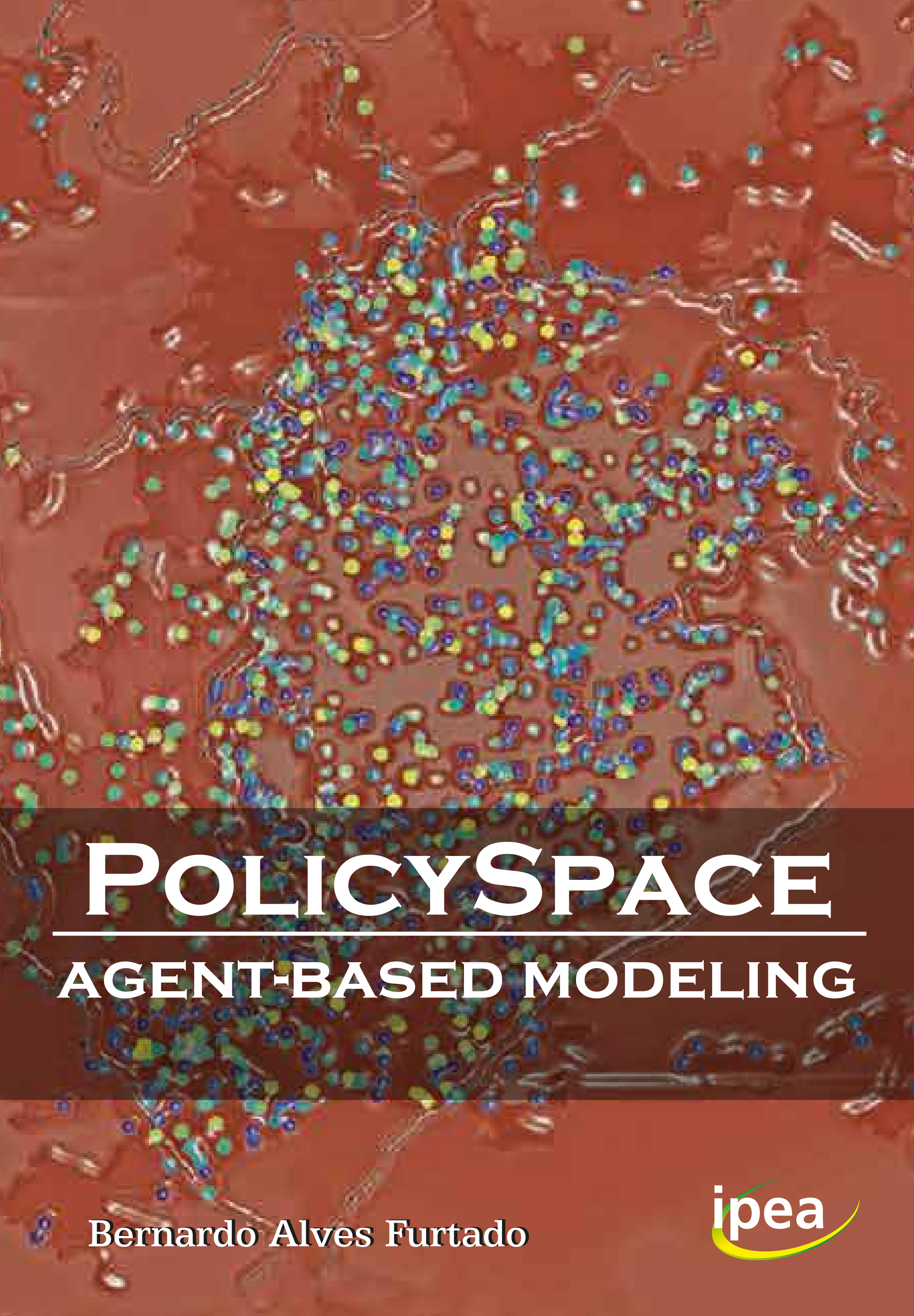


PolicySpace

Agent-Based Modeling

Bernardo Alves Furtado

ipea

The future is not predictable. It is created. Contextually and gradually, with small advances, understandings and adaptations, corrections upon corrections. The future is negotiated – and emerges as an inexorable result of social, institutional and political interactions.

Public policies – as a tenuous concept of planning, a vehicle of societal interests – balance, evaluate, contain visions (predominantly public) that guide, direct, drive optimal social development.

Future is a scenario of uncertainties. A scenario of social complexity in which policy proposals are made by the society, guided by the managers. Public managers, who seek to mediate policies and their multiplicity of effects, typical effects of complex systems, to the maximum extent that circumstances allow.

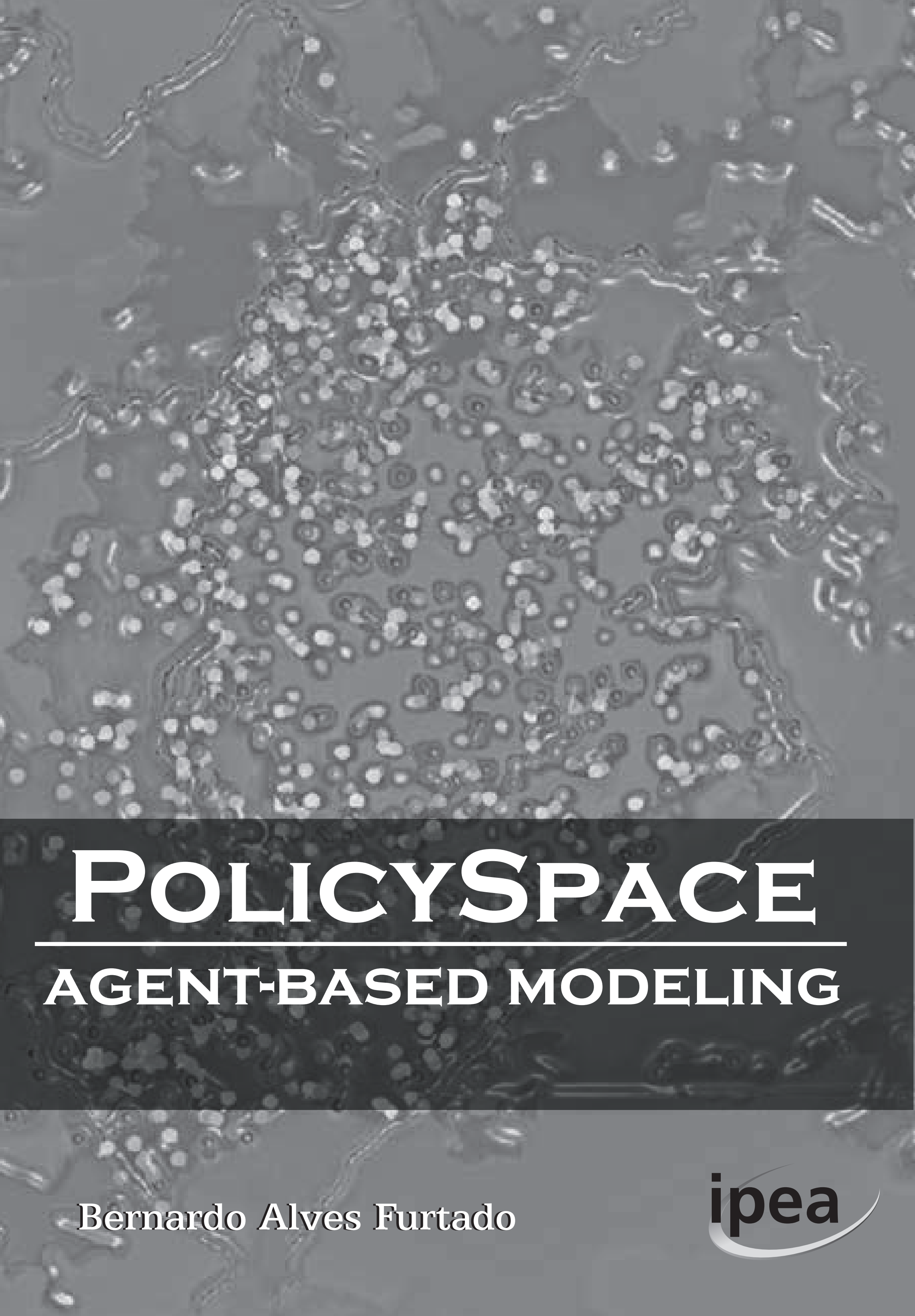
Given this context, the purpose of this book is to offer an additional tool of modeling to society in general and to managers of the *res publica* in particular. Not only conceptually, abstractly – but as a concrete tool, available and adaptable. In fact, PolicySpace is an empirical platform, adjusted for the case of 333 Brazilian municipalities located in the 46 population concentration areas. This set of metropolises is home to most of Brazil's socio-economic strength, but also its greatest challenges.

The core idea of PolicySpace's platform is to allow the analysis of alternatives – many alternatives – for the implementation of *ex-ante* public policies. That is, to anticipate, circumstantially, effects, developments, future results of changes, in the present; to analyze interactions between portions of society and institutions, in space and time.

PolicySpace is an agent-based model, including families, citizens, residences, businesses, markets, taxes, mobility, and municipalities, that allows "what-if" questions. It is an *in silico* laboratory, of extremely low relative cost. Yet, it is flexible, adaptable, that anticipates trajectories and, quantitatively, measures horizontal effects across sectors, places and times.

The book reviews the literature, explains concepts, and describes the methodology. It details the model, its parameters, and the full process. It validates the proposal and illustrates with applications.

Good reading. The future is not predictable – but neither it is dark.



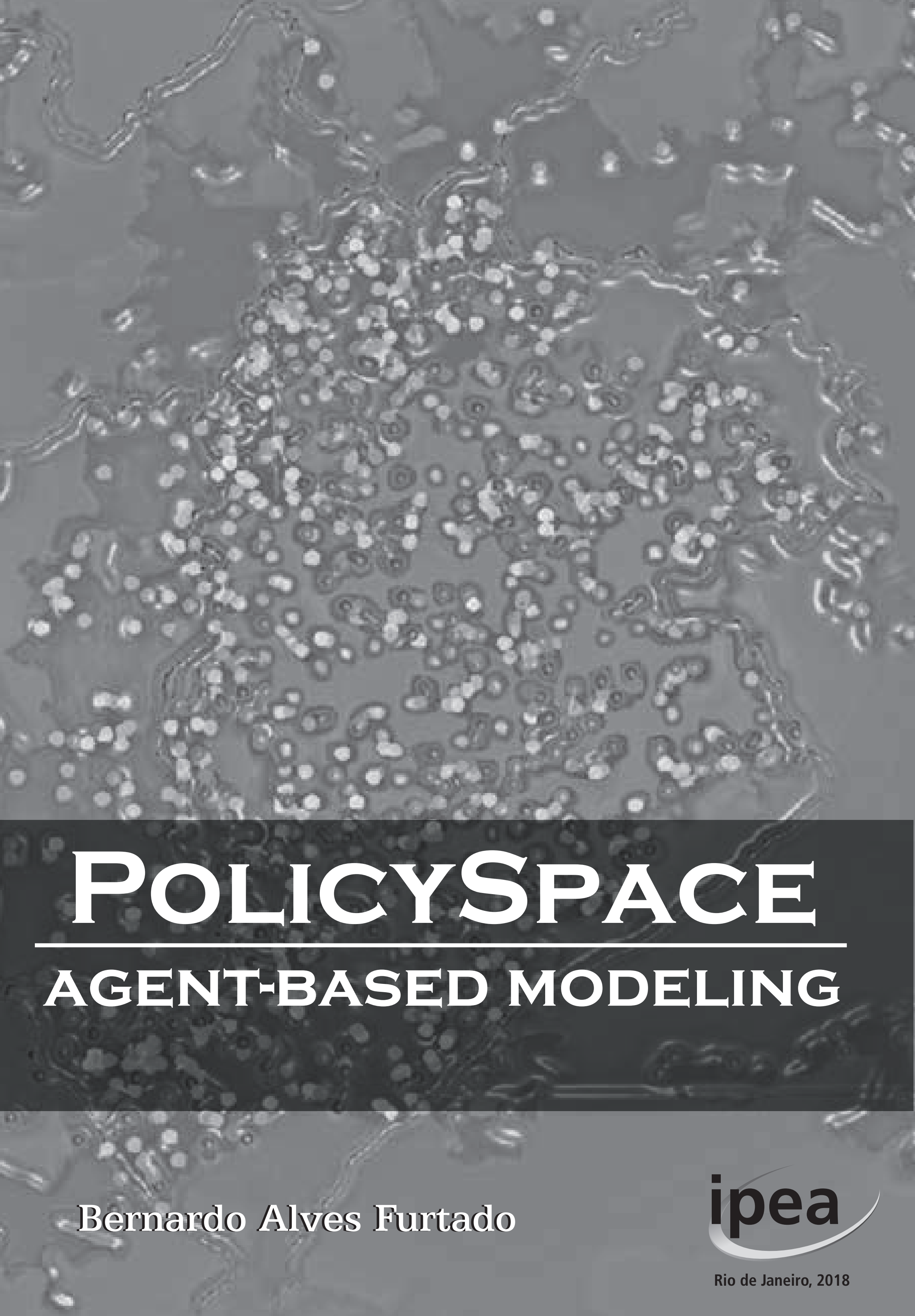
POLICYSPACE

AGENT-BASED MODELING

Bernardo Alves Furtado

ipea

URL: http://www.ipea.gov.br
Ombudsman: http://www.ipea.gov.br/ouvidoria



POLICYSPACE

AGENT-BASED MODELING

Bernardo Alves Furtado

ipea

Rio de Janeiro, 2018

Furtado, Bernardo Alves

Policy space : agent based modeling / Bernardo Alves Furtado. – Rio de Janeiro: Ipea, 2018.

121 p. : ill., graphs., maps color.
Includes bibliographical references.
ISBN 978-85-7811-326-1

1. Public policy. 2. Real estate market. 3. Research method. 4. Complex system. 5. Metropolitan region. 6. Brazil. I. Title.

CDD 003

CONTENTS

INTRODUCTION[1]

Public policies are the result of multiple, accumulated decisions. Public policies also generate heterogeneously consequences in time and space over parts of society. These temporal, spatial and sectoral effects combined reflect the intrinsic complexity of formulating, monitoring and eventually evaluating public policy.

Take a public policy, any one, and consider this. Most likely, the effects of implementing such policy will be specific to a given spatial boundary, either a municipality or a state jurisdiction, for instance. In addition, that same policy will be constrained by a unique timeframe, be it effective immediately or gradually over five or ten years. Further, a simple policy may affect distinctively, different sectors of the economy – services, industry; or different portions of society – the old or the young, the self-employed, the employee, or the employers, the rich or the poor. Each of the groups affected by the policy at hand would react to a greater or lesser degree, with more or less intensity and with distinct levels of capability over time.

Indeed, planning, management, and evaluation of public policies are issues that require detailed analysis, solid knowledge experience, and consideration of different perspectives.

Given that, it is fair to say that the analysis of public policies share at least two points in common with complex systems' approach: the focus on the agents involved and on the interaction among themselves and the environment.

Who are the agents in such context? Agents are the citizens, employees, institutions, social groups, all those who decide, manage, choose, affect, and are affected by policies. Interaction, in turn, is the dynamic and spatial clashes among agents. The actions and reactions that prevent that a given application of a policy is implemented linearly, as originally planned. There are resistances, suggestions, innovations, adaptations, and vetoes of the agents on the policies. In fact, interaction among agents is a central concept of complex systems.

1. The author thanks the reviewers Fábio Monteiro Vaz and Antônio Glauter Teófilo Rocha, and the comments of Bruno Cesar Pino Oliveira de Araújo, Dea Guerra Fioravante, Pedro Herculano Ferreira de Souza, Marcelo Medeiros Coelho de Souza, Fabiano Pompermayer and Rafael Henrique Moraes Pereira. We would also like to thank Cláudia Regina Mota Furtado for an English reading and suggestions. Errors and omissions are the sole responsibility of the author. This book is a translation of the Portuguese edition.

Therefore, it is reasonable to assume that the methodologies suitable for the analysis of complex systems may also benefit public policies studies. Agent-based modeling is just that, a complex systems' tool that contributes to the understanding of social interaction and its interfaces with policies.

Understanding that agent-based modeling allows *ex-ante* evaluation of public policies with very low relative costs may help policymakers anticipate insights and results before the actual policing effort of detailing, implementing, and even failing.

Usually, what we observe in practice, at least in Brazil, is that only occasionally proposed policies go through some kind of evaluation. Moreover, when they are evaluated, such analyses occur only after many years, financial resources, personnel, and mobilization of public administration have been wasted.

The diagnosis that public policies can benefit from other methodologies is not new. Such a notion is explored in a few books that seek to explicitly link the approach of complex systems and their modeling tools to the "art"[2] of politics.

Colander and Kupers (2014) make a first exercise focused on economics, arguing the inexorability of adopting concepts and practices beyond the *mainstream* framework in the economic arena. At the interface between the other social sciences, we have to highlight the efforts of Johnson *et al.* (2017), synthesized in the chapter of Edmonds (2017) and the contributions of Helbing (2012).

The work of Heppenstall *et al.* (2012) falls within the scope of the social sciences, but considering a geographic perspective. Finally, Geyer and Cairney (2015) bring 27 chapters on various topics in an attempt to coalesce the fields of political science and complex systems using arguments from international trade to education, crime and local government.

We list some examples of real cases of public policy specifically to *agent-based modeling* (ABM). ABMs have already been applied to real-time pricing (Kuhnlenz *et al.*, 2017). The authors tested if there were social beneficial effects in the energy market when the interaction among customers and distributors changed.

In transport studies, activity models allow the generation of demand and mobility analysis with high level of detail (Erath, Löchl and Axhausen, 2009; Horni, Nagel and Axhausen, 2016). In crime research, a case study associates the incidence of burglaries with urban renewal interventions using an agent-based model (Malleson, 2012). Finally, in political sciences, the dynamics of party influence and political positions are also examined with the use of ABMs (Laver and Sergenti, 2012).

2. More than one author refers to political practice as "art" in order to distinguish politics from "scientific practices".

The collections of cited studies are able to demonstrate both the policy side and the use of informative ABM framework as modeling alternatives. They also show that agent-based models can contribute to the anticipated effects of political intervention in society.

This book contribution is to make available a full agent-based modeling platform: *PolicySpace*. A platform that is modular and operational, validated and adjusted to the empirical analysis of public policies for the Brazilian case.[3, 4] *PolicySpace* is explicitly spatial, containing empirical data from municipalities and Brazilian metropolitan regions that allow a high level of interactivity between the model and the modeler. In fact, the modeler can change any of the *PolicySpace* parameters; choose modeled regions of interest, and adjust markets or procedures and rules according to his or her research needs. Automatically, *PolicySpace* generates and provides output data in text format containing general and regional indicators, data on citizens, firms, residences, as well as graphics that describe trajectories and behavior of the functional areas (PCAs)[5] and of each of the municipalities individually.

Another functionality provided by *PolicySpace* is the chance to test tax rates, observe impacts of changes on the model's indicators and evaluate the evolution of the quality of life of the municipalities in each of the tested scenarios.

In addition to the availability of the platform itself, the book makes a literature review conceptualizing and detailing ABM as a methodology. Thus, the goal of the book is to introduce, detail, validate and demonstrate applications of the proposed agent-based modeling platform - *PolicySpace* – that enables *ex-ante* public policy evaluation processes.

Concisely, the book presents a model based on operational and modular agents; justified, verified and validated by previous literature. In addition, brief results of selected applications are presented, including: *i)* impact of propensity to consume and productivity on macroeconomic variables; *ii)* firm's discrimination analyses by place of residence of the employees; and *iii)* how changes in collecting and distributing fiscal resources to municipalities in metropolitan areas affect individual municipalities.

Specifically, the model investigates alternative configurations of the distribution of taxes among the municipalities of a given PCA. Finally, empirical results for

3. Although implemented for the Brazilian case, simple changes in data entry (and, perhaps, in tax mechanisms, for example) can be adapted for the analysis of other countries.

4. *PolicySpace* open source code is available at: <github.com/BAFurtado/PolicySpace>.

5. PCAs are Areas of Concentrated Population, or functional areas, as defined by the Organisation for Economic Co-operation and Development (OECD). In Brazil, PCAs are defined by the Brazilian Institute of Geography and Statistics (IBGE, 2015, p.19).

various metropolitan regions of the country are presented, and multiple possibilities for expansion and adaptation of *PolicySpace* are detailed.

Initial results reinforce the validity and relevance of the Municipal Participation Fund (MPF) for better equalization of funds collected at the metropolitan level. In addition, there is a marginal gain for overall quality of life when the distribution of fiscal resources occurs as if the municipalities were one single entity. It should be noted that this gain is exclusively analytical and does not include the reduction of costs and the increase of efficiency that would be observed for the case of a single administration.

A further result of the simulations in *PolicySpace* is a brief analysis of the real estate market. It is clear from the simulation that when the real estate market is more dynamic the whole economy intensifies significantly. In such case, the indicators suggest a reduction in inequality, an increase in general production and in the profits, and capital of firms with higher household consumption, albeit with a reduction in families' saving levels.

Besides this introduction, chapter 1 conceptualizes ABMs, the design of the agents and their interaction with the environment. It also presents the arguments that favor the use of the methodology, emphasizing its adaptability and flexibility, but also stating some criticisms and their responses.

Chapter 2 specifies *PolicySpace* following the ODD protocol (*Overview, Design Concepts and Details*), recommended by the literature of ABMs. In doing so, it establishes the specific purpose of *PolicySpace*, that is, to offer a modeling platform empirically adequate to the study of the Brazilian metropolises, with emphasis on spatial and associated tax issues. The chapter then details the initial design of agents and the data used to construct them; explains the processes, the methods by which agents are updated discreetly in time and space. Finally, the most relevant algorithms are detailed in the submodule section. At the end of the chapter, the reader is expected to gain complete understanding of the model so that it may be replicated.

Chapter 3 presents *PolicySpace* validation. It compares simulated data and actual data and advances on the sensitivity analysis discussing the possibilities and scope of the model. Further, the chapter describes the intuition of the concepts of validation and sensitivity analysis.

We validate the model in three successive steps:

- initially, we illustrate the performance of the model with a comparison of macroeconomic behaviors which remain reasonably close to observed values;

- then, we show that the evolution of taxes used in *PolicySpace* remains within values and distributions that are compatible with the literature and those officially described; and
- finally, we test the parameters and rules to confirm the validity of the model.

Chapter 4 expands the sensitivity analysis to broaden the understanding of the possibilities of *PolicySpace* and its applications. Next, the chapter tests the parameters of productivity, propensity to consume, labor market and rules of price setting by the firms. It also tests general parameters of the simulation for each of the PCAs, such as average family size, percentage population sample in the model and residential vacancy.

Chapters 5 and 6 discuss two central applications of *PolicySpace*: the municipal fiscal analysis in the metropolitan scope and real estate markets. Chapter 7 illustrates simulation results for other Brazilian PCAs. Chapter 8 concludes the book with the summary of the contributions and possibilities of expansion of the presented platform.

Overall, the book details concepts and methodology (chapters 1 and 2), validates and guarantees the robustness of the model (chapters 3 and 4), and presents initial applications (chapters 5, 6 and 7). Chapter 8 indicates ways to expand the proposal.

Finally, it is worth mentioning that *PolicySpace* – albeit explicitly spatial and public policy oriented – is a tribute to *SugarScape*, the seminal model also developed in a book format, by Epstein and Axtell (1996). These scholars were able to systematize and make concrete the analysis of various topics of the social sciences through agent-based modeling in a relatively simple and affordable way over twenty years ago.

REFERENCES

COLANDER, D.; KUPERS, R. **Complexity and the art of public policy**: solving society's problems from the bottom up. Princeton: Princeton University Press, 2014.

EDMONDS, B. The room around the elephant: tackling context-dependency in the social sciences. *In*: JOHNSON, J. *et al.* (Ed.). **Non-equilibrium social science and policy**. New York: Springer Open, 2017. p. 195-208.

EPSTEIN, J. M.; AXTELL, R. **Growing artificial societies**: social science from the bottom up. Washington: Brookings Institution Press; Cambridge: MIT Press, 1996.

ERATH, A.; LÖCHL, M.; AXHAUSEN, K. W. Graph-theoretical analysis of the Swiss road and railway networks over time. **Networks and Spatial Economics**, v. 9, n. 3, p. 379-400, 2009.

GEYER, R.; CAIRNEY, P. **Handbook on complexity and public policy**. [s. l.]: Edward Elgar Publishing, 2015.

HELBING, D. **Social self-organization**: agent-based simulations and experiments to study emergent social behavior. New York: Springer, 2012.

HEPPENSTALL, A. J. *et al.* **Agent-based models of geographical systems**. New York: Springer, 2012. p. 746.

HORNI, A.; NAGEL, K.; AXHAUSEN, K. **The Multi-Agent Transport Simulation MATSim**. London: Ubiquity Press, 2016.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Ministério do Planejamento. **Arranjos populacionais e concentrações urbanas do Brasil**. Rio de Janeiro: IBGE, 2015.

JOHNSON, J. *et al.* **Non-equilibrium social science and policy**. New York: Springer, 2017.

KUHNLENZ, F. *et al.* **Implementing flexible demand**: real-time price vs. market integration, 2017. Available at: <https://arxiv.org/abs/1709.02667>. Accessed: Dec. 2017.

LAVER, M.; SERGENTI, E. **Party competition**: an agent-based model. Princeton: Princeton University Press, 2012.

MALLESON, N. Using agent-based models to simulate crime. *In*: HEPPENSTALL, A. *et al.* **Agent-based models of geographical systems**. New York: Springer, 2012. p. 411-432.

CHAPTER 1

METHODOLOGY

Before conceptualizing agent-based modeling specifically, it is worth emphasizing the notion of models in general. We consider model as a tool that contributes to the understanding of the modeled object, the "target" of the study, and thus allows inferences to be made on such target. Moreover, "different models usually optimize the understanding of different aspects of the target" (Magnani e Bertolotti, 2017, p. 15).

Models are artifacts, instruments that assist in the understanding of the phenomenon. Hence, the main objective of modeling in essence is to grasp further knowledge about the phenomenon.

This view is in tune with Johnson *et al.* (2017) and Page (2010) who argue that the analysis of public policy through agent-based modeling is complementary to studies of other disciplines in the social sciences and economics, and benefits from the notion that multiple models enlighten, so to speak, diverse aspects of reality and, consequently, of public policy.

1 AGENT-BASED MODELING – ABM

Agent-based modeling, or simply ABM, is a computer implementation of an artificial environment that contains agents that interact in time and space. Exactly that.

Of course, in a formal way, Epstein and Axtell (1996) characterize ABM[1] as a temporally discrete dynamic system, described by simultaneous generic equations.[2]

$$\begin{aligned} Agents_{t+1} &= f\,(Agents_t,\ Environment_t) \\ Environment_{t+1} &= g\,(Agents_t,\ Environment_t) \qquad (1) \end{aligned}$$

Therefore ABMs can be described as the discrete evolution in time and space of agents and environments through built-in rules in the processing functions f(.) and g (.).

1. ABM may refer to agent-based modeling, as a methodology, but it may also refer to a model. We use the term interchangeably throughout the book.

2. Another equivalent formula is available in Batty (2012, p. 40, 2.16).

Note that even this simple formalization implies that agents and their environment interact in time *t* so that the next moment *t* + 1 is a function of the agents and environment in the previous time. Intuitively, computational modeling in general and ABM in particular is the sole application of a deterministic known rule, formally described, to a set of input data that are "animated" to obtain results (Galan *et al.*, 2009).

The notation of equation 1 allows us to capture the *interactions* between agents, the evolution of this interaction between agents and the environment, or simply the autonomous transformations of agents and the environment.

Agents may be fixed or mobile, are transformed through autonomous functions or interactions with other agents and the environment. Agents' own intrinsic properties or relational conditionalities such as physical proximity or structural relationship may trigger changes and adaptations. These relationships in turn may be networks of connections, of friendships or a sense of belonging, for example.

The construction of an agent-based model with emphasis on interactions allows for infinite possibilities. Imagine the following example: at a given time *t*, at an specific location *s*, two specific agents x_i and x_j have their own methods of action and interact in a strictly deterministic way. We can assume that the results of the interaction between them may be distinct for the same agents x_i and x_j when they are at *t* + 1 or at *s* + 1. Further, other patterns of interaction may happen when there a third agent, say x_k is present. Even other results may be expected whether the environment *s* had a given probabilistic triggering attribute *p* for the same original time *t*.

In order to deepen the understanding of ABM, we present below some alternative definitions. Bilge defines ABM similarly to the one proposed in this book:

> the ABM approach is a 'bottom-up' modelling technique, sometimes resulting in unexpected, so-called 'emergent' behaviour. ABM is suitable for real world problems where there are a number of autonomous, heterogeneous and interacting agents. Often the development of an Agent Based Simulation (ABS) is in itself useful and educational where the end product does not necessarily predict the future, but is a virtual laboratory to experiment with ideas, and to test possible real-life scenarios in a safe environment" (Bilge , 2015, p.417).

In a somewhat simpler way, Hamill and Gilbert propose: "[ABM is] a computer program that creates an artificial world of heterogeneous agents and enables investigation into how interactions between these agents, and between agents and other factors such as time and space, add up to form the patterns seen in the real world" (Hamill and Gilbert, 2016, p. 4).

Following a similar pattern, Abdou, Hamill and Gilbert define ABM as "Agent-based modelling is a computational method that enables a researcher to create, analyze, and experiment with models composed of agents that interact within an environment" (Abdou, Hamill and Gilbert, 2012, p. 141).

Finally, Wilensky and Rand (2015) advocate that ABMs are especially suitable for non-deterministic studies, or for that phenomenon whose mechanisms are distributed or without effective central control.

ABMs are also referred as multi-agents modeling in computer science and especially within Agent-Based Computational Economics (ACE), whose comprehensive reference is the *site* maintained by Leigh Tesfatsion.[3]

In the context of ABM, agents have several attributes. They are heterogeneous and autonomous, interacting locally in explicit space. They act on (usually) simple rules and local information, so that they are not (necessarily) global optimizers (Epstein, 2006). Hamill and Gilbert (2016) prefer to qualify the agents as perceptive, that is, that they use sensors to perceive other agents, the environment and even their own *performance*. They may communicate, follow instructions, make decisions, and remember past decisions.

Crooks and Heppenstall (2012) add to these attributes the notion of proactive agents who work towards specific goals. Therefore, agents act, react, learn and adapt themselves considering both time and space.

As demonstrated by the generic form of ABM implementation (and equation 1), ABM methods are flexible and allow the analysis of a wide range of phenomenon and research objects. In fact, this methodological flexibility and permissiveness contribute to the fact that the field of ABM's application is still very open and lacking protocols and *benchmarking*. Such variety prevents the scientific community from achieving greater clarity in the advances and contributions made by each given model (Crooks and Heppenstall, 2012; Davidsson, Verhagen and Klügl, 2017).

A first attempt at better description and comparability is gaining adepts within the community is the proposal of Grimm *et al.* (2006, 2010). Denominated ODD (Overview, Design concepts and Details), the protocol standardizes the description of the proposed model, as well as its initial states, assumptions, limitations and characteristics, including the suggestion that the author should make available the source code and databases, so that the reader or user can replicate the exercise done. For *PolicySpace*, the computational algorithm is available as a GitHub repository at <GitHub.com/BAFurtado/PolicySpace>.

3. Available at: <http://www2.econ.iastate.edu/tesfatsi/ace.htm>.

In fact, the book Agent-based Modeling in Economics (Hamill and Gilbert, 2016), for example, provides all the code of the twenty models developed on the internet. This guarantees its replicability as well as the absence of a black box. That is, given that the modeler has invested the necessary effort, he or she should be able to check and test the full proposal.

1.1 Advantages of modeling and some criticisms

As a complementary methodological tool, ABM is relevant when the phenomenon under study possesses three central attributes: *heterogeneity*, *dynamics* and *interaction* (Gilbert and Hamill, 2016). Heterogeneity as an attribute of the phenomenon means that looking at the aggregate measures, the sum of individual parts simply, is not enough to understand the object (Anderson, 1972).

The dynamic attribute is relevant when the static comparison is not sufficient for adequate understanding of the phenomenon. Note that the dynamics, or reaction of other agents over time, is especially relevant to the case of public policies, in which agents (citizens, institutions) react actively to the new context of a given policy proposal. Dynamics may also be relevant from a spatial point of view. If a distinct location (at $s + 1$) transforms the agent so that its reaction or behavior is different from another location, then ABMs can contribute with a better understanding of the problem.

