	45
4.2.3	Performance	46
4.3	Discussion	51
4.3.1	The case against a lower SELIC	51
4.3.2	Possible Practical Limitations	52
4.3.3	A Brief User Guide	53
5	Conclusion	54

List of figures

Figure 3.1	High-level diagram of the framework.	27
Figure 3.2	Qualitative overview of the framework	28
Figure 3.3	Detailed overview of the framework	29
Figure 3.4	Genetic Algorithm Flowchart	35
Figure 4.1	Four-months-ahead unemployment: real data versus neural network's fitted values	40
Figure 4.2	Four-months-ahead inflation: real data versus neural network's fitted values	41
Figure 4.3	Eight-months-ahead unemployment: real data versus neural network's fitted values	41
Figure 4.4	Eight-months-ahead inflation: real data versus neural network's fitted values	42
Figure 4.5	Four-months-ahead unemployment: testing set versus out-of-sample predictions	43
Figure 4.6	Four-months-ahead inflation: testing set versus out-of-sample predictions	43
Figure 4.7	Eight-months-ahead unemployment: testing set versus out-of-sample predictions	44
Figure 4.8	Eight-months-ahead inflation: testing set versus out-of-sample predictions	44
Figure 4.9	Value function for the best individual in each generation	46
Figure 4.10	SELIC target: actual series versus solution generated by the genetic algorithm	46
Figure 4.11	SELIC target: actual series versus 12-month MA solution. First values were lost when calculating the MA.	47
Figure 4.12	Predicted 4-months-ahead inflation and actual inflation	48
Figure 4.13	Predicted 8-months-ahead inflation and actual inflation	48
Figure 4.14	Predicted 4-months-ahead unemployment and actual unemployment	49
Figure 4.15	Predicted 8-months-ahead unemployment and actual unemployment	49

List of tables

Table 4.1	Time series used in this study	38
Table 4.2	Set of parameters used in the grid search	39
Table 4.3	Set of tested parameter values for the genetic algorithm.	45
Table 4.4	Descriptive statistics of genetic algorithm solution as compared to actual CBB solution	50
Table 4.5	Descriptive statistics of genetic algorithm solution as compared to actual CBB solution	51

List of Abbreviations

ABM – Agent-Based Model
ADAM – Adaptive Moment
BFGS – Broyden–Fletcher–Goldfarb–Shanno
CBB – Central Bank of Brazil
CCB – Czech Central Bank
COPOM – Monetary Policy Committee
DSGE – Dynamic Stochastic General Equilibrium
FOMC – Federal Open Market Committee
GDP – Gross Domestic Product
HTAN – Hyperbolic Tangent
LDA – Linear Discriminant Analysis
LSIG – Logistic Sigmoid
MLP – Multilayer Perceptron
NK – New Keynesian
OLS – Ordinary Least Squares
PNADC – National Continuous Household Sample Survey
RBC – Real Business Cycle
RELU – Rectified Linear Unit
RMSE – Root Mean Square Error
SELIC – Special Settlement and Custody System
SGD – Stochastic Gradient Descent

1

Introduction

This work describes and discusses a novel tool to support the decision-making process of central banks. Monetary authorities around the world conduct policy in order to stabilise inflation and mitigate its deleterious effects. Some of them also take unemployment explicitly into account, trying to minimise the trade-off between this variable and inflation. Whether this is the case or not, the policy instrument is the short-term interest rate almost everywhere [Amstad and Martin, 2011], as defended by Michael Woodford in his influential book Woodford [2009].

Dynamic Stochastic General Equilibrium (DSGE) models have been pivotal in policy-making, despite their flaws. As attested by Blanchard and Galí [2007] and Tovar [2009], they were the favourite choice of central bankers before the 2007-2008 crisis, the event that brought the attention to the fact that DSGE's predictions were actually weak: nobody foresaw the upcoming crisis. According to many accounts, efforts to subdue these shortcomings have been insufficient, but in the absence of other options, DSGE models are still preferred [Christiano et al., 2018]. Therefore, to this date, there is an ongoing demand for building a new framework to support the study of policy implications and related decisions.

We propose a novel monetary policy strategy, aiming to provide a new tool and reduce the dependence on DSGE models. In order to achieve this, we first define an objective function that depends on future unemployment and future inflation only through interest rates, which are chosen to minimise the function's value. The objective function is derived from three empirical relationships that have long been established in economic literature: Okun's Law, the Phillips Curve, and liquidity effects. Using data from Brazil, we seek to minimise the value of this function by choosing interest rates via a genetic algorithm. Since the function is forward looking, we use a neural network to predict future values of unemployment and inflation.

Considering the ambitious and seemingly bold objectives of this work, three words of caution are needed. This is not an attempt to substitute DSGE models, but rather an effort to complement the set of frameworks that are available to central banks. Likewise, we do not claim that our strategy can

predict the next crisis, although we show that forecasts around the last financial crisis were satisfyingly accurate. At last, the strategy is general enough so as to accommodate data from any given country, but we only applied it to the case of Brazil. Future works should cover other scenarios.

1.1 Motivation

Unemployment is obviously harmful for individuals, while the reason why inflation is damaging is less clear - in the short-run, higher inflation usually comes along with growth, as a result of demand pressures. Nevertheless, inflation is always, and everywhere, a social problem¹. Even though individuals feel that rising price levels is detrimental to quality of life², the perception that inflation, either low or high, has a negative impact on the society as a whole is far from consolidated.

In fact, there are still some scholars, governmental authorities, and other public figures, that argue for the supposed benefits of inflation. They usually advocate that moderate inflation is useful to lower real wages without people realising it³.

However, even low inflation has the power of disorganising the economy and it is the poor who pay the higher costs. Unanticipated inflation (low or high) brings risks, and these eliminate some exchanges between consumers and firms that otherwise would have been attractive. Risks demand some kind of protection, usually in the form of indexed bonds or other financial instruments that only well-educated people are acquainted with and have access to. Since education is generally determined by income, the poorer end up having no protection against inflation. Besides, the proportion of consumption in the poorer's income is higher, implying that when prices rise, they suffer the most. Problems in indexation systems, adopted to change prices automatically in response to inflation, have similar unintended consequences.

In a country still marked by poverty⁴ and low levels of education⁵,

¹This sentence was inspired by a famous quote of monetarist economist Milton Friedman, who said: "Inflation is always, and everywhere, a monetary phenomenon".

²See, for instance, Gallup polls [Dornbusch et al., 1997]. People do consider inflation the most important problem – of all possible problems, including crime and drug addiction – facing the country when it is high, but forget about it when it is low.

³For example, if inflation was 5% but nominal wages increased by only 3%, people would not notice that their real wages decreased by 2%, or would feel better than in a situation where inflation is 0% and nominal wages, which then would be equal to real wages, are decreased by 2%.

⁴According to the Focus Economics [Focus Economics, 2018], Brazil is the 65th in the gross domestic product (GDP) per capita ranking.

⁵The United Nations [United Nations Development Programme, 2019] ranks Brazil at the 87th position in education attainments.

combating inflation is imperative. It is the Central Bank of Brazil (CBB) that has a mandate to keep inflation in acceptable levels and, to do so, their bankers rely on models from the DSGE family [De Castro et al., 2015]. These models have received harsh criticism for more than a decade now. In short, they are known to suffer from (i) lack of theoretical soundness [Solow, 2010, Romer, 2016]; (ii) identification issues or, in a nutshell, the problem of having more parameters to estimate than equations [Canova and Sala, 2009, Blanchard, 2016]; and (iii) unconvincing policy implications, specially in terms of welfare [Blanchard, 2018]. Hence providing new, more reliable, tools to assist in the decision-making process is vital.

Before explaining our specific objectives, it is important to clarify how the CBB operates nowadays and give a quick historical perspective. The Monetary Policy Committee (COPOM, from portuguese *Comitê de Política Monetária*), a Brazilian governmental body established under the CBB, has as its main goal maintaining price stability. Before 1999, COPOM pursued this objective within the framework of the currency band regime, establishing a band around which the exchange rate could vary freely. The Asian crisis in 1997 began to deteriorate the regime, starting the process of depleting the country's international reserves. An speculative attack in January 1999 and low bank reserves levels disrupted the exchange rate system and a flexible exchange rate was adopted. In June 1999, following the trend of the world's largest economies, Brazil adhered to the inflation targeting system by employing the SELIC rate ⁶ (the base rate of Brazilian economy) as an instrument to meet the inflation target ⁷.

Inflation targeting works as follows. If inflation is higher than the target the interest rate should rise, so as to discourage spending, since higher interest rates should stimulate savings. In contrast, if inflation is lower than the target the interest rate should decrease. In some countries, the central bank also monitors the unemployment rate, a control that tends to smooth interest rate choices. If, for instance, the central bank lowers interest rates, more jobs are likely to be created, fuelling spending further. But raising expenditures in most cases entails accelerating prices, so the central bank must be cautious. The method we develop builds upon these principles, which ultimately derive from well-established empirical evidence.

In Brazil, CBB's mandate is to tame inflation only, and concerns about the unemployment rate should be put aside, at least in principle. However, our strategy will take the unemployment rate into account, for two reasons.

⁶SELIC stands for Special Settlement and Custody System, from portuguese *Sistema Especial de Liquidação e de Custódia*

⁷For more information on the inflation targeting regime, see Bernanke et al. [1999]

First and foremost, our framework is supposed to be general, flexible enough to be used by any central bank. Secondly, it has long been speculated that Brazilian executive power's influence over the CBB has, in several occasions, steered the monetary authority towards considering the unemployment rate in its decisions [Afonso et al., 2016].

1.2

Objective

Our objective is to create a novel tool to support central banks in conducting monetary policy. From our perspective, it is essential for central banks to employ auxiliary tools to reduce their reliance on DSGE models. In our framework, the central bank works toward minimising future inflation and future unemployment rate by setting the base interest rate, having the option to change the relative weights of both inflation and unemployment rates, just as any central bank would do.

In contrast with the traditional approach, though, we do not build a model affected by *ad hoc* assumptions, identification issues, and questionable policy prescriptions. Rather, we use a genetic algorithm and a neural network, seeking to minimise a single objective function that depends on the SELIC rate through future inflation and future unemployment. We derive this objective function from three empirical relationships observed in a wide range of countries and periods: the Phillips Curve (negative association between inflation and unemployment), the Okun's law (negative correlation between the change in the unemployment rate and the output growth rate), and liquidity effects (higher interest rate impacts a country's total production negatively)⁸. Since the objective function is prospective, it is necessary to estimate future unemployment and inflation for every attempted solution (i.e. interest rate) generated by the genetic algorithm. Predictions are computed by a neural network, more specifically, a multilayer perceptron, whose inputs are the attempted SELIC rate and a large set of covariates.

1.3

Contribution

Given the importance of inflation and unemployment for any given country, and in particular for Brazil, where inflation has been historically high and unemployment is currently a major concern, this work's contribution intends to be of a practical nature above all. Nonetheless, it also adds to the literature on machine learning applied to monetary policy. There are few

⁸These relationships are detailed in Chapter 3.

papers relating both subjects and, to the best of our knowledge, this is only the second study trying to create a novel interest rate setting rule based on machine learning. The work by Kukal and Quang [2014] attempts to use fuzzy logic with the same purpose, but the results are not promising, as will be discussed in Chapter 2.

1.4

Organisation of the Thesis

This thesis is organised as follows. The second chapter contains a literature review on DSGE models, their problems, and possible alternatives or improvements, as well as a summary of the literature on machine learning applied to monetary economics. It also includes definitions, background, and developments of the three macroeconomic relationships we use throughout. The third chapter presents our proposed model, listing the time series we used and describing in detail the theoretical and empirical background of our strategy, the neural network structure, and the genetic algorithm to optimise monetary policy. In the fourth chapter we present and discuss the results and show how the model can be used in practice. The fifth chapter concludes this work.

2

Literature Review

2.1

Failure of DSGE Models and Proposed Alternatives

Before the 2007-2008 financial crisis, macroeconomists generally agreed that there was convergence in methodology: the DSGE framework was standard [Woodford, 2009, Galí and Gertler, 2007]. Olivier Blanchard even published a famous paper claiming that the state of the field of macroeconomics was good [Blanchard, 2009], although recognising that not everything was fine. Indeed, two years earlier, George Akerlof was already trying to cover up for this gap by exploring ways out of macro models' *ad hoc* assumptions [Akerlof, 2007]. He argued that the field had yet a lot to advance. Other leading economists, such as Peter Howitt, Alan Kirman, Axel Leijonhufvud, Perry Mehrling, and David Colander, advocated a more empirically and less model-based macroeconomics [Colander et al., 2008].

Then the crisis happened. It was a disruption in the financial sector comparable to that of 1929 and no macroeconomist saw it coming. Their models were evidently flawed. Even though the crisis was not predicted, the fact that macro models had severe shortcomings could have been anticipated. Economists were aware that DSGE models were built on unappealing assumptions to reach strange conclusions, and had been trying to mend it.

The family of DSGE models derives from Real Business Cycle (RBC) models proposed by Kydland and Prescott [1982] and Long Jr and Plosser [1983]. Some of RBC assumptions were: (i) rational expectations (economic agents do not commit systematic mistakes when planning ahead); (ii) no asymmetric information (individuals have perfect knowledge); (iii) representative firm (firms are all identical); (iv) representative household (households are all identical); (iv) infinitely lived and price-taking agents; and (v) perfectly competitive markets where prices adjust instantaneously.

Under these unrealistic assumptions, five neutralities [Akerlof, 2007] follow: (i) Life-cycle/Permanent Income Hypothesis (consumption is independent of current income); (ii) Modigliani-Miller Theorem (current profits are irrelevant for investment spending); (iii) Natural Rate Theory (inflation and

unemployment are independent in the long-run); (iv) Money Neutrality Theory (monetary policy is unable to stabilise the economy's total production); and (v) Ricardian Equivalence (taxes and budget deficits does not affect consumption). All of these conclusions are counter-intuitive. In fact, none of them was confirmed empirically and there are strong evidence against some of them [Colander, 2006].

