Universidade de Brasília
Faculdade de Economia, Administração e Contabilidade
Departamento de Economia

Learning and Bounded Rationality in Banking Crises

Bernardo Arraes Vinhosa

Brasília
2018

Bernardo Arraes Vinhosa

Learning and Bounded Rationality in Banking Crises

Tese apresentada ao Programa de Doutorado em Economia da Universidade de Brasília como requisito à obtenção do título de Doutor em Ciências Econômicas.

Universidade de Brasília

Faculdade de Economia, Administração e Contabilidade

Departamento de Economia

Orientador: Prof. Daniel Oliveira Cajueiro, PhD

Brasília

2018

Ficha catalográfica elaborada automaticamente,
com os dados fornecidos pelo(a) autor(a)

Vinhosa, Bernardo Arraes
VV784l Learning and bounded rationality in banking crises / Bernardo Arraes Vinhosa; orientador Daniel Oliveira Cajueiro. -- Brasília, 2018.
75 p.

Tese (Doutorado - Doutorado em Economia) -- Universidade de Brasília, 2018.

1. Economia bancária. 2. Modelos baseados em agentes. 3. Aprendizado adaptativo. I. Cajueiro, Daniel Oliveira, orient. II. Título.

Bernardo Arraes Vinhosa

Learning and Bounded Rationality in Banking Crises

Tese apresentada ao Programa de Doutorado em Economia da Universidade de Brasília como requisito à obtenção do título de Doutor em Ciências Econômicas.

Trabalho aprovado. Brasília, 17 de dezembro de 2018:

Prof. Daniel Oliveira Cajueiro, PhD
Universidade de Brasília
Orientador

Prof. Bernardo Pinheiro Machado Mueller, PhD
Universidade de Brasília

Prof. Vinícius Amorim Sobreiro, PhD
Universidade de Brasília

Ricardo Vieira Barroso, PhD
Banco Central do Brasil

Brasília
2018

Resumo

A análise estrutural de políticas regulatórias é útil para projetar políticas que almejem prevenir crises financeiras e bancárias. Em tempos de crise, as premissas de equilíbrio e expectativas racionais podem não valer. Para realizar análise estrutural de políticas sem elas, esta tese irá fornecer um exemplo de análise de política com modelos de aprendizado. O modelo de aprendizado de Atração Ponderada por Experiência Auto-ajustável é usado para construir uma simulação de sistema bancário que pode ser uado para análise de políticas regulatórias. Ele é então utilizado para realizar um exercício de calibragem no sistema bancário brasileiro. O resultado mais relevante da simulação é a reação excessiva dos participantes do mercado aos choques, devido ao aprendizado e à dinâmica de aquisição de informação associada.

Palavras-chave: Rede interbancária, regulação, modelo baseado em agentes, aprendizado adaptativo.

Abstract

Structural policy analysis is useful for designing policy aimed at preventing financial and banking crises. In times of crises, the assumptions of equilibrium and rational expectations might not hold. In order to perform structural policy design without them, this dissertation will provide an example of policy analysis with learning models. The Self-Tuning Experience Weighted Attraction learning model is used to build a banking system simulation which can be used in policy analysis. It is then used to perform a calibration exercise with data on the Brazilian Banking System. This simulation's most relevant result is market participants overreacting to shocks due to learning and the associated information acquisition dynamics.

Keywords: Interbank network, regulation, agent-based model, adaptive learning.

List of figures

List of tables

List of abbreviations

BCBS	Basel Committee on Banking Supervision
BIS	Bank for International Settlements
CAR	Capital Adequacy Ratio
EWA	Experience-Weighted Attraction
FRB	Board of Governors of the Federal Reserve System
FSB	Financial Stability Board
FDIC	Federal Deposit Insurance Corporation
ISO	International Organization for Standardization
LGD	Loss Given Default
PD	Probability of Default
PSF	Python Software Foundation
ROE	Return on Equity
RWA	Risk-Weighted Assets
STEWA	Self-Tuning Experience-Weighted Attraction
TBTF	Too Big To Fail

List of symbols

$A_i^j(t)$	Attraction update function of strategy j for agent i in t
ϕ	Attraction depreciation
δ	Foregone payoffs' weight
$N_i(t)$	Experience of agent i in t
ρ	Depreciation of experience
λ	Sensitivity to attractions
K	Bank's Capital
D	Deposits
i_d	Deposit interest rate
IB/IL	Interbank Loans
i_i	Interbank interest rate
L	Liquid assets
R	Real sector loans
F_b	Set of firms with loans in bank b
r	Real sector loans return
α	Ratio of capital to total liability
β	Ratio of liquid assets to total assets
T	Total assets or total liabilities
Π	Profit
w	Risk weight of firm's loans
δ_L	Illiquid assets depreciation
imp	Probability of impatient depositor
γ	Depositor's risk tolerance
U	Depositor's utility

c_d^t	Depositor's consumption in t
ls	Liquidity shortage
δ_I	Depreciation of insolvent bank's assets

Contents

Introduction

Financial stability, defined as a financial system's capacity to withstand shocks with minimal impact to the health of its institutions, plays a key role in citizens' welfare. Financial system fragility can lead to losses in savings, as well as a decrease in credit supply and financing, culminating in losses in output and employment (Schinasi, 2004). Government supervision and regulation thus play an important role in balancing the risks and benefits of financial intermediation.

In order to reduce the possibility of bank failures and mitigate its potential impacts, national governments agreed upon a comprehensive prudential regulation framework for banks in the aftermath of the Global Financial Crisis. This framework features liquidity requirements (BCBS, 2013a)(BCBS, 2014b), a leverage ratio (BCBS, 2014a) and capital requirements (BCBS, 2011), a countercyclical capital buffer requirement (BCBS, 2010) and capital surcharges for systemically important banks (BCBS, 2013b). Additionally, representatives of the Financial Stability Board member countries agreed upon a series of proposals related to bank recovery and resolution in order to increase the likelihood of successful resolution of large and complex financial institutions (FSB, 2014).

In order to preserve financial stability, specific anti-crisis policies are necessary, due to financial market participants' markedly different behavior during crises relative to normal times. Market forces such as bankers' own self-interest and market discipline exerted by investors can be stabilizing for a healthy bank but are often destabilizing for a distressed one (Eisenbach, 2017). Market discipline exerted on a distressed bank may lead to bank runs, which, in turn, could end up forcing the bank to fail for lack of liquidity. Distressed banks have incentives to gamble for resurrection (Tirole, 2006). The possibility of fire sales (Allen and Gale, 1998) and contagion (Allen and Gale, 2000) implies that risk must be assessed systemically. The importance of these phenomena in times of crises means they should not be ignored when designing anti-crisis policy.

Structural policy analysis is useful for designing policy aimed at preventing financial and banking crises. At the very least, it can add robustness to a reduced-form analysis, partially mitigating external validity concerns. Furthermore, it is the only available alternative when devising novel policy instruments.

Traditionally, structural policy analysis is performed with equilibrium models under the assumption of forward-looking expectation formation. However, the turbulence and unpredictability associated with financial crises make equilibrium and that kind of forward-looking behavior questionable assumptions. In order to perform structural policy design without relying on those assumptions, a plausible alternative is to use learning models instead.

In chapter 1, I present an argument for using learning models for policy design in banking crises, (as a complement to equilibrium models), and why a particular learning model – a variant of Self-Tuning Experience Weighted Attraction (STEWA) (Ho; Camerer and Chong, 2007) – is appropriate to that end. I describe why overreaction is a possible outcome when equilibrium is substituted for learning with information acquisition. Finally, I introduce a variation on the STEWA which can be useful as a robustness check when modeling banking crises.

For learning models to be useful for structural policy analysis and design in banking crises, a model of market participant behavior (banks, depositors and firms) is necessary, as well as a subsequent calibration of real-world data to this model. Both are performed in chapter 2. I present a multi agent banking model inspired by (Barroso et al., 2016), which is modified so agents learning according to Self-Tuning EWA. Afterwards, I use it to perform a calibration exercise with data on the Brazilian Banking System, taking into account its peculiarities. I present the result of the calibration exercise, so it can be used in policy design exercises.

Finally, in Chapter 3, I perform policy design exercises using the model and calibration data uncovered in Chapter 2. Those experiments serve as a proof of concept, to argue for the methodological appropriateness of the preceding chapters' methodological recommendations. Within the simulation, we are able to observe agents' reaction to shocks consistently with the financial and banking crisis literature. Additionally, simulations yield differing results for Self-Tuning EWA when compared to the original EWA model.

1 Policy analysis in financial crises

Banks and other similar financial intermediaries provide valuable services to society. At the same time, there are risks inherent to their activities: mainly, credit, market and liquidity risks (Freixas and Rochet, 2008). Those risks can occasionally materialize to such a degree that they compromise banks' ability to function normally. A significantly distressed banking system has the potential to compromise national and international economies, with dire consequences such as loss of savings and reductions in output and employment. Hoggarth, Reis and Saporta (2002) estimate banking crises produce a cumulative output loss of approximately 15% to 20% of annual GDP, while Reinhart and Rogoff (2009) point out that government debt increases 86% on average in real terms in the three years following banking crises.

The considerable impact of financial crises highlights the need for specific anti-crisis policies. Multi-agent systems with agents learning through Self-tuning Experience Weighted Attraction (Ho; Camerer and Chong, 2007) are a powerful tool to assess such policies. This modeling paradigm's emphasis on modeling agents individually rather than using a representative agent enables its models to reproduce unique characteristics of banking crises. A particular strength is the capacity to model contagion, both direct and indirect, caused by the formation of interbank loan networks or asset sales. These same phenomena also lead to increased complexity and reduced information availability, stressing the need for learning models. Learning is the key to enabling consistent choice-driven structural policy analysis in environments with limited information availability.

1.1 Learning and financial crisis policies

An important part of designing policy to prevent and mitigate the impact of banking crisis is forecasting financial market participants' reaction to policy changes. Models used to produce such forecasts should incorporate previous experience in the most objective way possible. It thus is important that these models be based on previously observed data and behavior, to the extent participants' past behavior during past crises can be informative to predict future behavior patterns in future crises.

Prominent features of banking crises compromise significantly the usefulness of reduced-form econometric models to forecast the impact of policies aimed at crisis prevention and response. Crises rarely occur within similar temporal or geographic contexts, so that unobservable but significant institutional factors cannot be assumed to be constant. Additionally, authorities normally apply crisis response measures to all banks under their jurisdiction, depriving policymakers of a control group. Most importantly, during crisis periods authorities sometimes resort to novel policy instruments with no prior history of use. It is thus paramount to be able to simulate unprecedented changes, highlighting the importance of out-of-sample prediction capabilities.

It is important for policymakers to use models that capture agents' decision-making process to improve their ability to forecast agents' future behavior based on past observations. The core principle is to model agents' behavior as a choice among their actions and observe what motivates that choice. Such structural policy analysis is most commonly performed within a framework sometimes referred to as "neoclassical", consisting of rational agents optimizing profits or utility in equilibrium (Harstad and Selten, 2013).

This neoclassical framework has desirable characteristics for policy design. It endows agents with choice and counterfactual analysis capabilities, so they can react to unforeseen circumstances. It also presents a consistent relationship between agents' actions and optimality of their results. Nevertheless, its reliance on equilibrium and explicit and tractable expectation formation present a problem when modeling banking crises. Consequently, there is demand for alternative policy frameworks that do not rely on such assumptions but retain choice modeling and consistency.

Policymakers, should not assume agents' expectations to be either constant or homogeneous during crisis episodes. Reinhart and Rogoff (2013) find that asset price bubbles, large capital inflows and credit booms typically precede banking crises, while there is a close association between unusually high credit growth and speculative bubbles (Kindleberger and Aliber, 2015). This evidence is consistent with market participants being likely to cast doubt upon and consequently review their beliefs and expectations during banking crises. Consequently, it is desirable that policy analysis aiming to prevent banking crises incorporate the effects not only of adverse events, but that of a period of diminished confidence followed by gradual uncertainty resolution. Agents would thus incorporate their own ignorance about their environment in their decision-making process.

During banking crises, policymakers can no longer consider the rate of change in market participants' economic environment slow relative to their learning capabilities, making the assumptions of full rationality and equilibrium less realistic (Kirman, 2011). Given equilibrium theory's inadequacy in such contexts, learning theory can offer alternative explanations on what kind of behavior to expect from agents (Fudenberg and Levine, 1998; Fudenberg and Levine, 2009).

Models featuring agents endowed with learning capacity can also help reconcile agents' more predictable and stable behavior in stable environments and less predictable behavior in unstable ones. Policy analysis aimed at financial crisis prevention and impact mitigation should be sensitive to the fact that agents in unstable environments place a high value on information acquisition, playing potentially sub-optimal strategies to learn more about their consequences. This behavior is called active learning (Fudenberg and

Levine, 2016). Furthermore, agents' information acquisition motive becomes weaker as their environment stabilizes and they build up confidence about their own knowledge of their context. This behavioral pattern typifies the explore-exploit dilemma (Sutton and Barto, 1998), with a typical example in economics being the multi-armed bandit problem (Berry and Fristedt, 1985).

Market participant's hard-to-predict behavior during financial crises is also related to increased complexity and information availability problems. This instability and lack of information necessary for expectation formation makes market participants more likely to deemphasize predictions relying on partial knowledge of the structure of the economy, instead emphasizing a more direct analysis of the relationship between their actions and payoffs. This direct relationship can be modeled by reinforcement learning methods (Roth and Erev, 1995). It is also likely that, to the extent possible, agents do not abandon completely counterfactual reasoning, in which case they can resort to belief-based learning (Fudenberg and Levine, 1998).

Finally, it is important to note that most agent-based models in economics are not choice-based. They implement agents making decisions through heuristics instead of choice. Dosi et al. (2017) provide an overview of heuristics-based models.

1.2 Reinforcement and belief-based learning

It is thus useful to endow agents with learning capabilities in order to make them proactive and capable of some counterfactual analysis without depending on an excessively strong degree of information processing capabilities and information availability. This would be inconsistent with bank runs and banking crises. For the same reason, agents' counterfactual reasoning capabilities should be subject to significant limitations. More specifically, they should be able to infer their payoffs in a limited set of counterfactual states of nature: those that differ from the factual state only relative the counterfactually reasoning agent's own behavior. This restriction is coherent with the following assumptions:

- Agents do not have enough information to deduct other agents' payoffs;
- Agents do not know other agents' beliefs, so they cannot predict the effects of their actions on other agents;
- Banks do not know if depositors are patient or impatient;
- Depositors cannot infer banks' asset quality and the interbank network structure;
- Banks cannot infer other banks' asset quality and interbank exposures.

Furthermore, representative agent models present diminished applicability for policy evaluation in scenarios featuring distressed banks. In such periods, the potential impact of asset sale and contagion dynamics between banks becomes significant. This heterogeneity among agents influences the formation of interbank loan networks and asset sales. These phenomena have a greater tendency to manifest themselves when banks differ relatively to their liquidity, with some banks lacking it and some having it in excess.