Interaction is characterized when: *i)* there is influence, power, networks or contagion among agents; *ii)* decisions are modified conditionally on decisions from neighbors (actual or from a network); *iii)* there are imitation or copy processes; and *iv)* agents act differently from what previously expected in a no-interaction environment.

Fundamentally, ABM as a methodology may contribute greatly when a research investigation contains objects that are composed of elements in which the heterogeneity, the dynamics and the interaction among the parts are central.

Another fundamental contribution of ABM – derived from the analysis of heterogeneous elements that interact dynamically and are consistent with the approach of complex systems – is that *we cannot expect linearity of the results of the policies*. This is very relevant in a public policy context. Given that agents react, learn, adapt and suffer influences from its neighbors and from the ever-changing context and environment they are in (Edmonds, 2017), the effects of public policy are unknown *a priori*. In fact, Geyer and Cairney (2015) illustrates that public policy evaluations revealed that in several public interventions relations of cause and effect were not clear.

This inability to forecast the consequences of public (or private) action, even with complete knowledge of the behavior of the system and agents (Squazzoni, 2014), implies the inability to control the process. Moreover, it implies that the measurement of policy results, i.e., the effective fulfillment of program and policy goals, is compromised from the start (Mueller, 2015).

The goal of policymaking presupposes the ability to control the process, which is false. The act of setting public policy goals should consider the fact that agents are reactive, have emerging behaviors, and are influenced by the environment, space, and time.

Effective goals, in order to function, should consider the development of the policy implementation as a process, with a series of marginal, subtle, adaptive, bottom-up adjustments.

ABMs cannot "predict" effects and results of public policies. However, its central feature of "artificial computer program" allows simulating alternatives, experimenting, making marginal amendments, changing parameters, rules, reactions, and interactions. Moreover, ABMs do all that in an inexpensive, relatively simple and agile way, so that scenarios can be tested (Furtado, Sakowski and Tóvolli, 2015). Thus, we argue that *insights* of the effects of policies can be obtained *in silico* before the actual execution of the policy itself.

Epstein rejects three central criticisms of ABM:

> as to the core indictment that agent models are non-mathematical, non-deductive, and ad hoc, the first two are false, and the third, I argue, is unimportant. Generative explanation is mathematical in principle; recursive functions could be provided. Ipso facto, generative explanation is deductive. Granted, agent models typically quantify over smaller sets than rational choice models and, as such, are less general. But, in many cases, they are more responsive to data… (Epstein, 2006, p.1602).

Crooks and Heppenstall (2012) discuss some contrasting aspects between ABMs and traditional models (TM). ABMs offer multiple futures, sometimes associated with probabilities, whereas TM predict a single future, or rather a trajectory. According to Crooks and Heppenstall, ABMs emphasize the mechanisms used to achieve the results (which may be non-unique). TM do not focus on explanations, but rather on results and correlations.

ABMs usually use more parameters than TM do. ABMs treat space (and time) explicitly, whereas TM usually do not include space, or do so in a reduced way. While in ABMs the environment is built (endogenously), TM present a given, exogenous environment. Finally, according to the aforementioned authors, in ABMs the modeler learns from the model, while in TM the researcher only reacts to the results.

Boero *et al.* (2015) add some other attributes to ABM, besides interactions, dynamics and heterogeneity. First, the modeler may choose the degree of realism for each model, according to the availability of data and knowledge about the object. Second, ABMs are better suited to identify and test the causality of phenomenon.

A third advantage pointed out by the same authors (Boero *et al.*, 2015) is the possibility of exploring "scalability", modularity and interdisciplinarity. In fact, given ABMs' own methodological flexibility, it is possible to scale the complexity of the models that are constructed. At the same time, you can add modules gradually to the main proposal, maintaining the consistency and coherence of the rest.[4]

Finally, Boero *et al.* (2015) reinforce the aspect of ABMs acting as a communication tool that allows interaction among professionals of different areas (Furtado, Sakowski and Tóvolli, 2015), mostly given its emphasis on visualizing output and trajectories. Niazi and Hussain (2011) make the argument in favor of growth of relevance and use of ABM in the various areas of science.

In short, it can be said with some ease that ABMs clearly bring a new perspective to modeling that, at least in theory, can contribute to illuminate more facets of the phenomenon being studied and to aggregate to their understanding.

1.2 ABMs as methodological tools

Terna (2015) suggests, supported by Ostrom's original work (1988), that computational simulation, such as ABM, would be a third way of doing science. Computational simulation adds – not necessarily competing, but perhaps in a complementary way – to the framework of traditional standard tools, namely: *i)* verbal argumentation, discourse; and *ii)* mathematical quantification through equations, statistics and econometrics, for example. It would thus be in a third general methodology: *iii)* the algorithmic one.

This notion of ABM as a singular methodological tool is older and goes back to Epstein's attempt to coin the term "Generative Science" (Epstein, 1999, 2006). Epstein argues that it is not enough to demonstrate an (observed) pattern, but to demonstrate how it was possible to arrive at a given pattern and to understand the underlying mechanisms. The scientist has to be able to identify what were the rules, attitudes, decision-making of which individuals or institutions that, together, in time and space, have led to a particular pattern observed.

> To explain a pattern, one must show how a population of cognitively plausible agents, interacting under plausible rules, could actually arrive at the pattern on time scales

4. PolicySpace is developed exactly like this. A central economic-spatial model that can be developed modularly, adding elements of demography, fiscal detail and mobility gradually.

of interest. The motto, in short, is [Epstein (1999)]: *If you didn't grow it, you didn't explain it* (Epstein, 2006, p.1587).

Finally, in a contemporary way, Nicolescu (1999) recalls that the human sciences, including politics and economics, do not follow some fundamental principles observed in the natural sciences, such as the existence of general, immutable laws. On the contrary, in the human sciences (as well as in politics) it is not possible to perform "experiments". Even tough any policy whose effects have not been analyzed, estimated, quantified *ex-ante* can be considered nothing more than an irresponsible experiment.

In addition, social phenomenon observe "... discontinuities, jumps and ruptures; [occur in] unique, discrete events ... [and are characterized by] uncertainties" (Furtado, Sakowski and Tóvolli, 2015, p.27). ABMs as a methodology may incorporate and test all of these phenomenon.

ABMs are distinct from some approaches that although similar are markedly different. Micro simulation (MS), for example, refers to the application of rules to simulate effects (Gilbert and Troitzsch, 2005). A researcher can simulate that considering only the first of the equations 1. This approach is useful, for example, to simulate the effects of a change in tax rates. Note, however, that MS does not allow interaction or heterogeneity (Hamill and Gilbert, 2016), that is, given a group of consumers and a given behavior, a change of tax rate, *ceteris paribus*, leads deterministically to an specific demand behavior. However, there is no disaggregated analysis of how the supply will behave.

Crooks and Heppenstall (2012) reinforce the differences between MS and ABM and claim that MS models unravels in only one direction and can only measure the effects of policies on citizens, but not their counterpart, i.e. the effects of citizens on each other and on policies.

In fact, Fujita, Krugman and Venables (1999) present an exercise in which they estimate the effects of agglomeration and disagglomeration. In this case, the analytical solution becomes very complicated, and it is necessary to simulate numerical results. The analysis, however, occurs in only one direction, and there is no dynamic interaction among the component agents of the system. Given the equations and the parameters, the results are calculated (iterated and numerically).

Considering these distinctions between MS and ABM, Birkin and Wu (2012) argue that MS studies the structure, whereas ABM is best suited to study behaviors. Although there are different focus, the authors emphasize that MS and ABM may reinforce each other and be used together as complementary tools.

Another approach that is very strong in the literature, easy to implement and that incorporates the notions of "systemic thinking" and *feedback* are Dynamic Systems (DS). Edmonds and Gershenson define them as follows:

> DS is an approach to modelling that represents a system in terms of a set of interconnected feedback loops (Forrester, 1971). It models these in terms of a series of flows between stocks plus additional connections between variables and flows. Crucially, it allows the representation of delays in such feedback and that the outcomes of some variables can control/affect the rate of other flows. These flows and relationships can then be simulated on a computer and (more recently) visualized. Its advantages are that a complex set of feedback relationships can be explored and hence better understood (Edmonds and Gershenson, 2015, p. 212).

Edmonds and Gershenson point out, however, that variables (stocks) represent abstract entities. Contrary to ABM, DS do not allow internal heterogeneity for each simulated stock. DS incorporate *feedback* and dynamism, however, DS treat each stage of the model as a single element that function as rates. Batty (2012) adds that DS also do not incorporate the spatial influence of the problem studied.

Moreover, the distinction between equation-based models (EBM) and ABM has a seminal contribution by Parunak, Savit and Riolo (1998), which is expanded by Sukumar and Nutaro (2012) and Cecconi and Campenni (2010). However, from these works it is not easy to distinguish explicitly the differences among the modeling techniques. The explanation that ABMs are solved numerically, recurrently, is also valid for General Equilibrium (DSGE) models or New Economic Geography models (Fujita, Krugman and Venables, 1999) which are EBM.

The fundamental distinction between EBM and ABM seems to be as follows: models whose communication is made exclusively through equations (EBM) or by means of algorithmic language (ABM) (Gräbner *et al.*, 2017).

Finally, it is worth dwelling a bit on the fundamental distinctions between DSGE and ABM models. DSGE models are based on the prevailing theory in contemporary economics, the so-called New Neoclassical Synthesis (NNS). The central assumptions are equilibrium, rationality, maximization, and representative agents. In contrast, ABMs[5] consider economics as a "a complex evolving system, i.e. as an ecology populated by heterogeneous agents, whose far-from-equilibrium interactions continuously change the structure of the system" (Fagiolo and Roventini, 2017, p. 1).

In particular, the 2009 financial crisis generated a series of criticisms to NNS and its inability to predict or understand crises, a recurring fact in the capitalist economy (Battiston *et al.*, 2016; Dosi *et al.*, 2012; Farmer and Foley, 2009). That

5. ABMs are also criticized by the difficulty to empirically validate models, excessive parametrization and calibration and lack of comparability among models.

incompleteness of crises within models occurs despite the addition of new "epicycles" that seek to "adjust" faulty models (Fagiolo and Roventini, 2017).

Fagiolo and Roventini (2017) list three main critics of DSGE models. The theoretical critical suggest that, since DSGE models are based on models of general equilibrium following the paradigms of Arrow-Debrew, DSGE models are subject to the same criticism of Arrow-Debrew models. According to Fagiolo and Roventini, you cannot prove the uniqueness and stability of the equilibrium, which would require infinite amount of information. The solution to this question is the use of representative agents, since they take the aggregation of rationality at the macro level (i.e., aggregate maximization) as something certain, although this is not necessarily true. In particular, the use of representative agents is a hurdle in the application of public policies, given the different reactions of each interest group. Fagiolo and Roventini (2017) go on to a full critic of DSGE including an empirical criticism that involves the computational process of solving the model, and the problems of identification, estimation and evaluation. Finally, they mention the assumptions of rational agent expectations, a full understanding of the economy by the agents, a complex and infallible problem solving ability, and unrestricted access to all information. All these assumptions together seem to be "Olympic" (Fagiolo and Roventini , 2017).

ABMs, in turn, base themselves on the literature referred to as Post-Walrasian Macroeconomics (Colander and Kupers, 2014) and evolutionary economics (Nelson Winter, 1982). The emphasis, in this case, is on the construction of empirically based models, supported by evidences and observed experiences. In such cases, the set of information available to the agents and the decision-making processes are closer to reality.

As mentioned before, there are numerous specific and recent manuals on agent-based modeling (Wilensky and Rand 2015), with emphasis on economy (Boero *et al.*, 2015; Hamill and Gilbert, 2016), social sciences (Edmonds and Meyer, 2013; Helbing, 2012), geography (Heppenstall *et al.*, 2012), spatial analysis (Batty, 2012) or political (Geyer and Cairney, 2015), which contain chapters detailing theory, methods and applications.

Numerous institutions around the world use ABM, or, more generally, the approach to complex systems,[6] which includes several other methodologies. Notably, ABM is a methodological tool at: *i)* the Bank of England (Baptista *et al.*, 2016); *ii)* the Organization for Economic Cooperation and Development (OECD) (OECD, 2009); *iii)* the Defense Advanced Research Projects Agency (Darpa);[7]

6. Available at: <https://www.complexityexplorer.org/explore/resources>.
7. Available at: <https://www.darpa.mil/news-events/2017-04-07>.

and *iv)* The Mitre Corporation (MITRE) and the European Union. Centers of excellence include the Santa Fe Institute, the University of Michigan and George Mason University, in addition to the Massachusetts Institute of Technology (MIT) and the New England Complex Systems Institute (NECSI).

In terms of *software* available to ABM specifically, a recent review (Abar *et al.*, 2017) listed more than eighty computer programs that can be used. Among the most well-known, we highlight: *i)* Netlogo, easy to learn; *ii)* Mason, repast and Swarm in Java; *iii)* Gama, for detailed spatial analysis; *iv)* anyLogic, commercial; *v)* Urbansim, for transportation and land use analysis; *vi)* Matsim for detailed studies, *real-time*, in transportation, and, of course, generic programming languages such as Python, Java, C, R and Matlab, to name a few.

All these resources, presented here as studies, institutions and *software* only to demonstrate that the subject is in open expansion and is currently in use in the advanced economies, although still with little use in Brazil.

REFERENCES

ABAR, S. *et al.* Agent based modelling and simulation tools: a review of the state-of-art software. **Computer Science Review**, v. 24, p. 13-33, May 2017.

ABDOU, M.; HAMILL, L.; GILBERT, N. Designing and building an agent-based model. *In*: HEPPENSTALL, A. *et al.* (Ed.). **Agent-based models of geographical systems**. London: Springer, 2012. p. 141-165.

ANDERSON, P. W. More is different. **Science**, v. 177, n. 4047, p. 393-396, Apr. 1972.

BAPTISTA, R. *et al.* Macroprudential policy in an agent-based model of the UK housing market. **Working Paper Bank of England**, v. 619, 2016.

BATTISTON, S. *et al.* Complexity theory and financial regulation. **Science**, v. 351, n. 6275, p. 818-819, Feb. 2016.

BATTY, M. A generic framework for computational spatial modelling. *In*: HEPPENSTALL, A. *et al.* (Ed.). **Agent-based models of geographical systems**. London: Springer, 2012. p. 19-50.

BILGE, U. Agent based modelling and the global trade network. *In*: GEYER, R.; CAIRNEY, P. **Handbook on complexity and public policy**. Cheltenham; Northampton: Edward Elgar Publishing, 2015. p. 414-431.

BIRKIN, M.; WU, B. A review of microsimulation and hybrid agent-based approaches. *In*: HEPPENSTALL, A. *et al.* (Ed.). **Agent-based models of geographical systems**. London: Springer, 2012. p. 51-68.

BOERO, R. *et al.* **Agent-based models of the economy**: from theories to applications. New York: Palgrave Macmillan, 2015.

CECCONI, F.; CAMPENNI, M. What do agent-based and equation-based modelling tell us about social conventions: the clash between ABM and EBM in a congestion game framework. **Journal of Artificial Societies and Social Simulation**, v. 13, n. 1, 2010.

COLANDER, D.; KUPERS, R. **Complexity and the art of public policy**: solving society's problems from the Bottom Up. Princeton: Princeton University Press, 2014.

CROOKS, A. T.; HEPPENSTALL, A. J. Introduction to agent-based modelling. *In*: HEPPENSTALL, A. *et al.* (Ed.). **Agent-based models of geographical systems**. London: Springer, 2012. p. 85-105.

DAVIDSSON, P.; VERHAGEN, H.; KLÜGL, F. Simulation of complex systems. *In*: MAGNANI, L.; BERTOLOTTI, T. (Ed.). **Handbook of model-based science**. London: Springer, 2017. v. 3. p. 781-795.

DOSI, G. *et al.* Income distribution, credit and fiscal policies in an agent-based Keynesian model. **SSRN eLibrary**, Jan. 2012.

EDMONDS, B. The room around the elephant: tackling context-dependency in the social sciences. *In*: JOHNSON, J. *et al.* (Ed.). **Non-equilibrium social science and policy**. Switzerland: Springer, 2017. p. 195-208.

EDMONDS, B.; GERSHENSON, C. Modelling complexity for policy: opportunities and challenges. *In*: GEYER, R.; CAIRNEY, P. **Handbook on complexity and public policy**. Cheltenham; Northampton: Edward Elgar Publishing, 2015. p. 205.

EDMONDS, B.; MEYER, R. (Ed.). **Simulating social complexity**: a handbook. New York: Springer, 2013.

EPSTEIN, J. M. Agent-based computational models and generative social science. **Complexity**, v. 4, n. 5, p. 41-60, 1999.

______. Remarks on the foundations of agent-based generative social science. *In*: TESFATSION, L.; JUDD, K. L. (Ed.). **Handbook of Computational Economics**: agent-based computational economics. North Holland: Elsevier, 2006. v. 2. p. 1585-1604.

EPSTEIN, J. M.; AXTELL, R. **Growing artificial societies**: social science from the bottom up. Cambridge: Brookings/MIT Press, 1996.

FAGIOLO, G.; ROVENTINI, A. Macroeconomic policy in DSGE and agent-based models redux: new developments and challenges ahead. **Journal of Artificial and Social Simulation**, v. 20, n. 1, 2017.

FARMER, J. D.; FOLEY, D. The economy needs agent-based modelling. **Nature**, v. 460, n. 7256, p. 685-686, 2009.

FORRESTER, J. W. Counterintuitive behavior of social systems. **Technological Forecasting and Social Change**, v. 3, p. 1-22, 1971.

FUJITA, M.; KRUGMAN, P.; VENABLES, A. **The spatial economy**: cities, regions and international trade. Cambridge: MIT Press, 1999.

FURTADO, B. A.; SAKOWSKI, P. A. M.; TÓVOLLI, M. H. A complexity approach for public policies. *In*: ______. **Modeling complex systems for public policies**. Brasília: Ipea, 2015. p. 17-36.

GALÁN, J. M. *et al.* Errors and artefacts in agent-based modelling. **Journal of Artificial Societies and Social Simulation**, v. 12, n. 1, Jan. 2009.

GEYER, R.; CAIRNEY, P. **Handbook on complexity and public policy**. Cheltenham; Northampton: Edward Elgar Publishing, 2015.

GILBERT, N.; TROITZSCH, K. **Simulation for the social scientist**. Great Britain: Bell and Bain Ltd., 2005.

GRÄBNER, C. *et al.* Getting the best of both worlds? Developing complementary equation-based and agent-based models. **Computational Economics**, p. 25, 2017.

GRIMM, V. *et al.* A standard protocol for describing individual-based and agent-based models. **Ecological modelling**, v. 198, n. 1, p. 115-126, 2006.

______. The ODD protocol: a review and first update. **Ecological Modelling**, v. 221, n. 23, p. 2760-2768, 2010.

GRIMM, V.; RAILSBACK, S. F. Designing, formulating, and communicating agent-based models. *In*: HEPPENSTALL, A. *et al.* (Ed.). **Agent-based models of geographical systems**. London: Springer, 2012. p. 361-377.

HAMILL, L.; GILBERT, N. **Agent-Based modelling in economics**. United Kingdom: Wiley, 2016.

HELBING, D. **Social self-organization**: agent-based simulations and experiments to study emergent social behavior. Berlin: Springer, 2012.

HEPPENSTALL, A. J. *et al.* **Agent-based models of geographical systems**. London: Springer, 2012.

JOHNSON, J. *et al.* (Ed.). **Non-equilibrium social science and policy**. Switzerland: Springer, 2017.

MAGNANI, L.; BERTOLOTTI, T. (Ed.). **Handbook of model-based science**. London: Springer, 2017. v. 3. p. 781-795.

MUELLER, B. Complex systems modelling in Brazilian public policies. *In*: FURTADO, B. A.; SAKOWSKI, P. A. M.; TÓVOLLI, M. H. **Modeling complex systems for public policies**. Brasília: Ipea, 2015. p. 261-278.

NELSON, R.; WINTER, S. **An evolutionary theory of economic change**. Cambridge: Harvard Business School Press, 1982.

NIAZI, M.; HUSSAIN, A. Agent-based computing from multi-agent systems to agent-based models: a visual survey. **Scientometrics**, v. 89, n. 2, p. 479, 2011.

NICOLESCU, B. **O manifesto da transdisciplinaridade**. São Paulo: Triom, 1999.

OECD – ORGANISATION FOR ECONOMIC CO-OPERATION. **Applications of complexity science for public policy**: new tools for finding unanticipated consequences and unrealized opportunities. Paris: OECD, 2009.

OSTROM, T. M. Computer simulation: the third symbol system. **Journal of experimental social psychology**, v. 24, n. 5, p. 381-392, 1988.

PAGE, S. E. **Diversity and complexity**. New Jersey: Princeton University Press, 2010.

PARUNAK, H. V. D.; SAVIT, R.; RIOLO, R. L. Agent-based modeling vs. equation-based modeling: a case study and users' guide. *In*: Multi-agent systems and agent-based simulation, INTERNATIONAL WORKSHOP, 1, 1998, Paris, France. **Anais**... Paris: MABS, 1998.

SQUAZZONI, F. A social science-inspired complexity policy: beyond the mantra of incentivization. **Complexity**, v. 19, n. 6, p. 5-13, 2014.

SUKUMAR, S. R.; NUTARO, J. J. Agent-based vs. equation-based epidemiological models: a model selection case study. *In*: ASE/IEEE INTERNATIONAL CONFERENCE ON BIOMEDICAL COMPUTING (BIOMEDCOM), 2012, Alexandria, United States. **Anais**… Alexandria: BioMedcom, 2012.

TERNA, P. From complexity to agents and their models. *In*: BOERO, R. *et al.* **Agent-based models of the economy: from theories to applications**. United Kingdom: Palgrave Macmillan. 2015. p. 10-30.

WILENSKY, U.; RAND, W. **An introduction to agent-based modeling**. Massachusetts: The MIT Press, 2015.

CHAPTER 2

POLICYSPACE: THE MODEL[1]

This chapter explains *PolicySpace*, describes its logic and its development. In addition to the intuition and the central facts that the model intends to mimic, here, we introduce *PolicySpace's* initial condition, its agents, markets, parameters and *schedule*, as well as its various processes.

1 INTUITION OF POLICYSPACE AND STYLIZED FACTS

The organizing fact of the proposal is the *urban economy* and the *real estate market*. We seek to investigate how prices of residences interact with the income level of inhabitants, with taxes, with places of work, and with the quality of the neighborhood. In addition, we want to gather better understanding of how the flow of revenues and expenses move dynamically and across space. We aim at better understanding the economic feedback effects within metropolitan municipal urban environments constrained by the distributive tax laws in effect, specifically for the case of Brazil.

Indeed, the heterogeneity and polarization between center and periphery (Furtado, Krause and França, 2013) observed for Brazilian metropolises is markedly present in *PolicySpace*. Such heterogeneity is included in the model via the initial design data, obtained from the *2000 Census*, but also from the actual rules and processes of the model.

Fundamentally, residences are cheaper in places where the quality of life is lower, where there are fewer companies, and the flow of tax investments is also lower (Furtado, Mation and Monasterio, 2013). These observed facts together generate a perpetuation of initial inequalities leading to distinct behavior of indicators across the municipalities of each modeled metropolitan region. In addition, the

1. The author acknowledges the contributions to previous development of the model, especially to Isaque Daniel Rocha Eberhardt (plots, automatic control of multiple simulations and general contributions in the project, from August 2015 to October 2016) and Alexandre Messa (discussions and reading) which have already been published (Furtado and Eberhardt, 2016, Furtado, Eberhardt and Messa, 2017a, 2017b). We would also like to thank initial comments by Davoud Taghawi-Nejad and Brais Alvarez-Pereira. In the period from August 2017 to January 2018, the project gained the contribution of Francis Tseng, who optimized and adapted the code, making it (much) more efficient and flexible. Finally, we thank the comments of an earlier revision of the model to referees Fabio Vaz and Fabiano Mezadre Pompermayer.

level of the central city – in several indicators – is disproportionately higher when compared to the other peripheral municipalities.

The main objective of *PolicySpace* is to investigate taxes collection and their spatial distribution among the municipalities of metropolitan agglomerations. The research hypothesis is that there is an optimal allocation of fiscal resources. Such allocation may lead to further development of the metropolitan cluster in terms of quality of life and economic strength.

In order to make such an analysis quantitatively, observable, and analytically understandable, we replicate the markets on which the government collects taxes. Once the markets are in place, alternative allocation of collected taxes may be investigated. Therefore, the proposal contains a basic labor market, with employees and employers, households and firms and a simple market of goods. Additionally, the model contains a real estate market. The capacity of the families to save determines the demand side. The hedonic calculus of the properties that considers their own characteristics and their location makes the offer side of the market.

Levels of tax collection affect investment of government on quality of life. Changes in quality of life implies changes in real estate prices across different locations. Families, in turn, adapt savings to their participation in real estate market, thus, generating the dynamicity of *PolicySpace*.

Together, this means that municipalities that collect more taxes raise real estate prices faster than their neighbors do. Households with highest savings can move to such municipalities, while families with low productivity – given the qualification of their members or with a high proportion of unemployed members – have to choose cheaper residences, probably in municipalities where prices are lower and quality of life worse. Typically, families are always seeking to live in the best places, but only a few of them can afford the investment costs (Furtado, 2009).

2 MODULARITY

One contribution of *PolicySpace* is that policy effects that happen across multiple areas of analysis may be examined in a connected way throughout the system. Illustratively, consider when several families move from one region to another within the city. Such move will influence land prices (in the region that loses families and in the region that receives them), urban mobility, and the goods market in both regions. However, the interconnectivity of these effects, the simultaneous study of real estate market, mobility and unemployment are difficult to pinpoint. These interactional effects, however, are inherent to the *agent-based modeling* (ABM).

Another contribution of *PolicySpace* as a tool is the process of analyzing public policies as a modular process. The implementation of these crossed effects does not have to occur simultaneously. The modeler may have primarily a goods market that works. Then he or she may implement location as a factor for the market, and then include families' mobility (the real estate market). With all these previous steps in place, you can then install a transport system that meets the demand in accordance with the location of families and firms.

Actually, that is exactly how *PolicySpace* was developed. The model was built gradually, integrated with previous working systems, but bringing new detailing and alternatives in the form of new modules. In order to help the reader have a better sense of the model, the stages of development as they happened are listed below.

The first implementation was of a temporal order. The `time_iteration` module[2] provides that days, months, quarters, and years elapse according to a given parameter of temporal duration that is typically set to 5,040 working days or twenty years. Such seminal module provided the platform on which daily, monthly, quarterly or annual operations could be individually added.

Soon, a first algorithm with abstract agents and municipalities was designed, tested, and published (Furtado and Eberhardt, 2016). There were one thousand agents and the municipalities consisted of seven square regions of different sizes. In that version, there were already three markets: of goods, labor, and real estate.

The next step was the transformation of abstract regions into actual municipalities, with real spatial configuration using *shapefiles* officially made available by the Brazilian Institute of Geography and Statistics (IBGE).

Yet another module read the correct municipal information of population, gender, and age group of each municipality to generate the agents of the simulation accordingly. At this stage, the parameter *number of agents: 1,000* became unnecessary and was replaced by real population. The configuration of the markets, however, was already in place and needed no change if applied to one thousand abstract agents or to the actual 31,022 inhabitants that correspond to 3% of the Population Concentration Area (PCA) of Natal in 2000 (IBGE, 2015).[3] Similarly, when we implemented the differentiation between urban or rural locations, the calculus of distance between place of residence and place of work needed no changes.

Gradually, other changes were included in *PolicySpace*. We included a module called *demographics* that processes agents' birthdays, mortality, and fertility monthly, and probabilistically according to IBGE official projection data. Such implementation means that *PolicySpace* starts with 31,022 agents,[4] but may have

2. In the current version this module moved to `simulation.py`

3. The PCA includes the municipalities Parnamirim, Extremoz, Macaíba, Natal, Nísia Floresta, São Gonçalo do Amarante and São José de Mipibu.

4. For the example of Natal's PCA with 3% of population.

only 29,980 inhabitants the following month, without the change of any other rule or procedures.[5] Finally, various graphs and statistics results were included as default of the simulation.

We use Python programming language to program, and a versioning GIT server, so that code changes are preserved and any previous configuration is immediately available.

Other contributions came from the introduction of the modules we called *controls*.[6] These control modules used to operate above the main simulation module and changed the number of times to run a given simulation, with changing configuration parameters and automatically processing results.

A major change in the development of *PolicySpace* was the inclusion of fiscal alternatives. Originally, the model contained only a consumption tax. A modular step then was the transformation of such tax into five taxes: on consumption, on wages, on profits, on housing property, and on transfer of property. Although more laborious, this change benefited from the existence of the markets on which taxes are levied. Yet, other aspects of *PolicySpace*, such as demographics or agents generation, remained intact.

Finally, policy test modules – alternative distribution of collected taxes – were added on top of the given structure.

All these steps are mentioned just to illustrate the modular construction thus far. The cumulative process benefits from all the advances made previously and are independent of each other.

3 ODD PROTOCOL: OVERVIEW, DESIGN AND DETAILS

As repeatedly suggested in the literature on agent-based modeling and observed more and more by the community, the description of ABMs is made following the protocol Overview, Design Concepts and Details (ODD) proposed by Grimm *et al.* (2006) and updated in 2010 (Grimm *et al.*, 2010). It consists of three main elements: *i) Overview* – general description; *ii) Design concepts* – which describe the proposal and its design; and *iii) Details* – detailing aspects of initialization, data and submodels. As far as possible, we will seek to follow the protocol proposed.[7]

According to the logic proposed by Grimm *et al.* (2010), the description occurs observing the following general idea:

5. *PolicySpace* currently does not include a migration module.

6. These modules have also been replaced and their functionality transferred to `main.py`. Actual control of alternatives now come from the command line as parameters. See details in section 4.

7. The authors explicitly suggest using the following text: "the description of the model follows the ODD protocol (overview, concepts and design and details)" (Grimm et al., 2006, p. 2765)

- overview;
- static description of the agents;
- temporal and spatial description – how the processes change the initial conditions of the agents and in what order;
- underlying concepts;
- design of the initial stage and description of the empirical data used; and
- detailing of relevant portions, for example, description of the markets in *PolicySpace*.

3.1 ODD: purpose

According to the ODD protocol, the first question is the purpose of the model. Simply put, the purpose is to provide a basic spatial platform that contains metropolitan areas and their municipalities, markets, residences, workers, and firms that replicate general macroeconomic behavior and indicate future trajectories from policy changes.

Specifically, the purpose of this research is to develop a model that is able to identify whether proposed changes in the distribution of fiscal resources between metropolitan municipalities are beneficial or not for all the metropolitan citizens in terms of quality of life and economic production within some reasonable time.

3.2 ODD: entities, states and scales

This item, according to the proposal of Grimm *et al.* (2010, p. 2764) shall include responses to the following question: "What kinds of entities are in the model? By what state variables, or attributes, are these entities characterized? What are the temporal and spatial resolutions and extents of the model?" The processes of updating variables and dynamics are described in the following subsections.

The model entities – their agents – are citizens, firms and municipalities. In addition, citizens organize themselves primarily in families when making consumption and residence decisions. The residence itself has its own characteristics independently of the family. Thus, the properties of each of these agents are detailed. Its methods and processes are in the following subsection.

3.2.1 Citizens

Formally, a class in the concept of Object Oriented Programming (OOP), citizens of the model have the characteristics of the table 1.

TABLE 1
Characteristics of citizens

Attributes	Values	Observations.
ID	*Integer*	Fixed. Unique.
Gender	*Male, female*	Fixed. Random.
Age	*Integer*	Reading proportion for each year by municipality of origin. Variable.
Birthday month	*Integer* [1, 12]	Fixed. Random.
Qualification	*Integer* [1, 15]	Fixed. Reading municipal data.
Money	*Float*	Initial drawn (50. 150). Variable.
ID firm	*Integer*	Initial: *none*. Unique. Variable.
Utility	*Float*	Cumulative sum of consumption.
Family	*Integer*	Family ID. Fixed.
Distance	*Float*	Variable. Distance home-firm when employed.

Author's elaboration.