In an effort to fix those results, macroeconomists began to add assumptions to the RBC core [Galí, 2015]. These were the so-called market imperfections and frictions, such as monopolistic competition and price stickiness (nominal rigidities). These new models were baptised as New Keynesian (NK) models and evolved into state-of-the-art DSGE models, Smets and Wouters [2007] being the most well-known among them. Although it delivers more plausible conclusions, there is still room for criticism.

For one, only the problem with money neutrality was solved, i.e. in NK models, monetary policy matters and affect output (total production). Even so, influential authors still insist in money neutrality [Fernández-Villaverde, 2010]. Secondly, the way NK models were conceived goes against RBC models' philosophy, which is that macroeconomic equations should originate from microeconomic behaviour. In economic jargon, RBC models intend to be *microfounded*. Once modellers implant mechanisms with no microeconomic justification, the models cease to be microfounded. These mechanisms (market imperfections and frictions) are the *ad hoc* assumptions we refer to throughout: there is no theory backing them, they were added to make macroeconomists feel more comfortable with their models' implications.

As other authors [Boianovsky et al., 2006, Romer, 2016], our view is that the issue with microfoundations has deeper roots, lying on RBC models themselves. Before RBC models became the rule, the literature was attempting to create a science based on heterogenous agents, a search that was closer to the real world than microeconomic theory was. Macroeconomists were trying to explain how recessions could happen knowing that economies do not behave like the perfectly competitive world of Walrasian/Arrow–Debreu general equilibrium theory [Arrow and Debreu, 1954]. Instead, they believed that markets could fail as result of lack of coordination among individuals [Boianovsky et al., 2006, Barbieri, 2013], who were prone to error and could adapt to the environment.

This pursuit basically came to an end with the appearance of RBC models. In this framework, fluctuations in economic output are due only to exogenous shocks, and the number and types of shocks increased in NK models (thus by extension in DSGEs). This means that in DSGEs, contrary

to even casual observation, market failures that culminate in recessions and unemployment are caused by imaginary forces that no economic agent was responsible for. In fact, there is only one agent of each kind (typically a representative household and a representative firm) each being affected by specific types of shocks. Right after the crisis, Robert Solow would point out that one important consequence of the representative agent assumption is that there are no conflicts of interest, no incompatible expectations, and no deceptions [Solow, 2010]. As a result, for instance, the only way that DSGE models can generate unemployment is to make it voluntary by adding external preference shocks (agents would suddenly and without justification prefer more leisure).

Another example of problems with the representative agent assumption is given by Kirman [2010]. According to him, the 2007-2008 financial crisis was an event where interconnection and interaction between banks played a central role. Nonetheless, these two features are not part of modern macroeconomic models. On the contrary, these models assume that individual banks optimise their profits in isolation without reacting to each other's decisions. He claims that the near-collapse of the world's financial system being caused by banks optimising profits individually and unaware of other bank's behaviour seems highly unlikely.

An obvious way out would be to recognise that market imperfections arise from individual's mistakes and lack of sensible coordination between them, and not due to imaginary forces. In other words, from the interactions among agents, seemingly random aggregate fluctuations in throughput emerge. Howitt [2012], Leijonhufvud [2014], and others point to agent-based models (ABMs) as a natural candidate framework that could be used to simulate and observe these phenomena. ABMs are composed by several kinds of heterogeneous agents that follow simple behavioural rules and have limited information [Tesfatsion, 2006]. They interact locally and macro phenomena arise from these micro-level activity.

However, difficulties to feed data into these models make them unsuitable for prediction purposes. In fact, a common method to take ABMs to data is to calibrate parameters with unknown values, just as DSGE modellers do and for which they are heavily criticised. As adverted by Canova and Sala [2009] and brought up years later by Hurtado [2014], Blanchard [2018], and others, DSGEs suffer from identification issues and thus many parameters have to be calibrated, i.e. modellers need to impute values to unknown parameters to be able to estimate the whole model.

Concretely, DSGEs' equations are estimated simultaneously but, quite

often, there are more parameters than equations. Hence the modeller has to guess the true value of some (usually most) parameters. The procedure to hide this problem takes many forms [Romer, 2016]. The researcher may choose the easiest path, which is to assume that a subset of parameters is zero, or three other more sophisticated approaches. First, the zero-value assumption could be hidden in some micro-level equation (one of those that lead to the final system of equations). Second, the mean or the median of time series related to the unknown parameters might be taken. For instance, if the undetermined parameter is the share of capital within the economy, it is set to be the historical mean of the share of capital in a given economy. Third, a Bayesian estimation could be carried and fed a range of values for the priors in until the desired outcome (e.g. the adequate set of impulse-response functions) is reached. The last case (searching for the right priors) is flagrantly *ad hoc* and non-scientific, and the other two (zero-value assumptions and data driven calculations) are also obviously prone to cherry picking.

Calibration in ABMs goes along the same lines, but a system of equations does not have to be estimated. Instead, zero-value assumptions and values taken from historical data are embedded in agents' behaviour. As a result, ABMs face much of the same criticism as DSGEs, although being more promising from a purely theoretical standpoint, since many stringent assumptions are not present and these models are powerful tools to create artificial economies that reproduce a strikingly high number of stylised facts about real economies [Delli Gatti et al., 2011].

In this work, as we are concerned with predictions for the real world, so we will not take the theoretical modelling path. As we intended to show throughout this section, this strand of research seems to have reached a dead end. Thus we attempt to develop a plainly statistical tool, the novelty of which is argued in section 2.3.

2.2

Macroeconomic Relationships

All the economic relationships we use to obtain our model are observed across almost all countries and at most periods for which historical data is available. In the next sections, we briefly discuss the literature on these facts and their validity in the case of Brazil, besides describing the correlations themselves.

2.2.1

Okun's law

Okun's law is an empirical negative relationship between changes in the unemployment rate and the output growth rate. Proposed by Okun [1963], this law should be seen as a rule of thumb rather than a scientific law because it is not based on a theory, although standard macroeconomic theory provides possible explanations for what is behind it. Simply put, a greater number of employed workers will produce more, hence increasing a country's total production, i.e. its gross domestic product (GDP). Recent studies, such as Ball et al. [2016], have shown that Okun's correlation still holds for most countries, and it is different for each country, reflecting differences in productivity per employed worker.

The general form of Okun's law is:

$$u - \bar{u} = F(y - \bar{y}) + \kappa \quad (2-1)$$

where u is the unemployment rate, \bar{u} is the natural unemployment rate, y is the GDP growth rate, κ is a random shock, \bar{y} is the natural GDP growth rate and F is a function such that $\frac{\partial F(y-\bar{y})}{\partial (y-\bar{y})} < 0$. Note that the difference $y - \bar{y}$ is the output gap. The reason why we do not assume any specific format for F will become clear in the following sections.

As many empirical studies show, Okun's law is still¹ valid for Brazil and robust to a range of methodologies employed to calculate \bar{y} , a non-observable variable. Tombolo and Hasegawa [2014], following the standard procedure, take F to be linear and estimate the coefficient of $y - \bar{y}$, often called Okun's coefficient. They apply two distinct methods, ordinary least squares (OLS) and the distributed lag model developed by Gordon and Clark [1984], and established the coefficient to be between -0.1878 and -0.2055 in the case of Brazil. Hence, $\frac{\partial F(y-\bar{y})}{\partial (y-\bar{y})} < 0$, and Okun's law holds. Similarly, and also through Gordon and Clark's method, Jin and Harper [2019] determined that the Okun's coefficient is about -0.140 in Brazil.

Both Tombolo and Hasegawa [2014] and Jin and Harper [2019] use the Hodrick-Prescott filter [Hodrick and Prescott, 1997] to calculate the potential GDP. Gois and Jorge [2018], on the other hand, tested the robustness of Okun's law in Brazil for other estimates of \bar{y} . Not only did they find that the relationship holds when \bar{y} is obtained from a linear trend, a quadratic trend, and a Beveridge-Nelson decomposition [Beveridge and Nelson, 1981],

¹In principle, nothing guarantees that it will continue to be valid forever. For instance, the original version of the Phillips Curve, as we shall see, did not hold for a long time.

but also showed that Okun's law is valid for three large metropolitan areas of the country.

2.2.2

Phillips Curve

The Phillips Curve concept was first presented in Phillips [1958]. He claimed that there was an inverse relationship between inflation and unemployment, which was persistent over time. His work was empirically-based, but he provided a theory to justify his findings. According to him, any fiscal stimulus would increase aggregate demand, driving up labour demand in particular, and subsequently decreasing the pool of unemployed workers. Companies would increase wages to compete and attract a smaller, more specialised, pool of applicants, and then pass the cost of higher wages to consumers by increasing prices. In short, Phillip's message was: changes in unemployment rate has a largely predictable effect on inflation.

However, as Milton Friedman foresaw, this would not hold in the long run. Between 1973 and 1975, the United States faced both high unemployment and high inflation, a phenomenon known as *stagflation*. Economists then proposed a slightly different theory: consumers and workers were concerned about *expected* inflation and unemployment, rather than inflation and unemployment *per se*. If agents can adapt their expectations about the future, their prophecies could become self-fulfilling and the Phillips Curve could only hold in the very short run.

Consider the *stagflation* episode. Theories that came afterwards posed that consumers and workers rationally expect inflation rates to increase as soon as they become aware that the monetary authority plans to embark on expansionary monetary policy, so they would adjust prices and wages before the policy is carried out. As prices and wages go up for reasons unrelated to demand (but rather to expectations of higher future prices and wages), consumption goods and labour demands decrease, counteracting the effects of an expansionary policy. This implies that interest rate setting would have only a limited effect in the short run.

To account for this theory, economists included statistical (rational) expectations of inflation and unemployment in their models. Roberts [1995] derived the version that is still in use in state-of-the-art DSGE models such as Clarida et al. [1999] and Smets and Wouters [2007]. However, as we argued, the assumptions on which these models rely are disconnected from empirical evidence. For this reason, we will use the New Classical version of the Phillips Curve, which reads as follows:

$$\pi - \bar{\pi} = H(u - \bar{u}) + \epsilon \quad (2-2)$$

where π is the inflation rate, $\bar{\pi}$ is the inflation rate associated with the long term steady state growth rate of output (the potential GDP, i.e. the GDP when the economy operates at full capacity) and ϵ is a stochastic variable. H is assumed to be a function such that $\frac{\partial H((u-\bar{u}))}{\partial (u-\bar{u})} < 0$.

First, note that the Phillips Curve's basic form (2-2) accounts for expectations, since we can consider that $E_t[\pi] = \bar{\pi} + \epsilon$, yielding the following expression:

$$\pi = E_t[\pi] + H(u - \bar{u}) \quad (2-3)$$

Evidence largely supports this version of the Phillips Curve or its equivalent, stated in terms of the output gap (i.e. $y - \bar{y}$, as defined in the previous subsection) instead of deviations of unemployment from its long term steady state, a form that can be easily derived by using Okun's law.

Strong empirical evidence across countries can be found in Coibion et al. [2019], which uses the form 3-1, and country-specific evidence abound. In the case of Brazil, Fasolo and Portugal [2004] estimated a convex Phillips Curve using a Kalman Filter and showed that a non-linear relationship between inflation and unemployment exists. Years later, Machado and Portugal [2014] estimated reduced-form Phillips curves for Brazil using the framework proposed by Harvey [1990]. They observe that inflation seems to have responded gradually less to measures of economic activity in recent years, providing some evidence that the Phillips curve has become flatter over the years, but the fundamental correlation still holds.

2.2.3

Liquidity effects

Until 1995, when the seminal paper by Christiano and Eichenbaum [1995] came out, there was little evidence supporting the theory that positive money supply shocks drive short-term interest rates down. These shocks are introduced in the economy when the central bank decides to produce money (by printing or coining) and such increase in its supply devalues the currency, thus decreasing its price (the short-term interest rate). Christiano and Eichenbaum [1995] presented a quantitative, general equilibrium model that reflects this hypothesis, besides evidence that this is consistent with the U.S. data. They argued that a positive money supply shock generates a large drop in the interest rate comparable in magnitude to what they find in the data.

From the quantity theory of money, which is perhaps the most-evidenced economic phenomenon on record [Friedman, 1987], we know that, holding

everything else constant, an increase in money supply has a positive impact on the GDP. This theory can be summarised by the following equation:

$$\begin{aligned} M \cdot V &= P \cdot Q \\ &= Y \end{aligned} \quad (2-4)$$

where M is the money supply, P is the price level, V is the velocity of money, i.e. how many times money exchange hands, Q is the real GDP (the GDP discounted for inflation), Y is the nominal GDP.

An expression in terms of the growth rate of GDP can be derived by rewriting equation 2-4 in terms of logs, where $m = \log(M)$, $v = \log(V)$, and $y = \log(Y)$:

$$m + v = y \quad (2-5)$$

Together with what Christiano and Eichenbaum [1995] demonstrated (that positive money supply shocks drive short-term interest rates down), we obtain a definition for liquidity effects:

$$y = m + v = G(i) \quad (2-6)$$

where i is the interest rate. We assume v is constant and G is a function such that $\frac{\partial G(i)}{\partial i} < 0$.

Minella [2003] investigated monetary policy and basic macroeconomic relationships involving output, inflation rate, interest rate, and money in Brazil. Vector autoregressive estimations suggested that positive interest-rate shocks are accompanied by a decline in money in the periods under study. In addition, the degree of inflation persistence is substantially lower in more recent periods. Thus, liquidity effects are prevalent in Brazil.

2.3

Machine Learning Applied to Monetary Policy

Literature relating machine learning to monetary economics is rather scarce and contributions are recent. Here, we review some recent applications of artificial intelligence to monetary problems and discuss a single paper that is closely related to our work. To the best of our knowledge, a policy rule based on fuzzy control Kukal and Quang [2014] is the only other work that attempts to design a new strategy for interest rate setting.

We begin by summarising the literature related to applications of machine learning to monetary policy, in which our work clearly fits. Rönqvist and Sarlin [2015] and Rönqvist and Sarlin [2017], for instance, built warning

models based on machine learning for the European Central Bank. Rönqvist and Sarlin [2015] attempted to gain insight into bank interconnections by using text-mining techniques on financial discourse. Rönqvist and Sarlin [2017] used neural networks in two stages to construct an indicator for financial distress. In the first stage, news articles were processed to filter those that contained meaningful information for financial stability. In the second stage, these inputs were processed into a signal: distress or tranquillity.