The combination of limited counterfactual capacity and learning-driven choice modelling is consistent with the combination of reinforcement learning and belief-based learning embodied in the EWA family of models (Camerer and Ho, 1999). Furthermore, agents' greater propensity to privilege information acquisition in less stable environments is indicative of the appropriateness of a specific kind of EWA model – Self-tuning EWA (Ho; Camerer and Chong, 2007).

The remainder of this chapter presents a proposal for a policy analysis framework meeting the aforementioned requirements. Given that a model incorporating both adaptive learning dynamics and heterogeneity among agents is complex enough to be analytically intractable, the proposal will be implemented and solved computationally, as is common in the heterogeneous agent literature (Borrill and Tesfatsion, 2011).

1.2.1 Experience-weighted Attraction

Experience-weighted Attraction (Camerer and Ho, 1999; Camerer, 2003), generalizes reinforcement learning (Roth and Erev, 1995) and belief-based learning (Fudenberg and Levine, 1998). Reinforcement learning makes successful (higher-payoff) factual actions more frequent. Belief-based learning, on the other hand, makes successful (higher-payoff) counterfactual actions more frequent. In fact, this approach even permits weighing differently the attractiveness of factual and counterfactual actions. Consequently, the use of this learning approach can endow agents with the necessary counterfactual analysis capabilities without placing excessive demands on agents' cognitive capacity and information availability.

The principles of EWA can be used for predicting how agents learn to play normal-form games when other agent's actions can't be deduced *ex ante*. The EWA implementation used in this dissertation is parametrized so it can be used to model a wide range of agent behavior, allowing agents to:

- vary how the relative values of payoffs affect the probability of playing the strategies that yield them;
- weigh differently how the payoffs from factual and counterfactual actions influence their behavior (for example, agents could place smaller weights on counterfactual payoffs if they don't fully trust their counterfactual analysis capabilities);

- weigh differently how past and current payoffs influence their behavior (for instance, placing a smaller weight on payoffs occurred further back in time to reflect that they took place in a different environment).

In order to specify more precisely how agents use the EWA principles to learn to play games, it is necessary to introduce its notation and formulae. Consider the case of a game with n agents, each indexed by $i (i = 1, \ldots, n)$. The strategy space of player i, S_i, encompasses m_i possible choices, i.e., $S_i = \{s_i^1, s_i^2, \ldots, s_i^j, \ldots, s_i^{m_i-1}, s_i^{m_i}\}$. Thus, the game's strategy space is $S = S_1 \times \ldots \times S_n$ and $s = (s_1, \ldots, s_n) \in S$ is a combination of strategies of all n players. In turn, $s_{-i} = (s_1, \ldots, s_{i-1}, s_{1+1}, \ldots, s_n)$ is the combinations of strategies of all players other than player i. Lastly, $s_i(t)$ is the factual strategy - the one actually chosen by i - at time t, while $\pi_i(s_i(t), s_{-i}(t))$ is the associated payoff.

The EWA model links strategies' payoff histories and their probability of being selected through a concept called *attraction*, updated each in time period where the agent has the opportunity to play. The attraction A of strategy j for agent i in t is given by:

$$A_i^j(t) = \frac{\phi \cdot N_i(t-1) \cdot A_i^j(t-1) + [\delta + (1-\delta) \cdot I(s_i^j, s_i(t))] \cdot \pi_i(s_i^j, s_{-i}(t))}{N_i(t)}, \tag{1.1}$$

where ϕ is the parameter used to discount past attractions, δ is used to weight foregone payoffs relative to the one resulting from the action effectively played and $I(s_i^j, s_i(t))$ is an indicator function, that assumes value of 1 when $s_i^j = s_i(t)$ and 0 otherwise. Finally, $N(t)$ is a measure of accumulated and discounted experience, given by:

$$N_i(t) = \phi \cdot (1-\kappa) \cdot N(t-1) + 1, \tag{1.2}$$

where κ represents the accumulation of experience. Taken together, κ and ϕ represent how agents discount past actions in response to a changing environment, either consciously or subconsciously.

Once attractions have been computed, the probability that player i chooses strategy j is calculated with the logit model:

$$P_i^j(t+1) = \frac{e^{\lambda \cdot A_i^j(t)}}{\sum_{k=1}^{m_i} e^{\lambda \cdot A_i^k(t)}}, \tag{1.3}$$

where a higher λ increases the probability of the player choosing actions with higher attractions.

The relationship between the EWA model's parameters and its associated learning dynamics can be described as follows:

- **weight of foregone payoffs** (δ): ranging from 0 to 1, it allows players to weigh differently factual and counterfactual play in case they consider the information regarding their foregone payoffs less reliable than the information regarding factual play. If the parameter's value is 0, the player totally ignores their foregone payoffs, with its learning simplifying to reinforcement learning. Conversely, when $\delta = 1$, the player places equal weight on factual and counterfactual data.

- **response sensitivity** (λ): positive, it controls how strongly higher attractions influence players' propensity to play certain strategies. In particular, when a given player's λ tends to zero, strategies' probabilities will tend to a uniform distribution irrespective of attractions, and when it tends to infinity, players will tend to play the strategy with highest attraction with probability tending to one.

- **rate of attraction growth** (κ): ranging from 1 to 0, it allows players to privilege certain strategies if their payoffs are consistently larger over time, for when $\kappa = 1$, attractions accumulate, and when $\kappa = 0$, attractions represent the weighted average of payoffs.

- **decay of previous attractions** (ϕ): ranging from 0 to 1, it allows players to place distinct weights on previous experience if the player considers its learning environment is subject to change. When $\phi = 0$, agents only take into account the most recent payoffs when deciding upon strategies' probabilities, whereas when $\phi = 1$, players value equally past and present payoffs.

1.2.2 Self-tuning experience weighted attraction

Ho, Camerer and Chong (2007) present a self-tuning variant of the EWA learning model, designed to mimic more closely humans' parallel learning and decision-making processes – namely, players' process of exploring alternative strategies, progressively focusing on strategies that yield higher payoffs consistently and reverting to a more exploratory behavior when changes to the environment affect actions' relative payoffs.

Not only did the authors bring the model closer to the theoretical foundations for learning behavior, but they also reduced the number of exogenous parameters in the model. This greater parsimony in terms of free parameters makes the new variant simpler to calibrate. Out of the four original parameters, the authors fix one (κ) and make other two dynamic and endogenous to the game's history (ϕ and δ), while only one remains exogenous (λ), still needing to be calibrated. This is relevant because statistical learning models customarily need data to calibrate free parameters and forecast behavior in novel games.

This kind of stability-sensitive behavior is implemented by endogeneizing ϕ, the decay parameter in the original EWA model. It represents the degree to which more recent information has greater influence in determining agents' behavior (if at all). The underlying principle is to make the determination of agents' behavior such that the similarity of the relative weights attributed to recent and distant information is commensurate with how stable their environment is.

To that end, this formerly exogenous parameter is replaced with a function of environment stability – in this model, taken to be the stability of other players' strategies. This metric is called player i's *surprise index* at time t, or $S_i(t)$, which acts as input to the *change detector function*, $\phi_i(t)$. The relative decay of information thus becomes a function of the players' experience within the game, and consequently dynamic and specific to each player. The surprise index is thus calculated by taking the squared deviation between other players' recent history of strategies and the strategy's expected frequency. The recent history $r_i^k(t)$ represents each players most recent observation of the strategies other players effectively played (attributing them a probability of 1 and 0 to all others). Each strategy k's expected frequency, in turn is represented by its historical average frequency, $h_i^k(t)$.

Mathematically, the surprise index is given by:

$$S_i(t) = \sum_{k=1}^{m_{-i}} (h_i^k(t) - r_i^k(t))^2 \tag{1.4}$$

And the change detector function is given by:

$$\phi_i(t) = 1 - \frac{1}{2} S_i(t) \tag{1.5}$$

Another change from the original EWA model to the self-tuning one is that, even within a single iteration, agents do not reinforce all strategies equally. They reinforce equally the strategy they played and a subset of counterfactual ones: those that yield a payoff greater than or equal to the factual, as observed *ex post*. This selective reinforcement attributes those strategies a δ of one. At the same time, they do not reinforce the remaining strategies, which receive a δ of zero. Consequently, if certain strategies yield higher payoffs more frequently, they will be reinforced more often and thus agents will focus on them to the detriment of others – tending to exploit consistently higher-payoff strategies rather than explore their strategy space. If the environment changes so that relative payoffs are altered then agents will revert to exploring the space.

The model's authors call this function the *attention function* due to one possible interpretation to why agents would reinforce some strategies to the detriment of others. The authors conjecture this distinction could come from limited attention subjects pay to counterfactual payoffs, ex post. It is plausible that subjects' attention is a limited cognitive resource and they only calculate payoffs for states where the payoffs are higher than the factual. Mathematically, the attention function is described by:

$$\delta_{ij}(t) = \begin{cases} 1 \text{ if } \pi_i(s_i^j, s_{-i}(t)) \geq \pi_i(t) \\ 0 \text{ otherwise} \end{cases} \tag{1.6}$$

The attraction accumulation parameter (κ) is set at zero, not allowing accumulation at all, instead making behavior depend on average payoffs. Empirically, the authors found κ did not seem to affect model fit significantly. Additionally, averaging instead of accumulating attractions produced more robust behavior predictions in some games.

Self-tuning EWA's most salient characteristic is the substitution of exogenous parameters by functions of each players' experience. This substitution has important implications. Conceptually, such a model can be useful to generalize apparently diverse behavior across distinct games where the principles of standard EWA are applicable, without the need to calibrate on a game-by-game basis. It thus helps to uncover a deeper structural similarity in agents' behaviors across games thus improving EWA's applicability.

When analyzing the predictive accuracy of self-tuning EWA across various games, Ho, Camerer and Chong (2007) find that it has a slightly worse fit than the original model when parameters are allowed to be calibrated on a game-by-game basis, but is superior when games are pooled together and parameters are restricted to be the same across games. When forecasting behavior for new games using parameters estimated for similar games, the parameter variation generated endogenously by the self-tuning functions produces better-fitting estimates than those based on standard EWA calibrated with separate parameters.

This is an important result. It has powerful implications because a model with the capacity to forecast behavior in novel situations is a valuable tool for policy analysis. At the very least, it makes the analysis of external validity more systematic. It can also make simulations of the potential effects of untested policy more credible.

1.2.3 Imitation-constrained counterfactuals in Self-Tuning EWA

Learning is a useful complement to equilibrium models for policy analysis in financial crises. This complementarity originates from a potential weakness some equilibrium models might have: it is more difficult to derive reliable predictions about the future in times of crises, thus hindering counterfactual analysis and the formation of forward-looking expectations. Self-tuning EWA, the proposed learning model, addresses this problem but still relies on agents' capacity to form a contemporaneous counterfactual. This is a less restrictive assumption but could nonetheless be questionable in some scenarios.

To a smaller extent, the complexity and lack of information availability that hinder the formation of forward-looking expectations could also be argued to hinder the formation of contemporaneous counterfactuals. This difficulty in contemporaneous counterfactual formation becomes stronger when the banking system features characteristics of a complex system (Battiston et al., 2016). This tends to occur when banks' payoffs depend on other banks' behavior (for example, when sell risky assets to each other or there is an interbank lending market) or are dependent on depositors' behavior exhibiting strategic similarity (as is the case when bank runs occur).

This problem can be partially averted if agents are capable of formulating counterfactual beliefs over their own actions by observing similar agents' behavior and outcomes. They would then interpret this information appropriately to generate their own counterfactual beliefs. Introspection is thus substituted for imitation. The model's authors themselves suggest, in more recent literature, that inferring foregone payoffs to generate counterfactuals in EWA can be done heuristically by imitating a similar player (Camerer and Ho, 2014).

A full model of counterfactual formation by imitation would involve each agent performing two steps:

- observing other agents' actions and payoffs to form a historical measure of agents' mutual similarities
- deriving each moment's counterfactual outcomes per strategy by observing other agents that played the strategy and weighing their results by similarity.

Such a procedure would significantly increase the EWA model's complexity.

In a more simplified manner, the adaptation suggested herein is that agents continue to formulate the same counterfactual as if they possessed introspection capabilities, but only reinforce strategies effectively played by some agent (themselves or other similar players). This procedure could be used as a type of robustness check on the original model. When analyzing a particular application of the EWA model, it is useful to compare the results obtained with and without this restriction. If results diverge, it would then be necessary to analyze whether the agents' capacity to form counterfactuals by introspection is a plausible assumption in the context of such application. If results are similar, then they can be considered robust to agents' potential incapacity for counterfactual formation, and the necessity for analyzing the plausibility of this assumption is significantly reduced.

This insight is an adaptation of an observation made by the model's authors when they first presented the Self-Tuning EWA model (Ho; Camerer and Chong, 2007). The authors suggest that attention is a limited cognitive resource that could limit the number of counterfactual strategies that a certain player analyzes. If limited attention indeed restricts agents from analyzing all possible counterfactual strategies, it is plausible that a similar agent playing a certain strategy could present an incentive for another agent to include such strategy in its counterfactual analysis.

1.3 Implications of Self-Tuning EWA

There are two main benefits to using learning models as complements to equilibrium models when equilibrium cannot simply be assumed, such as when performing banking crisis policy analysis. At the very least, a learning model that reaches similar results to an equilibrium model can serve as a robustness check. More importantly, there could be circumstances where learning models will result in agents behaving differently from their equilibrium counterparts. A relevant question is whether, this difference in agents' behavior could be significant enough to result in distinct policy implications from those derived from equilibrium analysis (either quantitatively or qualitatively). If there is a possibility results might differ, it is worth analyzing under what circumstances those differences in behavior might appear.

Intuitively, the slower are the changes to agents' environment, the more similar the results of equilibrium, regular EWA and self-tuning EWA should be. Equilibrium models and the EWA variants would then be more likely to yield distinct results when their environment is subject to rapid change, or shocks: unexpected, rapid and significant variations in one or more exogenous factors which impact the system.

Learning is largely a backward-looking process. Consequently, it is intuitive that learning models' expected outcome is that agents' response lag in relation to the changes in their environment. This lag can be mitigated if agents are able to observe their environments and privilege exploration in when in low-information or rapidly changing environments, as typified by the exploit-explore dilemma. This greater emphasis on exploration under such conditions leads to more random behavior immediately after a significant shock. The randomness brought about by the learning process could increase the probability of volatile behavior, that is, behavior that. In the time periods immediately following a shock there are two possibly non-exclusive possibilities for this enhanced probability of volatile behavior:

- a state of temporary greater volatility in behavior than the volatility associated with the long-term steady state ("overdispersion" in behavior);
- a state where some agents' expected behavior immediately after the shock temporarily distances itself from the pre-shock behavior even more than the post-shock long-term steady state does, before coming back to this long-term steady state (a phenomenon henceforth referred to as "overshooting", by analogy to the exchange rate phenomenon explained by Dornbusch (1976)).

Policymakers should thus be vigilant regarding the possibility that agents' more volatile behavior immediately after a shock can directly or indirectly cause banking system turmoil. Particularly impactful is the possibility that such unpredictability could provoke banking and financial system instability, compromising savings, payment system operations and credit allocation.