The model generates all citizens at once at the beginning of the simulation. Therefore, citizens have their own fixed identification and characteristics throughout the simulation, i.e. gender, month of birth and given family. Note that citizens' generation considers their municipality's age cohorts and level of qualification. In addition, qualification does not change in this version.

Monthly, the simulation alters citizens' amount of money on hand; their utility – given as their accumulated consumption; its firm ID (if any) at any given time, and the distance from the house the citizen currently lives in and the firm where he or she works.

3.2.2 Families

We consider families *collectives* (Grimm *et al.*, 2010) of citizens. They act as one when making decisions on consumption and choice of residence. That means that families decide together on how to use their financial resources.

Families also house new members when female members of childbearing age give birth to new members, following official IBGE calculated probabilities. In this version of the model, there is no generation of new families, unique na table 1 new members.

Family properties include, in addition to their unique and fixed ID: the balance with the sum of their members' financial resources, a savings account, the list of which members belong to the family, the residence the family currently occupies and its price, plus the families' average cumulative consumption (table 2).

The modeler decides on an average number of members per family at the beginning of each simulation. The parameter – as a population percentage –

determines the number of potential families. The initial allocation of members into families is another process that happens independently, so there are families with different numbers of members.

TABLE 2
Characteristics of families

Attributes	Values	Observations
ID	*Integer*	Fixed. Unique.
Balance	*Float*	Sum of monthly financial resources. Variable.
Savings	*Float*	Variable. Cumulative amount not consumed each month. Spent exclusively in real estate market.
Members	*Dictionary*	Contain the actual (programming) objects of the family members themselves.
House ID	*Integer*	Residence currently occupied by the family.
Consumption	*Float*	Cumulative sum of family consumption.

Author's elaboration.

3.2.3 Residences

Five characteristics are typical of residences, defined at the time of their creation, and do not change throughout the simulation: *i)* their address – given by real geographical coordinates (latitude and longitude) read from the official IBGE shapefiles; *ii)* residences' size, drawn between 20 and 120; *iii)* quality, categorical from 1 to 4; and *iv)* whether the residence belongs to an urban or rural area, according to the official proportion of each municipality (table 3).

TABLE 3
Characteristics of residences

Attributes	Values	Observations
ID	Integer	Fixed. Unique.
Address	Longitude, latitude	Fixed. Random.
Size	Integer [20, 120]	Fixed. Random.
Quality	Integer [1, 4]	Fixed. Random.
Property	Integer	ID of the owner family. Variable.

Author's elaboration.

In addition, although not recorded as a residence attribute, the price is calculated from the variables of size and quality (which are fixed) and a variable portion that depends on the *Quality of Life Index* (QLI) of the municipality of the residence.

Given the address, one can determine automatically the municipality of location of each residence. Finally, each home has a proprietary family ID that can vary throughout the simulation.

3.2.4 Firms

Firms also have addresses and fixed IDs once created. The number of firms by municipalities is determined by reading data from the Annual Social Information Report (Rais) of the Ministry of Social Development (MDS).

The attributes of the firms vary monthly (table 4).

TABLE 4
Characteristics of the firms

Attributes	Values	Observations
ID	*Integers*	Fixed. Unique.
Address	*Longitude, latitude*	Fixed. Random.
Balance	*Float*	Total amount of financial resources available. At the beginning of the simulation, drawn from a beta distribution (*alpha* = 1.5 and beta = 10) * 10,000.
Workers	*Dictionary*	They contain citizens (object) currently working at the firm.
Sold amount	*Float*	Accumulated.
Quantity Produced	*Float*	Accumulated.
Salaries paid	*Float*	Accumulated.
Current revenues	*Float*	–

Author's elaboration.

Firms produce a homogeneous product with the same technology. Employees' years of qualification determines firms' productivity.

3.2.5 Municipalities

Municipalities are agents of the type *spatial units*, according to Grimm *et al.* (2010). They follow official municipal boundaries, given by IBGE. Their processes and methods generate effects onto those agents (firms and residences) that are located within their limits.

In addition to geometry and name, municipalities record their own tax information and some aggregate values, such as gross domestic product (GDP), population, and an accumulated mobility indicator. Municipalities also retain information on all five categories of taxes and their total: tax on consumption, property, labor, firms and real estate transactions.

Municipalities also record their own amount of Municipal Participation Fund (MPF),[8] which is a fund constituted of financial transfers from fractions of taxes on labor and on firms. The modeler then may choose the rule that establishes how such funds are to be divided among municipalities.

3.2.6 Temporal and spatial extension

PolicySpace includes all PCAs as defined by IBGE (2015). These areas are the equivalents in practice to the “employment basins”, or functional urban regions, in the designation of the Organisation for Economic Cooperation and Development (OECD) (Ahrend *et al.*, 2014). PCAs are economic regions that comprises areas in which workers commute daily.

The 46 PCAs available in *PolicySpace* are: Aracaju, Belém, Belo Horizonte, Brasília, Campina Grande, Campinas, Campo Grande, Campos dos Goytacazes, Caxias do Sul, Cuiabá, Curitiba, Feira de Santana, Florianópolis, Fortaleza, Goiânia, Ilhéus-Itabuna, Ipatinga, João Pessoa, Joinville, Juazeiro do Norte-Crato-Barbalha, Juiz de Fora, Jundiaí, Londrina, Macapá, Maceió, Manaus, Maringá, Natal, Novo Hamburgo-São Leopoldo, Pelotas-Rio Grande, Petrolina-Juazeiro, Porto Alegre, Recife, Ribeirão Preto, Rio de Janeiro, Salvador, Santos, São José do Rio Preto, São José dos Campos, São Luís, São Paulo, Sorocaba, Teresina, Uberlândia, Vitoria, and Volta Redonda-Barra Mansa.[9]

The *PolicySpace* user can select whether to simulate a specific PCA or all of them together. In doing all of them, there is no interdependence between the PCAs. In fact, they are distinct simulations that happen in parallel, without cross-influence among them. Thus, their effects are restricted to each PCA individually.[10]

PolicySpace starts in 2000 in order to use data from the census as baseline. The user chooses the number of business days to simulate. The month is set to last 21 (working) days and the year 252 days. We set the model to run 5,040 days, which is equivalent to twenty years of simulation duration. That is, the base time horizon for the simulation runs from 2000 to 2020.

Tasks and interactions occur on a monthly basis, although daily, quarterly or yearly activities may be included. The only options currently available for quarters and years is the save data function. The default saving for data is monthly.

8. Fundo de Participação Municipal (FPM), in Portuguese.

9. The input of the PCAs in the model is with capital letters and without accents. Some PCAs are named after two or more cities. Five PCAs are formed by a single municipality: Campo Grande, Campo dos Goytacazes, Feira de Santana, Manaus and Uberlândia.

10. It is computationally possible to simulate two or more PCAs together. However, citizens will commute among them without considering time or resources wasted in transit.

3.3 ODD: processes and temporal execution

Grimm *et al.* (2010, p. 2764) suggest that this section should describe, "…who does what, in what order". The idea is to offer the general design of the processes and the temporal chain. Individual designs of more complicated processes are in subsection 3.7, under submodels.

Each simulation begins with the generation of the agents (or their reading from previous archived generations) made from official data (subsection 3.6); according to the parameter settings and the options of each individual simulation, as defined by the modeler (subsection 3.7).

Before the actual simulation begins, firms create their product (an instance object), with an initial stock of zero and starting price of one. Then, a pre-matching on the labor market runs so that the simulation starts with the average unemployment level for January, 2000: 8.6%.

The sequence of events happens as follows. Once more, following procedures for the ODD, the most relevant processes, the equations, the fundamentals in the literature and the *pseudocodes* are in the submodels subsection (3.7). This sequence is to give the user a sense of the development of the model.

1) Production: each firm, independently, given its staff and their qualification, updates its monthly production, according to the equation:

   $$Q_i = \sum_{i=1}^{L} l_i^{\alpha}$$

   In which Q_i is the quantity produced by firm i; l is the level of qualification of each employee and α is the parameter of productivity that varies between 0 and 1. That is, the greater α, the greater the productivity of the firms.

2) Demographics: for each federal state (in PCAs belonging to more than one state), for each municipality, for each citizen, the simulation probabilistically applies the demographic processes: birthday, mortality, and fertility. The odds of mortality and fertility are different from state to state and by calendar year, according to IBGE tables.

3) Consumption: monthly, each family participates once in the goods market with an initial demand weighted by a propensity to consume parameter (β). The financial amount not used in consumption each month goes to a savings account that is used exclusively in the real estate market. Families compare prices and distances for a number of firms, such number is given by a parameter (`size_market`). Firms are passive and meet their demands according to their supplying capacity. There is no municipal or

spatial restriction on the choice of firms, although distance is a criteria used by families. Taxes on consumption are collected at the time of sale.

4) Firms, wages, and prices: the firms then calculate their income, pay taxes on profits; pay their employees, and decide whether to raise prices.

5) Labor market: the first event of the labor market process occurs with the construction of a list of citizens who are of working age [16, 70] and unemployed. Each firm then evaluates their profits in the previous month. When, results are positive, firms hire one employee. Otherwise, they fire one. However, the frequency of firms entering the labor market – the frequency of this decision-making – is set by a parameter (`labor_market`). Finally, the simulation makes the *matching* of companies and employees. Firms that offer the highest salaries, choose first. The criteria choice to hire may either be employees' qualification, or, the firms' distance to the applicant's residence. This criterion is given by another parameter (`pct_distance_hiring`). Firms choose candidates among batches of one hundred candidates.

6) Real estate market: Monthly, all vacant houses are considered as homes for sale and enter the market. Then, a percentage of households determined by a parameter (`percentage_check_new_location`) also enters the market. Families are sorted according to available resources on their savings account. Houses are sorted by price. This procedure facilitates the *matching* between families and homes. Households with total savings below the cheapest house exit the market. Thus, the first family with enough savings to buy the most expensive home makes the offer. The price is calculated as the average between the offer side, the estimated real value of the house that is calculated considering its surface, quality level and the value of its region; and, on the supply side, the amount available in family resources recorded as the total savings. Property transaction taxes are collected when transactions actually happen. Although with a simple mechanism, the proposed market allows the mobility of families and the adjustment between saving capacity and real estate prices.

7) Taxes on properties: families owning homes pay a one-twelfth of the property tax, if they have resources.

8) Municipal investments: once the economic cycle has run out for the given month and the municipalities have collected all taxes, they weigh the resources by the current population and invest proportionately in the improvement of the quality of life (QLI) index.

9) Finally, monthly data indicators are calculated and saved.

3.4 ODD: general references

3.4.1 Basic principles

"Which general concepts, theories, hypotheses, or modeling approaches are underlying the model's design?" (Grimm *et al.*, 2010, p. 2764). The hypothesis of the model is that fiscal distribution among PCAs is not optimal. We conceptualize that in practice the core metropolis and the neighboring municipalities of a metropolitan region (PCA) constitute themselves an autonomous economic unit (Ahrend *et al.*, 2014; Firkowski, 2013). However, from a budgetary, decision-making and capacity-building perspective, municipalities (in Brazil) are constitutionally autonomous entities. There is, therefore, incompatibility between resources and power, on the one hand, and the need for planning, articulation and integration, on the other hand.

As PCAs bring together workers who produce predominantly in the central municipality of the metropolis, it would not make economic sense nor be fair to separate costs and benefits of the agglomeration process. In practice, when there is segregation between the metropolis and the periphery, the former receives the bonuses of agglomeration economies, leaving the burden of diseconomies, especially violence and traffic, to the periphery (Rodrigues, 2013).

In addition, there is strong evidence that scale is central to municipal bureaucracy, with larger municipalities being more efficient in collecting and providing public services (Gasparini and Miranda, 2011). Not to mention the inefficiency of having multiple municipal bureaucracy.

Finally, it is even difficult to define precisely what is local within a troubled and densely populated metropolitan region. In isolated municipalities, in the countryside of large states, autonomy over legislation and planning of land use, transportation, and housing is reasonable. In fact, the local citizens are the ones who are aware of their reality and who have local knowhow.

In metropolitan regions, however, the transport system needs to operate as a networked system. When the system is a compartmentalized way, the commuter ends up needing to use local services, make multiple transfers, even change from entirely different systems before making it to a final destination (Cervero, 2013, Pedroso and Lima Neto, 2013). The same need for integrated planning is present concerning land use and housing. Given the mobility of people in home-to-work journeys, the argument that activities are independent and autonomous and demands local government does not hold, as Royer (2013) suggests.

Understanding the metropolis as a single economic region that at the same time has political and budgetary fragmentation across independent municipalities, we want to quantify and qualify alternatives of public policies and their trajectories.

Hence, the research question is: what would systemically happen if the collection and distribution of taxes were centralized?

3.4.2 Emergency

What results emerge from the individual attributes of agents? (Grimm *et al.*, 2010). In the case of *PolicySpace*, the very interaction among the agents, given their rules and design, cannot be predicted *a priori*. Therefore, indicators such as unemployment, inflation and financial resources available to families are explicit emerging results. Nevertheless, the simulation model repeatedly shows that some indicators keep similar patterns of behavior over several simulations, though not with identical trajectories.

3.4.3 Adaptation

"What adaptive traits do the individuals have? What rules do they have for making decisions or changing behavior in response to changes in themselves or their environment?" (Grimm *et al.*, 2010, pp. 2764). The adaptation of agents among themselves and the environment occurs mainly through prices and not directly. For example, in the real estate market, families whose members have higher qualifications are more productive, get jobs more easily and, consequently, are more likely to purchase expensive homes in municipalities with better quality of life. Companies that are located closer to the demand market sell and can keep inventories down raise prices and wages paid, thereby producing more proportionately to other firms.[11]

Municipalities have their quality of life indexes relatively lowered, when they receive immigrants and increased when there is emigration. This pattern guarantees that the same amount of resources will be more or less effective depending on the increase (attraction) or decrease in population. These examples demonstrate how agents adapt within *PolicySpace*.

3.4.4 Objectives

Following the same line of reasoning from above, agents do not have a strict purpose such as maximization of utility. However, when conditions are fulfilled (family members employed, availability of significant resources in savings), families seek to improve the quality and location of their residence.

Firms, on the other hand, actively seek to maintain a team of employees adequate in size and qualification to their constantly changing demand. Given the conditions in which there is strong demand (perhaps when enough households consume based on distance rather than price), the firm actively pursues higher prices.

11. However, if they raise prices too high, they may eventually lose customers to the competition, if any.

3.4.5 Learning

Agents do not evolve their rules and behaviors. Nevertheless, as the environment (and prices) change continuously, results generated from the fixed behavior vary.

3.4.6 Prediction

This item refers to the ability of the agents themselves to make forecasts. This does not occur in *PolicySpace*. That is, the agents do not make decisions based on estimates or forecasts about the future.

The only agent attribute related to future times is families' savings. However, household savings are, in fact the residue of resources not consumed at any given month. Consumption is determined probabilistically from a beta distribution by the propensity to consume parameter, exogenously defined.

3.4.7 Sensing

Sensors, as proposed by Grimm *et al.* (2010), refer to messages and information that the agents provide to each other that affect their decision-making. In *PolicySpace*, there are no sensors as such. However, agents make use of other agents' information in order to interact.

For instance, information of wages paid every month by firms determines firms' hiring priority. In addition, within the labor market, information of workers' qualification or location of residence informs the hiring process. Firms also advertise prices and location so that consumers use this information to make their decisions.

Finally, real estate market participants use information from the municipal quality of life indicators when calculating sales price for their homes. Algorithmically, information on real estate market participants' savings account enters as input in the matching process for the market.

3.4.8 Interaction

Interaction among agents is not direct, but mediated by information they exchange. Following Grimm *et al.* (2010) concepts, there is *indirect interaction* that occurs via competition. For example, a family with a given level of savings purchases different houses according to the level of savings of the other families that are in the market in that given month. Likewise, a less-educated citizen may have to wait for subsequent moments in the job market to get a placement. Finally, firms that are paying smaller wages compete with firms that pay better wages in the selection of workers.

3.4.9 Stochasticity

PolicySpace makes use of stochasticity in a broad spectrum of decisions. Obviously, probabilities depend on the parameters and the distributions used. Whenever possible we keep parameters and distributions close to observed empirical data.

Generation of citizens include stochasticity in gender, birthday month, specific age within the cohort, qualification (gamma distribution),[12] initial available money (beta distribution), and allocation in the family. Families' allocation to a residence and the possibility of ownership of other residences are also stochastic. The residences themselves also include stochasticity in their size and quality, location and whether they are in an urban area or not. Finally, the initial balance of firms come from a beta distribution.

PolicySpace's processes also include stochastic decisions for mortality and fertility, drawn from official data. Families' decisions about which criteria to use in the goods market (price or proximity); when selecting firms to evaluate for possible consumption, and when deciding the amount to spent (beta distribution).

Firms' stochasticity is present when deciding how often to update prices and, in the labor market, when to choose which employee to fire or when to draw a sample of the candidates to participate in a selection process.

Finally, *PolicySpace* may run with a fixed *seed* so that results are exactly reproducible. Default configuration runs with random *seeds*. Hence, simulation results are distinct, but typically maintain most structural patterns. Yet, as recommended by ABM literature, we run simulations *n* times having variable *seed* so that results contain means of several runs.

3.4.10 Collectives

In this item, Grimm *et al.* (2010, p.2765) suggest the answer to the following question: "Do the individuals form or belong to aggregations that affect, and are affected by, the individuals?" In *PolicySpace*, municipalities are agents that affect residences' prices that are within, via quality of life index. Further, the dynamic population of a municipality affects the proportionality of the investment of the resources collected in the improvement of the municipal quality of life.

Indirectly, families within residences have access, albeit subjectively, to a better/worse quality of life of the municipality. Families are also collectives that financially affect and are affected by individual members within.

3.4.11 Observation

What observations and what data are collected and analyzed in the model? (Grimm *et al.*, 2010). There are options in the simulation mode choices that allow you to

12. When we do not mention distributions, it means values come from uniform distributions.

record information about citizens, families, firms, and households. By default, we always record general monthly data and municipalities' indicators. There is, however, the option to save them only on a quarterly or annual basis. In addition, data is in comma-separated values (CSV), with a semicolon as a separator and a dot as a decimal expression.

Information in the following order is recorded for firms: months, company ID, municipality ID, longitude, latitude, current cash, number of employees, quantity in stock, produced quantity, current price, sold quantity, revenues, profits, and wages paid.

For citizens, *PolicySpace* saves: months, municipality ID, gender, longitude, latitude, citizen ID, age, qualification, firm ID (if any, otherwise `None`), family ID, cash, cumulative consumption and distance from home to work (if applicable). Additionally, when saving agents option is on, data for citizens who have passed away are available as a separate file.

For residences, we record: months, residence ID, longitude, latitude, size, price, family ID (when applicable), and municipality ID.

In the case of families, the model saves: months, household ID, current house price, municipality ID, family savings, number of members and consumption decision strategy (whether price or proximity).

Observations recorded in all simulations are aggregated (for the whole PCA being simulated) or by municipality. In the case of municipalities: months, municipalities ID, sum of total trips from home to work, population, GDP, Gini, average of the values of homes, unemployment, quality of life index, GDP *per capita*, total tax revenues, and taxes on consumption, on residential property, on labor, on firms, on real estate transactions, and transfers from the MPF.

For the aggregated PCA, we record: months, the price index, GDP index, GDP growth, unemployment, average of workers per firm, mean available cash of households, average household savings, average firms' revenues, Gini, average of household accumulated consumption, inflation, and average QLI.

Optionally, the simulation settings allow the automatic elaboration of graphs of several of these indicators, especially by municipalities and PCA averages. You can also generate maps with some data from firms and families and their locations.

3.5 ODD: initialization

"What is the initial state of the model world, i.e., at time of a simulation run?" (Grimm *et al.*, 2010, p 2765). Once generated, the model saves the agents (in *pickle Python* format), so they are available for following simulations, on the condition that the four specific parameters of the simulation (percentage of population, PCA, percentage of vacant residences and average number of members in the family) are the same.

The first step in generating agents is to read IBGE database containing the official *shapefiles* (files that contain geometry, polygons, lines or points, linked to a cartographic projection). The reading process selects only the municipalities that compose the simulated PCA. PCAs that have municipalities on more than one state, such as the Integrated Development Region of the Federal District and Surrounding Areas (Ride/DF), contain municipalities from both states involved. As previously mentioned, PCAs are usually smaller than the metropolitan region as defined by IBGE (2015). Municipalities start the simulation with the value of the Human Development Index (HDI) in 2000 as their QLI. Initial GDP, accumulated commuting, and the tax resources are equal to zero.

Next, the *Generator* class generates citizens, households, families, firms, and regions (municipalities) in the form of dictionaries whose keys are agents' IDs and the values are the generated objects themselves.

In the process of populating the municipalities, the model follow the pseudocode of box 1.

BOX 1
Function *create_all*

```
Function: create_all
For each municipality among the municipalities of the simulation:
      Check whether the population will be simplified by cohorts:
      For each age or cohort:
           For each genre:
                Total number of citizens is determined1
                For each citizen in the selected total:
                      Select qualification2
                      When cohorts, select age
                      Selects financial resources
                      Select month of birthday
                      Creates agent object (citizen)
                      Add agent to the group of municipal agents
                           Update ID value
           Add agents to a general agent dictionary
           Calculate number of families
           Calculate number of residences
           Calculate number of firms3
           Call create_families function
           Call create_household function
           Call create_firms function
           Call allocate_to_family function4
           Include families, residences and municipal firms as general agents
           For each residence in the municipal residences:
                If the residence does not contain a family:
                      Select family and add family as property's owner
Return citizens, residences, families and firms
```

[1] By age, gender and municipality.
[2] According to municipality's data.
[3] For the three cases, according to the parameters established for the simulation.
[4] Given families and citizens, by municipality.

Very similar processes generate families, households and firms, in accordance with the initial values described in subsection 3.2. When generating residences, the proportion of urban and rural areas is observed, and longitudes and latitudes are drawn conditioned to the proportion. Allocation of agents in families is done randomly. For each citizen in the list, a family is chosen. This procedure guarantees the average number of citizens per family; however, there are families with variable number of members.

After the generation process – which is skipped when there is a saved generation available – a first round of the labor market is simulated in order to achieve average unemployment of 8.6% for the six metropolitan regions surveyed by the IBGE for January 2000. Finally, municipal populations are updated.

3.6 ODD: input data

A series of official data enters the model in the process of generating and updating the population of agents. The bases are detailed in this subsection.

Fertility

Fertility tables are available by states (UFs) and contain the average number (and the projections) of live births per woman, for the calendar years from 2000 to 2030, and by mother age from 15 to 49 years old. The information is from the Research Directorate of the Population Coordination and Social Indicators of the Management of Studies and Analyzes of the Demographic Dynamics of IBGE and are included in the *Projection of the Population of Brazil and Units of the Federation by Gender and Age for the Period 2000-2030*.[13]

Mortality

Similar to fertility and provided by the same source, mortality is incorporated into the *PolicySpace* by gender; age, from 0 to 110 years old; per calendar year, from 2000 to 2030; and by UF. However, in this case, some adjustments were necessary since IBGE only had mortality up to 90 years old and for periods of five years. Transformation to annual mortality was made by simple division. Mortality beyond 90 years old was made from the expansion of data from Castro (2015)[14] for 2008-2013, using the average for the years 2008-2010 to cover the previous period from 2000 to 2007 and the average for the period 2010-2013 to impute the period thereafter up to 2030.

13. Annexes numbers 12.891 and 12.892 received directly from IBGE.

14. Available at: <http://www.previdencia.gov.br/wp-content/uploads/2015/06/NOTA-TECNICA-ATUARIAL-EXTRAPOLACAO-DA-TABUA-IBGE-MPS.pdf>.

Census

The proportion of urban population, the number of citizens by gender, by age (from 0 to 100 years old) by municipality, and the *shapefiles* of municipalities and urban spots were obtained from the *Census 2000* (IBGE, 2003). The set of municipalities of each PCA was also obtained from IBGE (2015).

Education

Percentage of the municipal population by groups of years of study came from the *Demographic Census 2000* (IBGE, 2003). Small adaptations were made to the data: to the group with less than one year of study was assigned value 1; to the group between one and three years of study, value 2; to the group between five and seven years of study, value 6; and for the group of nine and ten years, value 10. Finally, to the group of undetermined years of study (1.2% of the population), we imputed value 9, which is close to the median value.

Firms

The number of establishments per municipality for 2000 comes from the Rais database of the Ministry of Labor (MT)

Municipal Human Development Index (HDI)

The HDI is available by municipality for 2000 and 2010 calculated by João Pinheiro Foundation (FJP) and the United Nations Development Program (UNDP). HDI is used as basis of comparison for *PolicySpace's* QLI.[15]

MPF

The real observed proportion of official distribution of MPF among municipalities provided by the National Treasury Secretariat (STN) is used in the model.

Taxes

Data on tax collection for the municipalities in the *PolicySpace* were also provided by STN.[16]

3.6.1 Data for other countries

The use of the platform for event analysis from other countries should contain equivalent data, namely: municipal (administrative region) data for population by gender, age, and years of study. Province or state data for mortality and fertility. Municipal administrative boundaries in shapefile format, including urban and rural polygons, number of firms, quality of life index (municipal HDI), fiscal

15. Available at: <http://atlasbrasil.org.br/2013/>.
16. Available at: <https://siconfi.tesouro.gov.br/siconfi/pages/public/consulta_finbra>.

distribution rules among municipalities and the aggregation of municipalities in functional regions. Validation would require regionalized data on unemployment, inflation, GDP, household consumption and tax revenues.

3.7 ODD: submodels

This subsection details the architecture of the most relevant parts of the model and includes literature and justifications for the decision-making implemented. It also refers to the validation and the tests in chapter 3. Hence, the subsection describes the three present markets in the model – goods, labor and housing. Further, we detail decision-making of firms for production and sales, pricing and wages mechanisms and household consumption behavior. Before the description of the submodels, we overview all the parameters used in the simulation. The alternative ways to run the simulation are in section 4.

3.7.1 Parameters

Models with fewer parameters are simpler. Moreover, simple is better than complicated, according to the Occam's razor principle. The model parameters presented below (table 5) serve more to analyze the variations of different configurations than for the construction of the model itself. In fact, several parameters can be set so as not to interfere with the simulation. The systematic analysis of the variation of the parameters, however, provides sensitivity analysis and insight into the internal functioning of the simulation.

Considering the goods market, five are the alternatives for the modeler.[17] The *alpha* parameter serves to vary the productivity of workers in relation to their qualification. When *alpha* is one, productivity equals the worker's years of study. When it is zero, all workers produce a unit. *Markup* is the percentage firms use when they decide to update their prices. The very frequency with which the firm verifies whether it will change prices is controlled by another parameter, *sticky_prices*.

On the families' side, two parameters control the magnitude of consumption (*beta*), in contrast to the savings, and the number of firms that the family evaluates when deciding the firm to consume.

17. Variations of parameter *production_magnitude* changed general behavior of the model very little, so we suggest the modeler to keep this parameter with fixed value.

TABLE 5
Default configuration parameters and values

	Parameters	Default values	Comments
#	Goods market – firms		
1	*Alpha*	0.24	Exponent in the function of productivity. It varies between 0 and 1, with 1 being the most productive worker.
2	*Markup*	0.15	Percentage used by firms to raise prices.
3	*Sticky_prices*	0.5	Factor that controls the frequency with which the firm decides to assess if prices increase. It varies from 0 to 1, the lower the value, the more often the firm evaluates.
4	*Production_magnitude*	76	It divides the effective monthly production of the firm, in order to make production compatible with other parameters (*markup*, prices).
	Goods market – families		
5	*Beta*	0. 7	Propensity to consume in the consumption function. It enters as a beta parameter: $(1 - \beta) / \beta$, in a random beta function varied in which $\alpha = 1$. Higher β, higher consumption, lower saving.
6	*Size_market*	10	Number of firms evaluated by families in the decision on cheaper product or nearer firm.
	Labor market – firms		
7	*Labor_market*	0.05	Frequency with which the firm assesses whether it participates in the labor market. The lower the value, the more frequent.
	Labor market – rules		
8	*Pct_distance_hiring*	0.17	Percentage of job applicants who will be evaluated by the distance residence-work criterion.
9	*Wage_ignore_unemployment*	False	*True*: Firms distribute all revenue collected. *False*: firms decrease wages paid in proportion to current unemployment.
	Real estate market – families		
10	*Percentage_check_new_location*	0.01	Percentage of families entering the housing market. Number of families that actually carry out transactions may be different.
	Taxes – rates		
11	*Tax_on consumption*	0.00039	Rate of consumption tax. Charged at time of sale.
12	*Tax on labor*	0.00013	Rate of salaries tax. Charged at the time of payment to employees.
13	*Tax_on estate_transaction*	0.0000015	Rate of transactions tax. Charged on the completion of the real estate purchase.
1 4	*Tax_on firms*	0.00044	Tax on firms' revenues less wages paid. Deduced monthly from the firms.
15	*Tax on property*	0.0000016	Rate on real estate property. Payment 1/12 monthly, conditioned to families' cash availability
	Taxes – rules		
16	*Alternative0*	True	Municipalities are independent entities concerning distribution of tax revenues. *False*: revenues are distributed equally among municipalities, weighted by current population.
17	*Fpm_distribution*	True	The rules and proportionalities of MPF are observed. False: MPF rules are not followed.
	Agents' generation		
18	*Members_per_family*	2.5	Average number of members per family.
19	*House_vacancy*	0.05	Percentage of vacant houses.
20	*Simplify_pop_evolution*	True	Simplifies population generation by aggregating age groups [6, 12, 17, 25, 35, 45, 65, 100] – groups can be modified. False: includes all ages, no cohorts.

Author´s elaboration.

The labor market has three parameters associated with its operating rules. *Labor_market* defines the frequency with which firms evaluate whether to hire or fire. *Wage_ignore_unemployment* is a boolean parameter that changes the rule that the company uses to decide on the value of wages offered. When true, the entire net revenue of the company is distributed in the form of salaries. When false, the company weighs the level of wages by the level of unemployment, paying less when unemployment is higher. Finally, *pct_distance_hiring* sets the percentage of candidates selected by location criteria. Note, when this parameter is zero, all candidates are hired using qualification as criterion.

The parameter of the real estate market only establishes the number of families that enter the market each month. Concerning the fiscal analysis, the parameters are the rates of taxes on consumption, labor, firms, property and real estate transactions. Additionally, there are two rules that together allow four configurations of the distributive tax system (see details in the description of the real estate market in subsection 3.7.7 and taxes and municipalities in subsection 3.7.9).

Other two parameters refer to families and residences generation. *Members_per_family* refers to the average number of members per family and *house_vacancy* determines the default percentage of vacant houses, or otherwise the number of residences to be set up over total number of families. Moreover, two parameters detail whether the population will be drawn in a simplified fashion by age groups (*simplify_pop_evolution*) and the actual age intervals for the groups (*list_new_age_groups*).

Two additional parameters refer only to internal adjustments of proportions relative to the magnitude of quantities used in *PolicySpace: production_magnitude*, which reduces the actual number of units produced, and *treasure_into_service*, mediates tax investment as QLI increments, in order to make the indices reasonably comparable to observed values.[18]

3.7.2 Firms: production function

Firms produce a singular, homogeneous product using fixed technology (Lengnick, 2013). The production function depends exclusively on the quantity and qualification of employees in a given month (Gaffeo *et al.*, 2008). The function itself is given by:

$$Q_i = \sum_{j=1}^{J} j^{\alpha}.$$

Where Q_j is the monthly quantity produced by the company i, j are the employees of the company and α is the parameter of productivity.

18. *Treasure_into_services* is set to one in the default configuration, therefore having null effect on the simulation.

3.7.3 Firms: decision on prices

Firms' decision to adjust prices is neither a trivial task nor a standard methodology (Blinder, 1994). Usually, the ability to impose prices correlates with the behavior of demand that reflects in their own inventory (Bergmann, 1974; Seppecher, Salle and Lavoie, 2017) and the firms' market share (Dosi, Fagiolo and Roventini, 2010). Pragmatically, Hamill and Gilbert (2016) suggest a parameter of variable price adjustment varying between 1% and 30%.

In *PolicySpace*, prices' increase depends on two steps. The first one, *sticky_prices*, determines whether the firm evaluates increasing prices or not, following Blinder (1994). If so, actual decision depends further on whether the quantity sold in the previous month is greater than the quantity produced in that month. When the decision has been made, prices are increased according to the fixed proportion of the *markup* parameter.