Chakraborty and Joseph [2017], in an effort to convince the economics researchers of the usefulness of artificial intelligence, conducted two case studies. The first models the occurrence of alert scenarios on the balance sheets of financial institutions using only a subset of the available information. They showed that, in this case, machine learning models (namely, naïve Bayes, k-nearest neighbours, decision trees, neural networks, and support vector machines) generally outperform conventional approaches. Second, they predicted United Kingdom's inflation by using the same machine learning techniques as before, demonstrating that a simple forecast combination (based on Timmermann [2006]) of these advanced models performs the best.

Hansen et al. [2017] took a step further and tried to create a tool to better understand monetary theory through United State's recent economic history. He used an unsupervised learning technique to analyse transcripts of Federal Open Market Committee (FOMC) meetings. The authors employed a linear discriminant analysis (LDA), a Bayesian factor model, to study how transparency, a key feature of central bank design, affects monetary policy makers' deliberations. They identified around 10,000 unique words across every transcript produced during Alan Greenspan's tenure as chair. Then, an LDA model was able to boil down into about 40 topics. Each FOMC meeting was then represented as the percentage of time spent on each topic. In this way, the authors were able to construct a dataset from which they found evidence for two effects: a positive discipline effect and negative conformity effect.

Other than the studies presented in the last paragraphs, there is only one more that we consider relevant, even more so because it is closely related to our work. Kukal and Quang [2014] proposed a new model for inflation targeting based on fuzzy control techniques. They applied their system to the context of the Czech economy from year 2000 to year 2011 and compared results with the actual monetary policy. Their model was able to replicate the actions of the Czech Central Bank (CCB) quite accurately. In our view, here lies the main shortcoming of Kukal and Quang [2014] work. Their rules merely automatise the CCB's actual strategy and they never pursuit a method that could improve its performance. As we shall demonstrate, our results suggest

that our framework could be of help in addressing solutions for at least lowering inflation, although (as expected) at the expense of unemployment.

3

Developing the Framework

Our framework is based on well-established empirical evidence. Much of the efforts in the dusk of modern macroeconomics were devoted to derive these empirical relationships from micro-level assumptions, providing microeconomic foundations for macroeconomic models. These assumptions include representative agents, perfect rationality, and perfect knowledge. DSGE models, which we criticised in the previous chapter, ultimately emerged from these studies. For this reason, in the spirit of Colander et al. [2008], we propose to depart from any stringent, *ad hoc*, assumptions and use only facts revealed by data.

The objective of our framework is to minimise future inflation and future unemployment by choosing an appropriate base interest rate. Broadly speaking, the method we propose to solve this problem consists of three steps. First, from three empirical relationships, we derive an objective function that depends on current and past interest rates through future inflation and future unemployment rates. Using data from Brazil, we seek to minimise the value of this function by choosing interest rates via a genetic algorithm. Since the function is forward looking, we use a neural network to predict future values of unemployment and inflation. Figure ?? shows the entire framework from a high-level perspective, with its components, inputs, and outputs.

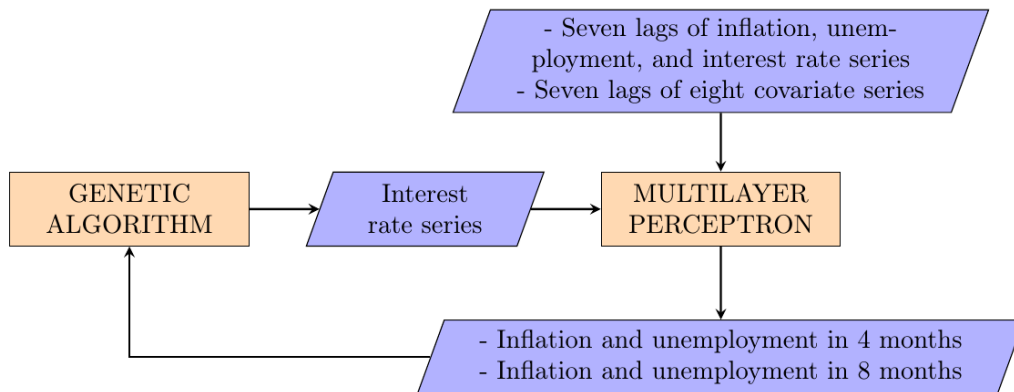


Figure 3.1: High-level diagram of the framework.

3.1

Overview

In this section we present a brief overview of the whole framework. The procedure of minimisation starts with the genetic algorithm generating potential solutions to the central bank's problem. Using these potential solutions (series of interest rates), a neural network predicts future inflation and unemployment rates. These two are necessary to evaluate the objective function for each potential solution and then pick the best among them (the ones for which the objective function are the lowest). If the number of iterations (or generations) has reached a maximum, the procedure stops and the solution is considered to be the one for which the objective function value is the lowest. Otherwise, the genetic algorithm generates new potential solutions from the ones that were previously picked. Figure 3.2 illustrates this procedure.

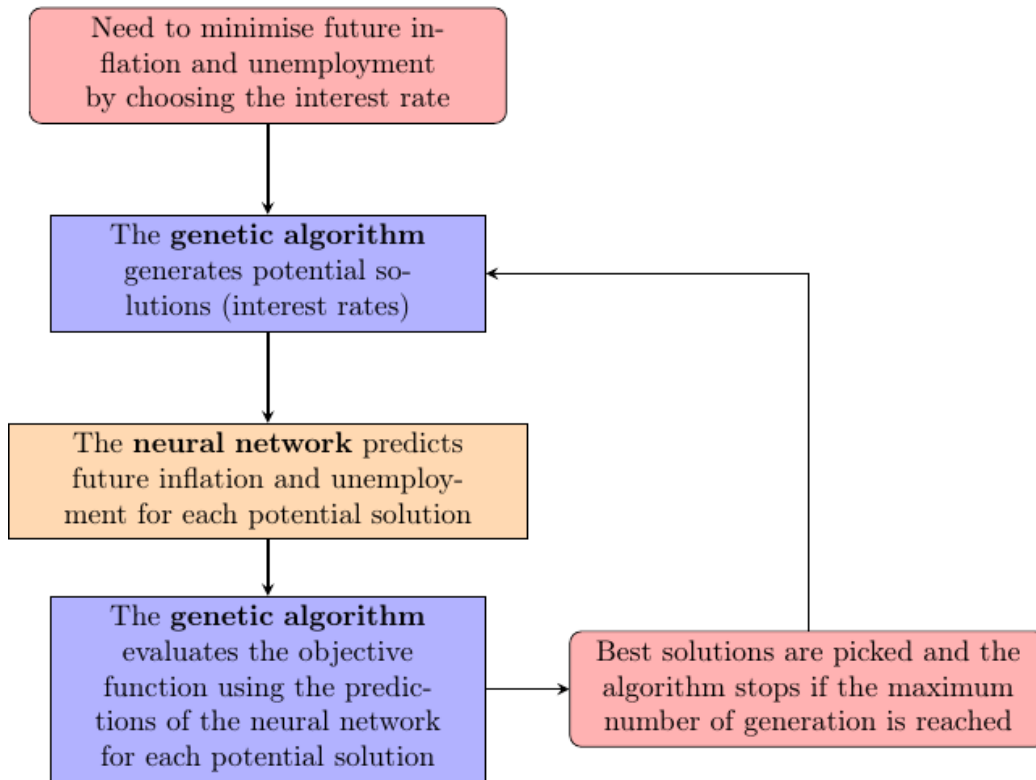


Figure 3.2: Qualitative overview of the framework

Now we detail the procedure a little longer. For clarity, the whole process is illustrated in Figure 3.3. From the three empirical relationships that were fully explained and justified in section 2.2, we derive in section 3.2 an objective function that has the following form:

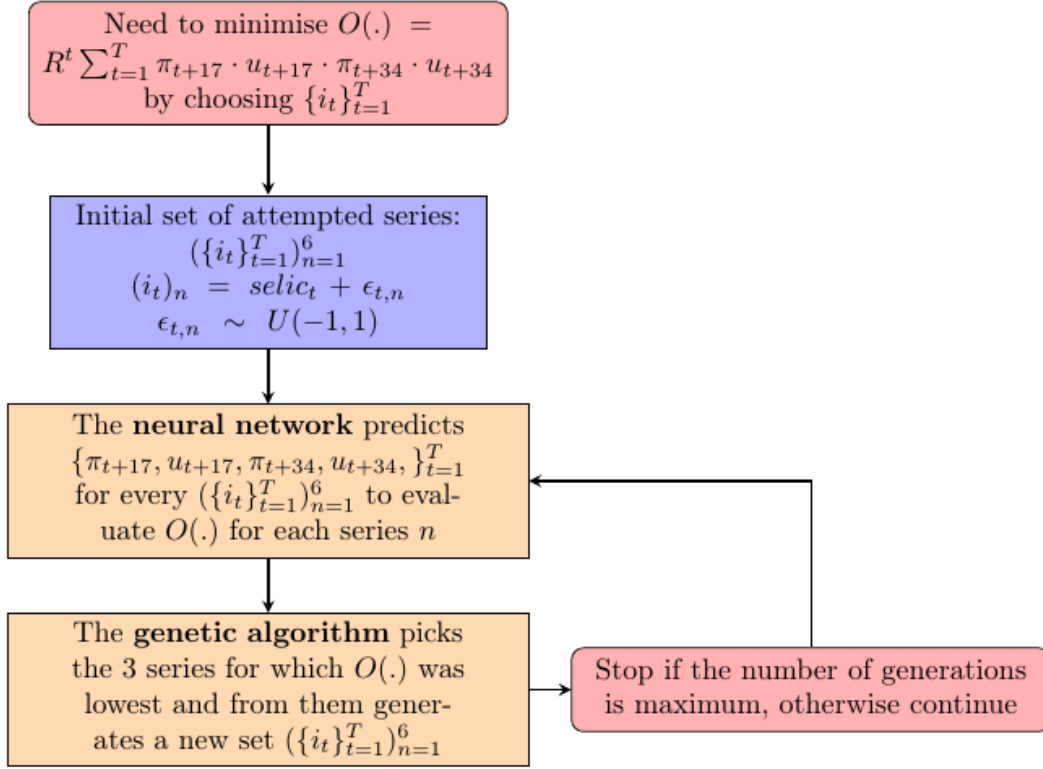


Figure 3.3: Detailed overview of the framework

$$O(.) = \sum_{t=1}^T R^t \cdot E_t[\pi_{t+17}] \cdot E_t[u_{t+17}] \cdot E_t[\pi_{t+34}] \cdot E_t[u_{t+34}]$$

Hence our objective function is a discounted sum of value functions $V(.) = E_t[\pi_{t+17}] \cdot E_t[u_{t+17}] \cdot E_t[\pi_{t+34}] \cdot E_t[u_{t+34}]$ where the discount factor is $R \in [0, 1]$, i.e. inflation and unemployment have more weight in periods that are closer to the current one. The arguments of $V(.)$ are the expected inflation and unemployment seventeen weeks (four months) ahead, as well as inflation and unemployment thirty four weeks (eight months) ahead, and these four variables are actually functions of the interest rate at t . This means that we need to minimise $O(.)$ by choosing a whole series $\{i_t\}_{t=1}^T$ as stated below:

$$\min_{\{i_t\}_{t=1}^T} \sum_{t=1}^T R^t \cdot E_t[\pi_{t+17}] \cdot E_t[u_{t+17}] \cdot E_t[\pi_{t+34}] \cdot E_t[u_{t+34}]$$

Since $O(.)$ is forward looking, we need to predict π and u for the two horizons we defined, i.e. to calculate expected values for future inflation and unemployment. We do so by using a neural network, specified in section 3.4.

3.2

Central Bank's Objective Function

We now derive the monetary authority objective function. First, we need to present some extensions to the fundamental empirical relationships we described in section 2.2.

Let η be a random shock and B a function such that $\frac{\partial B((y-\bar{y}))}{\partial (y-\bar{y})} < 0$. Then, by combining equations 2-1 and 2-2, we obtain a version of the Phillips Curve in terms of the output gap:

$$\begin{aligned}\pi - \bar{\pi} &= H(F(y - \bar{y}) + \kappa) + \epsilon \\ &= B(y - \bar{y}) + \eta\end{aligned}\tag{3-1}$$

As for liquidity effects, we know that the interest rates do not have an immediate impact on GDP [Boivin et al., 2010], so we consider a more general structure for G , allowing for arbitrary lags on i . We call this new function $G^L(i)$, where L is the number of lags on i and d is an integer representing the distance of the first lag:

$$y_t = G^L(i) \equiv G(i_{t-d}, i_{t-d-1}, \dots, i_L)\tag{3-2}$$

Now let us establish the central bank's problem in formal terms. At each time period, the central bank chooses an interest rate i to minimise a value function $V(\cdot)$, to be defined in the present section:

$$\min_i V(\pi, u)\tag{3-3}$$

where, as before, π (inflation) and u (unemployment) are functions of i and random shocks on the economy's fundamentals, but we drop the arguments for notation clarity.

As we anticipated in section 3.1, we actually solve a broader problem than the one represented by 3-3, choosing a whole series $\{i_t\}_{t=1}^T$ to minimise a discounted sum of value functions, where $R \in [0, 1]$ is the discounting factor.

$$\min_{\{i_t\}_{t=1}^T} \sum_{t=1}^T R^t V(\pi_t, u_t)\tag{3-4}$$

The reason why we decided to search for the whole series instead of, for instance, its last element, is that we need the algorithm to capture the whole dynamics of inflation and unemployment's response to interest rates over time. If we had optimised the objective function for only one period, the interest rate would simply be chosen to be as large as required for inflation to be zero, or as small as required for unemployment to be zero. In other words,

disregarding additional information about variables' behaviour would imply selecting unfeasible levels of interest rates, either too high or too low¹.