The possibility of obtaining significantly different results from equilibrium and learning models is worth exploring. It would provide an explanation for financial instability dynamics as part of the learning process of a procedurally rational agent. It would constitute an alternative explanation to instability not dependent on the agents presenting any kind of behavioral biases. Post-shock instability when agent learning is governed by Self-Tuning EWA is coherent with the mechanics of how the model works:

1. Newly preferred strategies get reinforced, increasing the probability that a previously low-frequency strategy will be chosen;

2. If such a previously low-frequency strategy is chosen, it will drive up the surprise factor, driving down the value of the change-detector function;

3. If the change detector function decreases in value, formerly high-payoff strategies tend to be less representative (lower relative attraction), while the new high strategy payoffs will tend to be chosen more frequently;

4. Following this temporary decrease in the change-detector function, agent behavior temporarily becomes more random and other strategies which were preferred neither in the pre-shock nor the post-shock period will see a temporary spike in their probabilities;

5. As they are chosen more frequently, the newly preferred strategies will tend to have higher expected attractions because they yield higher payoffs, which will lead them to be chosen more frequently, establishing a cycle;

6. By virtue of being chosen more frequently, those newly preferred strategies will tend to become less and less surprising, driving up down the surprise factor and driving up the change detector function, making behavior more stable and less random;

7. Once behavior and outcomes stabilize, the agents will play more often the newly preferred strategies to the detriment of the previously preferred strategies and strategies that were preferred in neither period.

1.4 Related works

In order to motivate the relevance of using the Self-Tuning Experience-Weighted Attraction learning model for banking crisis policy analysis, it is necessary to contextualize it within the literature highlighting how the existing literature supports it and in which points does it differ. In particular, STEWA will be contextualized within learning models, how other learning models have been applied to banking, and finally, the link between adaptive learning and financial stability.

1.4.1 Learning and bounded rationality models

Experience-Weighted Attraction is a comprehensive model from a conceptual viewpoint, combining belief-based learning and reinforcement learning. Self-Tuning EWA improves upon the original model by being more parsimonious and more easily generalizable across games – useful properties for policy analysis. It is not, however, the only bounded-rationality model to describe experimental results more successfully than game-theoretic concepts of equilibrium. The literature presents success cases other than EWA, among them Quantal Response Equilibrium, Level-k, Cognitive Hierarchy and models of sophistication and strategic teaching (Camerer and Ho, 2014).

One of the simplest models to experimentally outpredict game-theoretic equilibrium concepts is Quantal Response Equilibrium (QRE) (McKelvey and Palfrey, 1995). QRE is a family of non-learning game-theoretic model where agents respond with some noise to accurate beliefs, and higher-payoff actions are chosen more often. The absence of learning makes QRE a benchmark for learning models – that is, if a particular learning model does not fit experimental data significantly better than QRE, this is usually indication that learning is not a significant part of the phenomenon under study.

Level-k (Stahl and Wilson, 1995) and the closely related Cognitive Hierarchy (CH) (Camerer; Ho and Chong, 2004) and Generalized Cognitive Hierarchy (Chong; Ho and Camerer, 2016) models are appropriate for one-off games or as models of initial responses. Both models split agents into k "levels", representing the number of steps of finitely iterated strategic thinking an agent is able to perform. Level 0 agents might either play salient strategies or randomize over responses and level 1 agents believe all other players to be to level 0 agents, for example. This increases up until the most cognitively advanced players, in level k, who believe other players' levels range from k-1 to zero. Agents then best respond to what they believe to be the distribution of other players' levels within the population. The main difference between Cognitive Hierarchy and Level-k models is the modeling of other players' levels' distribution. In level-k models, an agent with a given cognitive level h supposes all other players are level h-1 players. In CH models, agent's true distribution of probabilities are modeled according to a Poisson distribution parametrized by τ, while their beliefs about other players' levels correspond to the true distributions normalized up to the level immediately below their own. Finally, in GCH models, agents' beliefs about others follow the rationale in CH, but differ in that they are affected by stereotype bias: more frequent levels are disproportionally represented in their subjective beliefs.

Other broad classes of learning models include teaching models, where a portion of agents (called "sophisticated") know the true structure of the game and consider the other players' learning process when determining their own behavior (Stahl, 2003; Camerer; Ho and Chong, 2002). This kind of model has limited appeal in modeling financial crises. For this kind of behavior to represent crisis dynamics, the sophisticated agents would have to know the true structure of the economic environment and the expected evolution of crisis dynamics. This assumption is not consistent with the behavior of market participants observed during crises.

1.4.2 Bounded rationality models in banking

This dissertation's application of self-tuning EWA to banking (and banking crises more specifically) is novel. Furthermore, the only works to apply standard EWA to banking are a series of interrelated works developed by Barroso Barroso (2011), Barroso (2014), Lima (2014) and Lucchetti (2016), summarized by Barroso et al. (2016). These, in turn, were influenced by the pioneering application of EWA to finance in (Pouget, 2007).

An example of a boundedly-rational evolutionary banking model in which equilibration is not instantaneous is (Temzelides, 1997). Depositors repeatedly play the bank run stage game and evolving their strategies to maximize their payoffs. Smith and Shubik (2014) present a bank run model with replicator dynamics which can, for certain parameter values, nest both Diamond and Dybvig (1983) and Morris and Shin (1998) global-game families of bank run models.

1.4.3 Adaptive Learning and Financial Instability

The importance of exploring the nexus between learning and financial instability is to present an explanation for crises that does not depend on irrational behavior. Instability motivated by learning instead of irrationality is easier to explore in policy analysis, because it exempts modelers from the need to argue in favor of including some biases to the detriment of others in the modeling exercise.

Other authors have already established a relation between learning dynamics and financial instability, going back at least to (Sethi, 1992). The existing literature tries to do so mostly within the context of Hyman Minsky's Financial Instability Hypothesis (Minsky, 1976; Minsky, 1986; Minsky, 1992). They often also refer to features of financial crises as described by Kindleberger and Aliber (2015), and sometimes try to explain the common features of financial crisis summarized by Reinhart and Rogoff (2009).

The Financial Instability Hypothesis (henceforth referred to as FIH) states that financial crises can have endogenous origins, being more probable after prolonged periods of stability that tend to breed excessive optimism about the future prospects of the economy, resulting in excessive indebtedness within the private sector. The incompatibility between the FIH and the Rational Expectations Hypothesis (Muth, 1961) is a recurrent theme in the literature (comparing) learning and financial stability. Sethi (1992) argues that, although Full Information Rational Expectations (FIRE) is inconsistent with the FIH, the latter does not need agents to behave irrationally. Guzman and Howitt (2016) argue that under FIRE, agents can identify what portion of shocks are temporary or permanent. Consequently, greater observed stability would not cause agents to believe the environment is less risky, a result incompatible with the FIH.

Another common theme the impact of discounting the relative influence of more distant information. Sethi (1992) points out that when expectation formation privileges more recent observations to the detriment of older ones, agents adjust their expectations more rapidly but there is a greater tendency for instability. Guzman and Howitt (2016) present a rule called *stochastic-gain learning*, in which the observed forecasting error determines the discounting parameter: larger errors are indicative of a regime change and drive agents to privilege more recent information, while smaller errors decrease the discrepancy in weights between more recent and more distant errors.

The main difference between the present work and the existing literature is that previous works model learning with explicit expectation formation, contrary to what happens in EWA and its variants - where agents' decisions depend on strategies' relative attractions. As previously argued, forward-looking expectation formation during financial and banking crises is an assumption that cannot be taken for granted.

Furthermore, another difference is that the existing works are representative agent models, not allowing for interactions or agent heterogeneity. The lack of interaction effects is especially significant. Even though part of the existing literature points out the importance of feedback loops – especially between expectations and leverage (Howitt, 2017) – the dimension of such feedback might be underestimated due a lack of modeling direct or indirect contagion phenomena.

1.5 Discussion

Heterogeneous agent frameworks incorporating learning dynamics constitute a promising approach for policy analysis geared towards preventing banking crises and mitigating their impact. In particular, the combination of reinforcement learning and belief-based learning featured in the EWA family of models can allow for a wide variety of behaviors, helping capture the peculiarities of market participants' behavior during crises. Specifically, the Self-tuning EWA can be useful in capturing market participants' information acquisition motive in rapidly changing or low-information environments.

Learning and heterogeneity can be useful features when modeling banking crises. Endowing agents with learning behavior allows for modeling boundedly rational agents with choice-driven behavior. Modeling agents individually, instead of resorting to representative agents is also important, in order to capture significant phenomena within financial crises, such as contagion and asset sales. A computational approach is necessary to perform policy analysis with these types of models, because the combination of learning and heterogeneity makes them analytically intractable.

When analyzing the possible dynamics of agents whose behavior follows Self-Tuning EWA, the possibility of agents overreacting to shocks emerges. Consequently, overreaction can be a rational response to financial shocks by unbiased agents that learn in a procedurally rational manner. This behavior is the result agents' tendency to explore by randomizing their actions in rapidly changing environments. It thus becomes necessary to investigate whether this phenomenon could result in significantly different policy implications in any relevant scenarios.

Finally, when self-tuning EWA is used due to the inappropriateness of assuming equilibrium *ex ante*, the same phenomena might make it less likely that agents are capable of reliably generating contemporaneous counterfactuals. If there is doubt regarding the robustness of such assumption, it is recommendable to perform robustness checks by restricting reinforcement to strategies played contemporaneously by other agents.

2 Modeling learning for banking crisis policy analysis

Structural policy analysis is an important instrument for designing policy aimed at preventing and reducing the impact of banking crises. It models market participants' choice to enable preference-consistent behavior under unforeseen circumstances. Moreover, learning models can enable structural policy analysis without having to assume equilibrium, which is useful for modeling agents' behavior during crises.

This chapter will present a banking model to be used in structural policy analysis. The banks featured within the model learn according to Self-tuning Experience Weighted Attraction (Ho; Camerer and Chong, 2007). This learning model is parsimonious yet powerful enough to be used as a framework for banking crisis policy analysis. The learning banking model presented herein refines and adapts an existing banking model based on regular EWA (Barroso et al., 2016) for use with Self-tuning EWA.

In order to generate a quantitative model for structural policy analysis, a simple calibration procedure is performed over data observed from the Brazilian banking system. Aside from the model itself, this exercise is also useful to demonstrate the viability of using a structural model based on EWA, indicating the model's potential for further development to be used in policy analysis.

Additionally, this chapter features a sensibility analysis of the calibrated response sensitivity parameter within a policy response exercise. The objective is to highlight the importance of properly calibrating the banking model. The banking simulation is run to find to what extent banks' capital buffers (the policy response) vary with capital requirements (the policy instrument). Results show that policy response varies significantly when the calibrated parameter varies.

2.1 Simulation structure

The simulation featured herein is structured as a repeated bank run game played multiple times independently. The simulation has two main control structures: repetitions (one for each time the repeated game is played) and cycles (corresponding to the stage game, played both factually and counterfactually).

Every cycle is contained within a repetition. Cycles are operationally independent (that is, player's payoffs depend only on the state of nature and players' strategies in the current cycle) but strategically dependent (players' probabilities for choosing each strategy at a given cycle depends on strategies' past payoffs in previous cycles). Repetitions are completely independent of each other, both in an operational and a strategic sense.

The learning process occurs within the cycles of a single repetition. There is no learning between repetitions: agents act as if they forgot everything they learned when a new repetition begins. At the end of each cycle, agents' preferences concerning their strategies are updated according to their attractiveness, which in turn, depend on their payoffs, as modeled in the EWA approach described in section 1.2.1.

Finally, EWA requires that banks assess payoffs for strategies they could have played. Accordingly, the bank run game is played several times in each cycle: once for the factual state of the world (the strategies the agents actually played), and once for each agent's possible counterfactual strategies. This makes it possible for agents to engage in belief-based learning.

2.1.1 Banks' behavior

Banks' states at a given point in time are characterized by their balance sheets. The balance sheet reflects banks' financial intermediation function, that is, their propensity to invest in long term illiquid assets (bank loans) while funding themselves with liquid liabilities, redeemable on demand.

Table 1 summarizes the bank's liabilities. They include:

1. Capital, K_b: This is the bank's shareholder's equity, consisting of the shareholder's initial paid-up capital and subsequent profits and losses;
2. Deposits, D_b: These constitute the bank's liquid liabilities. Bank b's total deposits add up to D_b. Banks pay depositor an interest rate of i_d;
3. Interbank loans—where bank b is acting as the borrower—IB_b. When necessary, banks borrow from other banks in $t = 1$, as a response to a liquidity shock; they mature in $t = 2$, thus, maturing in one period. Their interest rate is given by the interbank market, costing the borrower i_i.

Table 1 – Banks' liabilities

Symbol	Liability	Maturity	Cost
K_b	Capital	-	-
D_b	Deposits	$t+2$	i_d
IB_b	Interbank loans	$t+1$	i_i

Table 2 summarizes banks' assets. They include:

1. Liquid assets, L_b: They are cash or cash-equivalent securities, constituting the bank's liquidity reserves. They are held so banks can honor the depositors' withdrawal requests, and yield no return;

2. Interbank loans—where the bank b acts as the lender—IL_b: For every interbank loan in the liability side of one bank's balance sheet, there will be another bank that holds it as an asset. Consequently, a debtor bank's cost will be a creditor bank b's return of i_i;

3. Real sector loans, R_b. These are long-term (two-period) loans to firms in the economy's corporate sector. Banks originate these loans on $t = 0$, maturing on $t = 2$. Each bank b lends to the set F_b of firms—with the aggregate amount R_b corresponds to the total of those firms' loans.

Table 2 – Banks' assets

Symbol	**Asset**	**Maturity**	**Return**
L_b	Liquid assets	t	-
IL_b	Interbank loans	$t+1$	i_i
R_b	Real sector loans	$t+2$	r_b

When banks are faced with withdrawals from depositors, their responses vary with the intensity of this liquidity shock. If a bank's liquid asset holdings are enough to honor all its withdrawal requests, it uses liquid assets to repay depositors and no further action is needed. If a bank's liquid assets are not sufficient to repay withdrawals, banks then proceed to sell some of their illiquid assets (namely, their bank loans) in an attempt to raise enough cash to honor those requests. Finally, if even if the asset sale is not sufficient to repay depositors, then banks can, in some versions of the simulation, borrow from other banks with a liquidity surplus.

If the bank is unable to fully respond to the liquidity shock after trying to take the aforementioned measures, it is liquidated due to insufficient cash flow. Banks' loans to firms then mature. If they become insolvent when their loans mature, they are also liquidated for balance sheet insolvency. Finally, in the simulations where there are interbank loans, there could be another scenario: if a given bank lent to other banks, and these other banks fail, the lending bank could wind up insolvent depending on how many of their counterparts fail and how much these loans represent. In that case, they are also liquidated due to insolvency.