3.7.4 Goods market: families and consumption

Citizens consume together as a family. Salaries incomes received at each month are aggregated and family consumption decision is unique. The first step is to decide on a positive amount to consume. The choice is made between zero and the maximum available value, depending on the propensity parameter to consume β. β is the beta of the beta distribution, where α is one.

Next, families select the firms that will belong to their sample according to a number given by an exogenous parameter (*size_market*) and then decide whether to select the firm for prices or proximity, with a 50% chance for each decision. Among the firms on the sample, they choose the one with the lowest price or closest to their residence. The firm will sell the total demand of the family when available. Otherwise, change from lack of firms' offer go back to the family. Family resources not consumed in a given month are transferred to a savings account and can be accessed exclusively for participation in the real estate market.

3.7.5 Goods market: firms and sales

From the perspective of firms, since production and stocks are given and prices set, for each demand request, they simply check product availability, collect consumption tax and, if unable to supply the request completely, return the change to the families.

3.7.6 Labor market: matching

As the other markets in the model, the labor market is simple[19] (Hamill and Gilbert, 2016) and the goal is only to provide reasonable market dynamics so that

19. For a complete ABM labor market proposal that simulates the total set of workers for the United States and replicates stylized facts and dynamics, see Axtell (2013).

the application of taxes and the spatial distribution can be evaluated, that is, the purpose of the modeling fulfilled (Crooks and Heppenstall, 2012).

According to Neugart and Richiardi (2012), there is no consensus in the literature about the decision-making processes of firms in relation to the payment of wages. In any case, it is reasonable to suppose that a proportionate part of the revenues defines wages and that competition for new employees is given, at least in part, by the level of wages offered.

All citizens between 16 and 70 years old seek employment every month and make up the group of candidates. On the firm's side, labor market participation is assessed on a variable basis, depending on an exogenous parameter (*labor_market*). When the firm decides to participate in the market and obtains positive (or zero) profit in the previous month, the firm composes the group of firms hiring.

The first action that occurs in the market is a random shuffle of candidates. Then, given the exogenous decision parameter in the labor market by proximity of the firm or by qualification (*pct_distance_hiring*), candidates are separated into two lists proportional to the parameter. It is important to highlight that for the Brazilian case the place of residence is a relevant factor when selecting workers, as the legislation imposes on the employer the cost of transportation that exceeds 6% of the employee's salary. Therefore, especially for jobs with smaller pay and for less qualified works, the place of residence is used as a discriminating criterion. Even then, *PolicySpace* allows the simulation with the parameter equals to zero, which excludes the influence of the work commuting distance in the selection process.

In both lists, *PolicySpace* sorts firms by order of offered wages. Those firms paying higher wages, choose first. Wages are firm's gross revenue, discounted from taxes. In addition, the default setting also discounts the inverse of current unemployment (*wage_ignore_unemploment*: *False*). In other words, the salary mass offered by the company is given by gross revenue, deducted from taxes and multiplied by 1 minus the percentage of unemployed: wage-base = revenues * (1 - consumption tax) * (1 - unemployment rate). Internally to the firm, the salary mass is distributed proportionally to the productivity of each worker that is given by its qualification.

Candidates are also sorted by qualification. Hence, the firm that pays the highest salary chooses the most qualified candidate and so on. In the distance criterion, each firm "interviews" one hundred candidates and chooses the one whose residence is closest to the firm itself.

Note that each firm only hires one employee per round. If there are no more candidates or firms, the process stops and it initiates only the following month.

3.7.7 Real Estate Market

Jordan, Birkin and Evans (2012, p. 517) list seven reasons why families seek new homes. In the context of *PolicySpace*, two of these reasons are subjacent to the model. Families look for new residences because either they have changed their socioeconomic status or they are searching for better public service offer (education and transport, for example).

The number of vacant properties[20] comes from an exogenous parameter (*house_vacancy*). Every month all vacant homes and some families (also determined by an exogenous parameter: *percentage_check_new_location*) enter the market. Vacant residences that enter the market have their selling prices updated, given current municipality's QLI. We calculate Price following real estate hedonic pricing (Malpezzi, 2002; Rosen, 1974) which is a function of its size and quality – that are fixed – plus its location influence, which is variable.

More precisely, location influence is a function of the monthly QLI of the municipality. QLI is weighted by the municipal population so that QLI increment is the sum of the resources effectively distributed divided by the current number of inhabitants.

In order to speed the matching process, properties with prices above the family with greatest savings and families with savings below the minimum value of available residence leave the market.

Next, *PolicySpace* sorts families by savings and residences by prices. The matching starts with the family with greatest savings checking the most expensive property available (and necessarily smaller than its savings). Transaction price is the average between the price offered by the family and the price requested by the property (calculated). Thus, when all requirements are fulfilled the transaction takes place and transaction taxes are collected.

The next family checks if their savings is greater than the price of the next property. If not, the family searches for the next property they afford. And so on until there are no more families or residences on the market. Real estate market process finishes with the families' decision to change residence.

3.7.8 Decision to change residence

Immediately after acquiring a new residence, the family decides whether to move to the new home (or any other one they may also own). The criterion is simple, if everyone in the household is unemployed, the family moves to the cheapest house, releasing the most expensive residence(s) to market. Otherwise, the family chooses

20. See Nadalin and Igliori (2016) for an analysis of vacant houses for São Paulo.

the most expensive residence. Alternatively, the modeler may change the code so that the family moves to the house with the most quality (or size). Anyway, note that the price of residence is the linear combination of size, quality and location.

3.7.9 Taxes and municipalities

Investment of taxes collected during the month is the last stage of operationalization of *PolicySpace*. The model contains four alternatives for the distribution of resources collected through taxes among municipalities, via two parameters (table 6).

TABLE 6
Taxes distributed among the different alternatives
(In %)

Alternatives	(i) *Alternative* 0 = *True* FPM = *True*			(ii) *Alternative* 0 = *False* FPM = *True*		(iii) *Alternative* 0 = *True* FPM = *False*	(iv) *Alternative* 0 = *False* FPM = *False*
Taxes	Municipality	State	MPF	State	MPF	Municipality	State
Consumption	0.1875	0.8125		1		1	1
Labor		0.765	0.235	0.765	0.235	1	1
ITBI	1			1		1	1
Firm		0.765	0.235	0.765	0.235	1	1
IPTU	1			1		1	1
Criterion	Locally	Equally	MPF	Equally	MPF	Locally	Equally

Author's elaboration.
Note: ITBI – Tax on Real Estate Transaction; IPTU – Property Tax.

When both parameters (*Alternative0* and *FPM*[21]) are true, the process of taxes distribution among municipalities follows current practice, in a simplified way. This is equivalent to saying that each municipality receives the resources collected according to three principles: *i)* locally; *ii)* equally among municipalities; or *iii)* in accordance with the MPF rule.

Locally: ITBI and IPTU are local taxes, collected by the municipalities themselves, although with low consistency and values (Carvalho Junior, 2009). In the case of *PolicySpace*, the entire amount collected through ITBI and IPTU is reverted to the municipality of origin. In addition, part of the consumption tax also goes to the municipality. "Consumption" tax seeks to represent (in a simplified way) some of the concepts embedded in the Tax on Circulation of Goods (ICMS). The rules of ICMS indicate that 25% of the amount collected should

21. Throughout the book, we use the acronym MPF for the Municipal Participation Fund. However, the parameter's name comes from the original in Portuguese: FPM – Fundo de Participação Municipal.

return to the municipalities and three quarters of these 25% must be reverted to the generating municipality.

Equally: the taxes on labor and on firms seek to replicate the collection of income tax of citizens and of firms. The collected taxes are distributed by the states or the Union. That is, the money collected via these taxes is divided equally among the municipalities, weighted by the resident population.

MPF: the fund receives transfers from taxes on labor and on firms that is equivalent to 23.5% of each one. Total amount collected follows its own rule of distribution. Intuitively, this means that proportionately greater resources go for less populous municipalities. That is more beneficial to peripheral municipalities than the equal distribution of resources.

Operationally, given the specificity of MPF distribution coefficients that vary from state to state, capital and countryside, we use the observed data proportion of resource allocation and use this ratio as a distribution rule.

The alternatives (table 6) imply options about how resources are distributed (always weighted by population):

- one portion locally, other equally and a third portion in accordance with the rules of MPF;
- one portion equally and another according to the rules of MPF;
- all locally; or
- all equally.

Consider the hypothesis that the division among municipalities is spatially artificial and economically speaking they compose a functional region (Ahrend *et al.*, 2014) or PCAs (IBGE, 2015). That is, citizens commute and produce freely among municipalities and make no distinction between working city and living city. Thus, we can say in a normative manner that the option of equally distribution along with the MPF division rules is optimal. On the one hand, it allows workers to benefit from the investments of the taxes collected for their production or consumption, regardless of the municipality of residence. On the other hand, it distributes resources progressively through FPM so that poorer regions with lower QLI can reach higher levels of quality of life. Chapter 5 simulates exactly if this would be the case for this model.

BOX 2
Operationalization of the distribution of resources collected from taxes

```
# Intuitively. When there is FPM to distribute, the FPM portion of the firm and
labor taxes is deduced and distributed. Subsequently, the equal portion is deduced
and equally distributed. Then, the local portion is deduced and distributed. All
investments are weighted by the population. When this last local portion does
not apply, it is again collected and equally distributed. When the distribution
is all local, the operation is done only in the municipality itself. When the
distribution is all the equally type, all the resources are collected from
the municipalities and redistributed.

Function: invest_taxes
# Function called at the end of each month from module simulation.py
For each municipality among the municipalities of the simulation:
      Collect population (t – 1)
      Perform population count and update population
      Collect current population (t)
If ALTERNATIVE0 and FPM_DISTRIBUTION are true:
      Call function distribute_fpm
Otherwise, if ALTERNATIVE 0 is true:
      Call function locally
Otherwise, if FPM_DISTRIBUTION is true:
      Call function distribute_fpm (with parameter: False)
If not:
      Call function equally

Function: distribute_fpm
For each municipality among the municipalities of the simulation:
      Collects the portion relative to the FPM and the equal portion (the
      municipality removes from its vault and sends over to the centralization)
      # Within each municipality, the portions relative to the FPM are withdrawn,
      according to the percentage of taxes related to work and firm and returned
      as a sum of both percentages.
      # The portions corresponding to the equal share are then deducted from
      taxes on consumption, labor, and firm
The total value of the FPM is then distributed according to the state proportion
of FPM actually observed for that year.
Function: equally is called with the portions related to consumption, labor, and
firm taxes
If parameter loc is true:
      Calls function locally
If not:
      Calls function equally, with parameter clean = true

Function: locally
For each municipality among the municipalities of the simulation:
      QLIs are updated, using ITBI and IPTU values. Then, the info is saved and
      the values for these taxes become zero

Function: equally
If the value passed as parameter is empty:
      For each municipality
           The remaining values are collected from the municipal vault
For each municipality
      Distributes the available amount proportionally to the population. Then, the
      values are saved and the values for these taxes in the vault becomes zero
```

Author's elaboration.

4 OPERATIONALIZATION

PolicySpace's proposal is to establish itself as a platform for broad analysis of ex-ante policy analysis. Thus, we built into the platform some options to run varying simulations.

The main call <`python main.py run`> runs the simulation once, for any PCA, with parameters set to default mode, or any other desired parameter setting. Please, feel free to tweak with parameters (and PCAs).

Many times [option: `-n X`]: In addition, it is possible to choose a number of times (*X*) to run the model, so that the variation derived from the stochasticity is incorporated. The results are plotted as means of the results, along with the results themselves. The default is one simulation only.

Numerous processors [option: `-c` Y]: Moreover, you can choose the number of processors (*cores*) to be used (Y). The default is the use all processors of your machine. As an example, the line: <`python main.py -c 2 -n 100 run`> runs one hundred simulations, using two processors and the default parameters chosen.

Sensitivity analysis [*sensitivity*]: cumulatively, the sensitivity analysis is available for Boolean parameters (*True/False*) or numerical parameters, in which the modeler specifies the parameter name, the starting value, ending value and the number of intervals. For example: <`python main.py -c 2 n 4 sensitivity ALPHA:0:1:7`> simulates the program `main.py` four times for each of the seven intervals of `ALPHA`, from 0 to 1, using two processors. That is, the model runs 28 times, for this example. For Boolean values, just type: <`python main.py sensitivity WAGE_IGNORE_UNEMPLOYMENT`> to simulate once true, once false.

Numerous PCAs [*acps*]: given the construction of *PolicySpace*, the PCAs work individually, that is, the trips, the changes and the interactions occur only among the municipalities of their own PCA. The command <`python main.py acps`> simulates successively, each one of the PCAs, so that the modeler has the data set relating to all PCAs. The modeler can also type in PCAs to exclude from the simulation and still run all the other ones that are not excluded.

Alternatively, *PolicySpace* also allows the simulation with more than one PCA. In this case the chosen PCAs are considered as one large metropolis with the municipalities of all named PCAs in which citizens commute freely among them all. The modeler could choose two nearby PCAs, such as Sorocaba and São José do Rio Preto, for example.

Distributions [*distributions*]: The code <`python main.py distributions`> runs all four alternatives for the distribution of taxes, as described in subsection 3.7.9: taxes and municipalities.

Finally, we have a simple version with a web interface in which the modeler select alternatives and parameters, run the model and see the results via browser. This possibility extends the access of possible users who do not feel able to operate the programming directly, through the `Python terminal`. Although python should be running in the background.[22]

5 OUTPUT DATA

The production of output data is customized in the parameters of the simulation in order to save only the necessary files and avoid producing unnecessary data. In the default mode, the full set of parameters used is saved and the following plots with average aggregated indicators are saved, namely:

- average household consumption;
- average level of prices;
- mean QLI values ;
- average number of workers per firm;
- average financial resources available to families;
- capital of firms in absolute terms;
- average profit of firms;
- GDP in absolute terms and in percentage of growth;
- GINI average;
- average inflation;
- average unemployment; and
- collected amount of tax by type of tax.

In the configuration of plots in which individual municipalities are shown, the information available is: *i)* GDP; *ii)* GINI; *iii)* residential values; *iv)* GDP *per capita*; *v)* population; *vi)* QLI; *vii)* total taxes collected; *viii)* unemployment; and *ix)* evolution of the demand for transport.

In addition, values are saved in CSV format, separated by semicolon, including average monthly values to each of the municipalities separately. In this case, the following information is available:

- demand for total transportation of the inhabitants of the municipality;
- population;

22. In such case use <python main.py web>.

- GDP;
- municipal GINI;
- average value of municipal real estate;
- municipal unemployment;
- Municipal QLI;
- GDP *per capita*; and
- Taxes collected for each one of the taxes available in the model (total, consumption, property, real estate transactions, firms, labor, and transfers to PMF).

Additionally, we plot seven figures of localized and normalized spatial information to the full sample for the last month of the simulation:

- balance of companies;
- number of employees of the companies;
- values of residences;
- number of members per family;
- firms' profits;
- taxes paid by companies; and
- families' savings.

Finally, five other data files can be selected and saved monthly or quarterly or annually (families, agents and firms with a higher computational cost) as described below.

- Firms: ID, municipality, address (longitude and latitude), total balance, number of employees, quantity in stock, quantity produced, prices, quantity sold, revenues, profits, and wages paid.
- Agents: ID, gender, address (longitude and latitude), age, qualification, company ID (if applicable), family ID, financial resources, accumulated consumption, and distance to work (if applicable).
- Agents that die throughout the simulation constitute a separate base containing the same variables (but with interruption of update of the dynamic variables).
- Residences: ID, address (longitude and latitude), size, prices, ID of the occupant family (when applicable), and municipality.

- Families: ID, current house prices, current city, savings, number of family members, and monthly consumption strategy.

REFERENCES

AHREND, R. *et al.* **What makes cities more productive?** Evidence on the role of urban governance from five OECD countries. Paris: OECD Publishing, 2014. (OECD Regional Development Working Papers, 5).

AXTELL, R. **Endogenous firms and their dynamics**. Fairfax: George Mason University, 2013. Available at: <http://www.acefinmod.com/docs/esrc/axtell-firms.pdf>. Accessed: Mar. 2016.

BERGMANN, B. R. A microsimulation of the macroeconomy with explicitly represented money flows. *In*: BERG, S. (Ed.). **Annals of economic and social measurement**, v. 3, n. 3. [s.l.]: NBER, 1974. p. 475-489.

BLINDER, A. S. On sticky prices: academic theories meet the real world. *In*: MANKIW, G. (Ed.). **Monetary policy**. Chicago: The University of Chicago Press, 1994. p. 117-154.

CARVALHO JUNIOR, P. H. B. de. **Aspectos distributivos do IPTU e do patrimônio imobiliário das famílias brasileiras**. Rio de Janeiro: Ipea, 2009. p. 53. (Texto para Discussão, n. 1417).

CASTRO, L. G. **Nota técnica sobre a metodologia adotada pelo Ministério da Previdência Social na extrapolação das Tábuas de Mortalidade IBGE para as idades acima de 80 anos**. Rio de Janeiro: Ministério da Previdência Social, 2015.

CERVERO, R. B. Linking urban transport and land use in developing countries. **Journal of Transport and Land Use**, v. 6, n. 1, p. 7-24, Apr. 2013.

CROOKS, A. T.; HEPPENSTALL, A. J. Introduction to agent-based modeling. *In*: HEPPENSTALL, A. *et al.* (Ed.). **Agent-based models of geographical systems**. London: Springer, 2012. p. 85-105.

DOSI, G.; FAGIOLO, G.; ROVENTINI, A. Schumpeter meeting Keynes: a policy-friendly model of endogenous growth and business cycles. **Journal of Economic Dynamics and Control**, v. 34, n. 9, p. 1748-1767, 2010.

FIRKOWSKI, O. L. C. F. de. Metrópoles e regiões metropolitanas no Brasil: conciliação ou divórcio. *In*: FURTADO, B. A.; KRAUSE, C.; FRANÇA, K. (Ed.). **Território metropolitano, políticas municipais**. Brasília: Ipea, 2013. p. 21-52.

FURTADO, B. A. Modeling social heterogeneity, neighborhoods and local influences on urban real estate prices: spatial dynamic analyzes in the Belo Horizonte metropolitan area, Brazil. **Netherlands Geographical Studies**, v. 385, 2009.

FURTADO, B. A.; EBERHARDT, I. D. R. A simple agent-based spatial model of the economy: tools for policy. **Journal of Artificial Societies and Social Simulation**, v. 19, n. 4, p. 12, 2016.

FURTADO, B. A.; EBERHARDT, I. D. R.; MESSA, A. **Seal's operating manual**: a spatially-bounded economic agent-based lab. Brasília: Ipea, 2017a. (Research Report).

______. Governance for smart cities: a spatially-bounded economic agent-based lab (Seal) to foster urban analysis and policy evaluation. *In*: AIJAZ, R. (Ed.). **Smart cities movement in Brics**. New Delhi: ORF-Global Policy, 2017b. p. 63-71.

FURTADO, B. A.; KRAUSE, C.; FRANÇA, K. C. **Território metropolitano, políticas municipais**: por soluções conjuntas de problemas urbanos no âmbito metropolitano. Brasília: Ipea, 2013.

FURTADO, B. A.; MATION, L.; MONASTERIO, L. Fatos estilizados das finanças públicas municipais metropolitanas brasileiras entre 2000 e 2010. *In*: FURTADO, B. A.; KRAUSE, C.; FRANÇA, K. (Ed.). **Território metropolitano, políticas municipais**. Brasília: Ipea, 2013. p. 291-312.

GAFFEO, E. *et al.* Adaptive microfoundations for emergent macroeconomics. **Eastern Economic Journal**, v. 34, n. 4, p. 441-463, 2008.

GARCIA-MAGARIÑO, I.; LACUESTA, R. Agent-based simulation of real-estate transactions. **Journal of Computational Science**, v. 21, p. 60-76, 2017.

GASPARINI, C. E.; MIRANDA, R. B. Transferências, equidade e eficiência municipal no Brasil. **Planejamento e Políticas Públicas**, n. 36, 17 out. 2011.

GRIMM, V. *et al.* A standard protocol for describing individual-based and agent-based models. **Ecological modeling**, v. 198, n. 1, p. 115-126, 2006.

______. The ODD protocol: a review and first update. **Ecological Modeling**, v. 221, n. 23, p. 2760-2768, 2010.

HAMILL, L.; GILBERT, N. **Agent-based modeling in economics**. Guildford: Wiley, 2016.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Censo demográfico 2000**. Rio de Janeiro: IBGE, 2003.

______. **Censo demográfico 2010**. Rio de Janeiro: IBGE, 2013.

______. **Arranjos populacionais e concentrações urbanas do Brasil**. Rio de Janeiro: IBGE, 2015.

JORDAN, R.; BIRKIN, M.; EVANS, A. Agent-based modeling of residential mobility, housing choice and regeneration. *In*: HEPPENSTALL, A. *et al.* (Ed.). **Agent-based models of geographical systems**. London: Springer, 2012. p. 511-524

LENGNICK, M. Agent-based macroeconomics: a baseline model. **Journal of Economic Behavior and Organization**, v. 86, p. 102-120, 2013.

MALPEZZI, S. Hedonic pricing models: a selective and applied review. *In*: GIBB, K.; O'SULLIVAN, A. (Ed.). **Housing economics and public policy**. Oxford, United Kingdom: Blackwell Science, 2002.

NADALIN, V.; IGLIORI, D. Empty spaces in the crowd: residential vacancy in São Paulo's city center. **Urban Studies**, v. 54, n. 13, p. 3085-3100, Sept. 2016.

NEUGART, M.; RICHIARDI, M. **Agent-based models of the labor market**. Moncalieri: Labor, 2012. (Working Papers, n. 125).

PEDROSO, F. F. F.; LIMA NETO, V. C. Transportes e metrópoles: um manifesto pela integração. *In*: FURTADO, B. A.; KRAUSE, C.; FRANÇA, K. (Ed.). **Território metropolitano, políticas municipais**. Brasília: Ipea, 2013. p. 195-224.

RODRIGUES, A. L. Ingovernabilidade metropolitana e segregação espacial: receita para a explosão da violência. *In*: FURTADO, B. A.; KRAUSE, C.; FRANÇA, K. (Ed.). **Território metropolitano, políticas municipais**. Brasília: Ipea, 2013. p. 53-82.

ROSEN, S. Hedonic prices and implicit markets: product differentiation in pure competition. **Journal of Political Economy**, v. 84, n. 1, 1974.

ROYER, L. de O. Municípios "autárquicos" e região metropolitana: a questão habitacional e os limites administrativos. *In*: FURTADO, B. A.; KRAUSE, C.; FRANÇA, K. (Ed.). **Território metropolitano, políticas municipais**. Brasília: Ipea, 2013. p. 157-194.

SEPPECHER, P.; SALLE, I.; LAVOIE, M. **What drives markups?** Evolutionary pricing in an agent-based stock-flow consistent macroeconomic model. Paris: CEPN, 2017.

CHAPTER 3

VALIDATION

Models are useful only for what they were designed (and therefore can only be tested in such scope). This is in the spirit of the famous and useful citation of Box and Draper (2007, p. 414): "(...) all models are approximations. Essentially, all models are wrong, but some are useful". Moreover, as explained in the introduction: "(…) a model is only as useful as the purpose *for which it is constructed*". (Crooks and Heppenstall, 2012, p. 98, our emphasis)". Further, a model is only useful when checked and validated (Manson, Sun and Bonsal, 2012).

In order to test such usefulness, the evaluation process includes calibration, verification and validation (Manson, 2007). Calibration refers to the adjustment of magnitude of the parameters of the model in such a way that the model produces results that follow the levels observed empirically. Sensitivity analysis of the model helps with both the calibration and robustness of the simulation and at the same time triggers the beginning of policy alternatives studies. Validation or internal verification, in turn, ensures that the code executes the algorithms and the equations as indicated by the modeler.

At such times, "mistakes" can occur, according to Galán *et al.* (2009). Mistakes are inconsistencies that occur in the communication between the conception, the modeling proposal and the development of the model program itself. In the case of *PolicySpace*, there was no distinction among these roles, and therefore the potential to generate miscommunication was nonexistent.

Verification – within the computer sciences – indicates that the code itself is doing exactly what the programmer thinks the code is doing (Hamill and Gilbert, 2016). Note that it is not a question of checking the results of the program for processing errors, or only if the program works. This is to ensure that the algorithm implemented in a given program language which embeds the rules, decision making, assumptions, steps, and criteria of the model are the same in the design of the modeler and in the actual execution of the program, deterministically. Checking is ensuring that the results are exactly those predicted analytically.

Validation verifies whether the model can reproduce stylized facts, given a set of parameters (Fagiolo and Roventini, 2017). Therefore, this step demonstrates the consistency of the model and the available empirical data and it is a necessary requisite for the analysis of public policies, as Marks (2013, p. 41) states:

> although demonstrative simulation models are useful, not least at performing 'what if' exercises of exploration of different models, policy analysis requires validated, descriptive simulation models.

Simply put: "From a mathematical perspective, validation is the process of assessing whether or not the quantity of interest (QOI) for a physical system is within some tolerance – determined by the intended use of the model" (NRC, 2012, p. 5). In fact, Gatti *et al.* (2011) name their validation chapter as "empirical validation" and seek an "acceptable representation" of the observed system (Gatti *et al.*, 2011, p. 86).

This acceptable representation of results, however, must be free from "artifacts", which are changes in results given from ancillary assumptions. Those accessories may seem irrelevant, but may generate significant changes (Galán *et al.*, 2009). Hence, the modeler should know which assumptions generate which results.

Alternatively, Moss (2008) also supports validation by means of which he calls companion modeling as described by the French school. Companion modeling consults experts and stakeholders to validate both the process and the results of agent-based modeling. More recently, this same strand of literature is described as *participatory modeling* (Bommel *et al.*, 2014; Henly-Shepard, Gray and Cox, 2015).

Validation also refers to the ability to capture the key elements of the modeled phenomenon, and to guarantee that a model is useful, valid – in essence – for a certain question. Gräbner (2015) and Mäki (2009) argue that the model should not only represent reality in its simplified form, but also resemble reality. Thus, such resemblance enables exploration and consequently better understanding of the modeled phenomenon. Mäki (2009) calls these models surrogates. Those are different from the models Mäki calls substitute models. The latter's objective is to get to the results, independently and disconnected of the understanding of the mechanisms and of reality.

See and Ngo (2012, p. 183), in the line of Zeigler (1976), propose three categories of validation.

- Replicative: in which results are compared to actual data.
- Predictive: in which the model is able to make predictions about behaviors according to theory.
- Structural: wherein, in addition to replicating behaviors, it also replicates the mechanisms that produce the results.

Note that only replicative validation may not be sufficient. Gräbner (2015) suggests that the model can generate comparable conditions to those observed using incorrect mechanisms, or make wrong predictions using the correct mechanisms.

A broader view for validation shall ensure the reasonability and adequacy on three axes: *i)* input data; *ii)* the development process; and *iii)* the quality of the resulting descriptive data in relation to observed empirical data.

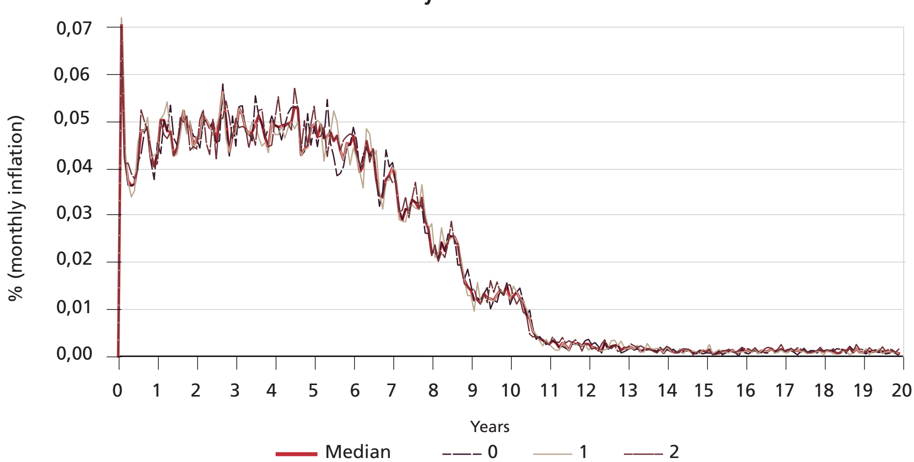
In the case of *PolicySpace's* proposal, official sources of empirical information provide the necessary input data; the development process and its assumptions align with the literature described in chapter 2. In the remainder of this chapter we present some results, their behavior and contrast them to the available empirical data. We also provide some comparative tests that include the set of all ACPs. In addition, chapter 4 presents sensitivity analyzes that reinforce the idea of robustness of the model, while illuminating aspects of the phenomenon studied.

1 ILLUSTRATION: MACROECONOMIC BEHAVIOR

The proposal does not have credit market nor does it model a system with interests. The macroeconomic analysis, therefore, is only illustrative and serves as a parameter of reasonability of the model. A complete macroeconomic model could follow the validation proposed by Guerini and Moneta (2017).

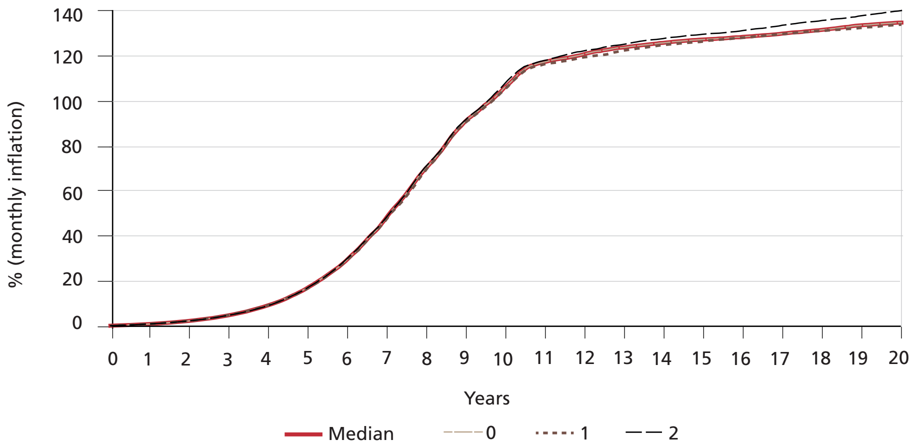
Inflation, for example, behaves better in the model when compared to the real case, with a monthly average of 0.02 and a standard deviation of 0.02 over the twenty-year period. Typically, the simulation shows greater variability in the first half of the analyzed period and soothes its behavior in the second period. In the case of the PCA of Belo Horizonte, with 2% of the population, see graph 1 for inflation and graph 2 for the price index.

GRAPH 1
Belo Horizonte: estimates of monthly inflation



Author's elaboration.
Note: Data for 2% of population and default configuration.

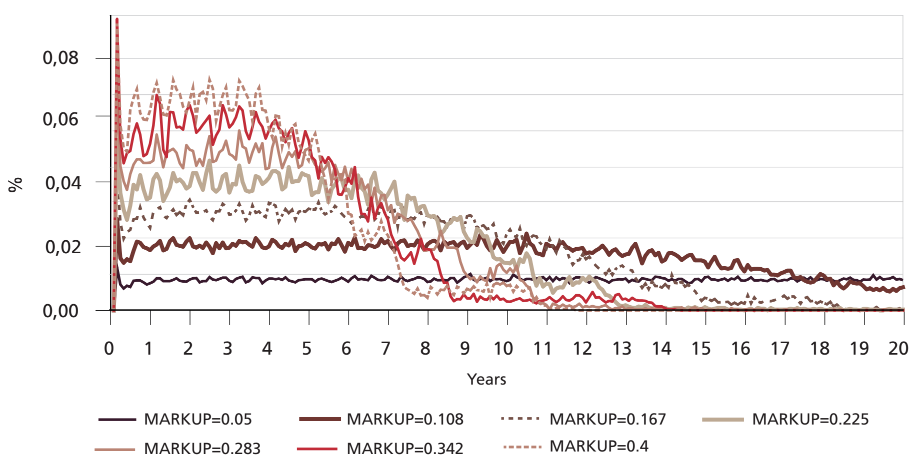
GRAPH 2
Belo Horizonte: estimates of price index evolution



Author's elaboration.
Note: Data for 2% of population and default configuration.