Now, by using the relationships exposed in sections 2.2.1, 2.2.2, and 2.2.3, we show that π and u are indeed functions of i and random shocks. First, we take the expression for liquidity effects (3-2) and use it in Okun's law (2-1) to determine an expression for u :

$$\begin{aligned} u &= \bar{u} + F(y - \bar{y}) + \kappa \\ &= \bar{u} + F(G^L(i) - \bar{y}) + \kappa \end{aligned} \quad (3-5)$$

Hence, the unemployment growth rate u is a function of two constants (the long-term steady state unemployment growth rate, \bar{u} , and the equivalent for the output growth rate, \bar{y}), shocks to unemployment, κ , and lags of interest rate i (recall that $G^L(i) \equiv G(i_{t-1}, i_{t-2}, \dots, i_L)$). Next, we substitute the expression for the unemployment growth rate (3-5) into the Phillips Curve (2-2) to obtain an equation for the inflation rate π :

$$\begin{aligned} \pi &= \bar{\pi} + H(\tilde{u} - \bar{u}) + \epsilon \\ &= \bar{\pi} + H(F(G^L(i) - \bar{y}) + \kappa) + \epsilon \end{aligned} \quad (3-6)$$

As a result, the inflation rate π is a function of two constants (the long-term steady state inflation rate $\bar{\pi}$ and its equivalent for the output growth rate, \bar{y}), shocks to unemployment, κ , shocks to inflation, ϵ , and lags of the interest rate i (again, recall definition 3-2).

From the last two equations (3-6 and 3-5), it follows that the central bank problem can be restated as a function of lags of the interest rate, constants from the economy, and stochastic variables:

$$\min_{\{i_t\}_{t=1}^T} \sum_{t=1}^T R^t V(\bar{\pi} + H(F(G^L(i) - \bar{y}) + \kappa) + \epsilon, \bar{u} + F(G^L(i) - \bar{y}) + \kappa) \quad (3-7)$$

We will assume that $V(\cdot)$ is a Cobb-Douglas, so it has the general form presented below:

$$z = x_1^{a_1} \cdot x_2^{a_2} \cdot \dots \cdot x_n^{a_n} \quad (3-8)$$

This family of functions possesses a number of important properties that have made it widely useful in the analysis of economic theories. First, returns

¹We tested the possibility of finding only the most recent value of the interest rate and, indeed, figures were as high as three digits or negative. Both would be absurd in any real economy: skyrocketing interest rates lead to severe recessions, while negative interest rates could possibly steer inflation towards an ascending spiral.

to scale is measured by the sum of its exponents i.e., $a_1 + a_2 + \dots + a_n$, hence it is easy to verify whether the central bank is responding aggressively (in which case returns to scale would be increasing) or passively (decreasing or constant). Secondly, marginal returns of each variable (i.e. x_1, x_2, \dots, x_n) depends on the levels of those variables, given the value of the others. This makes it easier to interpret how the value function varies with the inputs. In particular, the exponents represent the elasticities of each variable with respect to the value of the function.

If we plug in the Cobb-Douglas format and go back to the variables of interest the central bank problem becomes:

$$\begin{aligned} & \min_{\{i_t\}_{t=1}^T} \sum_{t=1}^T R^t \cdot \pi^{a_1} \cdot u^{a_2} \\ & \equiv \min_{\{i_t\}_{t=1}^T} \sum_{t=1}^T R^t (\bar{\pi} + H(F(G^L(i) - \bar{y}) + \kappa) + \epsilon)^{a_1} \cdot (\bar{u} + F(G^L(i) - \bar{y}) + \kappa)^{a_2} \end{aligned} \quad (3-9)$$

As explained in section 2.2.2, since after the *stagflation* phenomenon, economists (including central bankers) consider that consumers and workers are concerned about expected (future) inflation and unemployment. If this was not the case, the Phillips Curve would only hold in the very short-run. Our value function is fully compatible with this modern view: period by period, the algorithm chooses the interest rate that will minimise future inflation and future unemployment, because the function depends on lags of the interest rate, as expressed by $G^L(i)$.

To clarify this point, we simplify the notation a little. First, we forward the value function, such that it is easier to see that the algorithm is choosing the interest date *today* to influence inflation and unemployment *tomorrow*. We call the time period between today and tomorrow d , to match definition 3-2. Then, since there are stochastic components in π and u , we take their expected values:

$$\begin{aligned} V_t(\pi_{t+d}, u_{t+d}) &= E_t[\pi_{t+d}]^{a_1} \cdot E_t[u_{t+d}]^{a_2} \\ &= (\bar{\pi} + H(F(G^L(i) - \bar{y}) + \kappa) + \epsilon)^{a_1} \cdot (\bar{u} + F(G^L(i) - \bar{y}) + \kappa)^{a_2} \end{aligned} \quad (3-10)$$

Any choice of how far is the future (in our notation, it means how large is d) is good as any other, so here we define that $\pi \equiv \pi_{t+17}$ and $u \equiv u_{t+17}$ ², i.e.

²We tried several other horizons, but this one delivered the best results.

both inflation and unemployment four months ahead. Moreover, we add π_{t+34} and u_{t+34} to the value function with two purposes: (i) to increase the precision of the estimates, and (ii) to account for the fact that central banks also care about the long-term. Therefore, we will work with the following value function:

$$V_t(\pi, u) = (E_t[\pi_{t+17}]^{a_1} \cdot E_t[u_{t+17}]^{a_2} \cdot E_t[\pi_{t+34}]^{a_3} \cdot E_t[u_{t+34}]^{a_4}) \quad (3-11)$$

We lighten the notation further, by defining $x_{t+d}^e \equiv E_t[x_{t+d}]$:

$$V_t(\pi, u) = (\pi_{t+17}^e)^{a_1} \cdot (u_{t+17}^e)^{a_2} \cdot (\pi_{t+34}^e)^{a_3} \cdot (u_{t+34}^e)^{a_4} \quad (3-12)$$

Parameters a_1 , a_2 , a_3 , and a_4 reflect how much the central bank cares about mid-term inflation and unemployment compared to how much it cares about long-term inflation and unemployment. We assume $a_1 = a_2 = a_3 = a_4 = 1$, since we believe there are reasons to impose that the central bank should be more concerned with the mid-term indicators as opposed to long-term, or inflation as opposed to unemployment, both being crucially harmful for society as argued in section 1.1. This yields the final form of the objective function $O(\cdot)$ already presented in section 3.1:

$$O(\cdot) = \sum_{t=1}^T R^t \cdot \pi_{t+17}^e \cdot u_{t+17}^e \cdot \pi_{t+34}^e \cdot u_{t+34}^e \quad (3-13)$$

As we will explain in detail in the next section, the problem of minimising $O(\cdot)$ is solved by using a genetic algorithm.

3.3 Genetic Algorithm

The algorithm we use to minimise 3-13 by choosing $\{i_t\}_{t=1}^T$ is illustrated in figure 3.4 and works as follows. An individual (chromosome) is the whole SELIC rate time series, as explained in section 3.1. Naturally, a gene is considered to be an observation of the series. But in order to better understand the structure of the chromosome, we need to say something about the data. In a subsequent section (4.1), we will describe the data we used in detail. For now, it suffices to say that we will work with weekly data for inflation, unemployment, and covariates, and that there are 728 observations for each series (so $T = 728$). The algorithm solves for the whole SELIC time series, but there is no need to find 728 values. In fact, the CBB chooses the SELIC every seven weeks so that an individual actually has $N \equiv T/7 = 728/7 = 104$ genes.

Crossover operation is standard: new individuals are generated from $N/2$ random genes from two parents, using a one-point cut in each chromosome (individual). As for mutation, a percentage of the total number of genes are

randomly picked to be replaced by a random value in the interval $[-3, 3]$, spaced by 0.1, which are reasonable values for standardised interest rates. The percentage of mutated genes decay with generations: it starts at 7% and is reduced by a factor of 1.001 until it reaches the lower bound of 3%.

The initial population is composed of six individuals: the actual SELIC rate and five other individuals generated from the actual SELIC series by adding an stochastic component, namely, a uniform random variable between -1 and 1. By using the neural network that will be completely described in section 3.4, we predict the values of unemployment and inflation, both seventeen weeks (four months) and thirty four weeks (eight months) ahead, for each observation in each SELIC series (each individual). We then are able to evaluate the objective function $O(\cdot)$ for each one of the six initial individuals. The weekly discount rate (R) we use is 0.0012. This way, $(1 + R)^{52} \approx 6.5$, which is the mean annual SELIC interest rate over the period covered by the data. Using the base interest rate is standard in economic literature [Frederick et al., 2002].

Next, we pick the two individuals for which $O(\cdot)$ are the lowest and a random individual, the three survivors of the current generation. We apply the crossover and mutation operations to these survivors to obtain six new individuals (by combining all possible pairs of survivors while crossing over), the members of the new generation. If the lowest value $O^*(\cdot)$ of the objective function among all members of the new generation is lower than the lowest value for the previous generation, $O^{-1}(\cdot)$, we substitute the new generation for the old one. If not, we discard the new generation and apply the operators to the previous one again, restarting the cycle. The algorithm stops after two thousand generations, the reason being explained in the next chapter.

3.4

Neural Network Regression

At each week³, the central bank will estimate future unemployment and inflation for a given attempted solution i , in order to evaluate the objective function derived in section 3.2, as part of the algorithm defined in section 3.3. We use a multilayer perceptron (MLP) for this purpose. The inputs of the MLP are seven lags of the following variables:

- Unemployment rate
- Inflation rate
- Base interest rate (SELIC)

³In a subsequent section (4.1), we will describe the data we used in detail

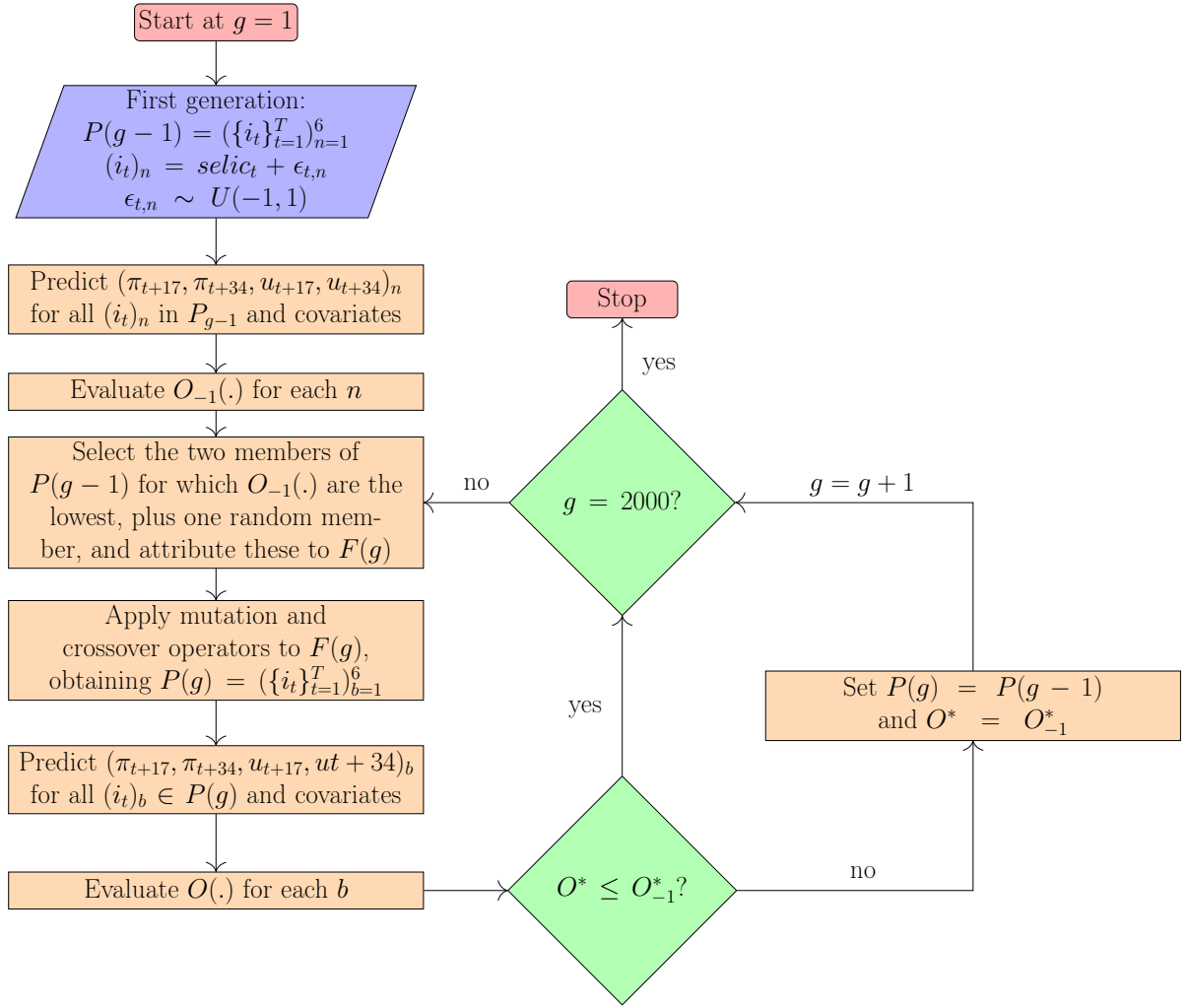


Figure 3.4: Genetic Algorithm Flowchart

- Effective interest rate, which is slightly different from the SELIC, since it is determined by the market
- Net government debt
- Government deficit
- Net current account
- Net capital account
- GDP
- Exchange rate with respect to US dollars
- United States official interest rate

There are four outputs: unemployment in seventeen weeks (four months ahead), inflation in seventeen weeks, unemployment in thirty four weeks (eight months ahead), and inflation in thirty four weeks. All values are normalised to increase convergence speed.

The reason why we chose a window of size eight is that the CBB makes its decision about the base interest rate every seven weeks. Therefore, we assume it considers current variables as well as their values on each of the six previous weeks. As for the variables themselves, they have long been established in economic literature to have predictive power over inflation and unemployment. Pointedly, we use a subset of variables that figure in the classical paper by James Stock and Mark Watson Stock and Watson [1999], which focus on forecasting inflation, and in a more recent work by Regis Barnichon and Paula Garda Barnichon and Garda [2016], about unemployment prediction.