When a bank is liquidated, its assets are sold at a discount of l % and the proceeds are divided among creditors, obeying a subordination hierarchy. First, legal and administrative costs are paid off, then bank loans are paid off (if any), then depositors and a finally If there is any money left, it is proportionally divided among shareholders.

The analysis of banks' strategic behavior is paramount to understanding the model's results. Each bank j is modeled with its strategies consisting of its initial capital and liquid assets as a proportion of its previously and exogenously determined asset size (respectively α^j and β^j). These parameters are sufficient to determine the bank's initial deposit base and loan amount, and consequently its initial balance sheet (because banks always start the stage game having neither borrowed to nor lent from other banks).

Banks' payoff functions are also crucial to characterize their strategic behavior. The payoff used to drive banks' learning behavior is their return on equity (RoE). This is consistent with a simplifying assumption, namely, not distinguishing between the bank manager and shareholder roles in order forego any analysis of agency issues. Furthermore, in case banks suffer losses, those losses are limited to shareholders' equity, that is, shareholders face only limited liability.

2.1.2 Depositors' behavior

Depositors are modeled as being subject to liquidity shocks in this simulation. At the beginning of every cycle, they necessarily deposit cash at their own bank. At a later moment, they have a positive probability of suffering a liquidity shock. In that case, they will withdraw the entirety of their deposits. Otherwise, they will wait until the deposit's maturity and then proceed to withdrawal. The amount they receive will depend on when they withdraw and whether the bank fails:

- a deposit held to maturity on a non-failing bank will earn the depositor interest, making his or her return positive;
- a deposit successfully drawn mid-cycle (that is, when the bank only fails at the end of the cycle or does not fail at all) will yield zero return (the depositor recoups the principal amount but receives no interest);
- finally, if a bank fails prior to or at the same moment the depositor withdraws, the depositor will receive a negative return, determined by the depositors' proceeds from the bank liquidation process.

2.1.3 Firms' behavior

Firms do not act strategically in any of the simulations featured in this dissertation. Accordingly, there are no associated strategies or payoffs for the simulation to keep track of. Firms' demand for credit is totally inelastic, meaning they borrow as much as the banks are willing to supply at whatever interest rate banks set. Their probability of default is exogenous, thus independent of the amount borrowed. A given firm f is characterized by the following parameters (heterogeneous among firms where applicable):

1. $R_{b,f}$, the amount bank b lent to firm f;
2. $i_{f,b}$, the interest rate paid by firm f to bank b on its loans;
3. PD_f, the probability of f not repaying its loan (probability of default);
4. LGD_f, firm f's loss given default, that is, the percentage of the loan's face value the bank will lose if f defaults on its loans.

2.1.4 Cycle timeline

Each cycle will consist of a bank run stage game, played once for the factual strategies and once for every counterfactual strategy agents want to learn from. Inspired by the seminal Diamond-Dybvig bank run game, the simulation's stage game will consist of three periods, $t = 0$ (the initial period, which can be referred to as "today"), $t = 1$ (the interim period, also referred to as the "short term") and $t = 2$ (the final period, also referred to as the "long term"). Within the stage game, the following events take place:

1. In the initial period, ($t = 0$), agents make their decisions and the whole model is configured for a new stage game according to the strategies:
   a) Banks choose their capital and liquidity ratios;
   b) The economy is configured according to agents' chosen strategies, with banks setting up their balance sheets accordingly (raising capital, capturing deposits and granting loans to firms).
2. In the interim period ($t = 1$), banks face a possible liquidity shock from depositors, and respond however possible:
   a) A portion of depositors suffer a liquidity shock and withdraw their deposits from their respective banks;
   b) Banks repay depositors' withdrawal requests with the liquid asset holdings immediately available;
   c) If banks' liquid asset holdings are not sufficient to honor their withdrawal requests, banks respond to the liquidity shock by selling the loans they originated;
   d) In some simulation scenarios, banks then proceed to fulfill any remaining unmet withdrawal requests with funds borrowed from other banks;

   e) If a bank still isn't able to honor all its withdrawal requests even after measures to respond to the liquidity shock, it will fail due to illiquidity and undergo a liquidation procedure.

3. In the final period ($t = 2$), banks face the risk of a solvency shock from firms' defaults:

   a) Banks' loans to firms mature, with firms possibly defaulting and consequently forcing the bank into liquidation caused by insolvency;
   b) In the event banks lent each other money, interbank loans then mature, with one bank's failure to repay possibly causing another bank's insolvency (a phenomenon also known as *direct contagion*);
   c) Deposits mature, with depositors receiving either their deposited amount back with interest if the bank did not go into liquidation, or the proceeds from the liquidation procedure otherwise;
   d) Banks observe their return on equity, and provide those payoffs as input to the learning algorithm.

After the end of the final period, all agents are reset for the next stage game (corresponding to a distinct counterfactual simulation). When all agents' possible counterfactual strategies have been played, the simulation proceeds to the repetition's next cycle. Finally, after the end of the repetition, the simulation proceeds to the next independent repetition.

2.1.5 Learning within the model

In order to adapt the existing banking model for use with the self-tuning EWA learning algorithm, some adaptations and modeling choices were necessary.

One of the main differences within from the self-tuning EWA model to the regular one is the presence of a change detector function $\phi_i(t)$. This requires that each agent i observe other agents' actions with the purpose of assessing to what extent their environment is changing in order to determine how informative past payoffs are as indicators of future actions' outcomes. This entails a modeling choice: to what agents does each self-tuning agent look in order to calculate the historical frequency of other players' strategies.

Banks' change detector function will look to other banks' played strategies. The motivation behind this choice is the following:

- banks cannot always look to depositors as indicators of a changing environment because in some variations of the simulations the depositors are zero-intelligence agents, with their behavior being exogenously determined;
- depending on the state of the game (and the parameters), banks could be significantly influenced by other banks (through interbank asset sale markets, or interbank loans);

- even in the absence of such phenomena, if other banks shift their strategies, it probably means the environment has changed significantly and past payoffs should be less influential in determining future payoffs.

Another minor adaptation to the learning algorithm is that no strategies are reinforced for a given agent when payoffs to all strategies are equal (resulting in no learning). In practice, this adaptation is not very relevant for banks' learning, because equal payoffs for different strategies will occur only in very exceptional circumstances. On the other hands, it might be useful if, in later versions of the framework, depositors are modeled as learning agents.

2.2 Calibrating model parameters

In order to perform policy analysis and design based on the banking model featured herein, it is necessary to estimate the model's exogenous parameters and calibrate the endogenous ones. The quantitative exercise will be performed by analyzing the Brazilian Banking System in the time period from January 2004 to December 2017. The data available for the Brazilian banking system has desirable features such as monthly bank accounting data frequency and daily aggregate deposits data. The time period was chosen to encompass two stress periods – one international, corresponding to the global financial crisis, and the other one corresponding to a domestic recession (Brazil's 2015-2016 recession).

A baseline scenario will be chosen to be illustrative of the end of the chosen period. Otherwise, when providing shocks or analyzing sensitivity to parameters, the range of variation in the chosen time period will serve as a basis regarding shock size and variability. The parameters will be the same across banks (but not necessarily across time) to keep the calibration exercise simpler. The main results regarding policy analysis require multiple agents in the system but not *ex ante* heterogeneity.

Table 3 summarizes the result of the calibration exercise, presenting the baseline scenario used for simulation. As will be shortly explained, some values were obtained from regulatory proxies, while most were estimated with available data. For the latter, the values correspond roughly to average monthly values for the year 2017.

2.2.1 Time period for measuring flow variables

Within this model, all time-dependent flow variables will take as reference a one-year horizon. Loans and deposit rates, default rates and deposit volatility are all measured relative to this chosen period.

Choosing a one-year time period is convenient for several reasons. Many rates, particularly expected default rates are computed with a one-year horizon in mind. Perhaps the most salient example is the regulatory use of the Vasicek model (Vasicek, 2002) for

Table 3 – Baseline scenario for simulation parameters

Symbol	**Name**	**Value**
wd_d	Mean probability of early withdrawal	7%
sd_{wd}	Standard deviation of probability of early withdrawal	5%
PD_f	Mean probability of firm default	6.5%
sd_{PD}	Standard deviation of firms' probability of default	2%
i_d	Interest rate on deposits	8.0%
r_b	Interest rate on loans	16.0%
i_i	Interbank interest rate	12%
δ_L	Haircut on sale of illiquid assets	15%
δ_I	Insolvency haircut on illiquid assets	40%
$\delta_{admlegal}$	Insolvency costs as proportion of assets	10%

allocation of capital against credit risk starting with the second version of the Basel Accord (BCBS, 2006). In prudential regulation, the one year horizon is often employed as a cut-off point below which liabilities are considered short-term, both in liquidity regulation (BCBS, 2014b) and in the definition of regulatory capital (BCBS, 2011).

2.2.2 Loans and deposit interest rates

Once the time period has been determined, it is possible to analyze rates for loans and deposits to parametrize the framework. Brazil has historically had high nominal and real interest rates relative to developed countries and most emerging economies. It is possible to see a long-term downward trend during the time period used to calibrate this model.

Both the deposit rate i_d and the loan rate r_b used to calibrate the model were taken from the overall aggregate rates in the Brazilian banking system, as measured by banks' accounting variables: the ratio between credit-related revenues and loans; and the ratio between funding costs and funding liabilities. Their values over time are represented in figure 1.

2.2.3 Probability of firm default

The probability of each firm's default, exogenous within the context of this simulation, is a parameter with important implication's to banks' strategies. Their values will be estimated by the proportion of banks' loans which are more than 90 days past due. In general, this corresponds to the expected future credit losses over a one-year period, thus justifying its use as a proxy for the loan portfolio's probability of default PD_f. Their aggregate values can be seen in figure 2.

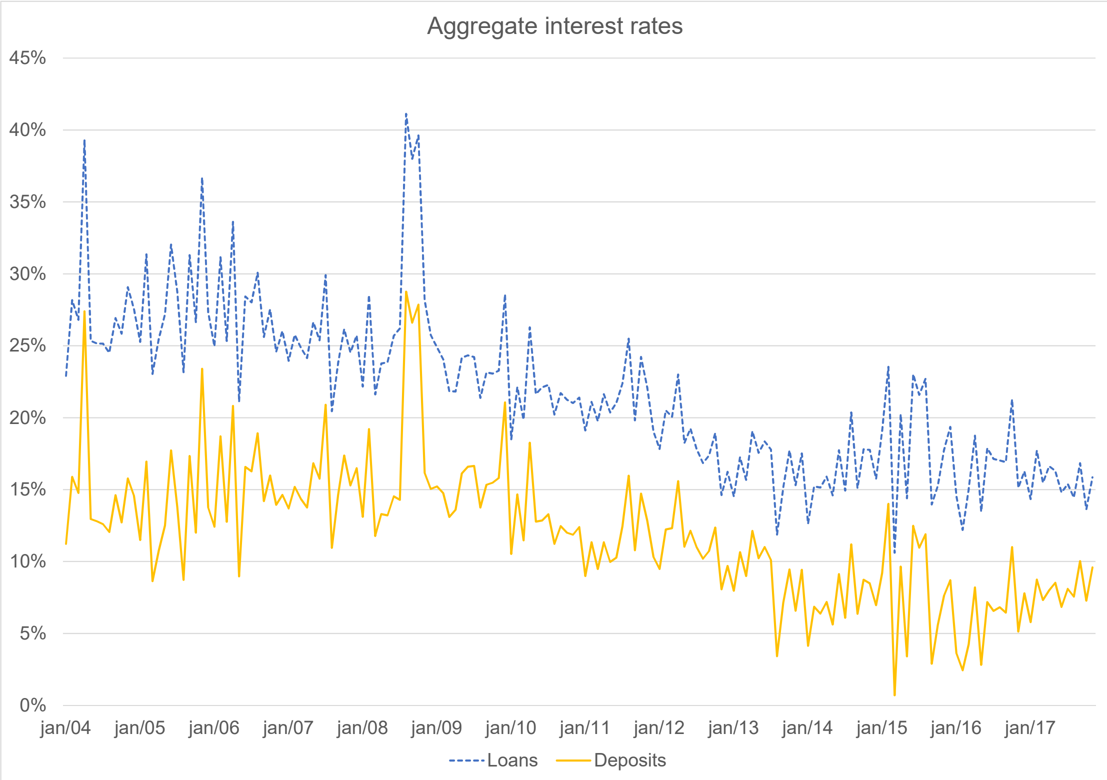

Figure 1 – Aggregate loan and deposit rates in the Brazilian banking system, 2004-2017