Yet, sensitivity analysis allows us to check how some parameters such as *markup*, for example, influence the behavior of prices. Notice how higher values of *markup* lead to higher inflation in the early years, however, a larger reduction in the final years compared to lower levels of *markup* (graphs 3 and 4).

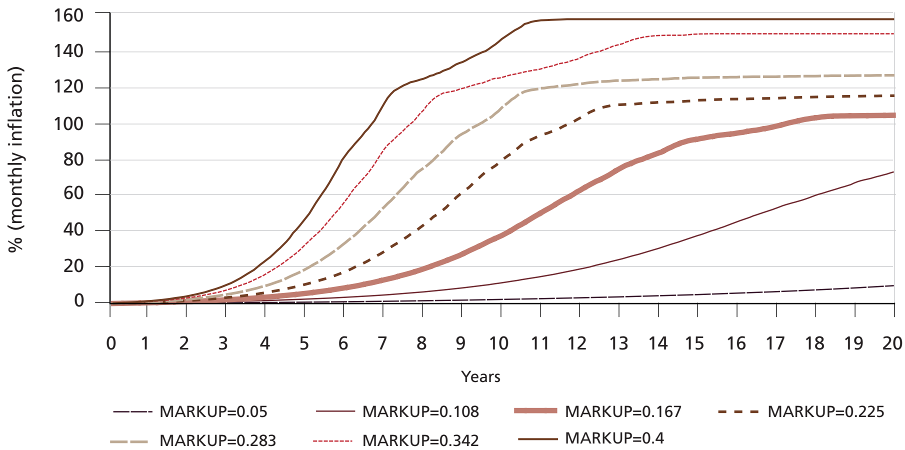
GRAPH 3
Brasília: sensitivity analysis with parameter *markup* – monthly inflation estimates



Author's elaboration.
Note: Data for 1% of the population and default configuration.

GRAPH 4

Brasília: sensitivity analysis with parameter *markup* – price index estimates

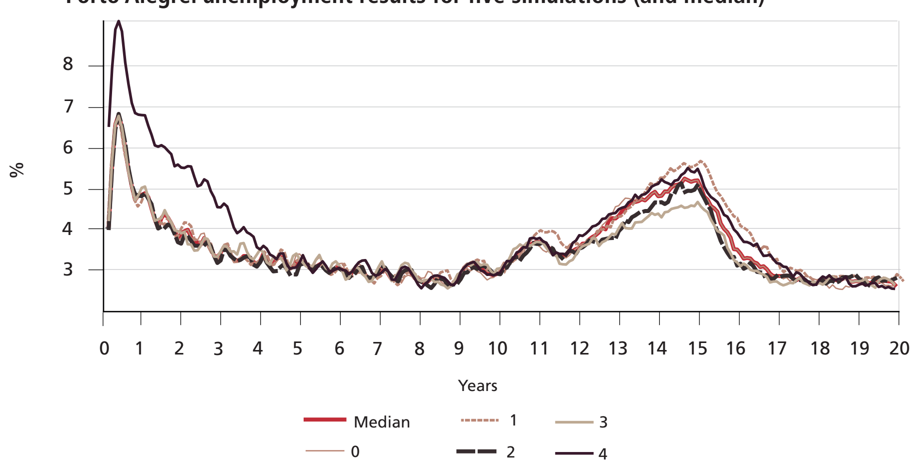


Author's elaboration.
Note: Data for 1% of the population and default configuration.

Unemployment in the simulation is much more detailed than the official information allows us to verify. In the simulation, unemployment is monthly and by municipalities. We observe significantly different values among the core municipality (by construction, with greater number of employers) and other municipalities. The results of the simulations indicate that unemployment remains consistent to expected values, varying between 3% and 10% over the period, with some relevant variability (graph 5).

GRAPH 5

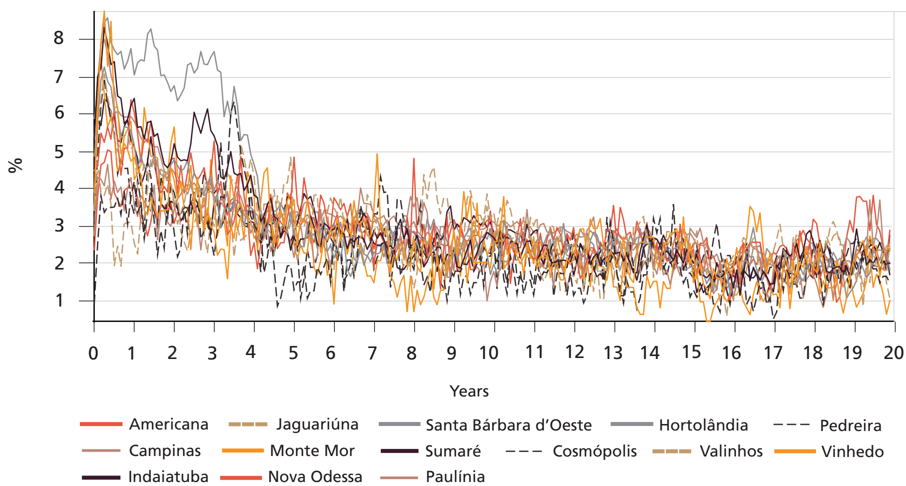
Porto Alegre: unemployment results for five simulations (and median)



Author's elaboration.
Note: Data for 2% of the population.

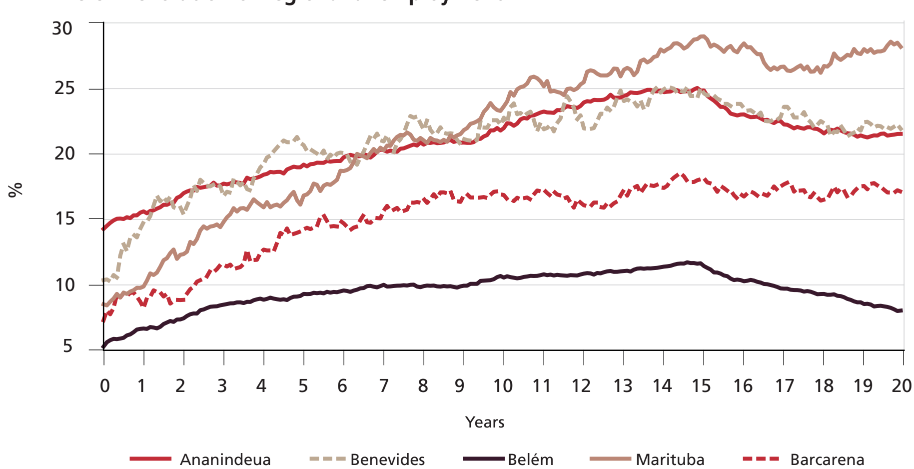
Considering all municipalities, there is greater variability and maintenance of unemployment at higher levels in less populous municipalities. Graph 6 illustrates this pattern for the case of Campinas, and graph 7, for the case of Belém.

GRAPH 6
Campinas: unemployment by municipality



Author's elaboration.
Note: Data for 2% of the population and default configuration.

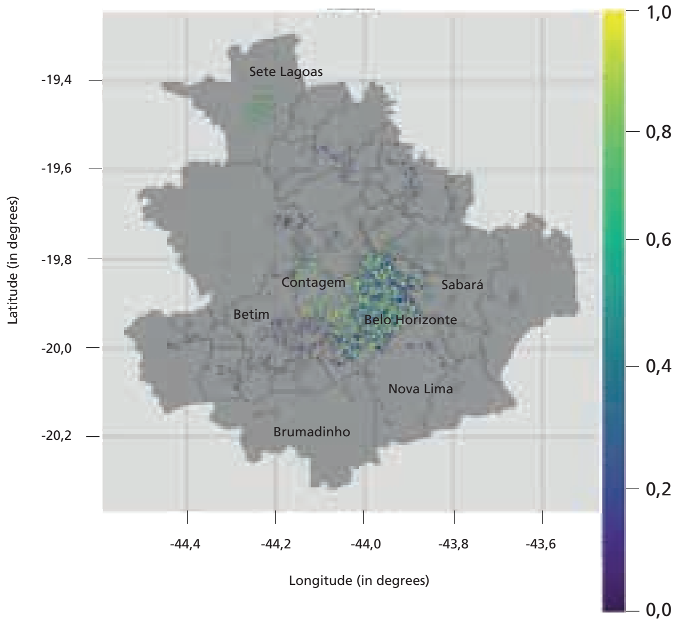
GRAPH 7
Belém: evolution of regional unemployment



Author's elaboration.
Note: Data for 2% of the population.

Alternatively, the job map (figure 1) shows that not only are there fewer companies in the peripheral municipalities, but these companies have relatively fewer employees than the others (purple balls in isolated urban spots and west of the main spot).

FIGURE 1
PCA of Belo Horizonte: normalized spatialization of jobs, by firms



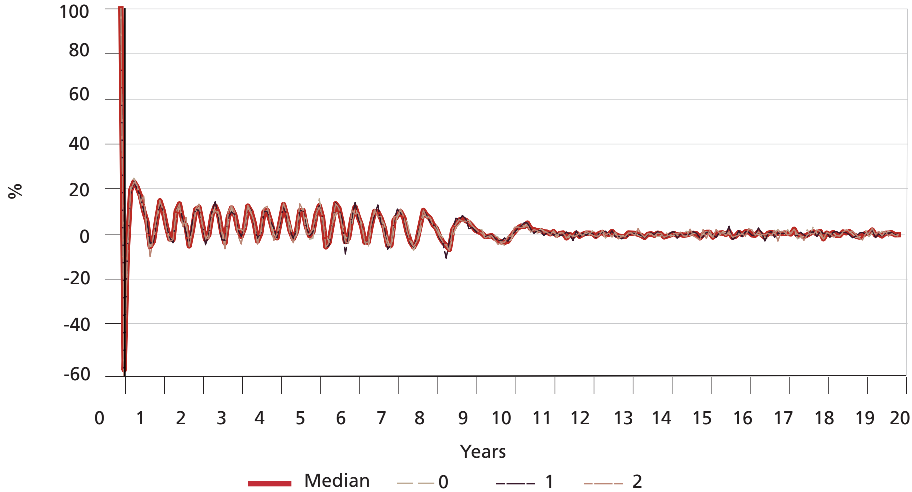
Author's elaboration.
Notes: 1. Data for 2% of the population.
2. The areas in dark gray are official urban spots.

1.1 Illustration: production

As mentioned, macroeconomic analysis is only illustrative. However, production growth is relevant to the central analysis of the research problem, namely, the analysis of tax behavior and tax collection on labor, goods and real estate markets. This subsection evaluates the simulated results of the gross domestic product (GDP) and the ratio of taxes to GDP and ration of taxes to total taxes collected.

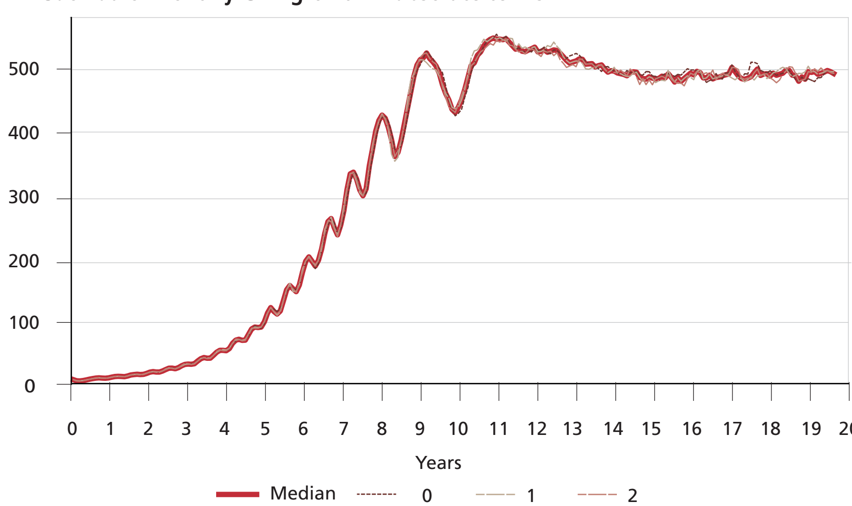
Brazilian GDP grows from 2000 onwards to 2009 losing strength in the following years. *PolicySpace* also simulates a more vigorous growth in the first ten years of the model which then becomes stable (graph 8 and graph 9). Graph 10 illustrates the difference between peripheral municipal GDPs in relation to capital cities, for the PCA of Belém.

GRAPH 8
São Paulo: monthly variation of GDP



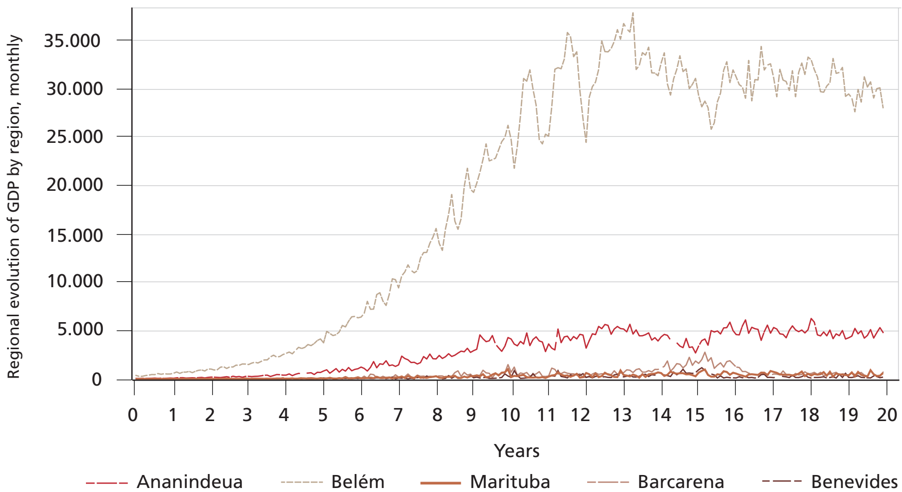
Author's elaboration.
Note: Data for 2% of the population, three simulations and default configuration.

GRAPH 9
São Paulo: monthly GDP growth in absolute terms



Author's elaboration.
Note: Data for 2% of the population, three simulations and default configuration.

GRAPH 10
PCA of Belém: evolution of municipal GDP



Author's elaboration.
Note: Data for 2% of the population.

2 TAXES

The model involves the simulation of five major categories of taxes. The Land Tax on Property (IPTU) is of immediate and generalized application to all residences in the model. In practice, municipalities have difficulties to operationalize and collect such taxes. There are relevant heterogeneities among cities – larger and smaller – as to which ones that are able to collect the taxes properly and those that collect them more erratically and sparsely (Carvalho Junior, 2016).

As the tax is independent of families' income or purchase, when families do not have disposable income at the time of the tax collection (all members unemployed at a given month), the tax is waived. This implementation is compatible with the observed heterogeneity of collection and effective payment capacity of households.

Tax on Property Transfers (ITBI) is charged at the time of the real estate transaction and it is mandatory for all transactions, with total collection capacity. Both ITBI and IPTU are collected (and distributed) in the municipality in the default configuration.[1]

Three other simplified groups complete the tax system. Their empirical basis is:

1. *Alternative0 and FPM_distribution are True.*

1) *Labor*: refers to the collection of tax on workers' income. It is an attempt to resemble Income Tax on Individuals (IRPF). According to Afonso, Soares and Castro (2013) the tax corresponds to 4.71% of the total in the form of wages and more 7.07% in the form of income and gains, representing 11.78% of the total collected (table 1).
2) *Firm*: Taxation on the payment of employers' salaries (20.3%) added to corporate income taxes (11.51%).
3) *Consumption*: taxation on goods and services, which together account for 44.98% of total revenues, according to Afonso, Soares and Castro (2013).

TABLE 1
Proportions of municipal taxes in relation to GDP and total collected

In relation to GDP	Afonso, Soares and Castro (2013)	Simulated	STN
Labor	0.0024	0.0105	
Firm	0.0065	0.0173	
Consumption	0.0093	0.0334	
Property (IPTU)	0.0007	0.0056	0.0044
Transaction (ITBI)	0.0008	0.0012	0.0012
Total collected	**0.0207**	**0.0679**	**0.0869**
FPM		0. 0534	0.0184
In proportion to the total			
Labor	0.1178	0.1522	
Firm	0.3181	0.2595	
Consumption	0.4498	0.4819	
Property (IPTU)	0.0356	0.0885	0.0919
Transaction (ITBI)	0.0399	0.0179	0.0119
FPM		0.4303	0.4120

Source: Afonso, Soares and Castro (2013, p.12) and STN.
Author's elaboration.

In addition, the tax system includes the Municipal Participation Fund (FPM), which is a municipal transfer from the Union calculated from 23.5% of the Tax on Industrialized Products (IPI) and 23.5%[2] of Income Tax (IR). In *PolicySpace*, the rate is applied on *Labor* and *Firm*. The FPM is not estimated by Afonso, Soares and Castro (2013), but it is available as real transfers from STN data.

Data of table 1 refer to the 333 municipalities of the 46 ACPs used for both simulated and STN data. Simulations refer to the average of three simulations and 2% of the population for all the ACPs.

2. The rate is 23.5% as of September 2007, in accordance with Constitutional Amendment No. 55. For simulation purposes, the rate is 23.5% for the entire period.

The main objective of the simulation with respect to taxes is to obtain an adequate proportion in relation to GDP. That is, the taxed proportion of the economy is compatible with that observed empirically. In table 1, we can observe not only the levels presented by Afonso, Soares and Castro (2013), but also the actual results – obtained through the STN – that compose exactly the same municipalities of the simulation.

The numbers suggest that the simulation replicates the tax collection to a reasonable level of similarity with empirical data, with about 7 % of GDP. In addition, property tax (IPTU) and transactions tax (ITBI) behave very closely to observed values.

Regarding the proportion of each tax among the five classifications used, similar numbers are obtained. Taxes on consumption comprise the large portion in which IPTU and ITBI have lower relevance. The values on the workers are close to those in the literature. The collected MPF in the simulation is also compatible to the observed ratio, according to the STN.

We understand that this comparison of empirical and simulated results is sufficiently similar in proportion to those observed. Thus, enabling the use of the simulation as a reasonable proxy for the set of taxes described in the model.

3 VALIDATION OF PCAS

We applied the Kolmogorov-Smirnov (KS) test to the set of ACPs. KS compares the equality between the accumulated probability distributions using as test statistic the longest distance between them. In the end, KS checks whether two independent samples can be of the same continuous distribution (Massey, 1952; Smirnov, 1939).

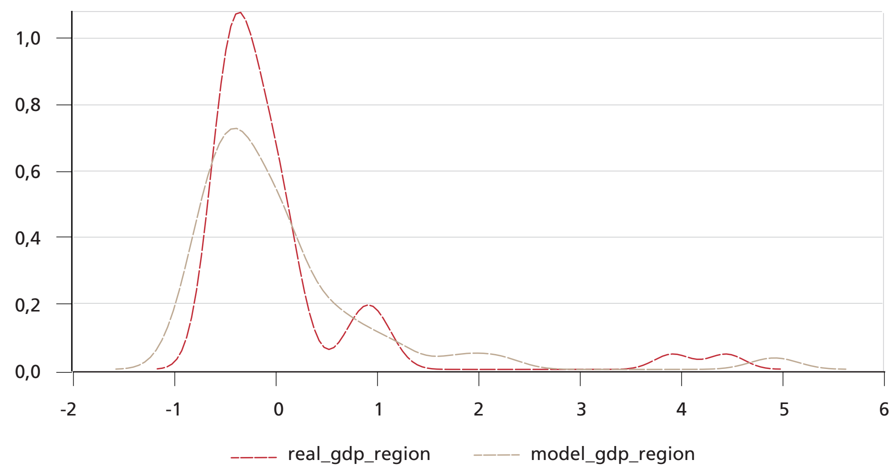
The rejection of the null hypothesis that the two samples belong to the same distribution occurs for *p*-values lower than 0.05, at 5%. The *p*-values greater than 0.05, therefore, indicate that one cannot reject that the samples come from the same distribution. In addition, the visual analysis and overlapping of the normalized distributions help to illustrate the differences between the two samples.

We compared the results for the set of all ACPs 46 simulated with 2% of the population, the average of three simulations for each PCA and with default configuration. The observed data comes from STN.

We were able to compare total tax revenue, GDP, the amount of transferred MPF, property taxes (IPTU) and real estate transactions (ITBI). GDP has a *p*-value of 0.32 and a certain likelihood between observed and simulated data. Empirical data, however, indicates a higher concentration of GDP from a few PCAs, which leads to a more elongated distribution when compared to the simulated data (graph 11).

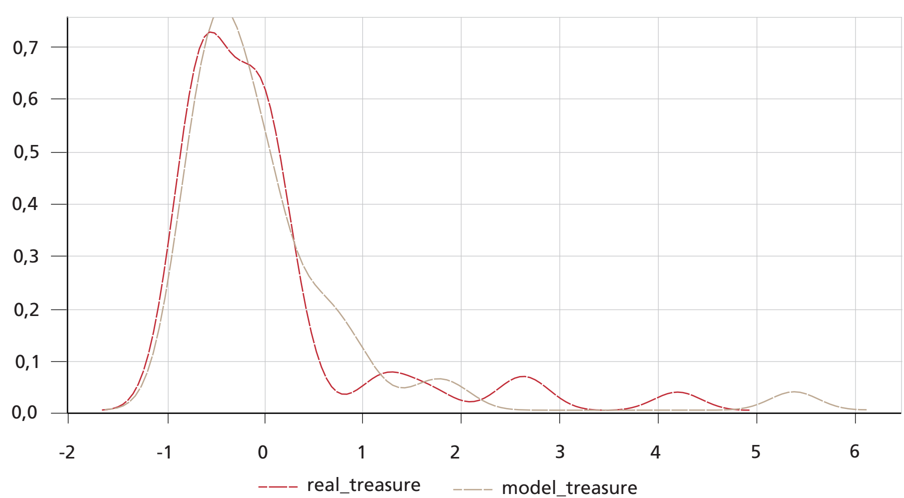
The total tax revenue, and the property taxes and FPM transactions, however, appear quite adjusted in the comparison between actual data and simulated data with high values of p -value for total collection, and FPM property tax and a little down to the ITBI (graph 12).

GRAPH 11
Comparison between observed and simulated GDP for all ACPs

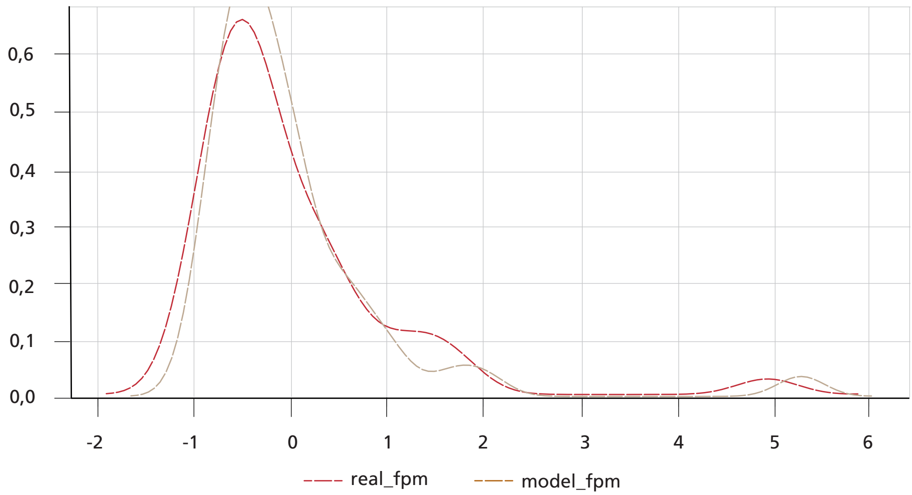


Author's elaboration.

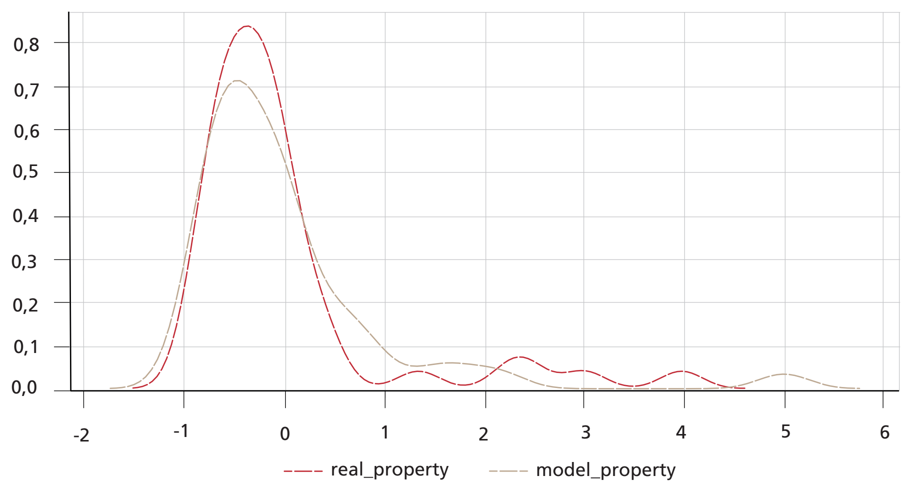
GRAPH 12
Comparison between observed and simulated values for all ACPs
12A – Total collected tax



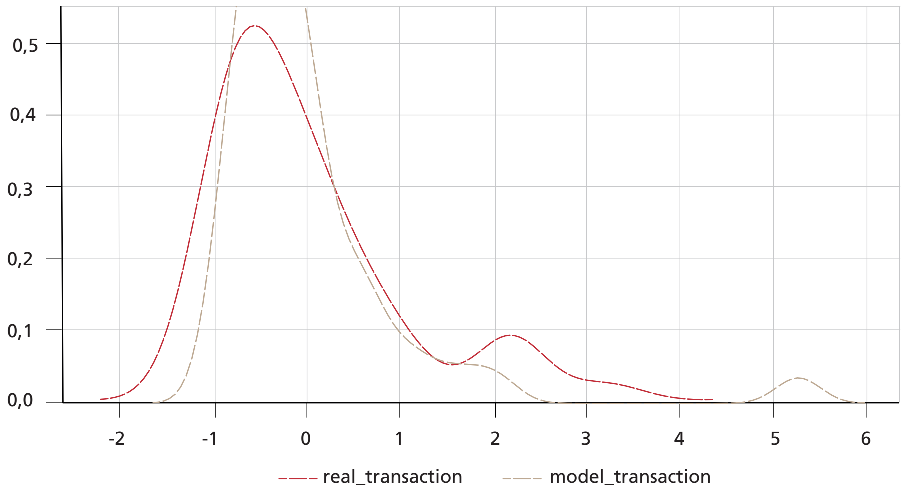
12B – FPM



12C – IPTU



12D – ITBI

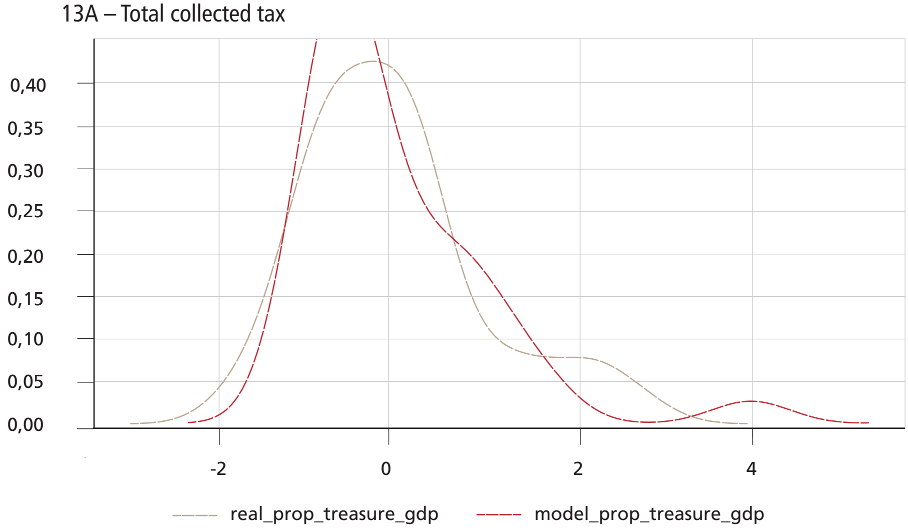


Author's elaboration.

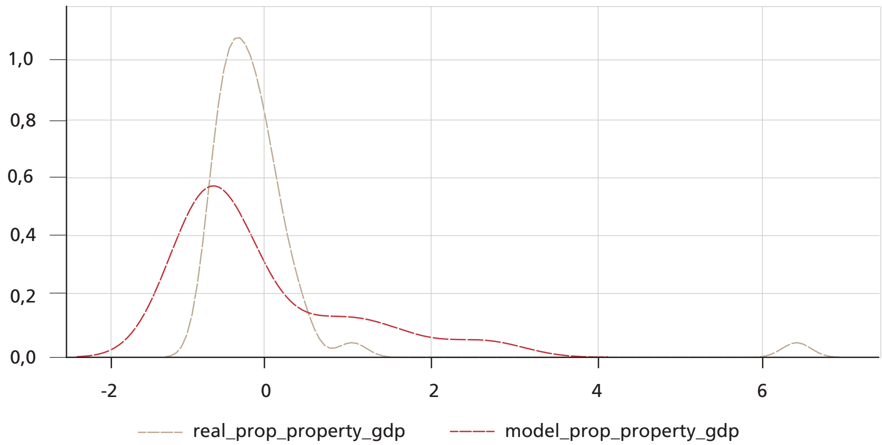
The comparison of taxes regarding the GDP presents a somewhat lower similarity. The proportion of the total tax collection presents a similar distribution between the real and the simulated one, with the observed results slightly more inclined to the right and significant in the KS test (graph 13). The FPM and ITBI proportions are slightly skewed to the right in relation to the observed data, ITBI *p*-value of 0.0583 boundary.

GRAPH 13
Comparisons between observed and simulated proportions in relation to GDP for all ACPs

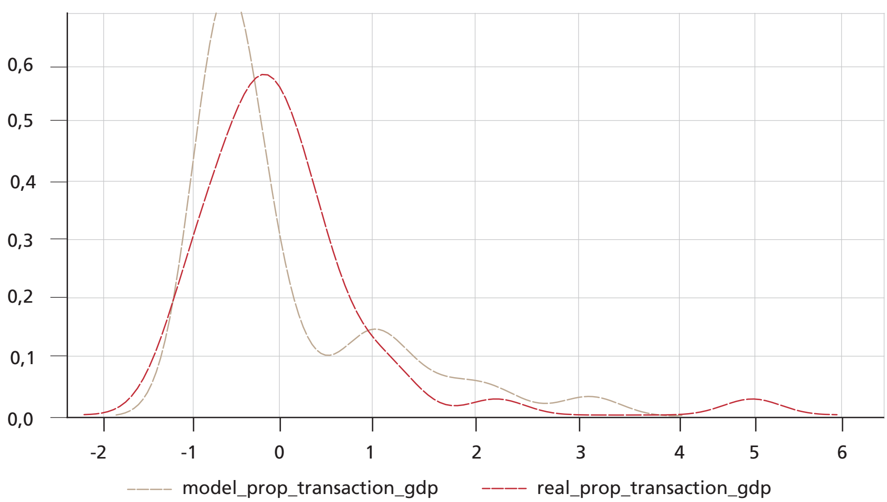
13A – Total collected tax



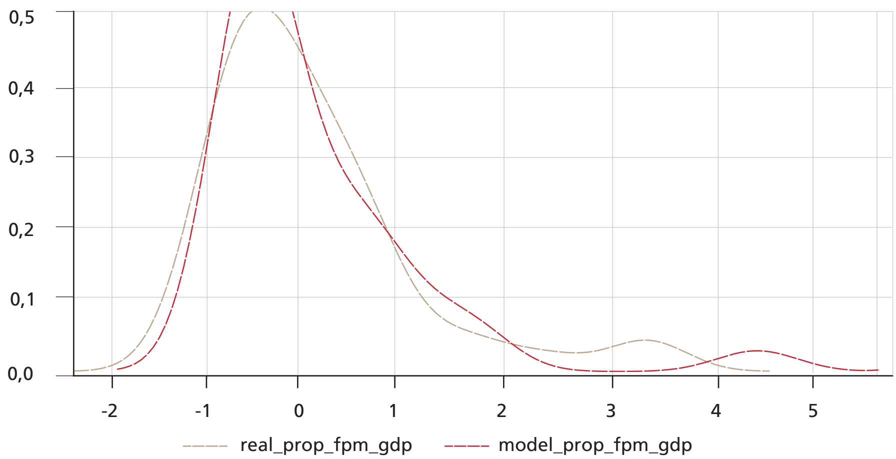
13B – IPTU



13C – ITBI



13D – FPM



Author's elaboration.