4

Simulations

In this chapter we present the time series we have used, the results of the simulations, and discuss their theoretical soundness, as well as matters of practical applicability. Results suggest that had the Brazilian central bank applied our strategy, all other economic conditions (apart from monetary policy) remaining equal, inflation could have been lower most of the time. However, the unemployment rate would have been higher quite frequently, and in our solution the interest rates series' persistence, as measured by autocorrelation, is possibly too weak, which goes against the best practices for monetary policy. We then propose that our method be used by central banks as yet another statistical tool to guide interest rate setting, rather than the main strategy, and delineate ways it could be used in practice.

4.1

Data

The time series we used were all extracted from the CBB website [SGS, 2019] using its publicly available API, with the exception of unemployment. In this case, National Continuous Household Sample Survey (PNADC, from portuguese *Pesquisa Nacional por Amostra de Domicílios Contínua*) series, which begins at 2012 and was too short for our purposes, was backward-extrapolated by Fundação Getúlio Vargas¹.

A complete list of the series, their frequencies and precise sources is shown in table tab:series. We observe that most series were originally monthly ones, but if we kept this frequency, there would be not enough data points to train the neural network. Since we also have daily series, we decided to work with weekly data. Values for monthly data were repeated for all the weeks of a given month, while data for daily series were collapsed in weekly averages. As a result, we obtained 728 observations per series, stretching from the 48th week of year 2001 to the 34th week of 2016.

¹We thank Fundação Getúlio Vargas for providing this series for our study.

Description	Frequency	Source
International reserves (US\$ millions)	Daily	BCB
Net public sector debt (%GDP)	Monthly	BCB
Public sector borrowing requirements (%GDP)	Monthly	BCB
United States official interest rates (%p.y.)	Monthly	FED
Net current account (US\$ millions)	Monthly	BCB
Net capital account (US\$ millions)	Monthly	BCB
Exchange rate US\$	Daily	BCB
National consumer price index (in 12 months)	Monthly	IBGE
SELIC target (%p.y.)	Daily	BCB
SELIC in annual terms (basis 252, %p.y.)	Daily	BCB
GDP accumulated in the last 12 months (current R\$ millions)	Monthly	BCB
Unemployment rate (PNADC, chained backwards)	Monthly	IBGE/FGV

Table 4.1: Time series used in this study

4.2 Results

First, we describe the MLP thoroughly and present its predictions (fitted values and forecasts for the testing set) for unemployment and inflation. Next, we show the genetic algorithm evolution and convergence. Finally, we provide evidence that our method might improve the central bank's performance.

4.2.1 Neural Network's Predictions

We have tested a large range of parameters using a grid search² to choose the most appropriate configuration for the MLP. The choice procedure consisted on two nested cross-validation loops, performed over 80% of random observations in our sample. This process, also known as double cross-validation, mitigates bias introduced by a fixed training set and works as follows:

1. In the outer (external) loop, the data set is divided into two subsets, training (70%) and validation (30%), and this is done ten times, so that there are ten disjoint validation sets.
2. The training sets are used in the inner (internal) loop for model building and model selection. They are repeatedly (also ten times) split into calibration (70%) and validation (30%) sets. Calibration sets are used to develop different models, whereas validation sets are used to estimate errors.

²The algorithm was coded by Pedregosa et al. [2011] and is available in *sklearn* package for Python. The neural network estimator was also obtained from this package.

3. In the inner loop, for each external training set, the model with the lowest prediction errors in terms of the mean root mean square error (RMSE) over the 10 internal folds is selected.
4. In the outer loop, for each selected model, the external validation set is employed to evaluate model performance.
5. Finally, the model with lowest RMSE is chosen.

The remaining 20% of observations have been used as a final testing set, again to avoid training bias.

The mean RMSE over the 10 outer folds was of 0.104 and the best configuration turned out to be the following: (i) the activation function is the hyperbolic tangent (HTAN); (ii) there are four hidden layers with five neurons each; (iii) the solving method is the limited-memory Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm, based on Fletcher [2013]. The accuracy of this network, as defined by the RMSE, is of 0.095 for the training set and of 0.149 for the testing set, so the fact that the preferred configuration for the MLP does not suffer from overfitting seems warranted.

The other options for the activation function were the rectified linear unit function (RELU) and the logistic sigmoid function (LSIG), while the other possibilities for the solving method were the standard stochastic gradient descent (SGD) and the stochastic gradient-based optimiser proposed by Kingma and Ba [2014], the adaptive moment estimator (ADAM). The whole set of parameters used in the grid search are shown in table tab:pars.

Parameter	Values
Size of the Hidden Layer	3 to 5
Number of Hidden Layers	2 or 4
Solving Algorithm	ADAM, SGD, or BFGS
Activation Function	RELU, HTAN, or LSIG

Table 4.2: Set of parameters used in the grid search

The preferred neural network delivers fairly accurate fitted values for a 4-month horizon, as Figures 4.1 and 4.2 show. As a whole, predictions for both variables were satisfactory. It is clear, though, that the network captured volatility much better in the case of inflation as compared to unemployment. For inflation, in addition, adherence was almost perfect in the first five and last three years, although some mid-term fluctuations were not reproduced with desirable accuracy.

For 8-months-ahead predictions, we see in Figures 4.3 and 4.4 that adherence of fitted values was weaker than in the latter case, with a shorter horizon. This was expected, since further values in the future are more difficult to predict accurately. In both cases, the network captured long-term trends significantly better than it did for short-term fluctuations. This is consonant with a network that does not suffer from overfitting. We consider these results to be adequate for our purposes, since 8-months-ahead inflation and unemployment can be thought of as being long-term economic variables, so that only average values (or trends) would matter in terms of policy. In other words, precise values would not have a significant impact in interest rate setting; in the long-term, the monetary authority is more concerned about broad movements in the variables of interest, instead of pursuing exact targets.

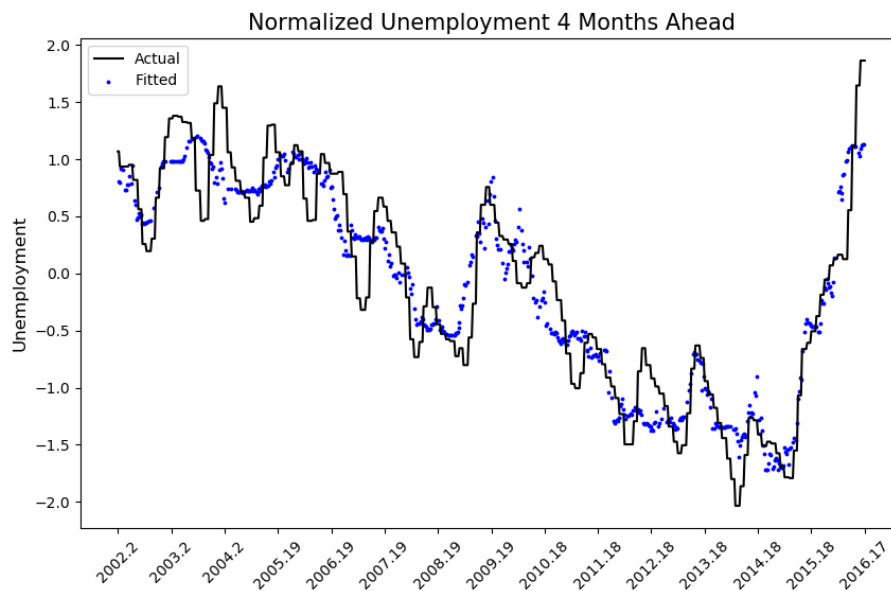


Figure 4.1: Four-months-ahead unemployment: real data versus neural network's fitted values

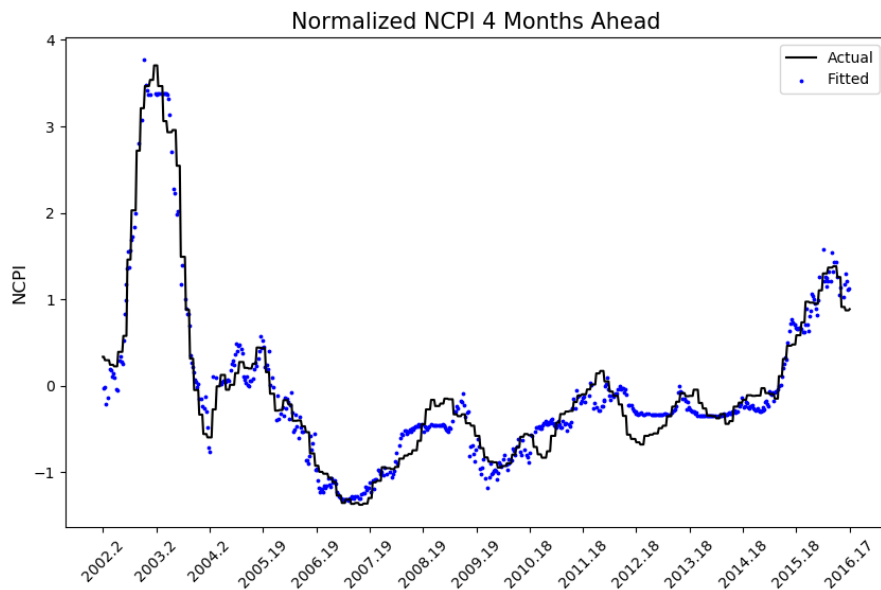


Figure 4.2: Four-months-ahead inflation: real data versus neural network's fitted values

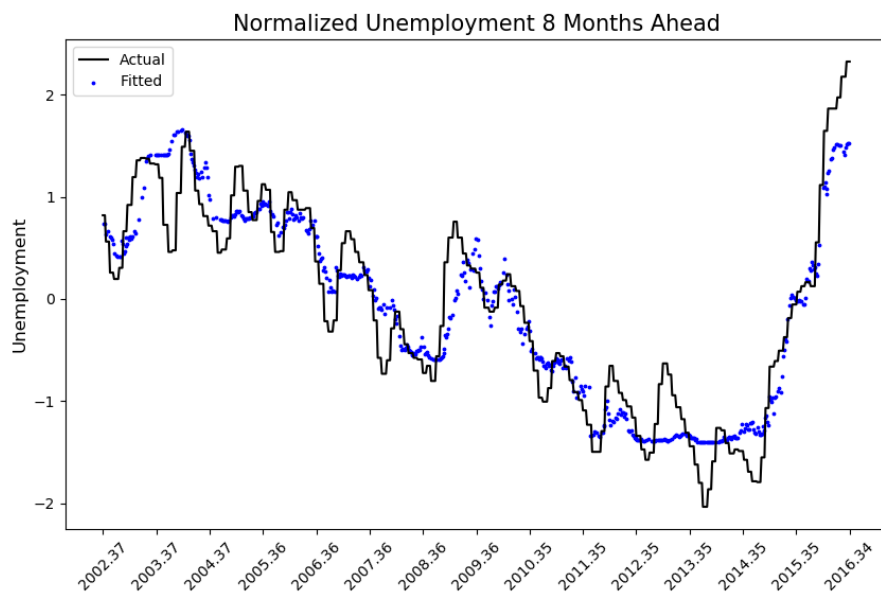


Figure 4.3: Eight-months-ahead unemployment: real data versus neural network's fitted values

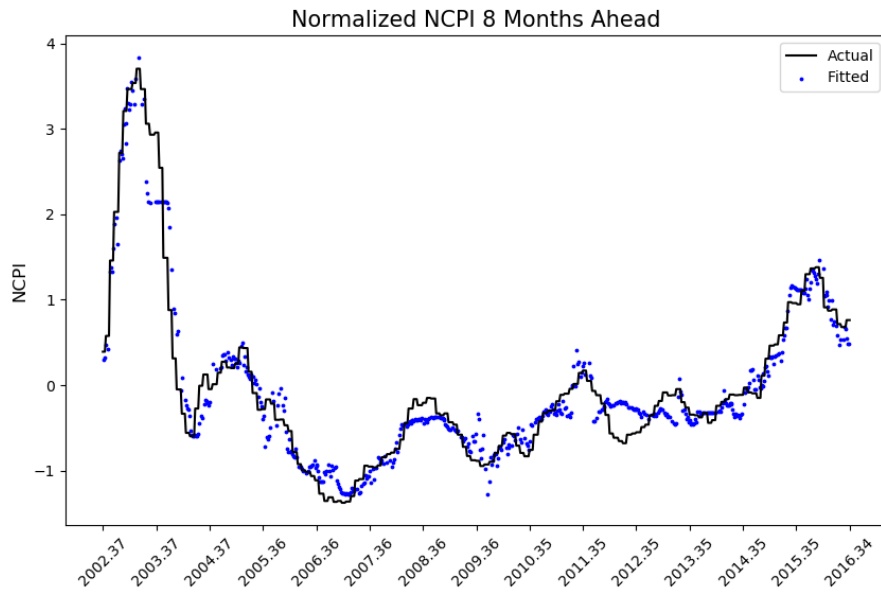


Figure 4.4: Eight-months-ahead inflation: real data versus neural network's fitted values

Network's predictions for the testing set were satisfactory too, as we can see in Figures 4.5 to 4.8 and could conclude from observing the low RMSE. However, errors in the case of inflation are generally lower relative to unemployment, for both horizons. As we shall explain, this might have had a consequence: simulations as a whole performed worse in terms of unemployment than in terms of inflation.