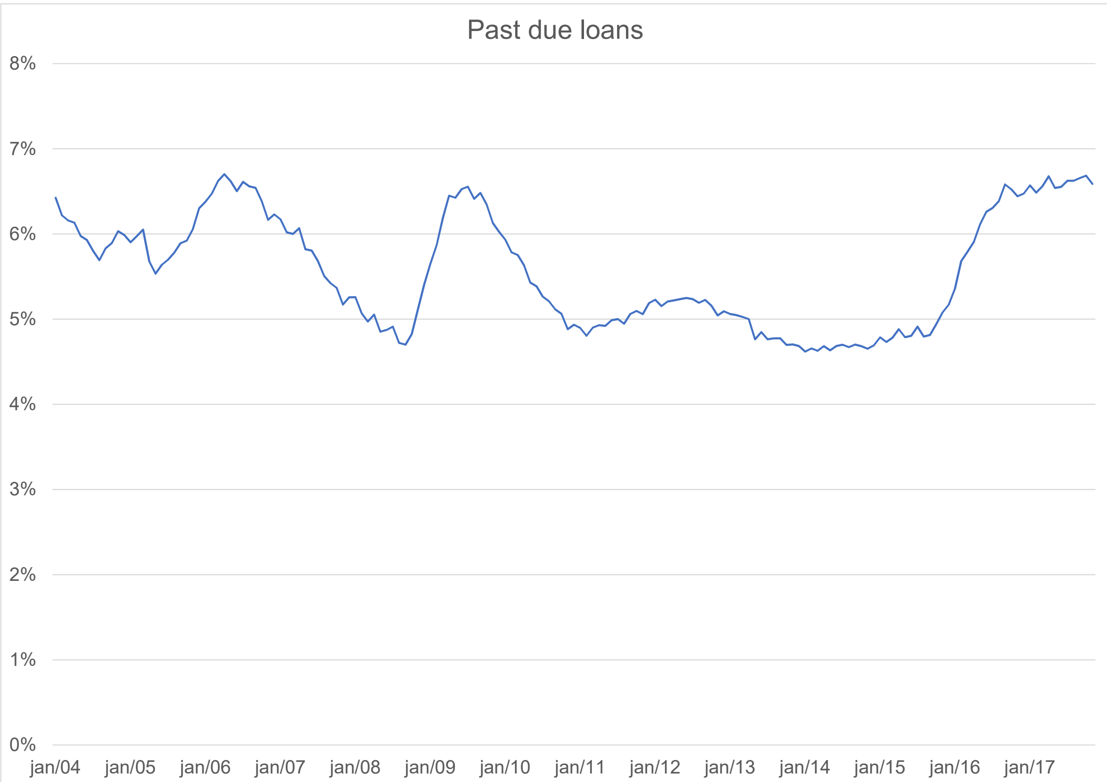

Figure 2 – Aggregate past due loans in the Brazilian banking system, 2004-2017

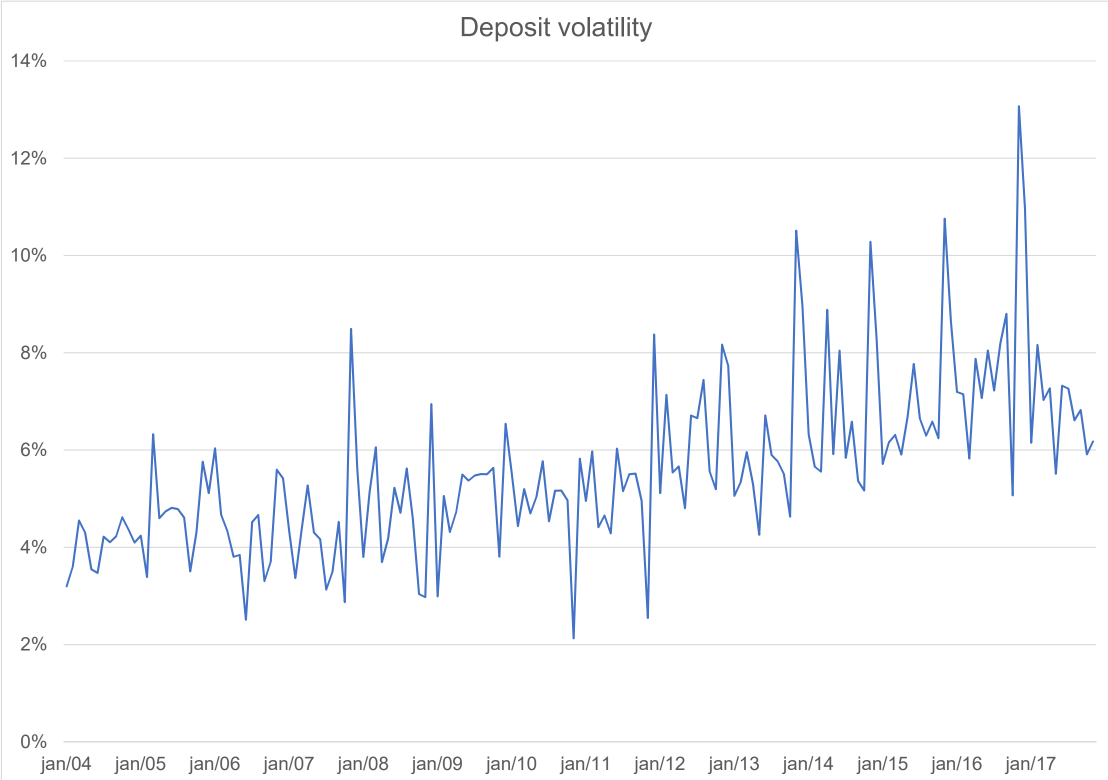

Figure 3 – Aggregate deposit volatility in the Brazilian banking system, 2004-2017

2.2.4 Deposit volatility

Deposit volatility is an important dimension within the framework. In the context of this model, it is the driving parameter behind each depositor's exogenous probability of withdrawal. Like in (Barroso et al., 2016), this simulation implements depositor behavior by stipulating that depositors withdraw their funds with a unique probability calculated each time according to a uniform probability distribution. Depositors' withdrawal probability will be defined as the proportion between each month's maximum aggregate withdrawals relative to the period's aggregate mean deposits. For example, if during a given month, the minimum daily value for deposits represents 97% of the mean value, deposit volatility for that month will be set at 3%.

The evolution of deposit volatility over time is represented in figure 3. The mean probability of early withdrawal wd_d will correspond to the expected value of aggregate deposit volatility over the calibration period, while the standard deviation of the probability of early withdrawal sd_{wd}

2.2.5 Regulatory haircuts

One of the way banks can deal with liquidity crises is by selling risky illiquid assets (synonymous with loans in this framework). Given their credit risk and illiquid nature, it is expected (and observed) that loans are sold with a haircut. A particularly difficult parameter to calibrate is the regulatory haircut for asset sales. There are no data available for the Brazilian banking system, so the haircut on sale of illiquid assets δ_L will be calibrated based on data from other jurisdictions, as well as regulatory haircuts. The default value for the haircut (that is, for non-crisis periods) will be set at 15%.

In practice, haircuts on illiquid assets will vary significantly with portfolio composition as well as market conditions. The Basel Committee's Liquidity Coverage Ratio (BCBS, 2013a) stipulates haircuts of at least 15% on high-quality debt securities (highly graded by rating agencies) corporate), 25 %, for residential mortgage backed securities, and finally 50% for equity and non-speculative grade corporate debt securities. Furthermore, during the Global Financial Crisis, the mean haircut in the organized U.S. market for Syndicated loans of 40% (Irani and Meisenzahl, 2017).

2.2.6 Liquidation values

There are no publicly available studies of resolution or liquidation costs for Brazilian banks. We resort to international studies – one comprehensive review of U.S. banks, and a more restricted one focusing on the largest OECD bank resolution episodes during the recent financial crisis.

Grimaldi et al. (2017) analyze costs associated with bank resolution episodes in OECD member countries during the most recent financial crisis. Most notably, the authors were able to find economies of scale in resolution costs. Applying their findings to Brazilian banks would yield resolution costs of approximately to 9% of total bank assets (for the largest banks) to 16% (for a representative bank in the Brazilian banking system).

Bennett and Unal (2014), Bennett and Unal (2015) present data on losses observed in bank resolution episodes overseen by the United States' Federal Deposit Insurance Corporation (FDIC). Typically, resolution costs vary between 22 to 33% of bank assets, also subject to economies of scale, consistently with the OECD study.

Given wide range of estimates and reduced impact in results, this simulation will follow (Barroso et al., 2016) in setting the insolvency haircut δ_I on illiquid assets to 40%. Furthermore, each insolvency episode will also result in legal and administrative costs corresponding to 10% of total assets.

2.3 Calibration results

The procedure for calibrating this banking framework consists in finding the value for response sensitivity λ which minimizes the distance between banks' observed choice variables (capital and liquidity levels) and the values for these same variables resulting from simulation. The distance metric to be minimized will be the sum of squared deviations of

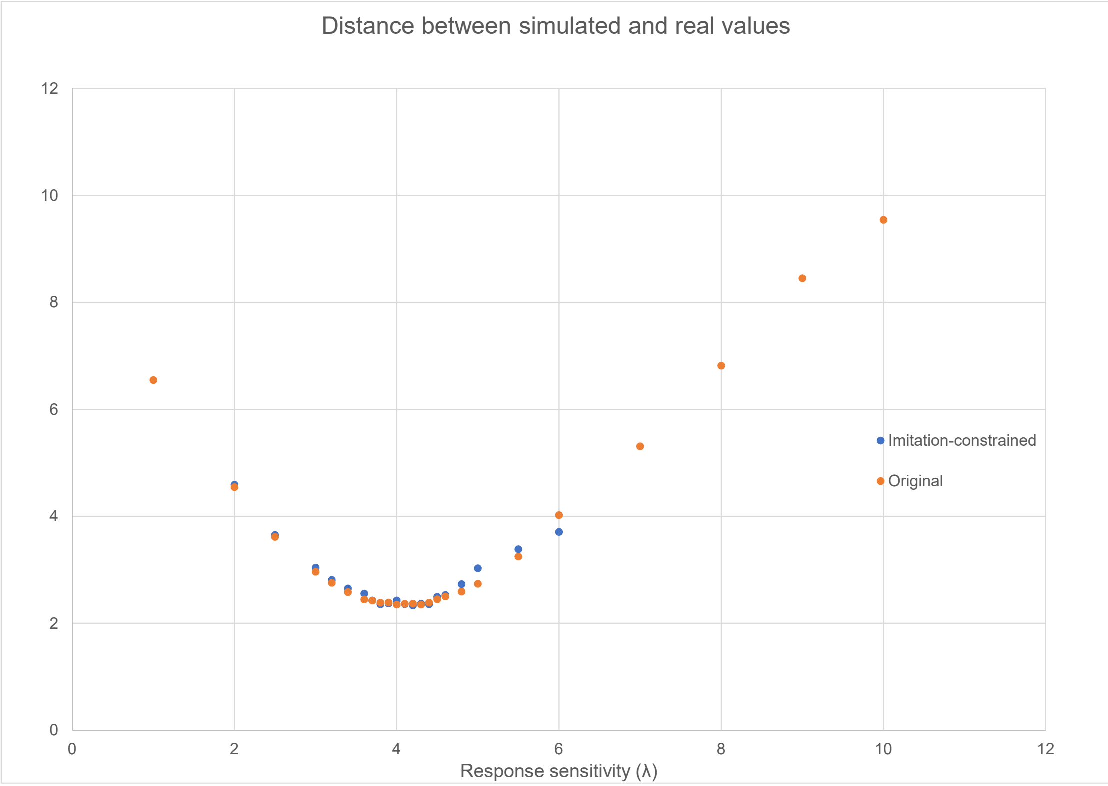

Figure 4 – Calibration results as a function of response sensitivity

simulated capital and liquidity relative to their observed values. The deviations in capital and liqudity for each period will be respectively weighted by the inverse of the variance of their observed values over the whole time period. The exogenous and outcome variables are aggregates over the Brazilian banking system observed monthly for the specified period.

The calibration procedure will iterate over several values of λ. For each possible value, the simulation iterates over each month of the time period. The simulation will then be run with the current value of λ and the exogenous variables' values for that month, according to the previous section's description. The resulting choice variables will be compared to the values observed at that time period.

Figure 4 shows the results for the calibration procedure - the distance between simulated and actual values as a function of the response sensitivity λ. It is possible to see that the function reaches its minima when the λ reaches approximately 4.

During crises, banking system behavior can be overrun with phenomena which increase its complexity and hinder banks' capacity to estimate counterfactual payoffs - even contemporaneous ones. For that reason, a useful robustness check is to verify whether the same results hold when banks only play strategies that at least one bank played - a proxy for counterfactual generation by imitation. Figure 4 shows the proximity of the results obtained in the calibration exercise for both the original procedure and the robustness check.

The minimal values obtained for the response sensitivity parameter as a result of the calibration procedure lie within the range of variation found by the authors of the STEWA model (Ho; Camerer and Chong, 2007) when estimating this parameter for archetypal games. Higher values of response sensitivity means banks' behavior exhibits a stronger best response characteristic. This situation is indicative of bank's payoffs having stronger ties to their chosen strategy.

Table 4 – Estimated response sensitivity from common games - adapted from (Ho; Camerer and Chong, 2007)

Game	Mixed strategies	Patent race	Median action	Pot games	p-Beauty contest	Price matching
λ value	4.13	9.24	5.64	7.34	2.39	10.17

2.4 Policy impact of response sensitivity calibration

This section features a sensibility analysis of the calibrated response sensitivity parameter λ within a policy response exercise – varying capital requirements to gauge banks' response in terms of capital buffers held. The objective is to hightlight the importance of properly calibrating the model by demonstrating banks' different responses to varying values of the parameter.

It is important that banks keep positive capital levels due to the loss absorption capacity of equity. Moreover, there is a discrepancy between how much capital banks need in times of crisis to protect against losses and their limited prospects for raising capital during those crisis periods. It is difficult to issue shares in times of crises due to investors' reduced appetite for equity issuance in such periods. Additionally, banks' possibilities for raising capital through earnings retention during crises are limited because of reduced or non-existent profits.

The theoretically optimal capital requirement for banks is the one that maximizes unconditional return on equity, given regulatory restraints. Higher capital levels lower banks' probability of default for a given shock probability through greater loss absorption capacity. On the other hand, they decrease banks' return on equity conditional on the bank not defaulting. Consequently, the optimal level of capital must be high enough to make failure improbable, but low enough so that banks' return on equity benefits from the additional investment opportunities made possible by leverage.

In practice, banks tend to hold capital levels significantly higher than regulatory minima. Even when banks' theoretical optimum coincides with the regulatory minimum, they face incentives to hold more than capital than this minimum. Banks hold capital buffers, for instance, in order not to face increased regulatory oversight, potentially increasing compliance costs, or still, to avoid being subject to dividend payout restrictions.

2.4.1 Experiment description and rationale

The banking simulation calibrated in the previous section is repeated for multiple values of the response sensitivity parameter λ, in order to gauge its impact on banks' response to policy change (different capital requirements). The following values of λ are used:

- 1.0: used by Barroso et al. (2016) in the banking model featuring standard EWA;
- 2.0: an approximate lower bound found by the authors of STEWA for archetypical games;
- 3.0: an approximate lower bound found in the calibration exercise;
- 4.0: the approximate optimal value found in the calibration exercise;
- 5.0: an approximate upper bound found in the calibration exercise;
- 10.0: an approximate upper bound found by the authors of STEWA for archetypical games.

Moreover, to gauge the impact of capital requirements, the experiment is repeated for multiple values of the minimum leverage ratio (equity as a proportion of total assets). Besides the original value of 3%, based on Basel Committee's leverage ratio requirement (BCBS, 2014a), experiments feature one lower value (1%) and two higher values (5 and 7%) for the requirement.

2.4.2 Results and analysis

Banks' liquidity and capital levels vary with both the policy variable (the minimum capital requirement) and the exogenous variable subject to calibration (the response sensitivity λ). Tables 5 and 6 feature, respectively, the mean capital levels (equity as a proportion of total assets) and the capital buffer – the difference between actual and required capital levels. Additionally 7 features banks' mean liquidity (liquid assets as a proportion of total assets) as a function of the requirement and λ.

Banks' reaction to λ in terms of leverage strategies is very clear: for each value of the capital requirement, both capital levels and buffers decrease with increasing response sensitivities. Given the fact that higher sensitivies drive agents' behavior away from randomness and towards best responses, this points towards banks' payoffs from more leveraged strategies being higher than payoffs resulting from less leveraged ones – at least within the range of strategies chosen in these experiments.

Banks' reaction to λ in terms of liquidity strategies is analogous: for each value of the capital requirement, banks reduce their liquidity holdings with increasing response

Table 5 – Banks' response to variations in policy and response sensitivity – capital levels

		Capital requirement			
		1%	3%	5%	7%
λ	1.0	0.0970	0.1131	0.1242	0.1347
	2.0	0.0880	0.1099	0.1231	0.1340
	3.0	0.0818	0.1063	0.1205	0.1326
	4.0	0.0770	0.1027	0.1189	0.1314
	5.0	0.0738	0.0990	0.1168	0.1303
	10.0	0.0592	0.0862	0.1075	0.1236

Table 6 – Banks' response to variations in policy and response sensitivity – capital buffers

		Capital requirement			
		1%	3%	5%	7%
λ	1.0	0.0870	0.0831	0.0742	0.0647
	2.0	0.0780	0.0799	0.0731	0.064
	3.0	0.0718	0.0763	0.0705	0.0626
	4.0	0.0670	0.0727	0.0689	0.0614
	5.0	0.0638	0.0690	0.0668	0.0603
	10.0	0.0492	0.0562	0.0575	0.0536

Table 7 – Banks' response to variations in policy and response sensitivity – liquidity levels

		Capital requirement			
		1%	3%	5%	7%
λ	1.0	0.1318	0.1352	0.1355	0.1365
	2.0	0.1252	0.1300	0.1314	0.1318
	3.0	0.1202	0.1246	0.1267	0.1289
	4.0	0.1155	0.1196	0.1226	0.1246
	5.0	0.1122	0.1156	0.1187	0.1208
	10.0	0.0993	0.1000	0.1026	0.1059

sensitivity. It is interesting to note that liquidity levels are less sensitive to λ than capital. This is expected, given that liquidity impacts banks' payoff less directly than capital levels – the denominator of return on equity.