IPTU did not appear significant in the test, indicating that the observed and simulated distributions are different. Note that the elongated distribution of the observed data reveals the concentration of the collection in specific ACPs, while the simulated data, flattened, explain the presence of the tax collection rule in a more general way among all ACPs.

Again, given the heterogeneity of municipalities and ACPs in the Brazilian case and given the objective of the proposal – which is to allow indications about the effect, direction and order of magnitude of fiscal public policy alternatives – we consider that *PolicySpace* has enough elements that replicate observed data to be considered valid.

REFERENCES

AFONSO, J. R. R.; SOARES, J. M.; CASTRO, K. P. de. **Avaliação da estrutura e do desempenho do sistema tributário brasileiro**: livro branco da tributação brasileira. Brasília: BID, 2013.

BOMMEL, P. *et al.* A further step towards participatory modelling: fostering stakeholder involvement in designing models by using executable UML. **Journal of Artificial Societies and Social Simulation**, v. 17, n. 1, p. 6, 2014.

BOX, G. E. P.; DRAPER, N. R. **Response surfaces, mixtures, and ridge analyses**. Hoboken: John Wiley and Sons, 2007.

CARVALHO JUNIOR, P. H. B. de. A administração tributária do IPTU e seu impacto na efetivação do estatuto da cidade. *In*: COSTA, M. A. (Org.). **O estatuto da cidade e a habitat III**. Brasília: Ipea, 2016. p. 207-232.

CROOKS, A. T.; HEPPENSTALL, A. J. Introduction to agent-based modelling. *In*: HEPPENSTALL, A. *et al.* (Ed.). **Agent-based models of geographical systems**. London: Springer, 2012. p. 85-105.

FAGIOLO, G.; ROVENTINI, A. Macroeconomic policy in DSGE and agent-based models redux: new developments and challenges ahead. **Journal of Artificial and Social Simulation**, v. 20, n. 1, 2017.

GALÁN, J. M. *et al.* Errors and artefacts in agent-based modelling. **Journal of Artificial Societies and Social Simulation**, v. 12, n. 1, Jan. 2009.

GATTI, D. D. *et al.* **Macroeconomics from the bottom-up**. London: Springer, 2011.

GRÄBNER, C. **Methodology does matter**: about implicit assumptions in applied formal modelling – the case of dynamic stochastic general equilibrium models vs agent-based models. [s.l.]: MPRA, 2015.

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

HAMILL, L.; GILBERT, N. **Agent-based modelling in economics**. Hoboken: John Wiley and Sons, 2016.

HENLY-SHEPARD, S.; GRAY, S. A.; COX, L. J. The use of participatory modeling to promote social learning and facilitate community disaster planning. **Environmental Science and Policy**, v. 45, p. 109-122, Jan. 2015.

MÄKI, U. Missing the world: models as isolations and credible surrogate systems. **Erkenntnis**, v. 70, n. 1, p. 29-43, Jan. 2009.

MANSON, S. M. Challenges in evaluating models of geographic complexity. **Environment and planning B**: Urban Analytics and City Science, v. 34, n. 2, p. 245-260, Apr. 2007.

MANSON, S. M.; SUN, S.; BONSAL, D. Agent-based modeling and complexity. *In*: HEPPENSTALL, A. *et al.* (Ed.). **Agent-based models of geographical systems**. London: Springer, 2012. p. 125-139.

MASSEY, F. J. Distribution table for the deviation between two sample cumulatives. **The Annals of Mathematical Statistics**, v. 23, n. 3, p. 435-441, Sept. 1952.

MOSS, S. Alternative approaches to the empirical validation of agent-based models. **Journal of Artificial Societies and Social Simulation**, v. 11, n. 1, Jan. 2008.

NRC – NATIONAL RESEARCH COUNCIL. **Assessing the reliability of complex models**: mathematical and statistical foundations of verification, validation, and uncertainty quantification. Washington: The National Academies Press, 2012.

SEE, L.; NGO, T. A. Calibration and validation of agent-based models of land cover change. *In*: HEPPENSTALL, A. *et al.* (Ed.). **Agent-based models of geographical systems**. London: Springer, 2012. p. 181-197.

SMIRNOV, N. V. Estimate of deviation between empirical distribution functions in two independent samples. **Bulletin Moscow University**, Moscow, v. 2, n. 2, p. 3-16, 1939.

ZEIGLER, B. P. **Theory of modelling and simulation**. Hoboken: John Wiley and Sons, 1976.

CHAPTER 4

APPLICATIONS AND SENSITIVITY ANALYSIS

This chapter presents some sensitivity analyses that implicitly serve as initial testing of the model and application illustrations. The modeler tests the parameters in search of those that together lead to a validated model and at the same time broadens his or her comprehension to the influence of each parameter on the simulation as a whole. Alternatively, in the words of O'Sullivan *et al.* (2016, p. 4):[1] sensitivity analysis "globally or sequentially tests each model parameter to measure its impact on model outcomes". Sensitivity analysis is the systematic variation of parameters and rules in order to understand the effects of small changes on results.

A robust model, at the end of the sensitivity analysis, is usually one whose behavior does not change disruptively with the small variation of a parameter or a behavior rule. Except, for those models that contain points of criticality, in which structural change is the expected behavior. As an example of the latter, think of the numbers of cars in a highway. A specific number of cars' density triggers road congestion and sudden speed decrease in the road. In these cases, agent-based modeling can contribute to the accurate identification of such parameter.

Usually, sensitivity analysis includes, in addition to the variation of the parameters themselves, the variation of the initial conditions and the inherent variability of the random processes of the simulation.

It is common knowledge that every process that includes random values (more precisely, pseudorandom) in computational simulations originates in an initial value – the *seed* – that triggers the deterministic pseudorandom process. The modeler who desires different random values for each simulation should allow *seed* variation.

In the case of *PolicySpace*, most initial conditions are given, since they originate in actual year 2000 data, according to the description (see chapter 2). The parameters are tested in the sensitivity analysis (as well as some rules of behavior), and the typically stochastic variability is tested through the analysis of multiple simulations, with variations in the *seed*, observing the median results (see chapter 3).

1. O'Sullivan, D. *et al.* Strategic directions for agent-based modeling: avoiding the YAAWN syndrome. *Journal of Land Use Science*, v. 11, n. 2, p. 177-187, Mar. 2016.

Sensitivity analysis contribute to the validation of the model. In addition, it can also contribute to the understanding of the phenomenon. That is, once the model is calibrated – the parameters are set to mimic a real situation – and the model has been verified and validated, the variation of parameters can be used to analyze answers to "what if" questions. For example, imagine a model calibrated for a given market (default configuration), in a specific time and space. Then changes in a parameter of a tax rate – property tax, for instance – generates changes in the results that can be compared with previous results. These exercises are performed for the *PolicySpace* case in this chapter and in the next one.

1 DISTANCE AS A SELECTION CRITERION IN THE LABOR MARKET

In order to contribute to the robustness of the model and verify that the results of a simulation are not spurious, it is possible to test the presence or absence of certain rules and assumptions. We describe two examples below. A third one is presented in the form of an application in chapter 5.

PolicySpace considers the distance from the residence of the job's candidate to the hiring firm as a criterion in the selection process. As stated earlier, this feature of the model depicts both the likelihood of candidates seeking positions near their homes, but also the fact that Brazilian companies have to legally take responsibility for transport expenses of workers whose commuting costs exceeds 6% of their salaries.

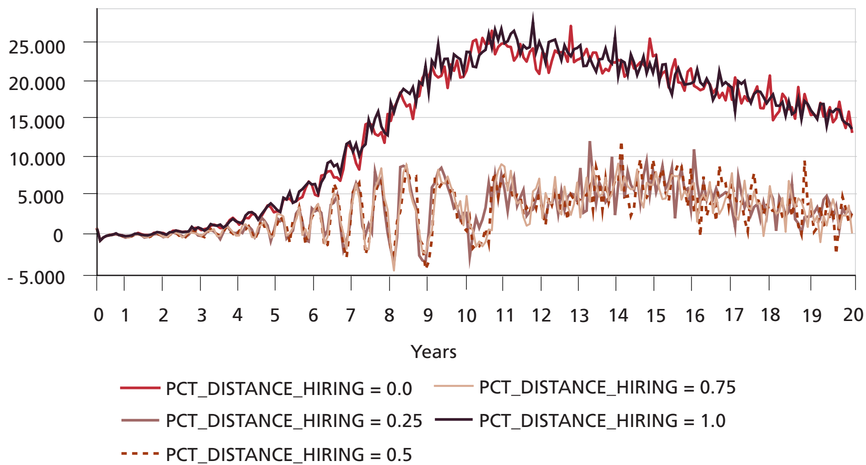
However, the modeler may consider this assumption irrelevant or inaccurate. Given that the rule is implemented in the model via a parameter (*pct_distance*), it is up to the modeler to decide how relevant he or she wants the rule to be. The modeler may decide, for instance, that no firm uses the criterion. Alternatively, he or she may set the rule so that all hiring occurs via distance and candidate's employee is not observed in the selection process.

The sensitivity analysis for the PCA of Belo Horizonte, with 2% of the population and average of three simulations per parameter tested, produced the following results (graph 1).

Note that the behavior for the maximum and minimum values of the parameter (only distance as criterion: *pct_distance*=1, or only qualification: *pct_distance*=0) are markedly different from the results of intermediate values of the parameters.

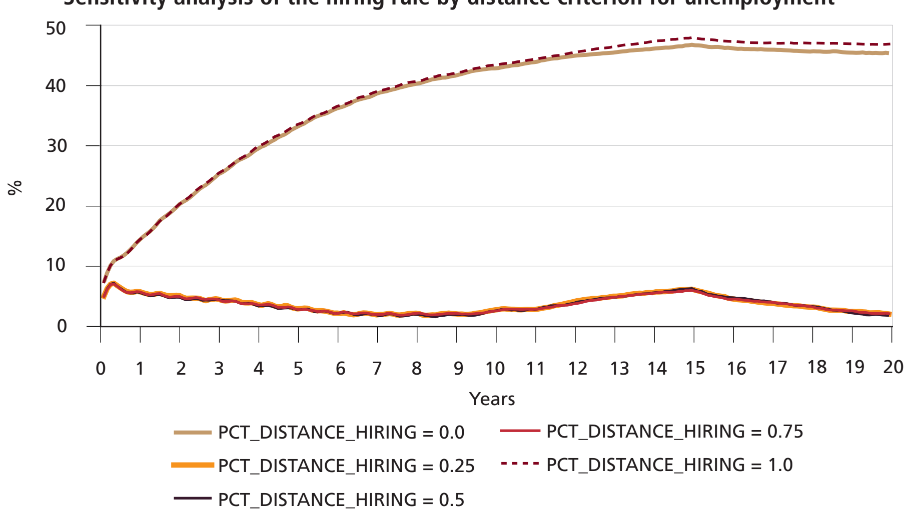
In fact, results seem to indicate that, although firms have higher profits (graph 1), unemployment remains at higher levels (graph 2) which leads to greater inequality (graph 3). Those are probably because only few citizens succeed in the selection process, either because they live nearby or because they are more qualified). With this mismatch, production (graph 4) and the *quality of life index* (QLI) are lower for both extremes, but even lower for the criterion by distance alone. Intermediate values (although closer to one or the other extreme) seem to allow greater general adjustment in the economy, generating results very close to each other. That suggests that the rule is relevant for the adjustment of the model, although the actual value of the parameter influences results just a little.

GRAPH 1
Sensitivity analysis of the hiring rule by distance criterion 'for firms' profit



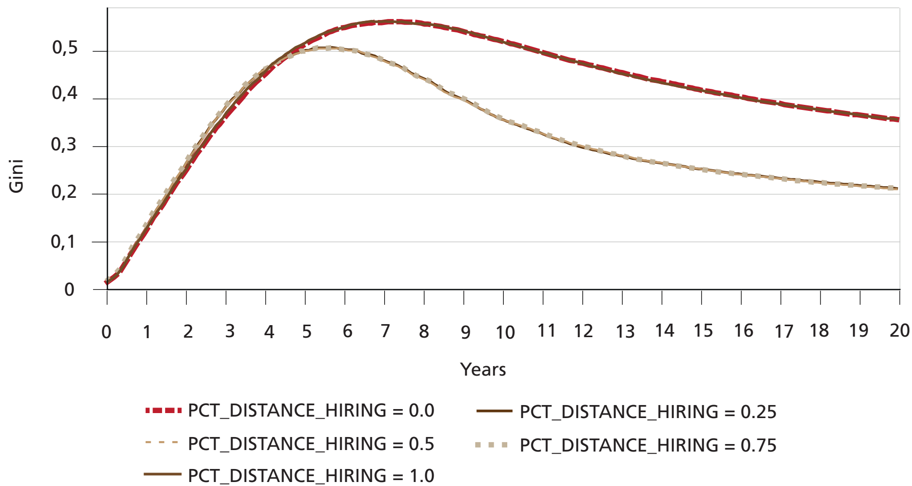
Author's elaboration.
Note: Results for Belo Horizonte, with 2% of the population, average of three simulations.

GRAPH 2
Sensitivity analysis of the hiring rule by distance criterion for unemployment



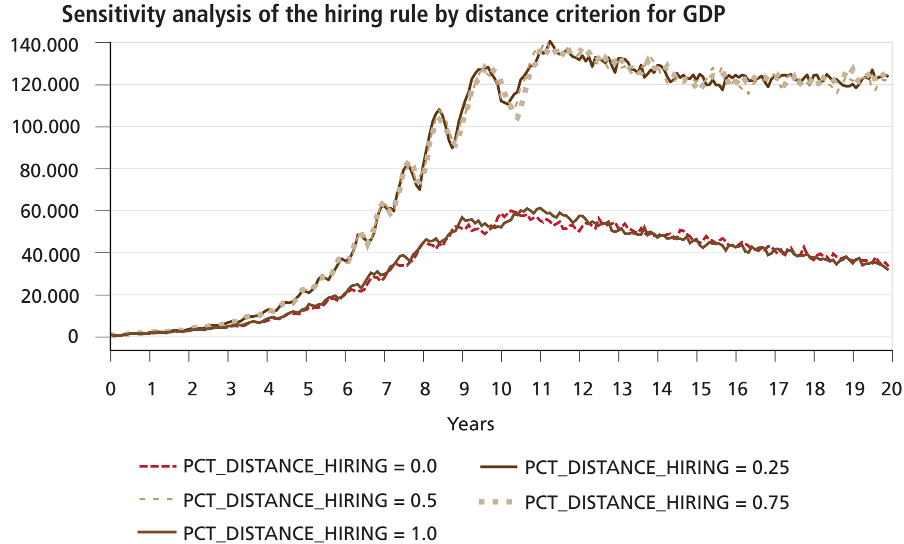
Author's elaboration.
Note: Results for Belo Horizonte, with 2% of the population, average of three simulations.

GRAPH 3
Sensitivity analysis of the hiring rule by distance criterion for the Gini indicator



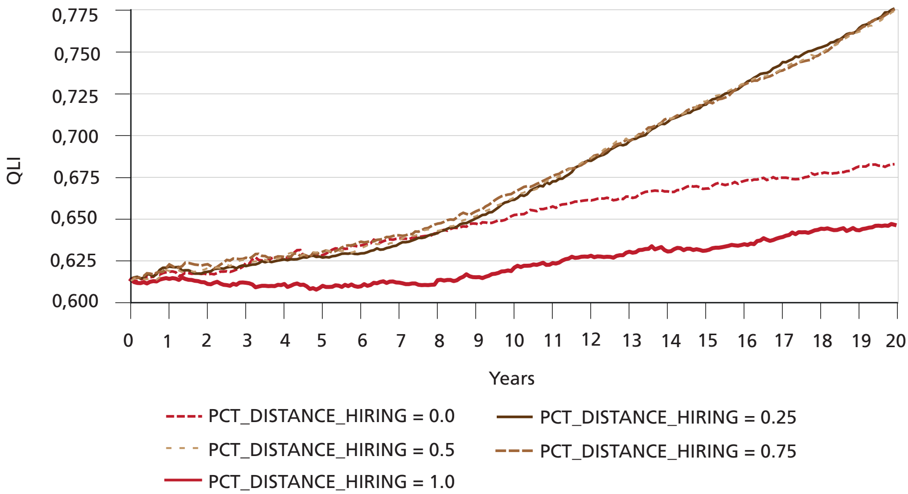
Author's elaboration.
Notes: 1. The Gini indicator is a measure of inequality that can be used for any distribution, although it is commonly used to measure the inequality of income distribution.
2. Results for Belo Horizonte, with 2% of the population, average of three simulations.

GRAPH 4
Sensitivity analysis of the hiring rule by distance criterion for GDP



Author's elaboration.
Note: Results for Belo Horizonte, with 2% of the population, average of three simulations.

GRAPH 5
Sensitivity analysis of the hiring rule by distance criterion for average QLI



Author's elaboration.
Note: Results for Belo Horizonte, with 2% of the population, average of three simulations.

2 UNEMPLOYMENT AS A CRITERION FOR WAGE SETTING

Another assumption of the model is that firms make decisions on wage setting weighted by observed unemployment in the PCA. The sensitivity analysis presents the results when this rule is ignored – that is when *wage_ignore_unemployment* is set to True in contrast to the default configuration of False.

Graph 6 confirms that when the option is True, firms pay in wages all of its revenues after taxes and therefore there is no increase of capital, which remains the same for all the simulation. On the contrary, when the parameter is False, the firm pays lower wages when there is unemployment, and much smaller when there is higher unemployment. As a result, the wage and tax deduction is less than the revenues on average, and the capital of the firms accumulates.

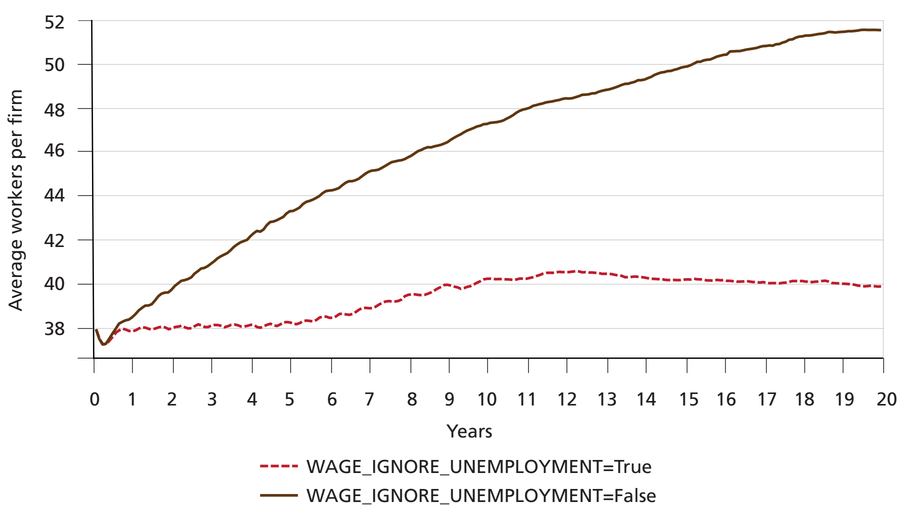
The higher pay of generalized wages leads to a lower concentration of employees per company (graph 7), lower profits for companies individually (graph 8), production of the larger region (graph 9), although with unemployment in higher levels (graph 10).

These two examples of adapting rules within *PolicySpace* serve simultaneously as a test of the mechanisms examined in the model and an indication that the rules bring greater reasonability and likelihood of the behavior of the economy to the model. Thus, making the model both more comprehensible and comparable to the observed reality. After these sensitivity exercises, we chose to use the default setting of 0.17 for the *percentage_distance_hiring* parameter, and False for the *wage_ignore_unemployment* parameter.

GRAPH 6
Sensitivity analysis of wage setting rule on firms' capital

Author's elaboration.
Note: Results for Belo Horizonte, with 2% of the population, average of three simulations.

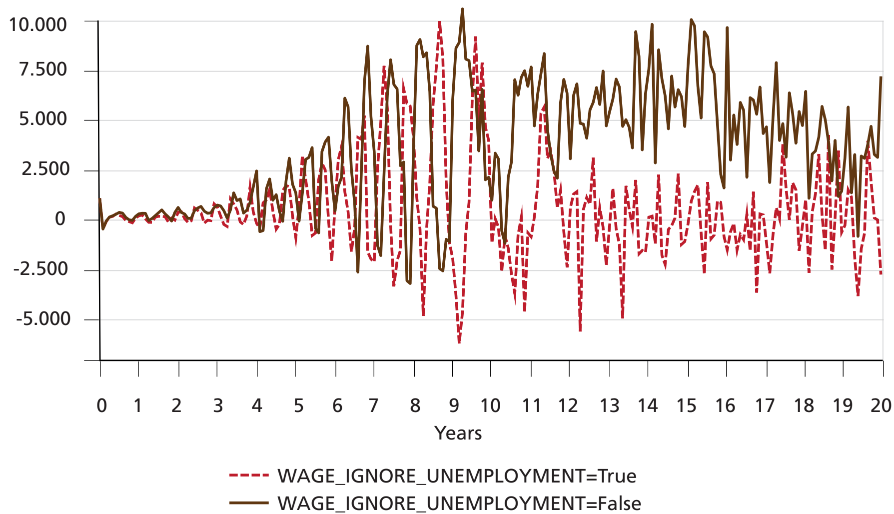
GRAPH 7
Sensitivity analysis of wage setting rule on average number of workers per firm



Author's elaboration.
Note: Results for Belo Horizonte, with 2% of the population, average of three simulations.

GRAPH 8

Sensitivity analysis of wage setting rule on firms' profit

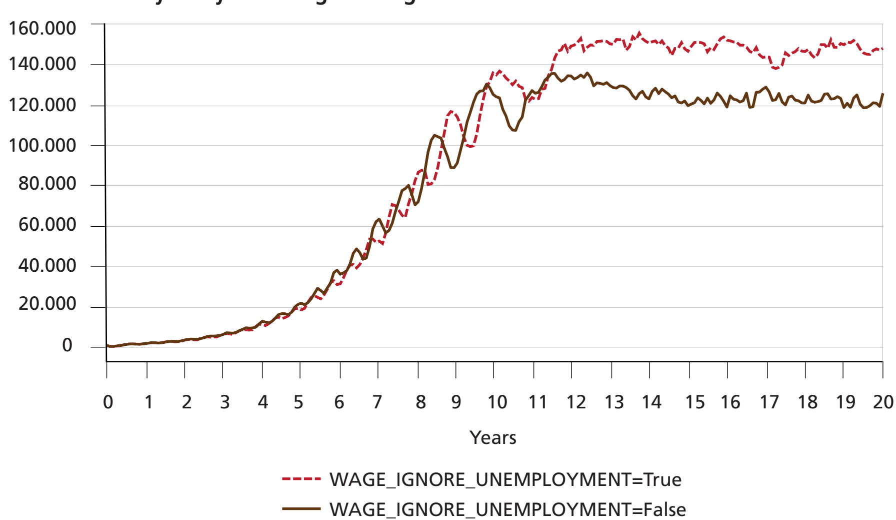


Author's elaboration.

Note: Results for Belo Horizonte, with 2% of the population, average of three simulations.

GRAPH 9

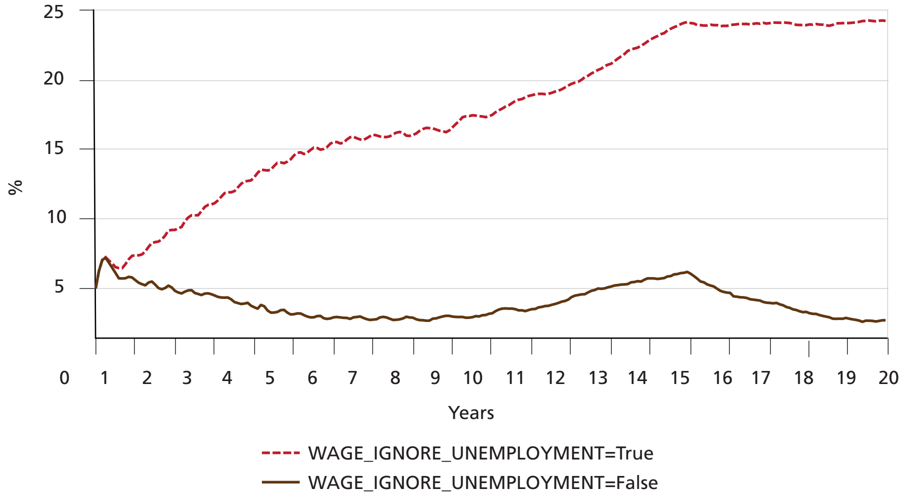
Sensitivity analysis of wage setting rule on GDP



Author's elaboration.

Note: Results for Belo Horizonte, with 2% of the population, average of three simulations.

GRAPH 10
Sensitivity analysis of wage setting rule on unemployment



Author's elaboration.
Note: Results for Belo Horizonte, with 2% of the population, average of three simulations.

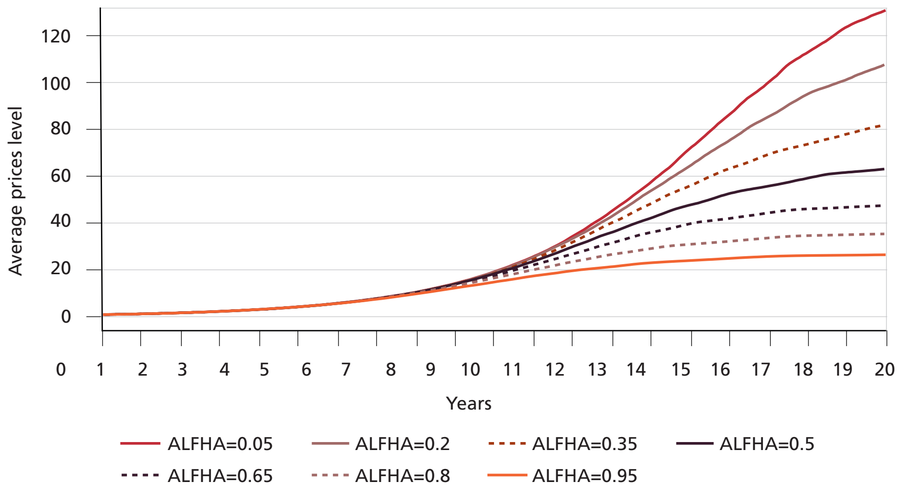
In the remainder of the chapter, we test:

- alpha that refers to workers' productivity;
- beta that is the families' propensity to consume parameter;
- family size;
- percentage of the population sample used in the simulation;
- the size of the real estate vacancy in the market;
- the frequency with which firms evaluate the labor market; and
- the relevance of firms' rigidity in changing prices.

3 PRODUCTIVITY

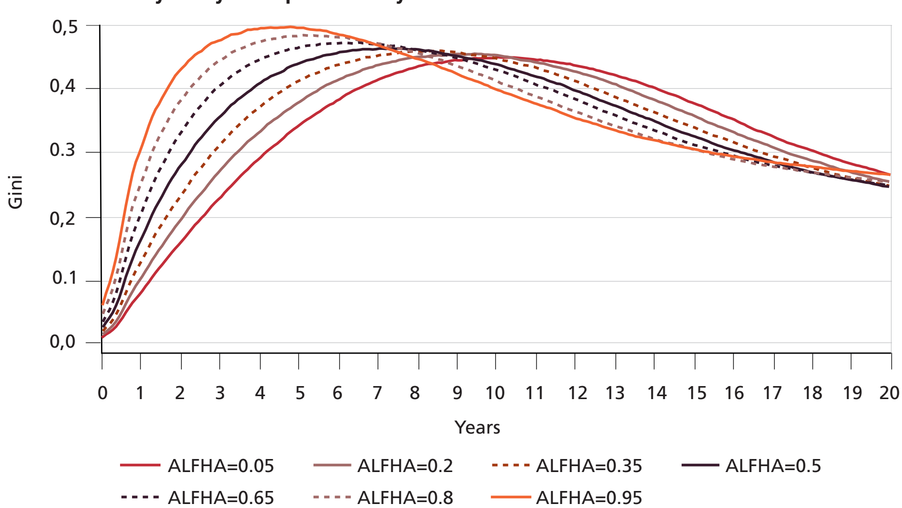
The highest worker productivity – measured by the *alpha* parameter – leads to lower prices (graph 11). Although with a medium-term impact, these lower prices do not lead to lower inequality (graph 12), higher GDP (graph 13) or lower unemployment (graph 14) at the end of the period. The higher consumption and dynamism of the economy, however, leads to higher accumulated levels of consumption and therefore tax collection and high QLI.

GRAPH 11
Sensitivity analysis of productivity on price levels



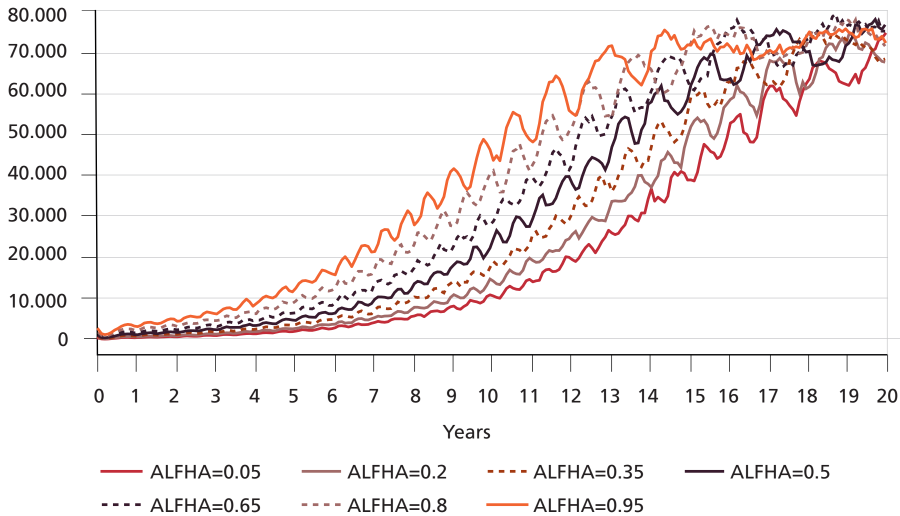
Author's elaboration.
Note: Results for Curitiba, with 2% of the population, average of four simulations.

GRAPH 12
Sensitivity analysis of productivity on the Gini indicator



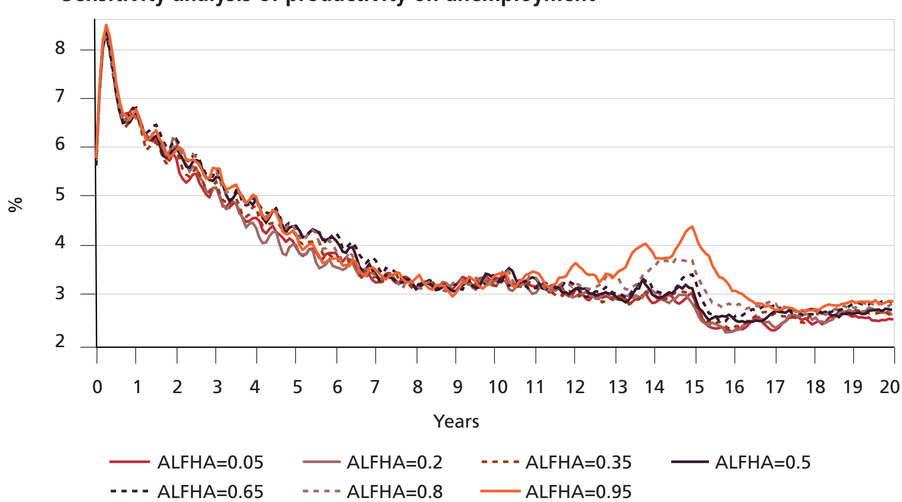
Author's elaboration.
Note: Results for Curitiba, with 2% of the population, average of four simulations.