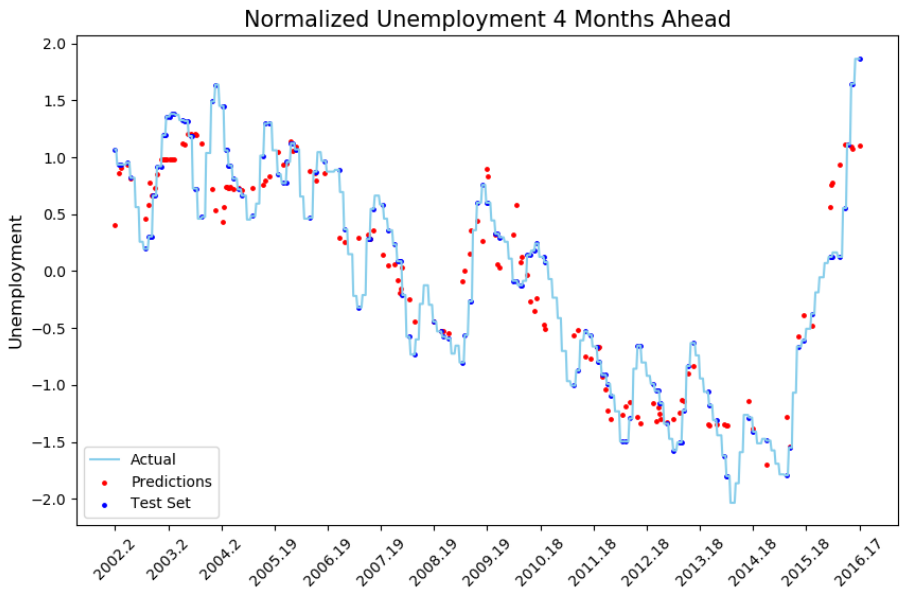


Figure 4.5: Four-months-ahead unemployment: testing set versus out-of-sample predictions

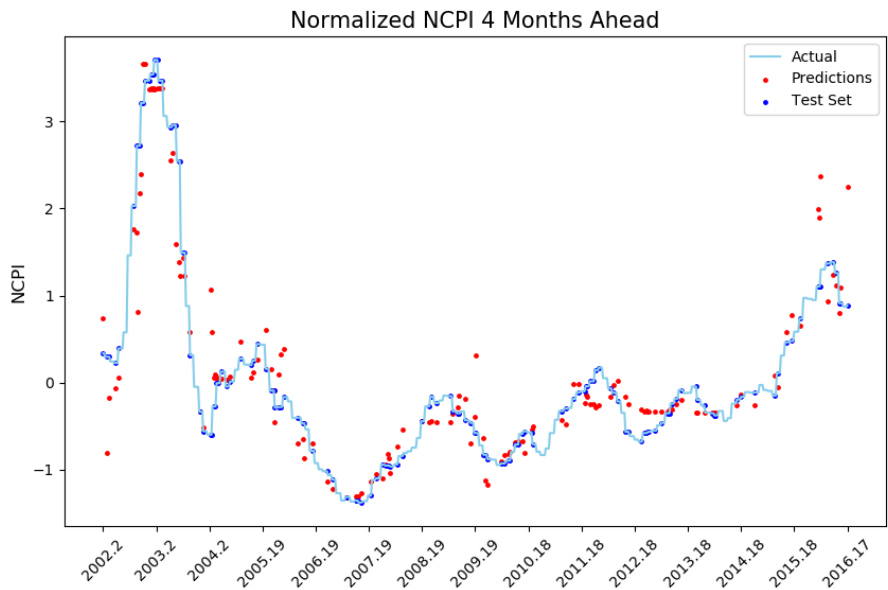


Figure 4.6: Four-months-ahead inflation: testing set versus out-of-sample predictions

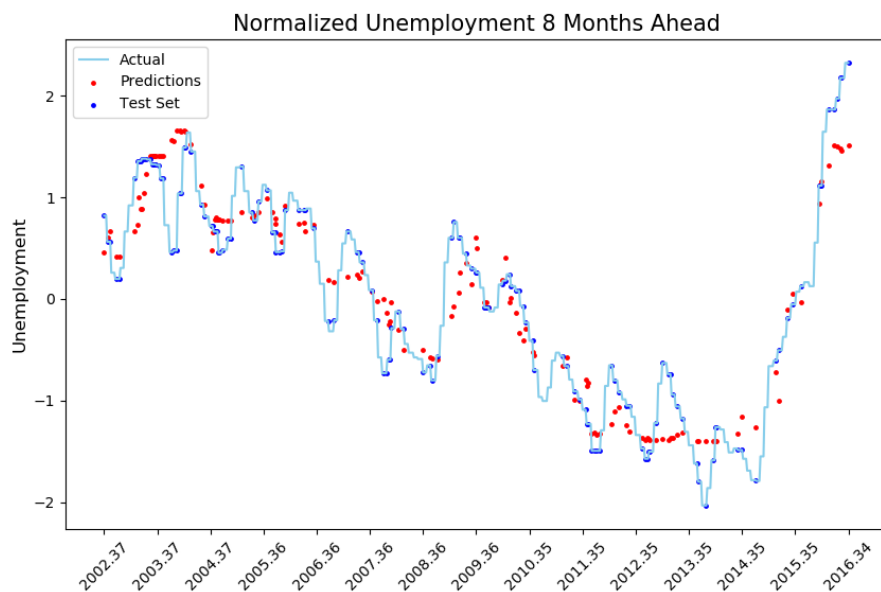


Figure 4.7: Eight-months-ahead unemployment: testing set versus out-of-sample predictions

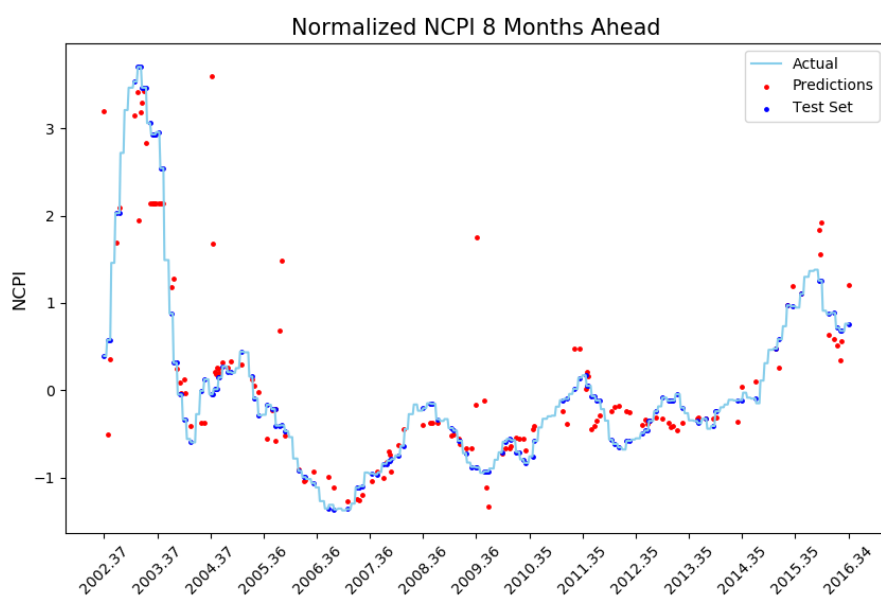


Figure 4.8: Eight-months-ahead inflation: testing set versus out-of-sample predictions

4.2.2

Genetic Algorithm Evolution

Testing a range of parameters for the genetic algorithm was a challenge, since each battery of 1000 generations took approximately five hours³ to run when the least number of survivors per generation (2) was set. When we added one more survivor, the running time almost doubled: eight hours per 1000 generations. With 4 survivors per generation the running time got prohibitive and we simply interrupted the algorithm. A similar running time hike happened when we increased the initial population from 6 to greater numbers, or when we changed the crossover operator from one-point cut to two-points cut. The full set of tested values for parameters are presented in table 4.3.

	Parameter	Values
Mutation	Rate	0.03, 0.04, 0.06, 0.07, or [0.03,0.07]
	Decay	1.001 or 1.0005
	Operator	Random choice or random sum
Crossover	Operator	One point
Survivors	Number	2, 3 , or 4
Initial Population	Number	6 , 12, or 18

Note: Preferred values are highlighted in blue.

Table 4.3: Set of tested parameter values for the genetic algorithm.

The other values were rejected on the basis of convergence speed. Fixed mutation rates reduced convergence speed, and the same was observed if the mutation rate was allowed to decrease from 7% to 3% at a rate of 1.0005, or if the mutation operator was adding to the chosen genes a random draw from a standard normal distribution (random sum) and not substituting random draws from a set of pre-determined values for these genes (random choice).

Figure 4.9 shows the value objective function for the best individual in each generation for our preferred configuration. Most of the evolution happened before 250 generations and convergence seems to have been achieved around the 1600th generation. We tested for 4000 generations and there were no significant changes, i.e. the best individual improved only by a few hundreds. However, 2000 more generations had a large cost in terms of time (10 hours, due to nested loops inside the crossover operator), so the improvement obtained with extra 2000 generations was relatively insignificant.

³The machine in which we ran the algorithm has the following specifications: processor Intel(R) Core(TM) i7-4700MQ, CPU 2.40GHz; memory 16gb; operating system Ubuntu 18.04.2 LTS

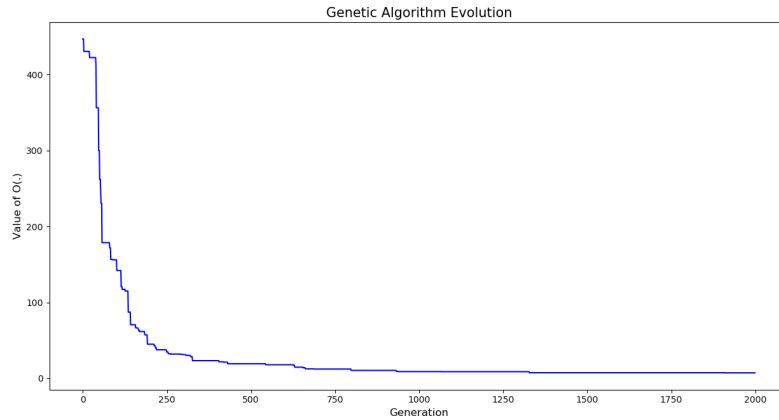


Figure 4.9: Value function for the best individual in each generation

4.2.3 Performance

We now assess our strategy's performance. First, we compare the actual (realised) SELIC series with the one generated by the genetic algorithm (the best individual after 2000 generations). Both are depicted in Figure 4.10.

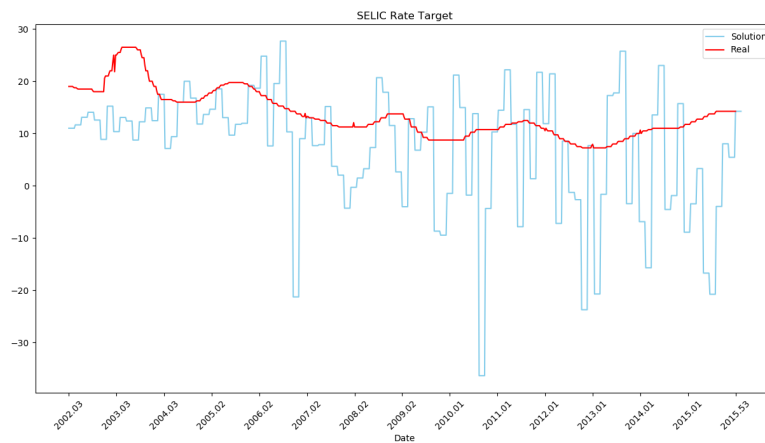


Figure 4.10: SELIC target: actual series versus solution generated by the genetic algorithm

We see that, according to the strategy, the SELIC rate should in general have been lower. This is especially true in the two extremes of the period covered by our model. Nonetheless, those were times of particularly high inflation and scholars often agree that the monetary policy was correct and should not have been softer (more expansionary), at least according to their models Abreu et al. [2014] and Giambiagi and Villela [2016]. Our results also

imply lower SELIC rates in the period in between, but the difference in this case is not so pronounced. The possible consequences of low interest rates in both sub-periods are discussed in section 4.3.1.

Moreover, it is clear that the generated SELIC series is significantly less persistent than the original series, albeit remaining in a shorter interval. Low persistence might pose a problem. According to several authors, for monetary policy to work, agents must be able to understand the central bank behaviour in order to adjust their expectations and make the decisions that, ultimately, policy-makers expect them to [Clarida et al., 1999, Woodford et al., 1999, Woodford, 1999].

In agreement with this view, the central bank must have credibility and be able to commit to a certain smooth path for the target interest rate. If the SELIC frequently suffers large shocks, agents may not be able to believe in the central bank's commitment, even if the institution anticipates these changes to the public. The policy would possibly be seen as discretionary, leading to negative effects on the economy, as we further explain in section 4.3.2.

To avoid these possible deleterious effects, we smooth the SELIC series⁴, calculating its 12-month moving average (MA). As we note in figure 4.11, results now seem much less likely to cause the problems discussed above.

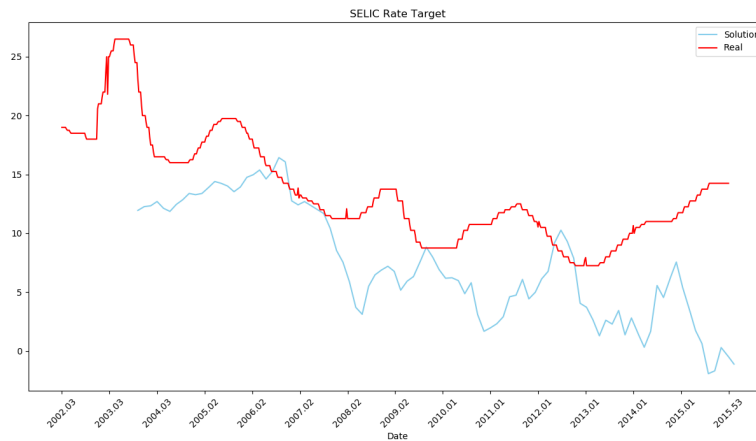


Figure 4.11: SELIC target: actual series versus 12-month MA solution. First values were lost when calculating the MA.

One last step is necessary to evaluate our strategy. We use the neural network to predict the values of unemployment and inflation for the MA version of our solution and compare it to actual inflation and unemployment. Figures 4.12 to 4.15 show comparisons for both variables and horizons. We see that

⁴We could also have used core inflation instead of inflation as an input, to try to smooth the interest rate *ex-ante*. However, we leave this for future works.

inflation, for both horizons, was quite frequently lower when predicted by our solution as compared to what was observed under CBB actual policy decisions, while the same cannot be said about the unemployment rate.