2.5 Implementation details

This dissertation's simulation was developed using the C++ programming language, in order to combine simulation performance and ease of development. The source code is based on the 2011 revision of the language standard (ISO, 2011), also known as *C++11*.

No compiler or platform-specific constructs were used to keep the source code portable to multiple hardware platforms and operating systems.

Silva (2018) discusses the relative merits of different programming languages and software platforms in the development of multi-agent systems for use in economic simulations. He also presents an implementation of Barroso et al.'s framework in the Python programming language (PSF, 2018). It was originally programmed in C++.

2.6 Literature Review

The most common way to calibrate Agent Based Models is for the modeler to define stylized facts which they want to match and assess which parameter combinations best fit these facts using Simulated Method of Moments. A recent and detailed example which also presents a historical perspective on the related literature can be found in (Chen and Lux, 2018).

Alternatively, recent methodological advances in model comparison and validation could be coupled with intelligent use of sampling and search algorithms in order to calibrate the models, to determine which parameter sets correspond to the best fit to reality. One could use, for example, the comparison methodology presented in (Guerini and Moneta, 2017), which attempts to measure similarity of real world and model generated time series causal structure using Vector Auto Regression, Vector Error Correction Models and Graphical Causal Models analysis. Another alternative is using the GSL-Div information theoretic criterion (Lamperti, 2018), which attempts to measure the distance between the dynamics of time series produced by simulation models and the empirically observable counterpart without needing to resort to the likelihood function.

Incorporating some of the methodological insights present in the aforementioned works could significantly increase the complexity of the calibration exercise (and consequently the analysis of results). Additionally, this would possibly lead to greater computational complexity and increased run times. As a possible antidote to such complexity, we can mention Lamperti et al.'s research on Machine Learning Surrogates for Agent-Based Models (Lamperti; Roventini and Sani, 2018), in which the calibration exercise is performed not on the model itself, but on a simplified surrogate version built by machine learning techniques.

This surrogate model is built to provide an accurate approximation of the original model, but with significantly faster execution times. The approximation is performed by a machine learning model called extreme gradient boosted trees (XGBoost) (Chen and Guestrin, 2016). It is calibrated by feeding it the original model's results over a reduced sample of the original model's parameter space. The authors justify the use of such model due to the usually non-linear nature of the Agent Based Models' response to their inputs (parameters and initial conditions).

The main motivation behind the simple procedure used to calibrate this banking model is the reduced applicability of sophisticated ones to the particular banking model presented herein. The complexities associated with interbank interactions render implausible the derivation of closed-form solutions and, consequently, likelihood functions – the approach taken by the authors when presenting the original EWA (Camerer and Ho, 1999) and Self-Tuning EWA (Ho; Camerer and Chong, 2007) models. Moreover, the applicability of state-of-the-art calibration methodologies discussed above is limited because this model is not structured to generate dependencies between results in different time periods.

2.7 Discussion

This chapter introduces a multi-agent banking model with learning. It is calibrated based on data obtained from the Brazilian banking system. In order to make it a choice-based model to enable its use in structural policy analysis, it features the Self-Tuning EWA learning model. The use of the Self-tuning variety of EWA instead of the original one is important because it allows information-acquisition dynamics to influence agents' behavior with relevant implications for financial stability.

The calibration exercise performed in this chapter provides parameter values for use in bank crisis policy analysis – for use in both baseline and crisis scenarios. It is worth noting that the robustness check performed by partially restricting reinforcement of counterfactual strategies does not yield significantly different results compared to those obtained without the restriction.

The framework featured herein can also be used in other contexts. It can be adapted to assess future regulatory modification proposals, regarding issues such as capital requirements, liquidity requirements and deposit insurance. It can also be calibrated for use in other jurisdictions.

3 Linking Learning and Instability

Banking crises can have disastrous consequences, notably lost output and fiscal deterioration (Reinhart and Rogoff, 2009). Such undesirable outcomes can be the result of extreme exogenous shocks but can also come as the result of more moderate shocks amplified within the banking system. This shock amplification phenomenon is frequent and significant enough within crises to be a cause for concern among policymakers (Kindleberger and Aliber, 2015).

A potential source of banking system instability is the possibility of market participants overreacting to shocks: when their immediate response to shocks is stronger than the one observed over the long term. Overreaction can result in qualitatively different outcomes when agent's response is strong enough to result in explosive non-equilibrium dynamics to shift the system to a different equilibrium, if it exists. Even if the outcome of agents' overreaction does not scale up to full-blown instability, this phenomenon can nonetheless have important consequences. It could possibly alter the relationship between shock magnitude and outcome severity to one significantly different from that observed in non-crisis periods.

There are many amplification mechanisms, such as fire sales and bank runs, whose analysis can benefit from the use of simulations that model participants individually instead of as representative agents. These multi-agent simulations can be particularly useful when agent behavior is driven by learning. Learning allows the system under simulation to temporarily distance itself from equilibrium states, while keeping the desirable properties of optimization-driven frameworks: agents' behavior being driven by choice and consistent with preferred outcomes.

Given the potential impact of overreaction to banking system instability and the desirability of learning-based multi-agent simulations to banking policy analysis, it is important to investigate whether learning can lead to overreaction within multi-agent models of banking systems and, in that case, whether this overreaction can result in instability. In order to answer these questions, this chapter features a series of computational simulations based on a multi-agent banking system model, where agents learn according to the Self-tuning Experience Weighted Attraction (STEWA) model (Ho; Camerer and Chong, 2007). The results point to overreaction by market participants, influencing their response in quantitative terms, but not intensely enough to make results qualitatively different. The experiments also show that factors affecting overreaction in the STEWA learning model are consistent with market participants' behavior described by the literature on banking crises. Those results constitute evidence in favor of the methodological adequacy of Self-tuning EWA as a learning model for banking crisis simulation and policy design.

3.1 Overreaction during crises

The literature on financial instability points out that learning can be a source of instability, at least since (Sethi, 1992), and more recently in (Guzman and Howitt, 2016) and (Howitt, 2017). The existing models, however, are representative agent models, making it difficult to simulate important phenomena in banking crises, such as contagion and fire sales. Consequently, it is important to analyze whether learning can cause overreaction, and particularly, lead to instability within the context of multi-agent systems. An important possibility for instability through overreaction in learning-driven MAS is through learning models that include an information acquisition motive, such as Self-tuning EWA. Another distinction that separates STEWA from models featured in the literature that links learning and financial instability is that in the self-tuning model, exhibits some randomness in choice behavior instead of restricting agents to playing expected best responses.

These peculiarities come into play to make agents overreact to certain shocks, doing so in a manner consistent with the literature on banking crises. This means the significance of overreaction should depend on how surprising the shock is, making agents' behavior more unpredictable. Larger shocks over a shorter period of time should trigger more intense overreaction, especially if agents' environment exhibited low volatility prior to the shock.

If changes of similar magnitude spread out over longer periods of time, agents will have more time to adapt to such changes. The slower the changes, the better approximation equilibrium becomes (Kirman, 2011). If changes are so slow that the agents can form an accurate mental model of their environment, they can be completely rational and forward-looking, behaving optimally, which precludes overreaction.

This is consistent with the STEWA model. A more gradual change to agents' environment would mean, within the context of this model, that it occurs over multiple learning cycles. There is thus a larger probability that the strategies most frequently played in a given round were also frequently played in previous rounds, driving down the surprise factor.

Another important aspect is that financial crises often occur shortly after a period of low volatility resulting from high investor confidence. This can be caused by an increase in asset prices and associated increase in leverage (Minsky, 1992). This increase in confidence can also occur as a result of a period of sustained financial market innovation, making market participants ignore warnings about the buildup of debt (Reinhart and Rogoff, 2009).

In the STEWA model, this is reflected in the surprise factor being inversely proportional to the volatility observed in other players' strategies. This will drive up the historical frequency of a greater number of strategies, making them less surprising when played, resulting in a lower surprise factor. Thus, whenever the volatility in agents' economic environment makes their relative payoffs more volatile, this will be reflected in a lower surprise factor.

3.2 Experiment design

In order to examine the hypotheses stated in the previous section regarding banks' overreaction to shocks during crises, a series of experiments will be run using the banking model presented in this dissertation's previous chapter. The first will establish the overreaction phenomenon. Each subsequent experiment will examine a particular factor influencing how this overreaction behavior varies.

The experiments will follow a common structure. For every independent repetition, the simulation is run for a number of cycles prior to the shock. During this pre-shock period, banks will gradually stabilize their strategies. After this initial period, a shock to the banking system will occur: either instantaneous or spread out over multiple cycles. Then, there will be a relatively short "shock" window during which the banks' choice variables (capital and liquidity) can quickly distance themselves from the pre-shock steady-state, and then, after reaching a peak, converge to the post-shock long run state, representing a partial reversion towards the original state. This difference between this short-term peak and the long-run post-shock steady state ("overshooting") will be shown to be significant in some experiments and is the main object of this analysis.

In order to gauge the extent to which the overshooting phenomenon is significant, it can be compared to two references:

- The difference between pre-shock and post-shock steady states: if the overshooting deviation is not significant compared to this benchmark, it will not have practical significance for market participants' and authorities' behavior;

- The post-shock volatility: if the overshooting deviation is not significantly greater than the post-shock volatility, any temporary initial overreaction can be just a manifestation of increased volatility, not necessarily a distinct phenomenon.

Finally, in order for a peak in capital or liquidity to be considered an overreaction, it must take place within a short time period (an order of magnitude smaller than the time needed for convergence). Extreme values occurring after this interval will be understood to be a manifestation of natural variability in banks' behavior.

To analyze banks' overreaction to shocks, the metrics of interest are all related to their choice variables (capital and liquidity). For those choice variables, the following metrics of interest will be analyzed:

- overshooting window: a relatively short number of iterations during which the maximum value is assessed, starting with the period in which the shock begins;
- overshot value: the maximum value the metric of interest reaches within the overshooting window;
- post-shock mean: the mean of the metric's observed values after the overshooting window;
- overshooting variation: the difference between the overshot value and the post-shock mean;
- overshooting percentage: the proportion between the overshooting variation and the shift in steady-state means (pre- and post-shock).

For all experiments, the overshooting window will be [10] cycles long – that is, for the purpose of detecting overshooting, the value to be considered will be the maximum reached within that interval.

3.3 Initial experimental configuration

The possibility of banks overreacting to financial shocks is important for all financial market participants, such as depositors, authorities and other banks. In order to better investigate this phenomenon, a first step is to establish an initial scenario where it takes place. The financial and banking crisis literature show that shocks are most dangerous when banks are highly levered relative to long-term trends (Minsky, 1992; Reinhart and Rogoff, 2009), which usually happens in periods of low default rates or high asset prices. An overreaction to a shock is thus more probable when a shock strikes a banking system in a benign condition. Consequently, the initial scenario prior to the shock will be one where probability of default and deposit volatility are unusually low. Furthermore, given the benign condition of the banking system, banks can fund themselves at low prices. The combination of all these factors points towards a scenario where banks are profitable and are probably not risk-averse, leading them to become highly levered and hold a low stock of liquid assets.

The initial pre-shock scenario for this simulation is inspired by the baseline scenario presented in this dissertation's previous chapter. This means each bank will be parametrized similarly *ex ante*, according to the Brazilian banking system's aggregate parametrization.

In particular, banks will have their learning processes governed by the response sensitivity value found in the last chapter's calibration exercise ($\lambda = 4.0$).

This initial scenario will differ from the previous chapter's baseline in order to configure a particularly benign situation for the banking system. More precisely, banks will face a lower probability PD_f of firm default (2.5% versus 6.5%), will face lower expected deposit volatility wd_d (4% versus 7%) and will be able to fund themselves paying out a lower interest rate i_d to depositors (6% versus 8%). Additionally, the benign scenario will also manifest itself in decreased volatility, by means of lower standard deviation for the probability of firm default sd_{PD} (1% vs 2%) and early withdrawal sd_{wd} (1% vs 5%). Table 8 synthesizes the pre-shock parametrization for the initial experiment.

Table 8 – Parametrization for the initial experiment

Symbol	**Name**	**Value**
wd_d	Mean probability of early withdrawal	4%
sd_{wd}	Standard deviation of probability of early withdrawal	1%
PD_f	Mean probability of firm default	2.5%
sd_{PD}	Standard deviation of firms' probability of default	1%
i_d	Interest rate on deposits	6.0%
r_b	Interest rate on loans	16.0%
i_i	Interbank interest rate	12%
δ_L	Haircut on sale of illiquid assets	15%
δ_I	Insolvency haircut on illiquid assets	40%
$\delta_{admlegal}$	Insolvency costs as proportion of assets	10%
λ	Response sensitivity	4.0

The shock applied to the banking system in this initial experiment will correspond to increases in probability of firm default, deposit volatility and a decrease in the haircut practiced in the sale of illiquid assets. These variables were chosen because they impact significantly bank behavior within the model and because they vary significantly in most banking system crises. Rapid increases in probability of default are often described as an important cause of banking crises, when latent risk built up in banks' loan portfolio materializes. Deposit volatility also increases in banking crises. It is very common for banks to lose access to some of their funding sources during crises, forcing them to either to draw upon their stock of liquid assets or sell illiquid assets. Additionally, an increase in the haircut on risky assets is also common, given that these assets are usually subject to credit or market risk. This diminished funding puts pressures on banks to sell risky assets, increasing their supply and driving down their price. Finally, an increased perception of risk by market participants will drive banks' funding costs up, cutting into their profit margins. Those parameters will be subject to shocks:

- the deposit rate i_d will return to its baseline value (8%);
- the probability of firm default PD_f will increase to 8.5%;
- the deposit volatility parameter wd_d will be set to 15%, which is slightly less than two standard deviations above the original mean. As a reference, the regulatory run-off factor for wholesale funding subject to deposit insurance is 20% (BCBS, 2013a);
- the haircut on risky assets δ_L will be set to 40%, which was approximately the mean haircut in the U.S. market for syndicated loans during the global financial crisis (Irani and Meisenzahl, 2017). As an additional reference, non-speculative grade corporate debt securities receive a 50% regulatory haircut when being accounted for in liquidity requirements (BCBS, 2013a).