GRAPH 13
Sensitivity analysis of productivity on GDP



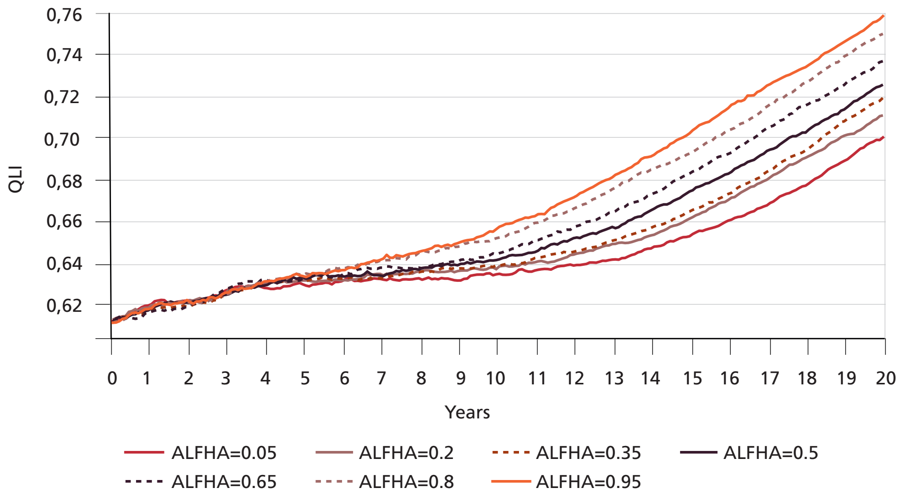
Author's elaboration.
Note: Results for Curitiba, with 2% of the population, average of four simulations.

GRAPH 14
Sensitivity analysis of productivity on unemployment



Author's elaboration.
Note: Results for Curitiba, with 2% of the population, average of four simulations.

GRAPH 15
Sensitivity analysis of productivity on QLI



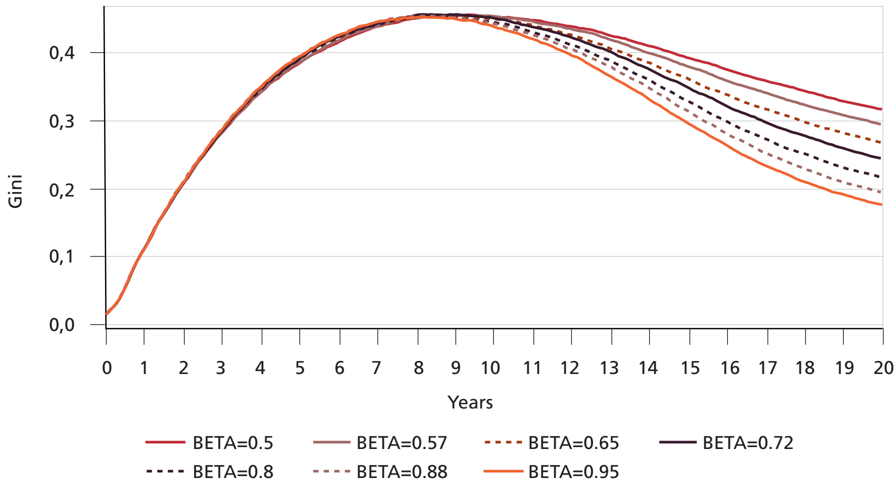
Author's elaboration.
Note: Results for Curitiba, with 2% of the population, average of four simulations.

4 PROPENSITY TO CONSUME

Families' propensity to consume – measured by the *beta* parameter – divides the families monthly available income into a consumption part and a savings deposit that is eventually used in the real estate market.

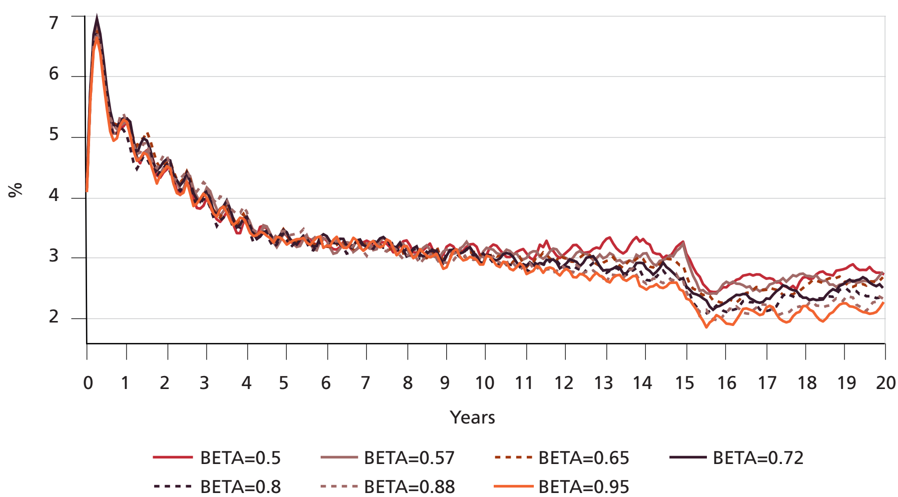
Higher levels of consumption reduce inequality (graph 16), perhaps because unemployment is maintained at slightly lower levels (graph 17). However, higher consumption also leads to a significant increase in prices (graph 18) and, therefore, to a greater variability in firms' profits (graph 19).

GRAPH 16
Sensitivity analysis of the propensity to consume on the Gini indicator



Author's elaboration.
Note: Results for Porto Alegre, with 2% of the population, average of four simulations.

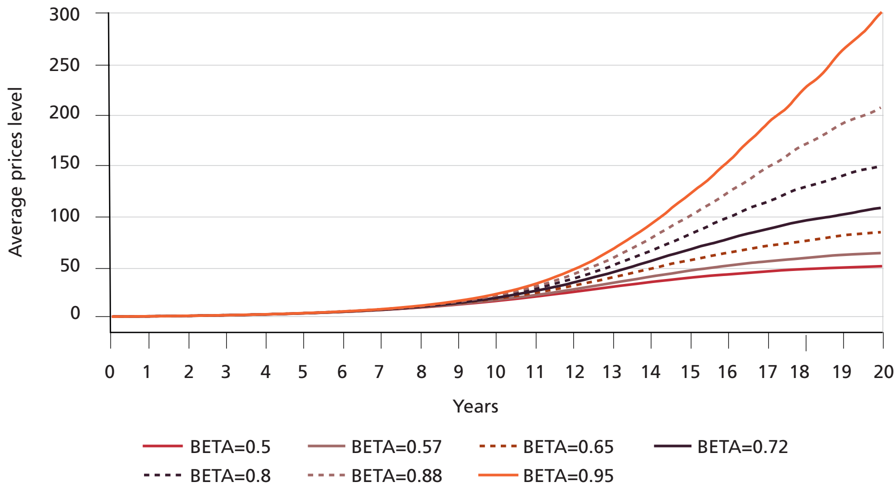
GRAPH 17
Sensitivity analysis of the propensity to consume on unemployment



Author's elaboration.
Note: Results for Porto Alegre, with 2% of the population, average of four simulations.

GRAPH 18

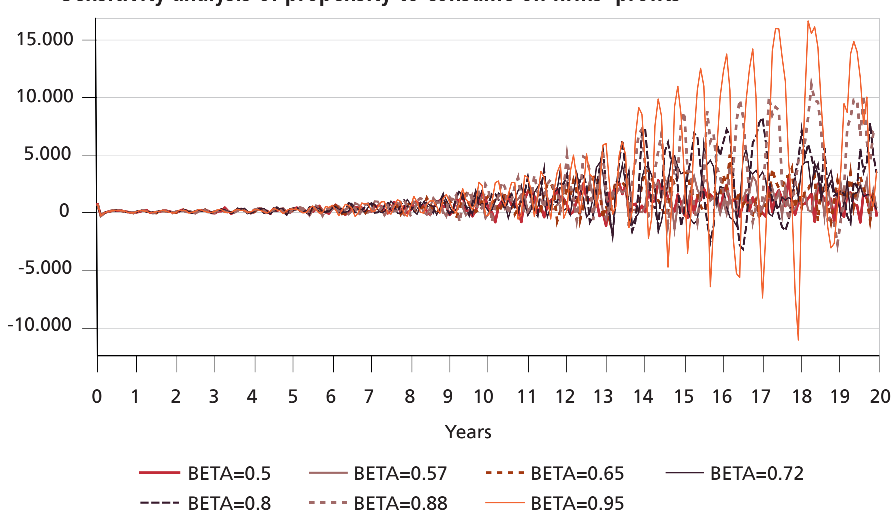
Sensitivity analysis of the propensity to consume on price levels



Author's elaboration.
Note: Results for Porto Alegre, with 2% of the population, average of four simulations.

GRAPH 19

Sensitivity analysis of propensity to consume on firms' profits



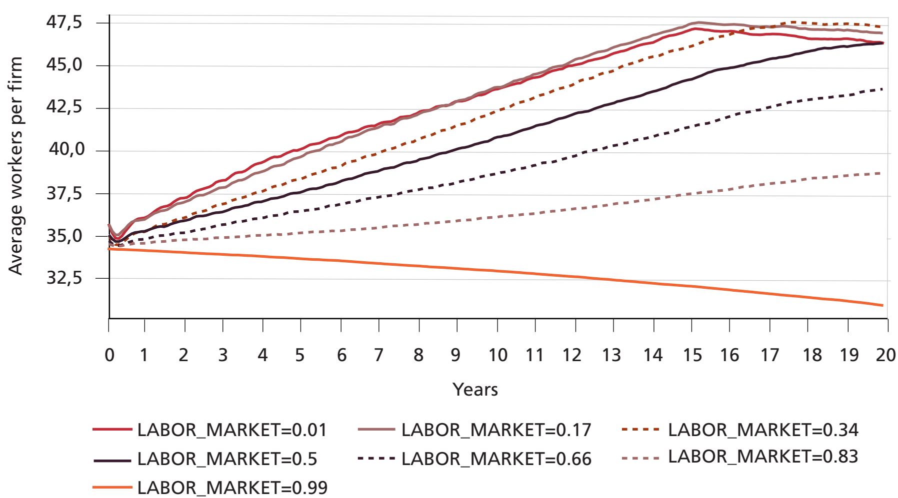
Author's elaboration.
Note: Results for Porto Alegre, with 2% of the population, average of four simulations.

5 FREQUENCY OF FIRMS IN THE LABOR MARKET

Labor_market is a parameter that controls the frequency with which firms enter the market. When the parameter is low, firms enter the market – hiring or firing, often – when, on the contrary, the value is very high, firms almost do not enter the labor market, and there is greater rigidity in work behavior.

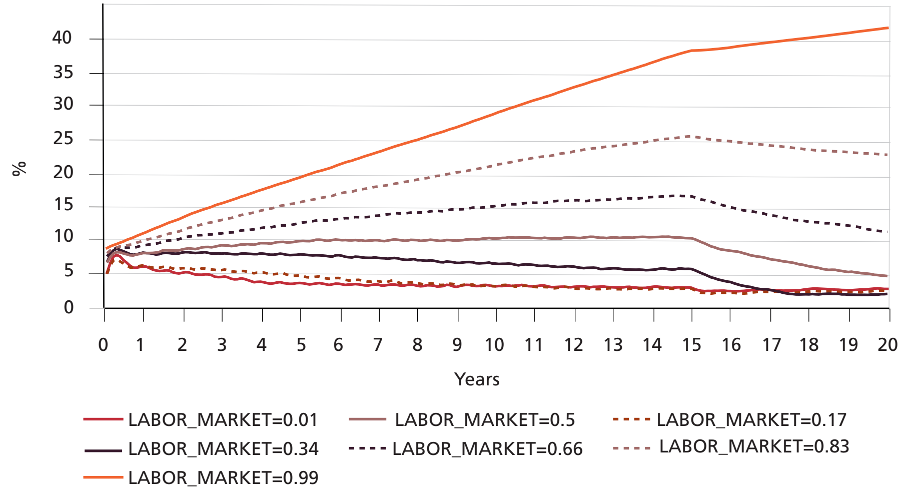
When the labor market is more dynamic and the entry of firms happen on a regular basis, there is a greater number of workers per firm (graph 20). Such concentration of workers (possibly better qualified) in firms that pay higher wages leads to smaller unemployment (graph 21) e inequality (graph 22). Together these factors lead to a higher absolute GDP in the period of 20 years when the market is dynamic and smaller GDP when there is rigidity (graph 23).

GRAPH 20
Sensitivity analysis of the dynamism of the labor market on the average number of workers per firm



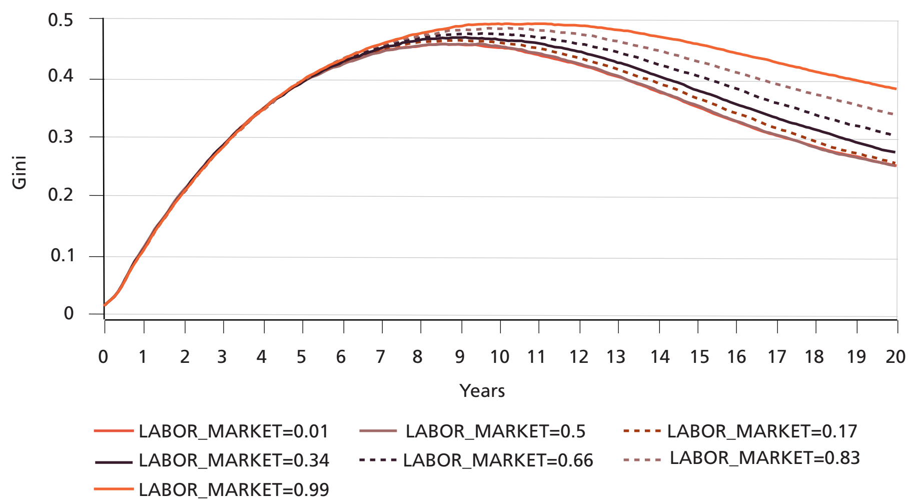
Author's elaboration.
Note: Results for Porto Alegre, with 2% of the population, average of four simulations.

GRAPH 21
Sensitivity analysis of the dynamism of the labor market on unemployment



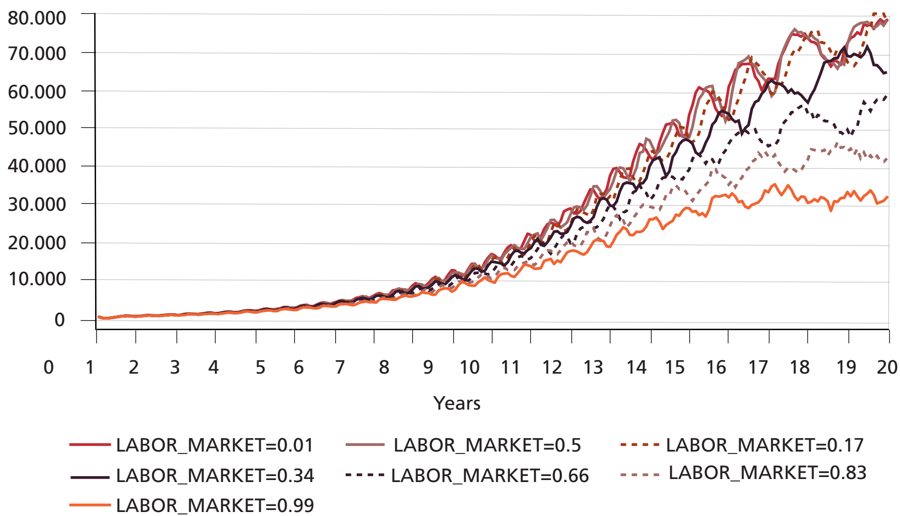
Author's elaboration.
Note: Results for Porto Alegre, with 2% of the population, average of four simulations.

GRAPH 22
Sensitivity analysis of the dynamism of the labor market on the Gini indicator



Author's elaboration.
Note: Results for Porto Alegre, with 2% of the population, average of four simulations.

GRAPH 23
Sensitivity analysis of the dynamism of the labor market on GDP



Author's elaboration.
Note: Results for Porto Alegre, with 2% of the population, average of four simulations.

6 RIGIDITY IN PRICE ADJUSTMENT

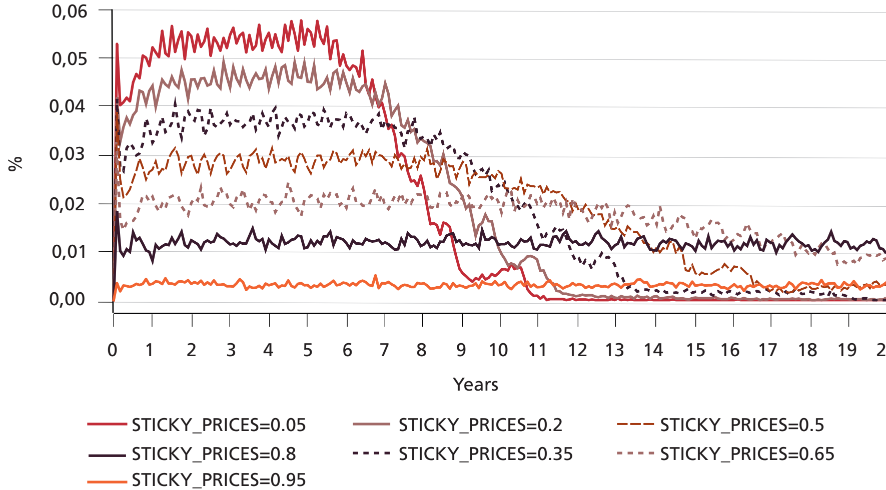
The *sticky_prices* parameter regulates how often firms check the need to raise prices. High values of the parameter indicate lower frequency of adjustments; and lower values, more constant verification by firms.

The parameter is relevant with strong effects on the economy as a whole. When firms adjust prices more regularly inflation rises early in the period, and then becomes stable. On the contrary, when firms maintain prices and do not make their adjustments so frequently, inflation remains at lower levels, but stable until the end of the period (graph 24).

Although the overall effects on inflation do not differ much from the standard behavior, the effects on GDP are strong, with very low values when the parameter is greater than 0.65. That is, when firms check to adjust prices, on average, only once every three months, production in the economy suffers for the given configuration (graph 25). Possibly due to the impact of price adjustment on profits (graph 26). Such effect on profits in turn impacts the firms' hiring and firing strategy, affecting employment levels (graph 27). Households maintain higher levels of savings (graph 28) because they do not participate in the consumer market with such intensity. Finally, all these factors together seem to indicate that more frequent adjustments of prices lead to higher levels of consumption and quality of life (graph 29).

GRAPH 24

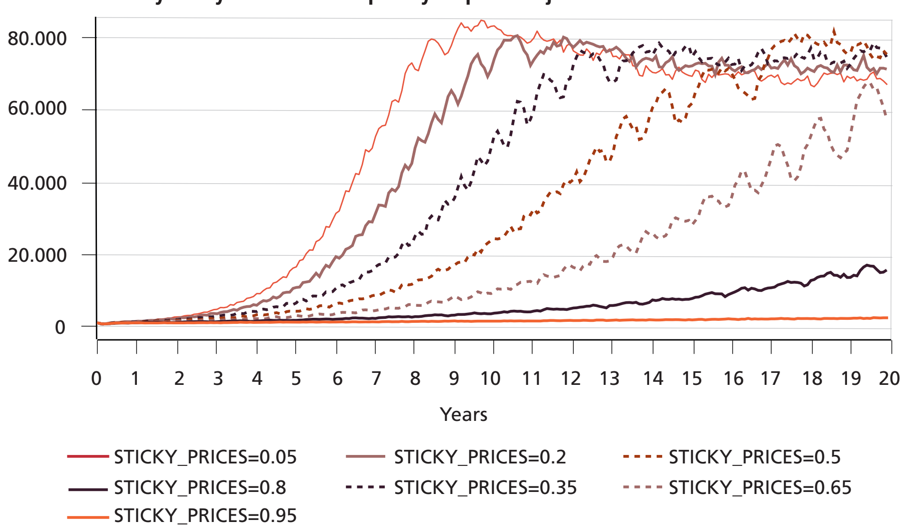
Sensitivity analysis for the frequency of price adjustment of firms on inflation



Author's elaboration.
Note: Results for Porto Alegre, with 2% of the population, average of four simulations.

GRAPH 25

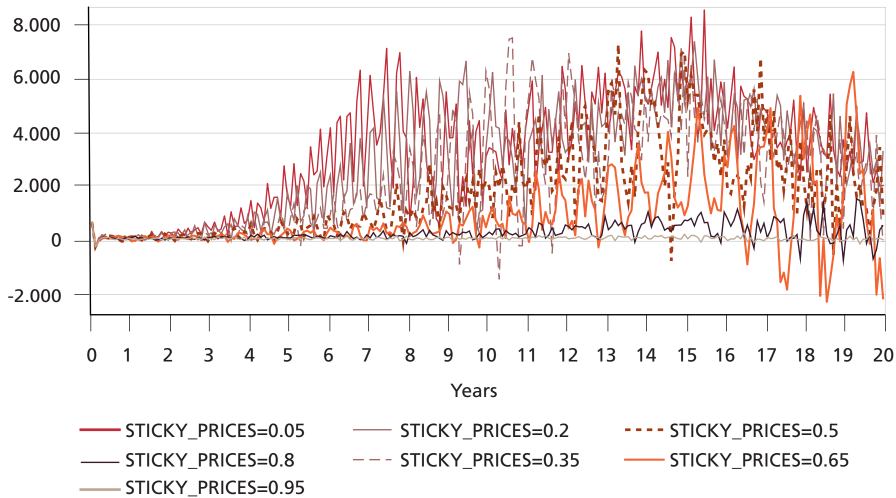
Sensitivity analysis for the frequency of price adjustment of firms on GDP



Author's elaboration.
Note: Results for Porto Alegre, with 2% of the population, average of four simulations.

GRAPH 26

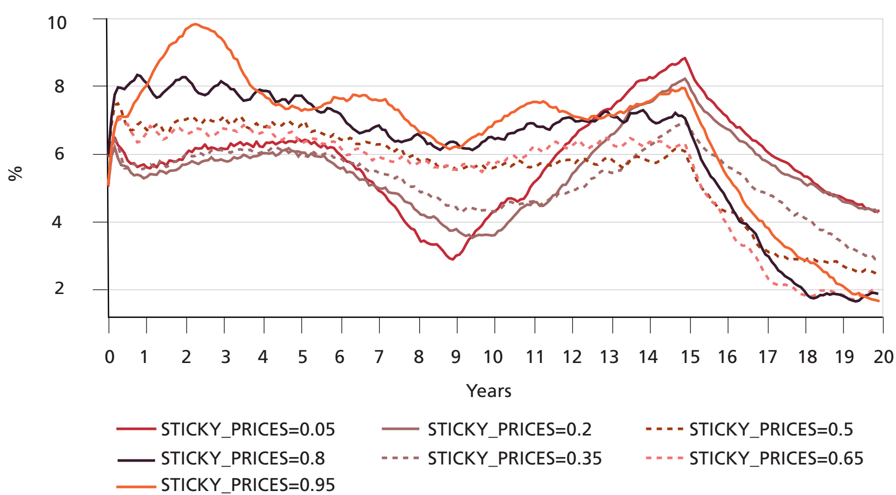
Sensitivity analysis for the frequency of price adjustment of firms on profits



Author's elaboration.
Note: Results for Porto Alegre, with 2% of the population, average of four simulations.

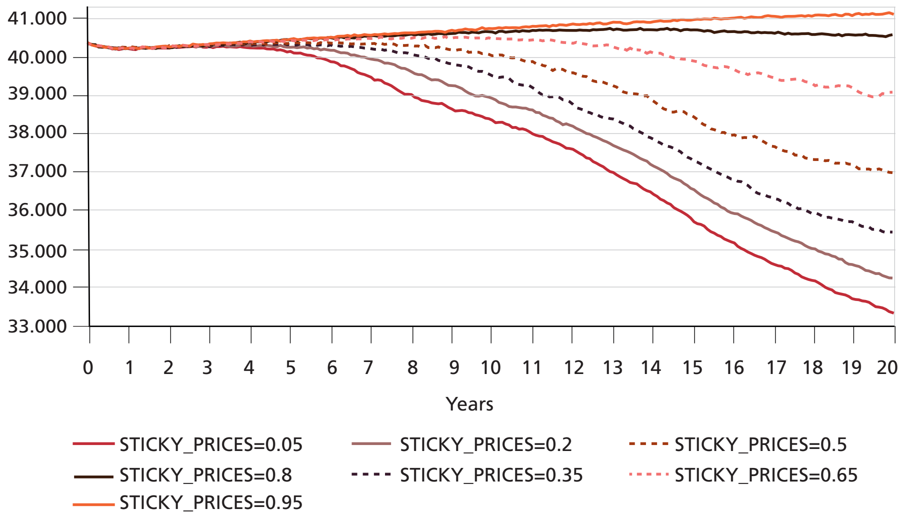
GRAPH 27

Sensitivity analysis for the frequency of price adjustment of firms on unemployment



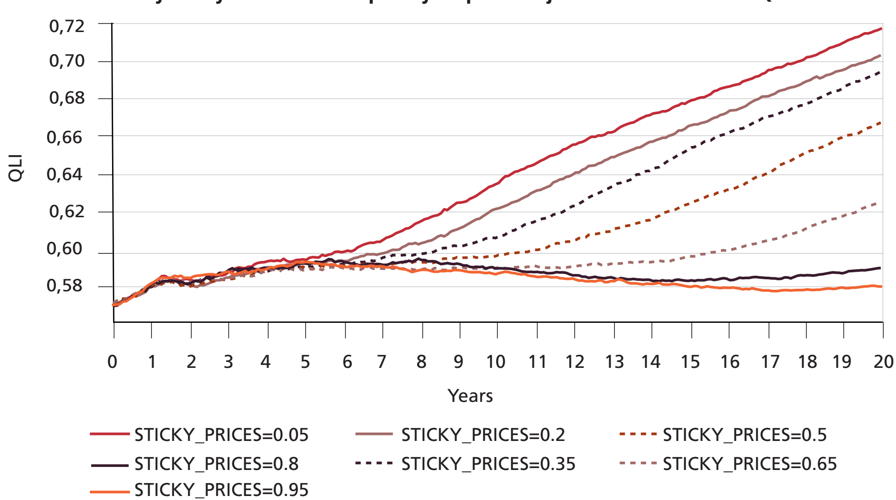
Author's elaboration.
Note: Results for Porto Alegre, with 2% of the population, average of four simulations.

GRAPH 28
Sensitivity analysis for the frequency of price adjustment of firms on households' savings



Author's elaboration.
Note: Results for Porto Alegre, with 2% of the population, average of four simulations.

GRAPH 29
Sensitivity analysis for the frequency of price adjustment of firms on QLI



Author's elaboration.
Note: Results for Porto Alegre, with 2% of the population, average of four simulations.

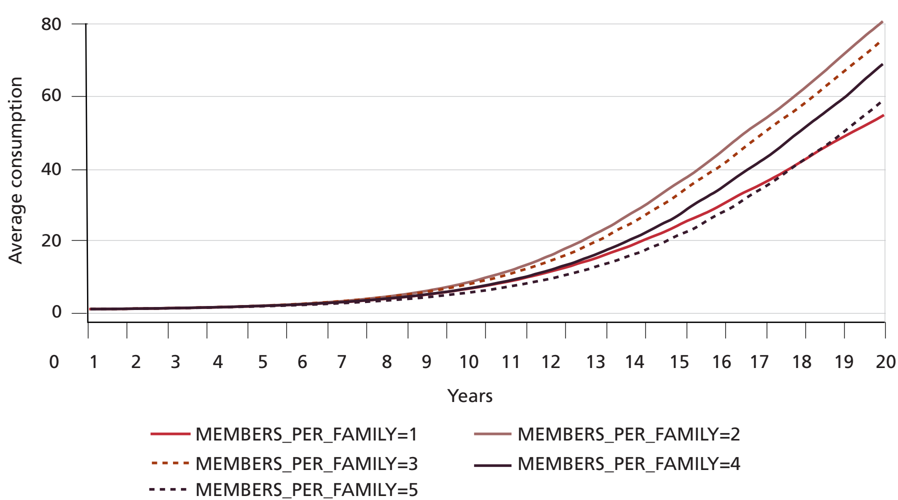
Besides the adjustments for each simulation configuration, there is another set of parameters that regulate the general conditions of the agents of each simulation (*run*). They refer to more general indicators of agent composition, such as the average number of family members, the sample size of the PCA population to be used, or the magnitude of the number of vacant houses in the PCA.

7 DEMOGRAPHY: FAMILY SIZE

For the family size case, we tested five basic possibilities: larger families, with an average number of five members, in descending order, to families with an average of only one individual.

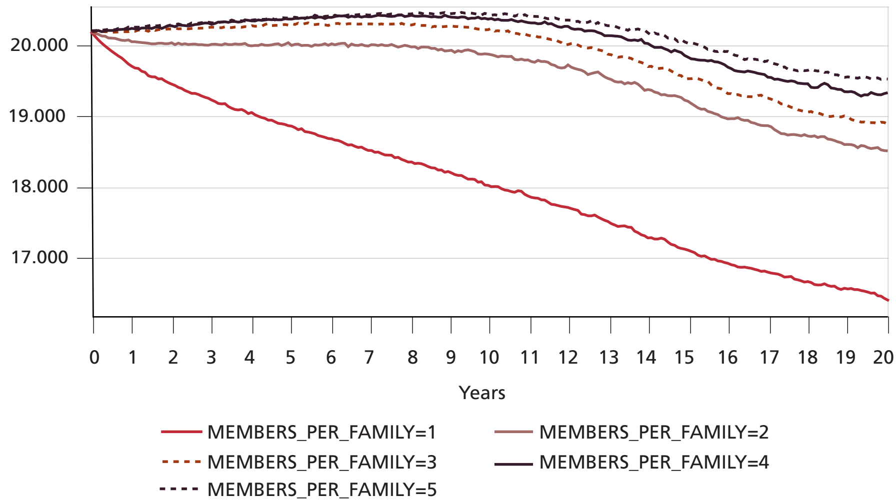
Given that household consumption is shared and, especially, that age of members is also random (although within probabilistic values for each municipality), families with fewer members are more likely to have no working-age individuals. Thus, the average consumption is lower when the average number of members per family is only one individual (graph 30). However, the relationship between cumulative consumption and number of family members is not linear, and there seems to be a better fit in families with an average of three members. In terms of savings, households with more members – and consequently more workers – are able to maintain higher levels of savings (graph 31).

GRAPH 30
Sensitivity analysis for the average household size on average consumption



Author's elaboration.
Note: Results for Brasília, with 1% of the population, average of three simulations.

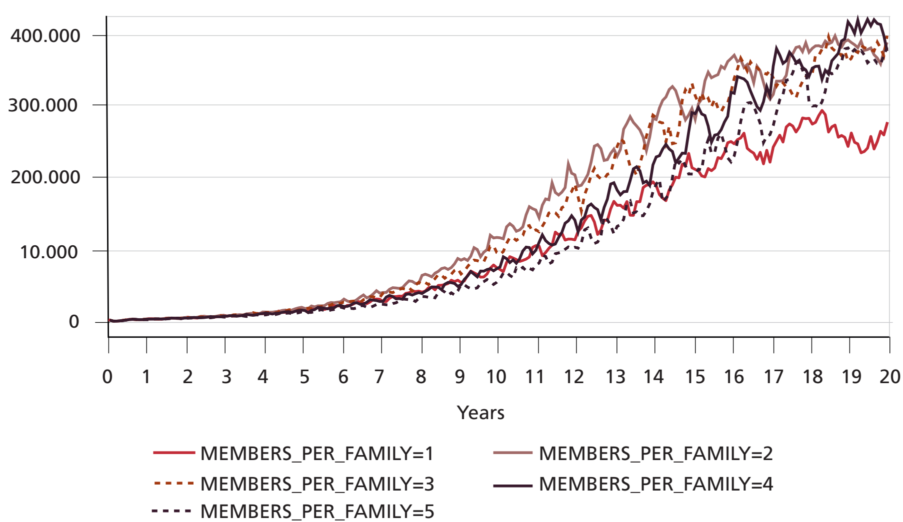
GRAPH 31
Sensitivity analysis for the average household size on savings



Author's elaboration.
Note: Results for Brasília, with 1% of the population, average of three simulations.