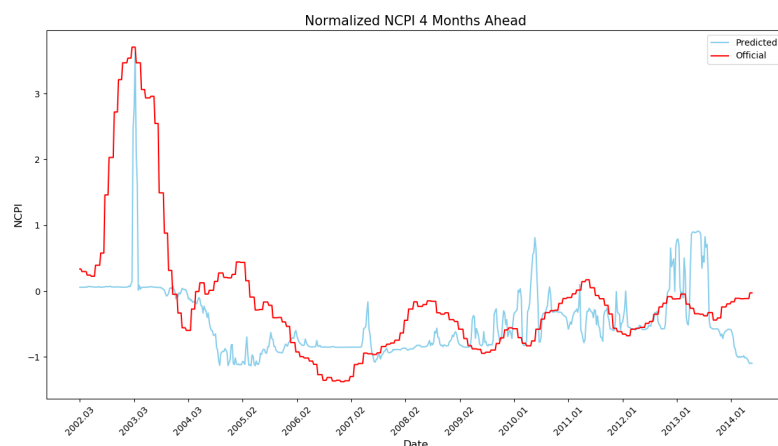


Figure 4.12: Predicted 4-months-ahead inflation and actual inflation

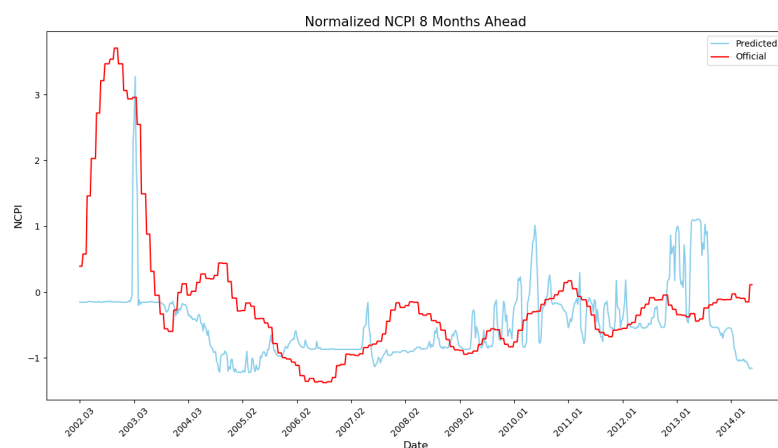


Figure 4.13: Predicted 8-months-ahead inflation and actual inflation

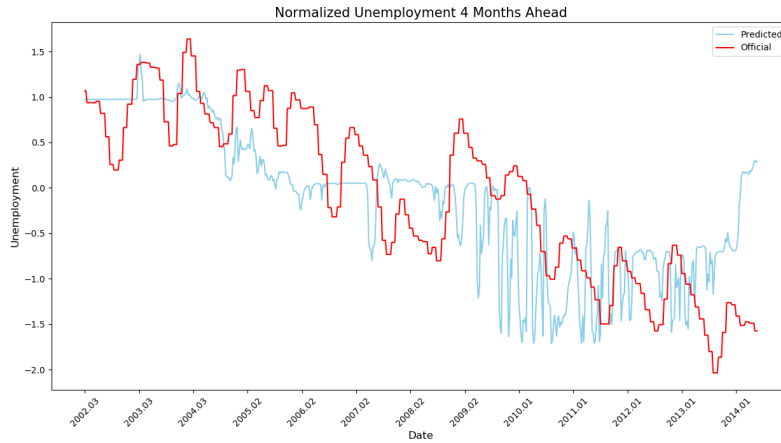


Figure 4.14: Predicted 4-months-ahead unemployment and actual unemployment

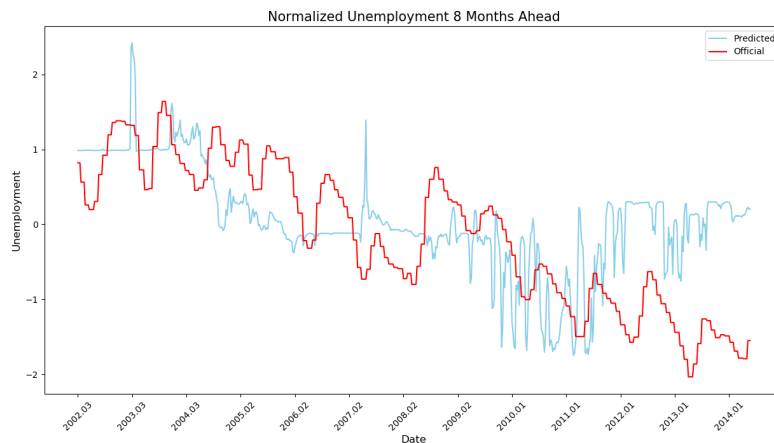


Figure 4.15: Predicted 8-months-ahead unemployment and actual unemployment

Table 4.4 summarises what has been shown in the figures. On average, predicted inflation was lower than actual inflation for 62.48% of the time. Predicted unemployment, however, was lower only for 39.69% of weeks. This result might seem undesirable, but we need to stress, as we did in section 2.2.2, that there is a trade-off between the expected inflation and the unemployment rate, so in fact the outcome could be considered reasonable. It is possible, then, that the objective function has another minimum, where predicted unemployment rate would be lower at most of the periods, as opposed to predicted inflation.

Another set of results is worth noting. As a rule, predicted inflation and unemployment rate were less volatile than actual inflation and unemployment. This means that both variables remained bounded in a smaller interval, i.e.

	Lower (%)	Relative Volatility	Relative Persistence
NCPI 4 mo	58.6	-0.07	-0.059
Unemployment 4 mo	45.27	-0.24	0.005
NCPI 8 mo	66.36	-0.08	-0.063
Unemployment 8 mo	34.11	-0.45	-0.009
SELIC	94.88	0.11	0.001

Table 4.4: Descriptive statistics of genetic algorithm solution as compared to actual CBB solution

our strategy reduced the variance of both variables for the two horizons. This, of course, is a desirable feature of prospective inflation and unemployment, as it implies less uncertainty in the economy Bachmann et al. [2013].

In spite of that, jumps were more frequent, with persistence, as estimated by an order one auto-regressive model (AR(1)), being lower in all cases but the 4-months ahead unemployment and the SELIC. The MA version of the SELIC series generated by our strategy is, in its turn, lower than the actual SELIC for 94.88% of the time, which is a positive outcome in general, but there are arguments against a reduced SELIC for some periods covered by our data, as we detail in the next section (4.3.1). In addition, by using the same criteria as before, the MA of the SELIC series found by our algorithm is relatively more volatile, albeit almost as persistent as the CBB SELIC.

It should be pointed out that weaker results for unemployment rates could have been a consequence of the worse performance of the neural network with regard to this variable. To check the robustness of predicted unemployment rates, we made a minor change to the structure of our strategy, seeking to improve the neural network's forecasts. We did so by training separate networks for mid (4 months ahead) and long-run (8 months ahead).

Configurations were chosen by following the same procedure described in section 4.2.1 (nested cross-validation). The network used to predict long-term variables turned out to be equal to the network we used previously, whereas the one intended to forecast short-term values was slightly different, with two five-neurons hidden layers, instead of four. Performance as measured by RMSE over the testing set's prediction errors was better in both cases (the mean was 0.057), but simulations ended up performing poorer. Results are reported in table 4.5 and suggest that original outcomes were not adversely impacted by the neural network.

	Lower (%)	Relative Volatility	Relative Persistence
NCPI 4 mo	56.9	0.16	-0.073
Unemployment 4 mo	46.98	-0.21	-0.002
NCPI 8 mo	51.63	0.02	-0.088
Unemployment 8 mo	45.89	-0.11	0.002
SELIC	77.05	0.23	0

Table 4.5: Descriptive statistics of genetic algorithm solution as compared to actual CBB solution

4.3

Discussion

4.3.1

The case against a lower SELIC

As already noticed, our strategy suggests that the SELIC rate should have been lower for more than 94% of the covered weeks, specially towards the beginning and the end of the period. However, scholars often agree that, specifically around these two periods, monetary policy should not have been looser [Giambiagi and Schwartzman, 2014, Abreu et al., 2014, Giambiagi and Villela, 2016]. In 2002, the CBB was forced to implement a SELIC hike, because the market was still fearful of the upcoming policy by left-wing elected president Luiz Inácio Lula da Silva. Besides, a high SELIC was necessary to combat inflation, which started the year at 13%.

In contrast, loosening monetary policy (lowering the SELIC target) did not bring benefits in terms of GDP in 2012/2013, but inflation went up. SELIC was at its lowest historical level and not even that was capable of reversing the downward trend in GDP. Often, this is attributed to an acceleration of public spending and consequent demand pressures. Eventually, the CBB had to raise the SELIC to compensate investors from risk: public debt was increasing fast. In 2015, the fiscal situation was so deteriorated that the then Minister of Finance, Nelson Barbosa, proposed a controversial fiscal band to limit government spending. Since a band is unpredictable by definition, the market did not see it with good eyes. Not surprisingly, inflation during Lula's successor, president Dilma Rousseff, years was consistently above the centre of the target.

Therefore, lower interest rates in the end of the period covered by data would almost surely lead to higher inflation only, not preventing the economy from shrinking and unemployment to rise. Apart from these two periods, there is no reason to speculate that the SELIC should have been higher and we could in fact expect that lower levels would have had positive impacts on GDP and

employment with no dissipation through inflation. Furthermore, even though the monetary authority might decide not to set the actual SELIC equal to the one computed by the model, it would still have a useful tool for advice.

4.3.2

Possible Practical Limitations

The originally generated SELIC series is considerably less persistent and more volatile than the actual series set by the CBB, although the MA version corrected the persistence problem (4.2.3). In that section, we contended that high volatility and/or persistence could lead to unfavourable economic results, since large fluctuations could damage the perception of CBB's commitment to a clear rule. This type of argument is found throughout the literature [Clarida et al., 1999, Woodford et al., 1999, Woodford, 1999, Obstfeld and Rogoff, 2009, Galí, 2015, Woodford, 2011, Taylor, 2013, 2012].

In particular, Taylor [2012] concludes that that rule-based policies have clear advantages over discretion. He compares two distinct periods in the United States: from 1985 to 2003, when he maintains that policy was rule-based and economic performance was generally good, and from 2003 onward, when he sustains authorities did not follow a predictable approach and economic performance was poor. In Taylor [2013], Obstfeld and Rogoff [2009] the authors blame Federal Reserves's (the US central bank) interventionist, discretionary monetary policy for the 2007-2008 financial crisis, although Obstfeld and Rogoff [2009] enumerate and test other possible reasons.

The explanation goes along the following lines. Transparent commitment gives the public the sense that institutions are reliable, enhancing the ability of market participants to plan rationally. Under discretion and the belief that the rules are not steady, market behaviour is distorted, for agents are driven towards inefficient short-sighted responses. Quite often, besides the episodes we already mentioned, the popularity of discretionary Keynesian fiscal policy is cited as the main factor behind the deep recessions of the seventies.

The argument above, although not confirmed empirically (evidence is anecdotal), is largely accommodated by DSGE models, in which representative agents are forward-looking and usually rational. According to our evaluation of these models in section 2.1, this does not make the case for pro-commitment argument's validity. In fact, Bullard and Mitra [2002], Evans and Honkapohja [2003], and Evans and Honkapohja [2006] show that the assumption of rational expectations is not innocuous. Under learning, i.e. when representative agents are allowed to adapt to new information, there are certain implementations of discretionary policy that achieve the rational expectations equilibrium (the

optimum). In line with these theoretical findings and building upon a series of other papers, Orphanides [2002] follows an empirical approach and claims that US policy decisions during the seventies were systematic and forward-looking (implying commitment), but nonetheless lead to poor economic performance.

In summary, despite most central banks all over the world preferring strategies involving commitment, there is no academic consensus about the superiority of commitment over discretion.

4.3.3

A Brief User Guide

In principle, given new data (new observations of the covariates), the monetary authority could re-estimate the whole model to find another series from which it would look only at the last element, the current period's interest rate computed by the model. However, this could be rather time-consuming and the output series (the solution) could have an undesired low persistence, undermining the central bank's credibility. Moreover, there is no guarantee that the previous solution would change significantly with the addition of one new gene. So we suggest two other, more feasible ways central banks could use the strategy we develop.

First, the optimisation problem could be reduced to finding a smaller portion of the interest rate series, holding the rest of the solution fixed. This could improve running time significantly, since the crossover operation would be quicker. Second, the authority could regress the generated interest rate series on predicted inflation, predicted unemployment and covariates, and then save the coefficients (or the trained structure) to find the current optimal interest rate. Even though the bank might decide not to set the actual SELIC equal to the one suggested by the model, it would still have a useful tool for advice.

5 Conclusion

This work has presented in detail a novel tool to back the decision-making process of central banks. We have answered to an ongoing demand for building a new framework to study policy implications, a challenge that has been in place since the failure of DSGE models in predicting the 2007-2008 financial crisis. Our strategy solves the central bank's problem of setting the interest rate without resorting to those kind of models, and has been applied to the case of Brazil.

Our framework was based on well-established empirical evidence. The first step was to obtain an objective function for the central bank, which we derived from three empirical relationships: Okun's Law, the Phillips Curve, and liquidity effects. Then we designed a genetic algorithm to minimise this objective function by choosing the interest rate. Since the function is forward looking, we used a neural network to predict future values of unemployment and inflation.

More specifically, predictions were made by a MLP whose inputs are the attempted solution (a SELIC target series) and a large set of covariates. As for the minimisation of the objective function, we adopted a genetic algorithm solution where an individual is the whole SELIC rate time series and a gene is an observation of the series. The best individual, i.e. the series that minimises the objective function, as compared to the original SELIC rate series, was shown to imply lower values for inflation for 62.48% of the time. Predicted unemployment, however, was lower only for 39.69% of covered periods, as it faces a trade-off with inflation.

The generated SELIC target series is considerably less persistent than the actual series set by the CBB. Thus applying our strategy could lead to unfavourable economic results, since large fluctuations could damage the perception of CBB's commitment to a clear rule. However, we argued that there is no consensus about the superiority of commitment over discretion.