This shock to the banking system leads banks to overreact relative to their choice of liquidity reserves (the proportion between liquid asset holdings and total assets). Figure 5 shows a sudden and permanent rise in banks' capital strategies (the proportion between equity and total assets) immediately after the shock. Figure 6, on the other hand, demonstrates a sudden rise in banks' liquidity before a partial reversion to a lower level. Table 9 illustrates the significance of overreaction relative to shift in long-term strategies.

Table 9 – Leverage and liquidity shock reactions in initial experiment

Metric	Value (leverage)	Value (liquidity)
Pre-shock steady state	0.0570	0.1019
Post-shock steady state	0.1069	0.1210
Post-shock volatility	0.0026	0.0027
Overreaction deviation	0.0033	0.0085
Deviation/shift	6.68%	44.27%
Deviation/volatility	1.2475	3.113

3.4 Variations on the initial experiment

After providing evidence of overreaction in bank's behavior as a consequence of the learning process, it is important to investigate what can influence the intensity of this phenomenon. This is important in order to check for robustness with regards to assumptions and parameters. Also, it is useful as a check of whether bank' overreaction behavior is consistent with the literature on banking crises. This analysis consists in altering the initial simulation to space the shock out over many periods. Furthermore, the initial simulation will also be modified to run with the original EWA model, in order to

Figure 5 – Pre- and post-shock leverage, initial experiment

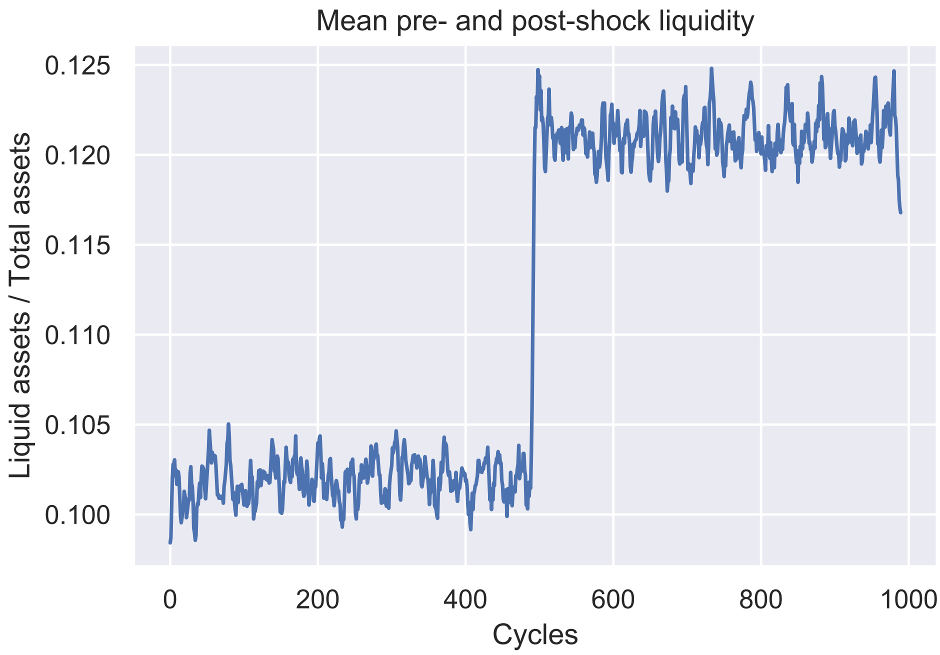

Figure 6 – Pre- and post-shock liquidity, initial experiment

gauge the role of the Self-tuning EWA model's distinctive information acquisition dynamics in agents' overreaction.

3.4.1 Original EWA

Given that the original EWA model is already used in banking simulations (Barroso et al., 2016), it is important to compare the results obtained by the Self-tuning EWA learning model and by its predecessor, highlighting in what situations the results differ the most. In order for the models to be comparable, the simulations have to account for a fundamental difference in the models: the decay of previous attractions is an exogenous parameter in EWA, while in the self-tuning variant, it is endogenously determined by the change detector function. In order to ensure that any variation in results is attributable to the differences in the model — and not to parameter values — the self-tuning model must be run first. Then, the value of the change detector function prior to the shock should serve as input to be used it the original EWA model.

Figures 7 and 8, as well as Table 10 show the banks' response to shocks under the original EWA model. Contrary to the initial experiment, there is no significant evidence of overshooting behavior

Table 10 – Leverage and liquidity shock reactions in the original EWA model

Metric	Value (leverage)	Value (liquidity)
Pre-shock steady state	0.0574	0.1054
Post-shock steady state	0.1076	0.1189
Post-shock volatility	0.0028	0.0028
Overreaction deviation	0.0029	0.0013
Deviation/shift	5.76%	9.96%
Deviation/volatility	1.0474	0.4869

3.4.2 Shock over multiple cycles

When a banking system is subject to a shock of a given intensity, the longer the shock takes to come into full effect, the smaller the magnitude of the overreaction phenomenon. A slower transition to another state alleviates banks' information acquisition motive, reducing their need to explore of the strategy space. Within the STEWA model, this is captured by a slower change in strategies' relative payoffs, reducing the surprise factor.

In order to demonstrate this property, the baseline scenario will be repeated with the previously instantaneous shock being spread out over 20 cycles. The results are shown

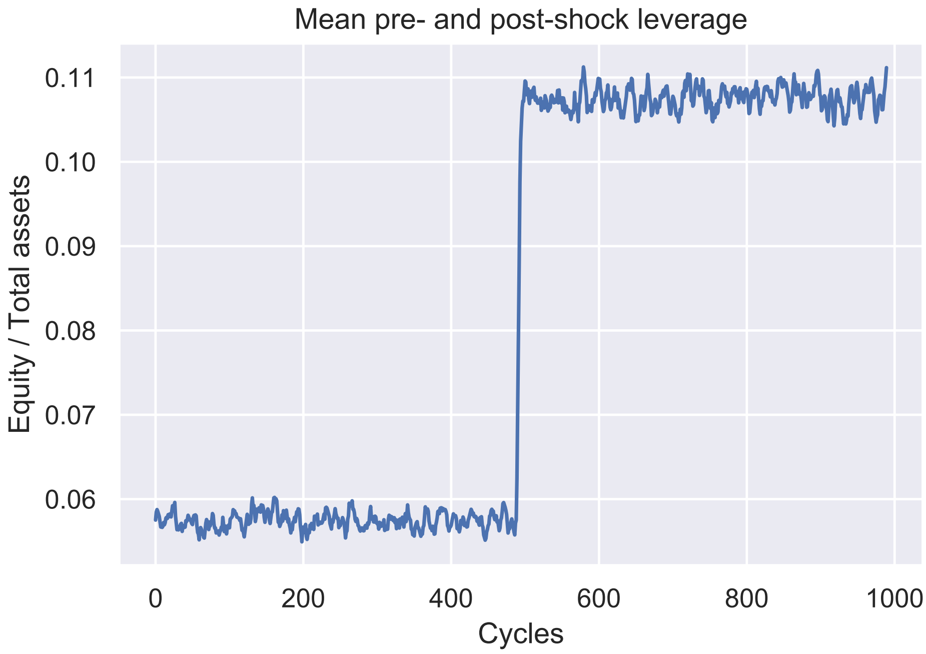

Figure 7 – Pre- and post-shock leverage, original EWA

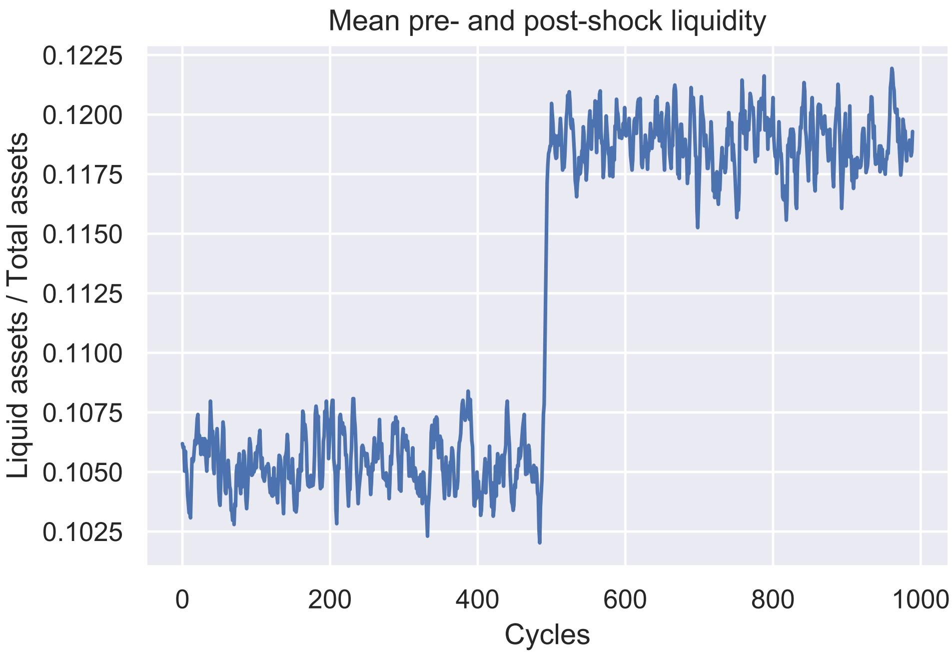

Figure 8 – Pre- and post-shock liquidity, original EWA

Figure 9 – Pre- and post-shock leverage, shock over multiple cycles

in Table 11 and Figures 9 and 10. From those results, it is possible to infer that a slower shock curbs banks' tendency to overreact.

Table 11 – Leverage and liquidity shock reactions when shock is spread out over multiple cycles

Metric	Value (leverage)	Value (liquidity)
Pre-shock steady state	0.0571	0.1021
Post-shock steady state	0.1072	0.1209
Post-shock volatility	0.0027	0.0027
Overreaction deviation	0.0033	0.0020
Deviation/shift	6.62%	10.39%
Deviation/volatility	1.2376	0.7172

3.5 Discussion

The results obtained by means of this chapter's experiments show banks overreacting to shocks. The magnitude of this overreaction depends on shock speed and information acquisition dynamics. This points to market participants' behavior observed during banking crises being consistent with the main features of the Self-tuning EWA learning model. These initial results on market participants' overreaction signal the relevance of continuing development of banking models using learning - in particular, to try to discover if there are crisis scenarios where this overeaction leads to unstable behavior.

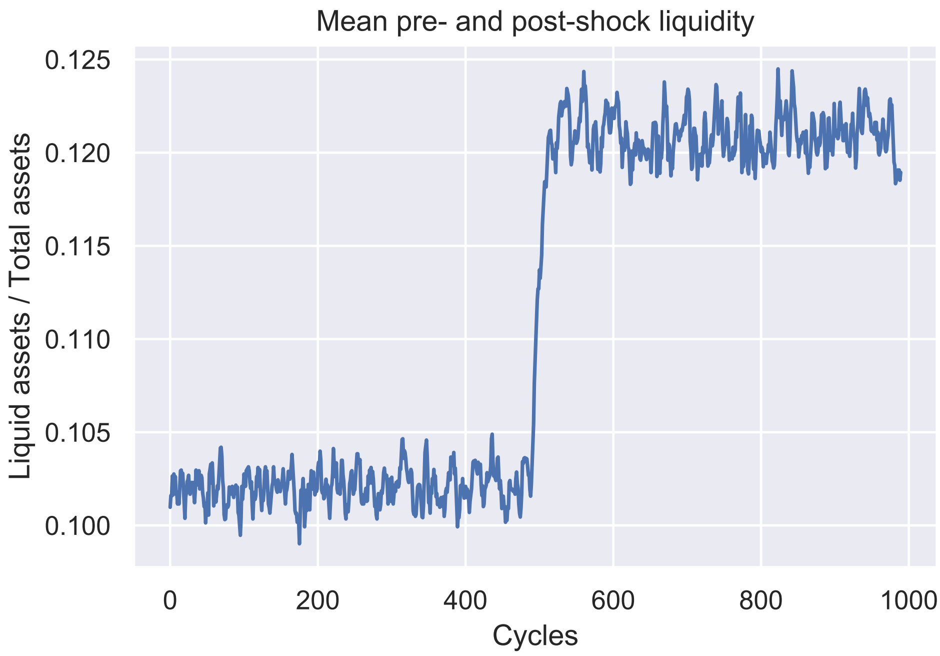

Figure 10 – Pre- and post-shock liquidity, shock over multiple cycles

The implications of overreaction for banking crisis modeling can be significant. If banks overreact to crises by trying to decrease their leverage or increasing liquid asset holdings, it makes them more likely to try to sell risky assets simultaneously. This would depress those assets' prices, increasing the probability of a fire sale, triggering a liquidity spiral (Brunnermeier and Pedersen, 2008).

Going forward, policymakers can now have at their disposal another tool for performing structural policy analysis in financial crises. Its use in structural policy analysis for banking crises is a promising application of this class of models. The overreaction dynamics present in these simulations can complement the results obtained by more traditional structural models that rely on equilibrium and forward-looking expectation formation.

Finally, an important opportunity for improvement is to modify the simulation to encompass multiple opportunities for bank runs and bankruptcies within a single cycle. This modification would enable more sophisticated regulatory analysis, widening the range of problems to which the banking model can be applied. Relevant examples include time-varying capital requirements, timely bank resolution and stress testing.

Conclusion

The usefulness of the results and tools featured in this dissertation extends beyond policymaking. Bank managers can use results as input to improve their bank's crisis management. They will also be able to better prepare their banks for the effects of other banks' potential distress, improving their own capacity to infer such events' possible impacts. Market participants, such as depositors and investors can also benefit from simulating crisis scenarios, using the results as input for more reliable risk pricing and the exercise of market discipline upon banks.

This dissertation also features methodological contributions. The simulation framework can also be used to perform structural analysis of future regulatory modification proposals, regarding issues such as capital requirements, liquidity requirements and deposit insurance, in a way which encompasses the potential impact caused by crisis episodes.

Simulation frameworks such as the one featured in this dissertation can be a useful tool for performing sensitivity analyses and reverse stress tests. Reverse stress testing, in particular, is a risk management activity that has strong synergies with crisis management and recovery planning. Authorities suggest that reverse stress tests "[. . .] *can be seen as a starting point for developing scenarios to test the effectiveness of the menu of recovery options* [. . .]" (FSB, 2012, p. 10).

Additionally, the modeling of bank crises such as done in this dissertation can contribute to improve the explanatory power of stress tests. A possible change in practice could be the substitution of the metric used to indicate individual or systemic stability, from a fixed capital target determined ex ante to the probability of triggering recovery or resolution procedures. A fixed capital target can lead to a false sense of security if set too low, or induce needless panic if set too high. Moreover, the same capital ratio can represent varying degrees of individual or systemic distress depending on why and how capital reserves were depleted.

Furthermore, it could prove fruitful to incorporate features from other EWA-based bank simulations, notably (Barroso et al., 2016). The introduction of an intelligent central bank choosing policy variables (for example, minimum capital requirements) could help to shed light on their optimal levels - possibly influencing the overreaction dynamics. Modifying depositors to behave intelligently by endogeneizing their decision of whether to run could impact bankins system stability and should be a direction for further study.