Another possible problem with our solution is that it suggests that the SELIC rate should have been lower for more than 94% of the covered weeks. Nevertheless, it is possible that lower interest rates in the end of the period covered by data would almost surely lead to higher inflation only, not

preventing the economy from shrinking and unemployment to rise. At the same time, scholars often agree that, around the beginning of the period, monetary policy should not have been looser

Nonetheless, apart from these two periods, we contended that there is no reason to speculate that the SELIC should have been higher and that we could expect lower levels to lead to a higher GDP and employment with no dissipation through inflation. Furthermore, even though the monetary authority might decide not to set the actual SELIC equal to the one calculated by the model, it would still have a useful tool for advice.

We suggest two ways central banks could use our framework. First, the optimisation problem could be reduced to finding a smaller portion of the interest rate series, holding the rest of the solution fixed. Second, the authority could regress the generated interest rate series on predicted inflation, predicted unemployment and covariates, and then save the coefficients (or the trained structure) to find the current optimal interest rate. We leave implementation details for future works.

Further contributions could also benefit from three other ideas. Neural Networks predictions could be compared to predictions compiled in the Focus Bulletin, a CBB publication which compiles predictions for inflation collected among market players and consultancy firms. Furthermore, core inflation could be used instead of inflation as an input to our framework, in an attempt to smooth the interest rate *ex-ante*, without resorting to the moving average *ex-post*. Finally, each future variable (inflation and unemployment) in each time horizon (mid-term or short term) could be predicted by a different neural network.

Bibliography

- [1] Marcelo de Paiva Abreu et al. A ordem do progresso: dois séculos de política econômica no brasil. *Editora Campus–Elsevier*, 2014.
- [2] José Roberto Afonso, Eliane Cristina Araújo, and Bernardo Guelber Fajardo. The role of fiscal and monetary policies in the brazilian economy: Understanding recent institutional reforms and economic changes. *The Quarterly Review of Economics and Finance*, 62:41–55, 2016.
- [3] George A Akerlof. The missing motivation in macroeconomics. *American Economic Review*, 97(1):5–36, 2007.
- [4] Marlene Amstad and Antoine Martin. Monetary policy implementation: common goals but different practices. *Current Issues in Economics and Finance*, 17(7), 2011.
- [5] Kenneth J Arrow and Gerard Debreu. Existence of an equilibrium for a competitive economy. *Econometrica: Journal of the Econometric Society*, pages 265–290, 1954.
- [6] Rüdiger Bachmann, Steffen Elstner, and Eric R Sims. Uncertainty and economic activity: Evidence from business survey data. *American Economic Journal: Macroeconomics*, 5(2):217–49, 2013.
- [7] Laurence Ball, Davide Furceri, Daniel Leigh, and Prakash Loungani. Does one law fit all? cross-country evidence on okun’s law. In *documento presentado en el IMF-OCP Workshop on Global Labour Markets, París*, volume 1, 2016.
- [8] Fabio Barbieri. Complexity and the austrians. *Filosofía de la Economía*, 1(1): 47–69, July 2013.
- [9] Regis Barnichon and Paula Garda. Forecasting unemployment across countries: The ins and outs. *European Economic Review*, 84:165–183, 2016.
- [10] B. S. Bernanke, T. Laubach, F. S. Mishkin, and A. S. Posen. *Inflation Targeting*. Princeton University Press, 1999.
- [11] Stephen Beveridge and Charles R Nelson. A new approach to decomposition of economic time series into permanent and transitory components with

- particular attention to measurement of the 'business cycle'. *Journal of Monetary economics*, 7(2):151–174, 1981.
- [12] Olivier Blanchard. The state of macro. *Annual Review of Economics*, 1(1): 209–228, 2009.
- [13] Olivier Blanchard. On the future of macroeconomic models. *Oxford Review of Economic Policy*, 34(1-2):43–54, 2018.
- [14] Olivier Blanchard and Jordi Galí. Real wage rigidities and the new keynesian model. *Journal of money, credit and banking*, 39:35–65, 2007.
- [15] Olivier Blanchard. Do dsge models have a future? Policy Brief PB 16-11, Peterson Institute for International Economics, 2016.
- [16] Mauro Boianovsky, Roger Backhouse, et al. Whatever happened to microfoundations? In *Anais do XXXIV Encontro Nacional de Economia [Proceedings of the 34th Brazilian Economics Meeting]*, number 76. ANPEC, 2006.
- [17] Jean Boivin, Michael T Kiley, and Frederic S Mishkin. How has the monetary transmission mechanism evolved over time? In *Handbook of monetary economics*, volume 3, pages 369–422. Elsevier, 2010.
- [18] James Bullard and Kaushik Mitra. Learning about monetary policy rules. *Journal of monetary economics*, 49(6):1105–1129, 2002.
- [19] Fabio Canova and Luca Sala. Back to square one: Identification issues in dsge models. *Journal of Monetary Economics*, 56(4):431–449, 2009.
- [20] Chiranjit Chakraborty and Andreas Joseph. Machine learning at central banks. Bank of England working papers 674, Bank of England, September 2017.
- [21] Lawrence J Christiano and Martin Eichenbaum. Liquidity effects, monetary policy, and the business cycle. *Journal of Money, Credit and Banking*, 27(4): 1113–1136, 1995.
- [22] Lawrence J Christiano, Martin S Eichenbaum, and Mathias Trabandt. On dsge models. *Journal of Economic Perspectives*, 32(3):113–40, 2018.
- [23] Richard Clarida, Jordi Gali, and Mark Gertler. The science of monetary policy: a new keynesian perspective. *Journal of economic literature*, 37(4):1661–1707, 1999.
- [24] Olivier Coibion, Yuriy Gorodnichenko, and Mauricio Ulate. Is inflation just around the corner? the phillips curve and global inflationary pressures. In *AEA Papers and Proceedings*, volume 109, pages 465–69, 2019.

- [25] David Colander. *Post Walrasian macroeconomics: Beyond the dynamic stochastic general equilibrium model*. Cambridge University Press, 2006.
- [26] David Colander, Peter Howitt, Alan Kirman, Axel Leijonhufvud, and Perry Mehrling. Beyond dsge models: toward an empirically based macroeconomics. *American Economic Review*, 98(2):236–40, 2008.
- [27] Marcos R De Castro, Solange N Gouvea, Andre Minella, Rafael Santos, and Nelson F Souza-Sobrinho. Samba: Stochastic analytical model with a bayesian approach. *Brazilian Review of Econometrics*, 35(2):103–170, 2015.
- [28] Domenico Delli Gatti, Saul Desiderio, Edoardo Gaffeo, Pasquale Cirillo, and Mauro Gallegati. *Macroeconomics from the Bottom-up*. Springer Milan, 2011.
- [29] R Dornbusch, S Fischer, and R Startz. *Macroeconomics*. richard d, 1997.
- [30] George W Evans and Seppo Honkapohja. Expectations and the stability problem for optimal monetary policies. *The Review of Economic Studies*, 70(4):807–824, 2003.
- [31] George W Evans and Seppo Honkapohja. Monetary policy, expectations and commitment. *Scandinavian Journal of Economics*, 108(1):15–38, 2006.
- [32] Angelo Marsiglia Fasolo and Marcelo Savino Portugal. Imperfect rationality and inflationary inertia: a new estimation of the phillips curve for brazil. *Estudos Econômicos (São Paulo)*, 34(4):725–776, 2004.
- [33] Jesús Fernández-Villaverde. The econometrics of dsge models. *SERIEs*, 1(1-2):3–49, 2010.
- [34] Roger Fletcher. *Practical methods of optimization*. John Wiley & Sons, 2013.
- [35] Focus Economics. The poorest countries in the world, 2018. URL <https://www.focus-economics.com/blog/the-poorest-countries-in-the-world>.
- [36] Shane Frederick, George Loewenstein, and Ted O'donoghue. Time discounting and time preference: A critical review. *Journal of economic literature*, 40(2):351–401, 2002.
- [37] Milton Friedman. Monetary policy: Tactics versus strategy. *The search for stable money*, pages 361–82, 1987.

- [38] Jordi Galí. *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications*. Princeton University Press, 2015.
- [39] Jordi Galí and Mark Gertler. Macroeconomic modeling for monetary policy evaluation. *Journal of economic perspectives*, 21(4):25–46, 2007.
- [40] Fabio Giambiagi and Alexandre Schwartzman. *Complacência*, volume 1. Elsevier Brasil, 2014.
- [41] Fabio Giambiagi and André Arruda Villela. *Economia brasileira contemporânea*. Elsevier Brasil, 2016.
- [42] Rodrigo Melo Gois and Marco Antonio Jorge. Investigating the okun's law in brazil and three of its metropolitan areas. *Revista de Economia Mackenzie*, 14(1), 2018.
- [43] Robert J Gordon and Peter K Clark. Unemployment and potential output in the 1980s. *Brookings Papers on Economic Activity*, 1984(2):537–568, 1984.
- [44] Stephen Hansen, Michael McMahon, and Andrea Prat. Transparency and deliberation within the fomc: a computational linguistics approach. *The Quarterly Journal of Economics*, 133(2):801–870, 2017.
- [45] Andrew C Harvey. *Forecasting, structural time series models and the Kalman filter*. Cambridge university press, 1990.
- [46] Robert J Hodrick and Edward C Prescott. Postwar us business cycles: an empirical investigation. *Journal of Money, credit, and Banking*, pages 1–16, 1997.
- [47] Peter Howitt. What have central bankers learned from modern macroeconomic theory? *Journal of Macroeconomics*, 34(1):11–22, 2012.
- [48] Samuel Hurtado. Dsge models and the lucas critique. *Economic Modelling*, 44:S12–S19, 2014.
- [49] Zhenhu Jin and Alan Harper. Okun's law: an empirical test using brazilian data. *Research in Business Economics Journal*, 13, 2019.
- [50] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- [51] Alan Kirman. The economic crisis is a crisis for economic theory. *CESifo Economic Studies*, 56(4):498–535, 2010.

- [52] J. Kukal and T. Van Quang. A monetary policy rule based on fuzzy control in an inflation targeting framework. *Prague Economic Papers*, 2014.
- [53] Finn E Kydland and Edward C Prescott. Time to build and aggregate fluctuations. *Econometrica: Journal of the Econometric Society*, pages 1345–1370, 1982.
- [54] Axel Leijonhufvud. Economics of the crisis and the crisis of economics. *The European Journal of the History of Economic Thought*, 21(5):760–774, 2014.
- [55] John B Long Jr and Charles I Plosser. Real business cycles. *Journal of political Economy*, 91(1):39–69, 1983.
- [56] Vicente da Gama Machado and Marcelo Savino Portugal. Phillips curve in brazil: an unobserved components approach. *Estudos Econômicos (São Paulo)*, 44(4):787–814, 2014.
- [57] André Minella. Monetary policy and inflation in brazil (1975-2000): a var estimation. *Revista Brasileira de Economia*, 57(3):605–635, 2003.
- [58] Maurice Obstfeld and Kenneth Rogoff. Global imbalances and the financial crisis: products of common causes. 2009.
- [59] Arthur M Okun. *Potential GNP: its measurement and significance*. Cowles Foundation for Research in Economics at Yale University, 1963.
- [60] Athanasios Orphanides. Monetary-policy rules and the great inflation. *American economic review*, 92(2):115–120, 2002.
- [61] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
- [62] Alban W Phillips. The relation between unemployment and the rate of change of money wage rates in the united kingdom, 1861–1957 1. *economica*, 25 (100):283–299, 1958.
- [63] John M Roberts. New keynesian economics and the phillips curve. *Journal of money, credit and banking*, 27(4):975–984, 1995.
- [64] Paul Romer. The trouble with macroeconomics. Technical report, Stern School of Business, New York University, 2016.

- [65] Samuel Rönqvist and Peter Sarlin. Bank networks from text: interrelations, centrality and determinants. *Quantitative Finance*, 15(10):1619–1635, 2015.
- [66] Samuel Rönqvist and Peter Sarlin. Bank distress in the news: Describing events through deep learning. *Neurocomputing*, 264:57–70, 2017.
- [67] SGS. Sistema gerenciador de séries temporais, 2019. URL <http://www4.bcb.gov.br/pec/series>.
- [68] Frank Smets and Rafael Wouters. Shocks and frictions in us business cycles: A bayesian dsge approach. *American economic review*, 97(3):586–606, 2007.
- [69] Robert Solow. Building a science of economics for the real world. *House Committee on Science and Technology Subcommittee on Investigations and Oversight*, 20, 2010.
- [70] James H Stock and Mark W Watson. Forecasting inflation. *Journal of Monetary Economics*, 44(2):293–335, 1999.
- [71] John B Taylor. Monetary policy rules work and discretion doesn't: A tale of two eras. *Journal of Money, Credit and Banking*, 44(6):1017–1032, 2012.
- [72] John B Taylor. *Getting off track: How government actions and interventions caused, prolonged, and worsened the financial crisis*. Hoover Press, 2013.
- [73] Leigh Tesfatsion. Agent-Based Computational Economics: A Constructive Approach to Economic Theory. Staff General Research Papers Archive 12514, Iowa State University, Department of Economics, January 2006.
- [74] Allan Timmermann. Chapter 4 forecast combinations. volume 1 of *Handbook of Economic Forecasting*, pages 135 – 196. Elsevier, 2006. doi: [https://doi.org/10.1016/S1574-0706\(05\)01004-9](https://doi.org/10.1016/S1574-0706(05)01004-9). URL <http://www.sciencedirect.com/science/article/pii/S1574070605010049>.
- [75] Guilherme Alexandre Tombolo and Marcos Minoru Hasegawa. Okun's law: evidence for the brazilian economy. 2014.
- [76] Camilo E Tovar. Dsge models and central banks. *Economics: The Open-Access, Open-Assessment E-Journal*, 3, 2009.
- [77] United Nations Development Programme. Human Development Report 2019. Technical report, 2019. URL <http://hdr.undp.org/en/2019-report/download>.
- [78] Michael Woodford. Optimal monetary policy inertia. *The Manchester School*, 67:1–35, 1999.

- [79] Michael Woodford. Convergence in macroeconomics: elements of the new synthesis. *American economic journal: macroeconomics*, 1(1):267–79, 2009.
- [80] Michael Woodford. *Interest and prices: Foundations of a theory of monetary policy*. Princeton University Press, 2011.
- [81] Michael Woodford et al. Commentary: How should monetary policy be conducted in an era of price stability? *New challenges for monetary policy*, 277316, 1999.