Finally, porting this framework's source code implementation from C++ to other programming languages would increase chances for collaboration within the academic community. Python is the most likely target language for porting, given its gentle learning curve and recent growth among economics professionals. Silva (2018) presents a more in-depth discussion the regarding pros and cons of different programming languages for developing multi-agent simulations in economics, with an emphasis on Python.

Bibliography

ALLEN, F.; GALE, D. Optimal financial crises. *The Journal of Finance*, Wiley Online Library, v. 53, n. 4, p. 1245–1284, 1998. Cited on page 19.

ALLEN, F.; GALE, D. Financial contagion. *Journal of political economy*, The University of Chicago Press, v. 108, n. 1, p. 1–33, 2000. Cited on page 19.

BARROSO, R. V. *Modelo Dinâmico Computacional de Rede de Bancos*. Master's Thesis — Universidade de Brasília, 2011. Cited on page 33.

BARROSO, R. V. *Avaliação de Políticas Regulatórias e de Estrutura do Mercado Financeiro em um Modelo Dinâmico de Sistema Bancário com Aprendizado*. Doctoral Dissertation — Universidade de Brasília, 2014. Cited on page 33.

BARROSO, R. V. et al. Interbanknetwork and regulation policies: an analysis through agent-based simulations with adaptive learning. *Journal of Network Theory in Finance*, Incisive Media, v. 2, n. 4, p. 53–86, 2016. Cited 8 times on pages 20, 33, 37, 46, 47, 50, 62, and 67.

BASEL COMMITTEE ON BANKING SUPERVISION. *Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework – Comprehensive Version*. 2006. Available at: <https://www.bis.org/publ/bcbs128.pdf>. Cited on page 44.

BASEL COMMITTEE ON BANKING SUPERVISION. *Guidance for national authorities operating the countercyclical capital buffer*. 2010. Available at: <https://www.bis.org/publ/bcbs187.pdf>. Cited on page 19.

BASEL COMMITTEE ON BANKING SUPERVISION. *Basel III: A global regulatory framework for more resilient banks and banking systems*. 2011. Available at: <https://www.bis.org/publ/bcbs189.pdf>. Cited 2 times on pages 19 and 44.

BASEL COMMITTEE ON BANKING SUPERVISION. *Basel III: The Liquidity Coverage Ratio and liquidity risk monitoring tools*. 2013. Available at: <https://www.bis.org/publ/bcbs238.pdf>. Cited 3 times on pages 19, 47, and 60.

BASEL COMMITTEE ON BANKING SUPERVISION. *Global systemically important banks: updated assessment methodology and the higher loss absorbency requirement*. 2013. Available at: <https://www.bis.org/publ/bcbs255.pdf>. Cited on page 19.

BASEL COMMITTEE ON BANKING SUPERVISION. *Basel III leverage ratio framework and disclosure requirements*. 2014. Available at: <https://www.bis.org/publ/bcbs270.pdf>. Cited 2 times on pages 19 and 50.

BASEL COMMITTEE ON BANKING SUPERVISION. *Basel III: the net stable funding ratio*. 2014. Available at: <https://www.bis.org/publ/bcbs295.pdf>. Cited 2 times on pages 19 and 44.

BATTISTON, S. et al. Complexity theory and financial regulation. *Science*, American Association for the Advancement of Science, v. 351, n. 6275, p. 818–819, 2016. Cited on page 29.

BENNETT, R. L.; UNAL, H. The effects of resolution methods and industry stress on the loss on assets from bank failures. *Journal of Financial Stability*, Elsevier, v. 15, p. 18–31, 2014. Cited on page 47.

BENNETT, R. L.; UNAL, H. Understanding the components of bank failure resolution costs. *Financial Markets, Institutions & Instruments*, Wiley Online Library, v. 24, n. 5, p. 349–389, 2015. Cited on page 47.

BERRY, D. A.; FRISTEDT, B. Bandit problems: sequential allocation of experiments. *Monographs on statistics and applied probability*, Chapman and Hall, London, v. 5, p. 71–87, 1985. Cited on page 23.

BORRILL, P. L.; TESFATSION, L. Agent-based modeling: the right mathematics for the social sciences? *The Elgar companion to recent economic methodology*, Edward Elgar Publishing, v. 228, 2011. Cited on page 24.

BRUNNERMEIER, M. K.; PEDERSEN, L. H. Market liquidity and funding liquidity. *The review of financial studies*, Society for Financial Studies, v. 22, n. 6, p. 2201–2238, 2008. Cited on page 65.

CAMERER, C.; HO, T. H. Experience-weighted attraction learning in normal form games. *Econometrica*, Wiley Online Library, v. 67, n. 4, p. 827–874, 1999. Cited 2 times on pages 24 and 53.

CAMERER, C. F. *Behavioral game theory: Experiments in strategic interaction*. Princeton, NJ: Princeton University Press, 2003. Cited on page 24.

CAMERER, C. F.; HO, T.-H. Behavioral game theory experiments and modeling. In: *Handbook of game theory with economic applications*. [S.l.]: Elsevier, 2014. v. 4, p. 517–573. Cited 2 times on pages 29 and 32.

CAMERER, C. F.; HO, T.-H.; CHONG, J.-K. Sophisticated experience-weighted attraction learning and strategic teaching in repeated games. *Journal of Economic theory*, Elsevier, v. 104, n. 1, p. 137–188, 2002. Cited on page 33.

CAMERER, C. F.; HO, T.-H.; CHONG, J.-K. A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, MIT Press, v. 119, n. 3, p. 861–898, 2004. Cited on page 32.

CHEN, T.; GUESTRIN, C. Xgboost: A scalable tree boosting system. In: ACM. *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. [S.l.], 2016. p. 785–794. Cited on page 52.

CHEN, Z.; LUX, T. Estimation of sentiment effects in financial markets: A simulated method of moments approach. *Computational Economics*, Springer, v. 52, n. 3, p. 711–744, 2018. Cited on page 52.

CHONG, J.-K.; HO, T.-H.; CAMERER, C. A generalized cognitive hierarchy model of games. *Games and Economic Behavior*, Elsevier, v. 99, p. 257–274, 2016. Cited on page 32.

DIAMOND, D. W.; DYBVIG, P. H. Bank runs, deposit insurance, and liquidity. *Journal of political economy*, The University of Chicago Press, v. 91, n. 3, p. 401–419, 1983. Cited on page 33.

DORNBUSCH, R. Expectations and exchange rate dynamics. *Journal of political Economy*, The University of Chicago Press, v. 84, n. 6, p. 1161–1176, 1976. Cited on page 30.

DOSI, G. et al. *Rational Heuristics? Expectations and Behaviors in Evolving Economies with Heterogeneous Interacting Agents*. [S.l.], 2017. Cited on page 23.

EISENBACH, T. M. Rollover risk as market discipline: A two-sided inefficiency. *Journal of Financial Economics*, Elsevier, v. 126, n. 2, p. 252–269, 2017. Cited on page 19.

FINANCIAL STABILITY BOARD. *Recovery and Resolution Planning: Making the Key Attributes Requirements Operational*. 2012. Available at: <http://www.fsb.org/wp-content/uploads/r_121102.pdf>. Cited on page 67.

FINANCIAL STABILITY BOARD. *Key Attributes of Effective Resolution Regimes for Financial Institutions*. 2014. Available at: <www.fsb.org/wpcontent/uploads/r_111104cc.pdf>. Cited on page 19.

FREIXAS, X.; ROCHET, J.-C. *Microeconomics of banking*. Cambridge, MA: MIT press, 2008. Cited on page 21.

FUDENBERG, D.; LEVINE, D. K. *The theory of learning in games*. Cambridge, MA: MIT Press, 1998. Cited 3 times on pages 22, 23, and 24.

FUDENBERG, D.; LEVINE, D. K. Learning and equilibrium. *Annuual Review of Economics*, Annual Reviews, v. 1, n. 1, p. 385–420, 2009. Cited on page 22.

FUDENBERG, D.; LEVINE, D. K. Whither game theory? towards a theory of learning in games. *Journal of Economic Perspectives*, v. 30, n. 4, p. 151–70, 2016. Cited on page 23.

GRIMALDI, M. B. et al. Estimating the size and incidence of bank resolution costs for selected banks in oecd countries. *OECD Journal: Financial Market Trends*, Organisation for Economic Cooperation and Development (OECD), v. 2016, n. 1, p. 1–36, 2017. Cited on page 47.

GUERINI, M.; MONETA, A. A method for agent-based models validation. *Journal of Economic Dynamics and Control*, Elsevier, v. 82, p. 125–141, 2017. Cited on page 52.

GUZMAN, M.; HOWITT, P. Learning, expectations, and the financial instability hypothesis. In: *Contemporary Issues in Macroeconomics*. [S.l.]: Springer, 2016. p. 50–60. Cited 2 times on pages 34 and 56.

HARSTAD, R. M.; SELTEN, R. Bounded-rationality models: tasks to become intellectually competitive. *Journal of Economic Literature*, v. 51, n. 2, p. 496–511, 2013. Cited on page 22.

HO, T. H.; CAMERER, C. F.; CHONG, J.-K. Self-tuning experience weighted attraction learning in games. *Journal of Economic Theory*, Elsevier, v. 133, n. 1, p. 177–198, 2007. Cited 11 times on pages 11, 20, 21, 24, 26, 28, 29, 37, 49, 53, and 55.

HOGGARTH, G.; REIS, R.; SAPORTA, V. Costs of banking system instability: some empirical evidence. *Journal of Banking & Finance*, Elsevier, v. 26, n. 5, p. 825–855, 2002. Cited on page 21.

HOWITT, P. Learning, leverage and stability. *International Journal of Computational Economics and Econometrics*, Inderscience Publishers (IEL), v. 7, n. 3, p. 265–279, 2017. Cited 2 times on pages 34 and 56.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. *ISO/IEC 14882:2011 Information technology – Programming languages – C++*. 2011. Available at: <https://www.iso.org/standard/50372.html>. Cited on page 51.

IRANI, R. M.; MEISENZAHL, R. R. Loan sales and bank liquidity management: evidence from a US credit register. *The Review of Financial Studies*, Oxford University Press, v. 30, n. 10, p. 3455–3501, 2017. Cited 2 times on pages 47 and 60.

KINDLEBERGER, C.; ALIBER, R. *Manias, panics, and crashes: a history of financial crises*. London, UK: Palgrave Macmillan, 2015. Cited 2 times on pages 22 and 55.

KIRMAN, A. Learning in agent-based models. *Eastern Economic Journal*, Springer, v. 37, n. 1, p. 20–27, 2011. Cited 2 times on pages 22 and 56.

LAMPERTI, F. An information theoretic criterion for empirical validation of simulation models. *Econometrics and Statistics*, Elsevier, v. 5, p. 83–106, 2018. Cited on page 52.

LAMPERTI, F.; ROVENTINI, A.; SANI, A. Agent-based model calibration using machine learning surrogates. *Journal of Economic Dynamics and Control*, Elsevier, v. 90, p. 366–389, 2018. Cited on page 52.

LIMA, J. I. *Um arcabouco computacional para estudo do setor bancário atraves de modelos baseados em agentes*. Master's Thesis — Universidade de Brasília, 2014. Cited on page 33.

LUCCHETTI, A. H. *Interbank network and regulation policies: an analysis through agent-based simulations with adaptive learning*. Master's Thesis — Universidade de Brasília, 2016. Cited on page 33.

MCKELVEY, R. D.; PALFREY, T. R. Quantal response equilibria for normal form games. *Games and economic behavior*, Elsevier, v. 10, n. 1, p. 6–38, 1995. Cited on page 32.

MINSKY, H. P. *John Maynard Keynes*. New York: Columbia University Press, 1976. Cited on page 33.

MINSKY, H. P. *Stabilizing an Unstable Economy*. New Haven, CT: Yale University Press, 1986. Cited on page 33.

MINSKY, H. P. *The Financial Instability Hypothesis*. Working Paper 74, Levy Institute, 1992. Cited 3 times on pages 33, 56, and 58.

MORRIS, S.; SHIN, H. S. Unique equilibrium in a model of self-fulfilling currency attacks. *American Economic Review*, JSTOR, p. 587–597, 1998. Cited on page 33.

MUTH, J. F. Rational expectations and the theory of price movements. *Econometrica: Journal of the Econometric Society*, JSTOR, p. 315–335, 1961. Cited on page 34.

POUGET, S. Adaptive traders and the design of financial markets. *The Journal of Finance*, Wiley Online Library, v. 62, n. 6, p. 2835–2863, 2007. Cited on page 33.

PYTHON SOFTWARE FOUNDATION. *Python 3.6.4 documentation*. 2018. Available at: <https://docs.python.org>. Cited on page 52.

REINHART, C. M.; ROGOFF, K. S. *This time is different: Eight centuries of financial folly*. Princeton, NJ: Princeton University Press, 2009. Cited 5 times on pages 21, 34, 55, 56, and 58.

REINHART, C. M.; ROGOFF, K. S. Banking crises: an equal opportunity menace. *Journal of Banking & Finance*, Elsevier, v. 37, n. 11, p. 4557–4573, 2013. Cited on page 22.

ROTH, A. E.; EREV, I. Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and economic behavior*, Elsevier, v. 8, n. 1, p. 164–212, 1995. Cited 2 times on pages 23 and 24.

SCHINASI, G. J. *Defining Financial Stability*. Working Paper 187, International Monetary Fund, 2004. Cited on page 19.

SETHI, R. Dynamics of learning and the financial instability hypothesis. *Journal of Economics*, Springer, v. 56, n. 1, p. 39–70, 1992. Cited 3 times on pages 33, 34, and 56.

SILVA, C. M. *Framework colaborativo de simulação computiacional para o estudo de regulação bancária*. Master's Thesis — Universidade de Brasília, 2018. Cited 2 times on pages 52 and 68.

SMITH, E.; SHUBIK, M. Runs, panics and bubbles: Diamond–dybvig and morris–shin reconsidered. *Annals of Finance*, Springer, v. 10, n. 4, p. 603–622, 2014. Cited on page 33.

STAHL, D. *Sophisticated learning and learning sophistication*. 2003. Mimeo. Available at: <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=410921>. Cited on page 33.

STAHL, D. O.; WILSON, P. W. On players′ models of other players: Theory and experimental evidence. *Games and Economic Behavior*, Elsevier, v. 10, n. 1, p. 218–254, 1995. Cited on page 32.

SUTTON, R. S.; BARTO, A. G. *Reinforcement learning: An introduction*. Cambridge, MA: MIT press, 1998. Cited on page 23.

TEMZELIDES, T. Evolution, coordination, and banking panics. *Journal of Monetary Economics*, Elsevier, v. 40, n. 1, p. 163–183, 1997. Cited on page 33.

TIROLE, J. *The theory of corporate finance*. Princeton, NJ: Princeton University Press, 2006. Cited on page 19.

VASICEK, O. The distribution of loan portfolio value. *Risk*, v. 15, n. 12, p. 160–162, 2002. Cited on page 43.