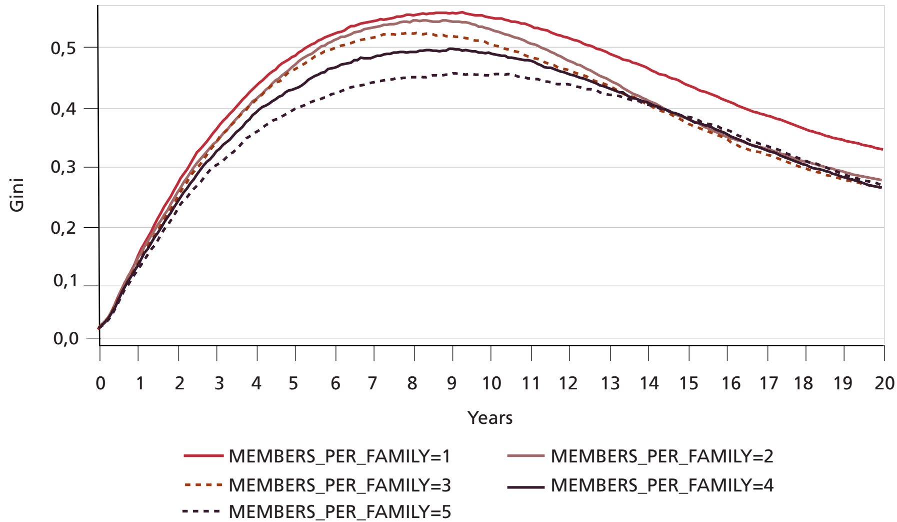
GDP and inequality measured by the Gini indicator are not so different for families of three, four or five members (graphs 32 and 33).

GRAPH 32
Sensitivity analysis for the average household size on GDP



Author's elaboration.
Note: Results for Brasília, with 1% of the population, average of three simulations.

GRAPH 33
Sensitivity analysis for the average household size on the Gini indicator

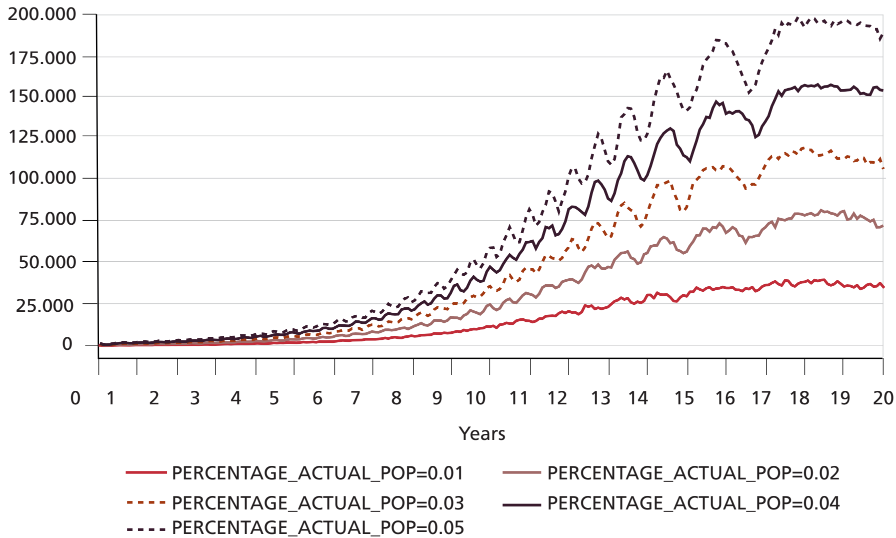


Author's elaboration.
Note: Results for Brasília, with 1% of the population, average of three simulations.

8 SAMPLE OF THE POPULATION

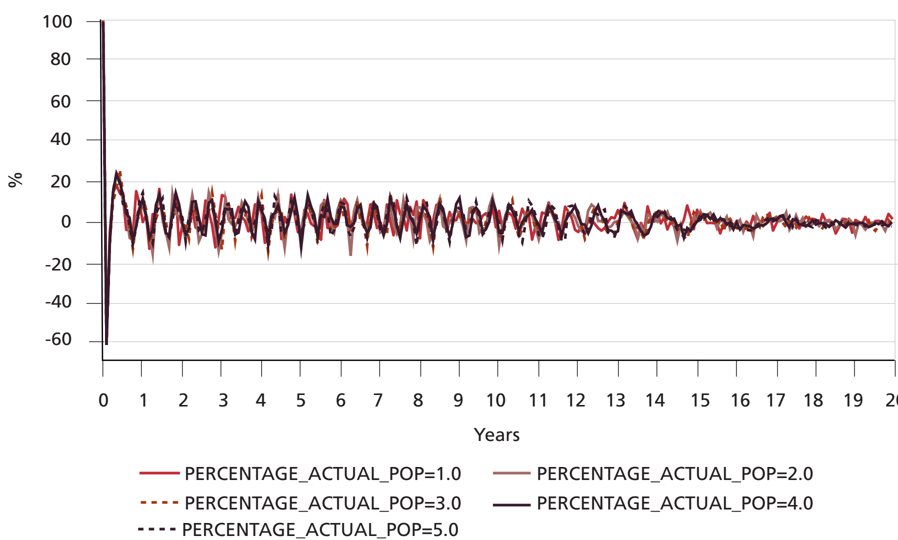
The use of various samples of the population for different simulations indicated that for a few indicators there is level change and some increase in variability, for example, on GDP and household savings level (graphs 34 and 35). However, there is no significant variation in price levels, unemployment, inequality, concentration of employees per firm, nor household consumption.

GRAPH 34
Sensitivity analysis of sample size of simulation population on GDP



Author's elaboration.
Note: Results for Brasília, with 1% of the population, average of three simulations.

GRAPH 35
Sensitivity analysis of the sample size of the simulation population on GDP's variation



Author's elaboration.
Note: Results for Brasília, with 1% of the population, average of three simulations.

9 HOUSE VACANCY

The intensity of house vacancy seems to affect very slightly the behavior of the model as a whole. Only subtly, there is a decline in the cumulative consumption of households, as housing vacancy increases (graph 36). This consumption, in turn, also slightly affects inequality that is higher when there is greater vacancy.

GRAPH 36
Sensitivity analysis of the intensity of residential vacancy on the accumulated consumption of families

Author's elaboration.
Note: Results for Brasília, with 1% of the population, average of three simulations.

CHAPTER 5

APPLICATIONS: FISCAL MUNICIPAL ANALYSIS AND FEDERALISM

As shown in the Taxes and municipalities section (subsection 3.7.9 of the chapter 2) there are four alternative tax distribution in *PolicySpace*. We analyze the variability in the effects among the options (see also table 6, chapter 2). They are: the presence or absence of the Municipal Participation Fund (MFP) rule (True or False) and the fusion of all municipalities of the PCA into one for the purpose of tax distribution (Alternative0 = False) or not (True).

The theoretical discussion itself suggests that the current distributive alternative[1] – with the existence of MFP and its implicit progressivity – is beneficial to the poorer municipalities of the metropolitan regions. According to the previous discussion (see chapter 2, section ODD: general concepts), the most progressive alternative would be the one that would maintain the presence of MFP and implementation of the fusion of municipalities for the purposes of tax distribution (i.e. MFP=True and Alternative0=False). The least progressive alternative of all is the combination of the lack of MFP and the segregation of municipalities in the distribution of taxes, which disregards the fact that the Population Concentration Areas (PCAs) operate as a single productive entity (i.e.: MFP=False and Alternative0=True).

In practice, the monetary amount collected that is redistributed from each rule is relatively small, if the modeler maintains the levels of proportionality in relation to GDP and in the total of taxes collected, as shown in the validation of taxes (section 2 of chapter 3). Thus, the repercussions throughout the model are relatively small for the tested configurations. It should be emphasized that the analysis carried out is of a structural nature and does not incorporate at any moment the gain with the reduction of the bureaucratic apparatus necessary for the maintenance of municipal structures, nor the increase of efficiency associated to gains of scale for municipalities higher as indicated in the literature (subsection 3. 7.9, chapter 2).

Anyway, all exercises we did suggest that MFP by itself is fundamental to the progressivity in the redistribution of taxes observed (graph 2). The fusion of

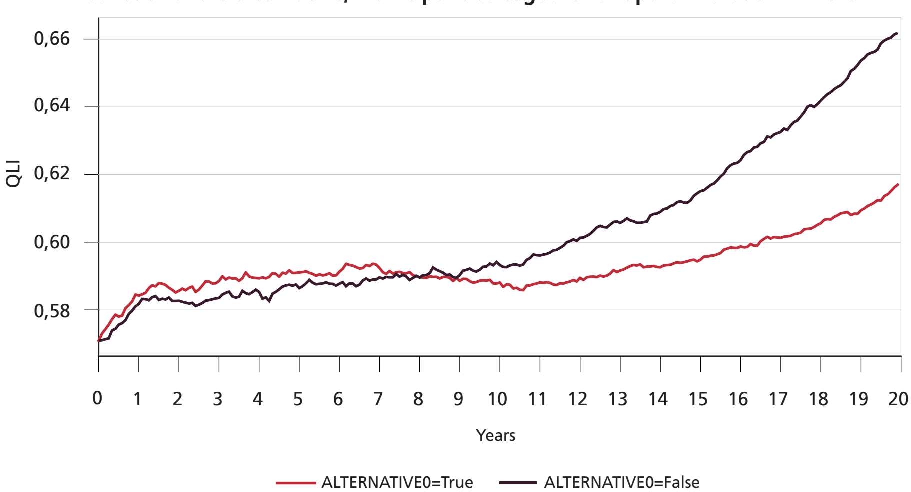
1. *Status quo*: MFP=True and Alternative0=True.

the municipalities for distributional effects is only marginally higher for the quality of life index indicator (QLI) (graph 3).

In contrast, when the option of municipalities, together or apart (Alternative0), is tested with the MFP = False parameter, that is, in the absence of the MFP distribution rule, then the fusion of municipalities (Alternative0=FALSE) indicates results much higher than the option with municipalities apart (graph 1).

This indicates that Alternative0 – municipalities together or apart – produces similar distributive effects to MFP. In the absence of the distributive effect of MFP, the alternative would have similar effects. However, given that MFP rule is in effect, fusion or not of municipalities generates only marginal effects.

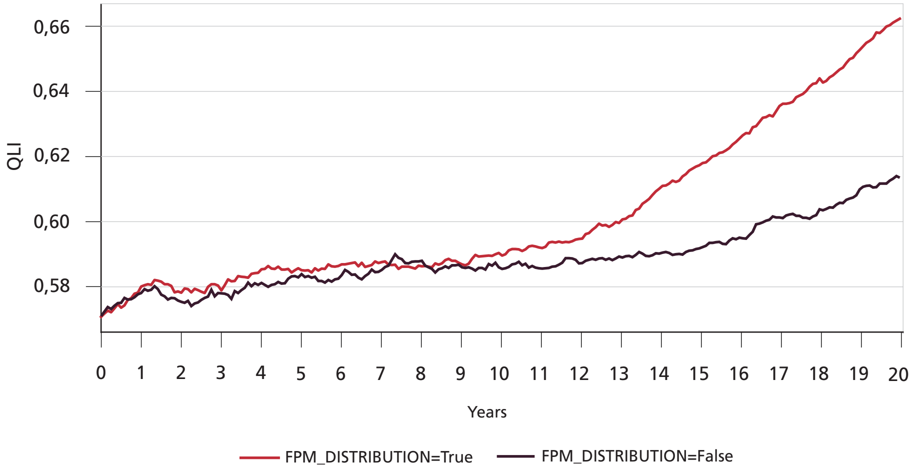
GRAPH 1
Distributive rule alternative, municipalities together or apart without MPF rule



Author's elaboration.
Notes: 1. The average QLI is always weighted by the population. For this simulation MFP=FALSE.
2. Results for Brasília with 1% of the population, average of four simulations.

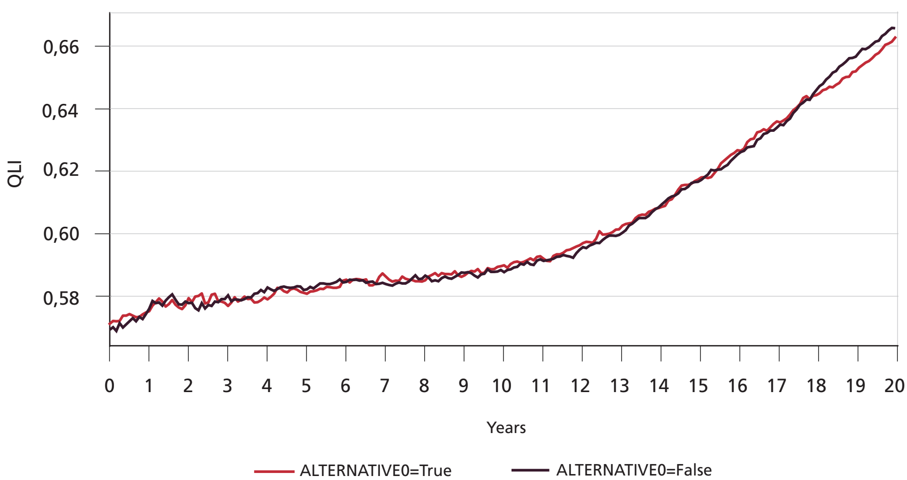
When viewed together, for several tested PCAs, the results are consistent with the theoretically more progressive option (black dotted lines in the graphs) with slightly higher QLI values comparatively to the current configuration (red) for both cases in Rio de Janeiro (graph 4), Ipatinga (graph 5) or Curitiba (graph 6).

GRAPH 2
Distributive rule alternative, presence or absence of MFP



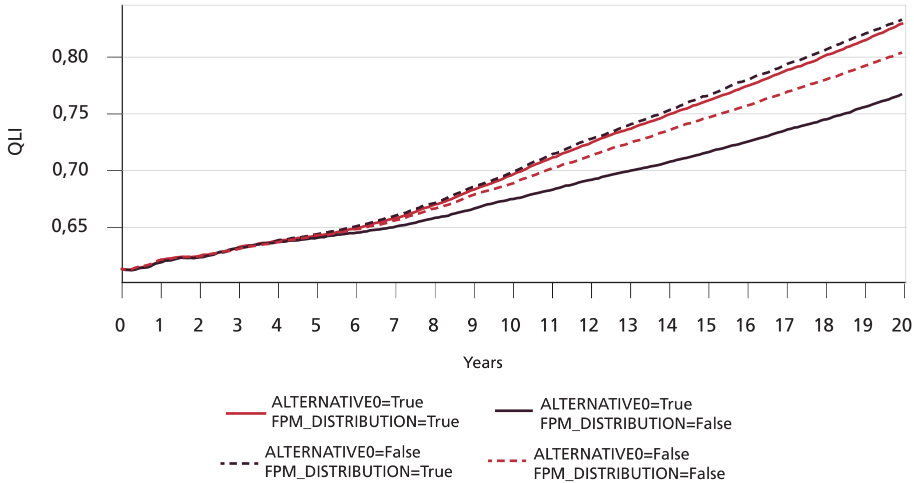
Author's elaboration.
Notes: 1. The average QLI is always weighted by the population. For this simulation ALTERNATIVE0=TRUE.
2. Results for Brasília with 1% of the population, average of five simulations

GRAPH 3
Distributive rule alternative, municipalities together or apart with MPF rule



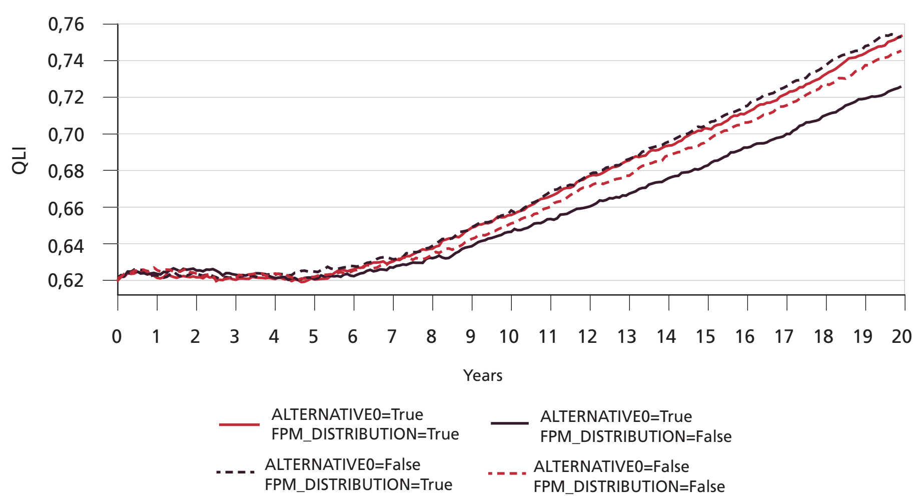
Author's elaboration.
Notes: 1. This simulation follows the default setting in which MFP=TRUE.
2. Results for Brasília with 1% of the population, average of five simulations.

GRAPH 4
Distributive rule alternatives



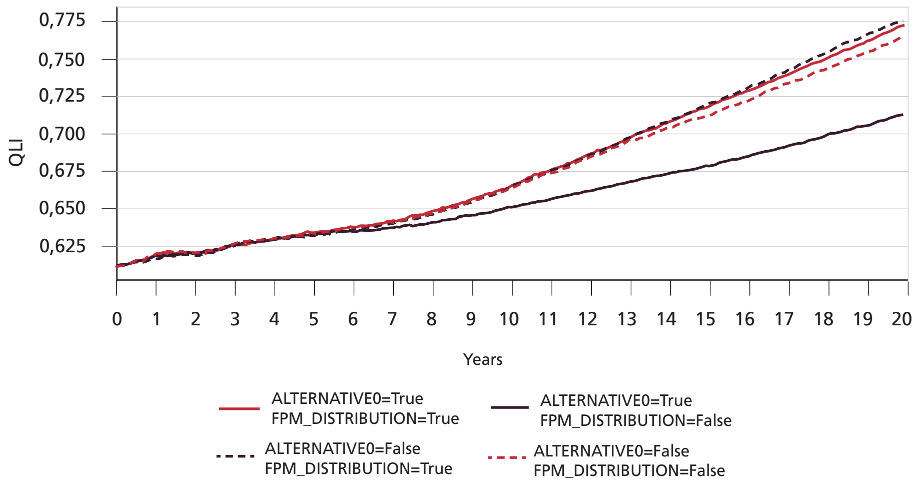
Author's elaboration.
Note: Results for Rio de Janeiro with 2% of the population, average of ten simulations.

GRAPH 5
Distributive rule alternatives



Author's elaboration.
Note: Results for Ipatinga with 2% of the population, average of ten simulations.

GRAPH 6
Distributive rule alternatives



Author's elaboration.
Note: Results for Curitiba with 2% of the population, average of ten simulations.

The analysis strongly indicates that the maintenance of progressive rules of the MFP are essential to the progressive distribution of taxes among municipalities with improved quality of life across the region in general, weighted by population. The alternative to fusion municipalities for the purposes of tax distribution, although positive, in the modeled configuration was shown with only marginal structural influence. If such measures were adopted, this slightly positive marginal results would be added to efficiency gains, cost reduction and less operational fragmentation. In addition, it is worth mentioning that the municipalities' fusion option would generate effects similar to those generated by the MFP, when it is not present, suggesting that a rule would be close substitute for the other, in terms of distribution of tax revenues.

CHAPTER 6

REAL ESTATE MARKET

In *PolicySpace*, the real estate market gives a boost to the economy by injecting household savings into the market. In fact, tests with no real estate market (no families entering the market each month) indicated much lower growth comparatively to tests with a small percentage of families participating in the market. Overall, without the real estate market the whole economy behaves differently, with inflation close to zero, structurally higher unemployment, and greater inequality.

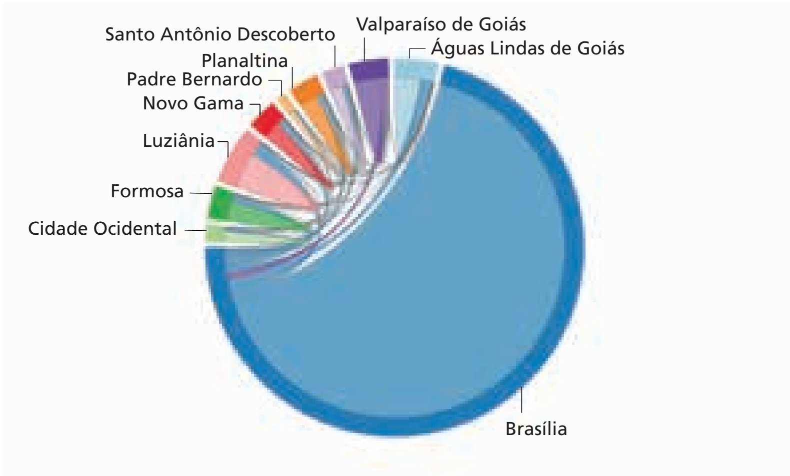
As an illustration, we simulate the case of Brasilia (and Porto Alegre) with 2% of the population and average of four simulations.

In Brasilia, the change in the *percentage_check_new_location* parameter, that is, the number of families entering the real estate market monthly, from 0.4% to 2%, resulted in:

- maintenance at the same levels of unoccupied houses at the end of the period (13.2%);
- marginal increase in the value of houses;
- increase in families who actually moved (from 28% to 50%); and
- decrease of families who have moved to more expensive homes (90% to 81%) at the expense of families who have moved to cheaper homes (from 10.3% to 18.7%).

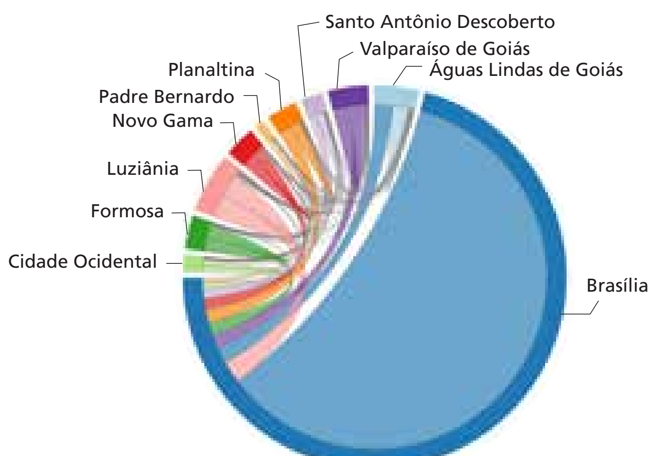
When the real estate market is less dynamic, the changes in Brasília predominate and the flow is always slightly higher from Brasília to the neighboring municipalities, except for Valparaíso de Goiás (graph 1). In contrast, when the market is more dynamic and the number of families moving is larger, the flow reverses with the migration of families from the peripheral municipalities to Brasília, except Águas Lindas de Goiás (graph 2). In this case, the internal migration to Brasilia is also reduced.

GRAPH 1
Flow of changes of families among municipalities for the PCA of Brasília with a less dynamic market



Prepared by: Cayan Atreio Portela Bárcena Saavedra, with simulation data.
Note: Data for 1% of population and *percentage_check_new_location* = .003.

GRAPH 2
Flow of changes of families among municipalities for the PCA of Brasília with a more dynamic market



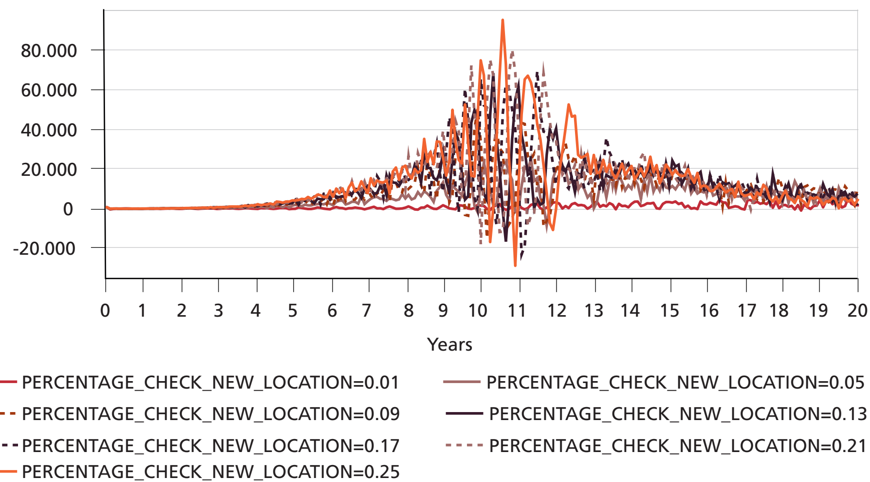
Prepared by: Cayan Atreio Portela Bárcena Saavedra, with simulation data.
Note: Data for 1% of population and *percentage_check_new_location* = .017.

Illustratively, for the case of Porto Alegre, we conducted an exercise with a significant change in the increase of families entering the housing market, varying the percentage from 1% to 25% of the families. The relevance of the real estate market in making the economy dynamic as a whole can be seen in the following graphs.

The profit of firms (graph 3), for example, shows a significant increase in its variability with the increase of population participating in real estate.

GRAPH 3

The real estate market analysis: results[1] of firms' profit for Porto Alegre



Author's elaboration.

Note: For 2% of the population for Porto Alegre and average of four simulations.

In addition, by construction, there is greater and proportional decline in the level of household savings, when more of them participate in the market (graph 4), although it should be noted some convergence at the end of the period.

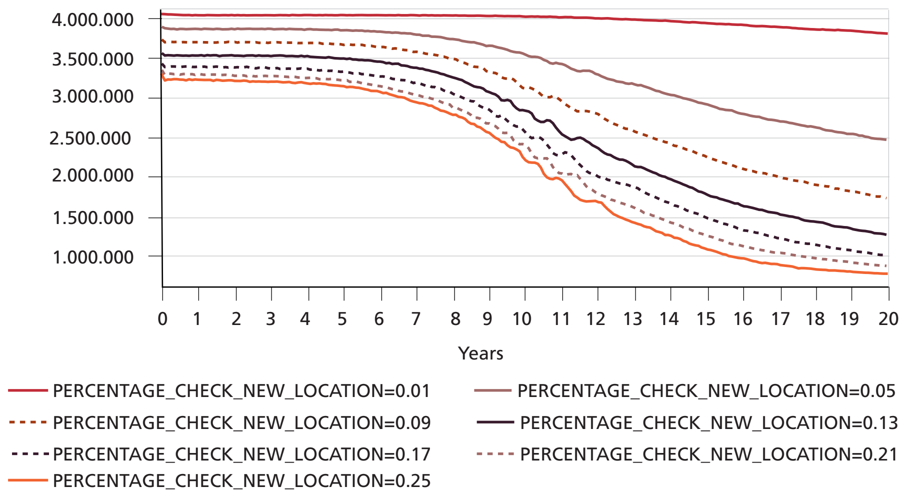
Inequality, as measured by the Gini coefficient indicates two distinct patterns, one higher, when the market is relatively dynamic (*percentage_check_new_location*=0.01), and another, lower, converging, when there is greater dynamism in the real estate market (graph 5).

Unemployment, on the other hand, presents three different behaviors (graph 6). Unemployment at higher levels which then settles into lower values for the less dynamic market; a second pattern that remains stable with a small increase for the 15th year of the simulation; and a third for the most dynamic market simulations that starts with fairly low unemployment, close to zero, and then stabilizes between 2% and 4%.

In general, the dynamism of the real estate market above a given level of household participation leads to higher levels of production, although there is convergence at the end of the period, and even reversion of the trend when the intensity is very high (graph 7).

GRAPH 4

The real estate market analysis: results[1] of families' savings for Porto Alegre

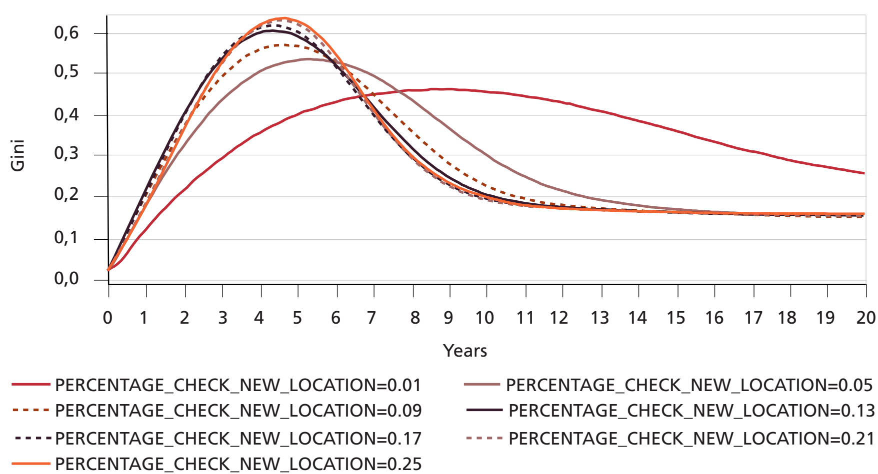


Author's elaboration.

Note: For 2% of Porto Alegre population and average of four simulations.

GRAPH 5

Analysis of the real estate market: results[1] of the Gini indicator for Porto Alegre

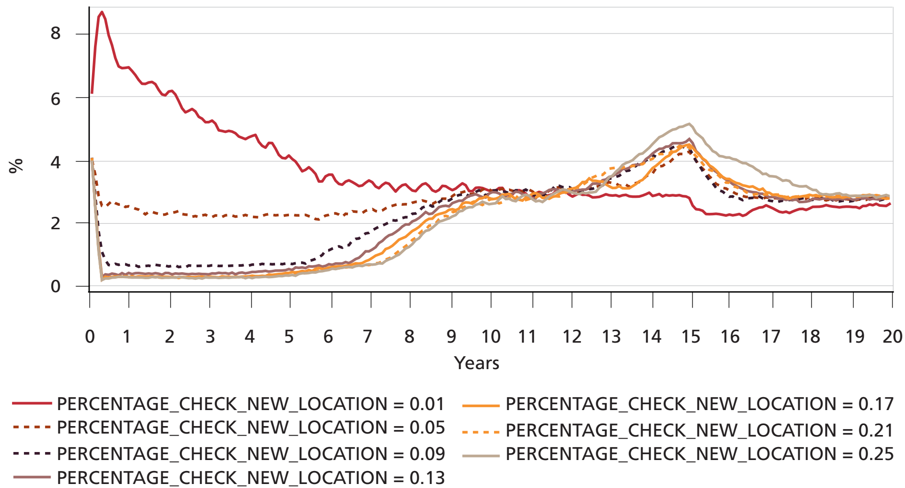


Author's elaboration.

Note: For 2% of the population of Porto Alegre and average of four simulations.

GRAPH 6

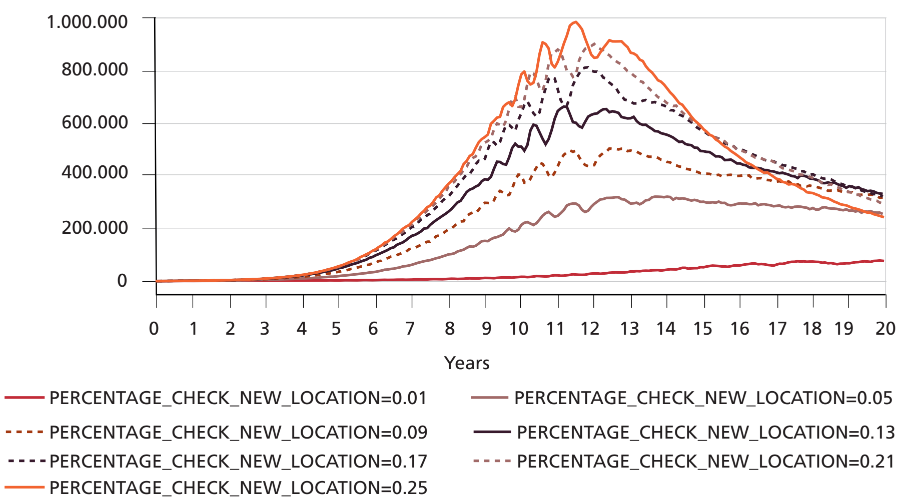
Real estate market analysis: unemployment results[1] for Porto Alegre



Author's elaboration.
Note: For 2% of the population of Porto Alegre and average of four simulations.

GRAPH 7

Real estate market analysis: results[1] of GDP for Porto Alegre



Author's elaboration.
Note: For 2% of the population of Porto Alegre and average of four simulations.

The results presented in this chapter indicate that a more dynamic real estate market is associated with an average improvement of families at higher production and lower inequality levels and more families obtaining access to better homes. The examples also indicated the possibility that a more dynamic market is to greater access to cities with better quality of life.

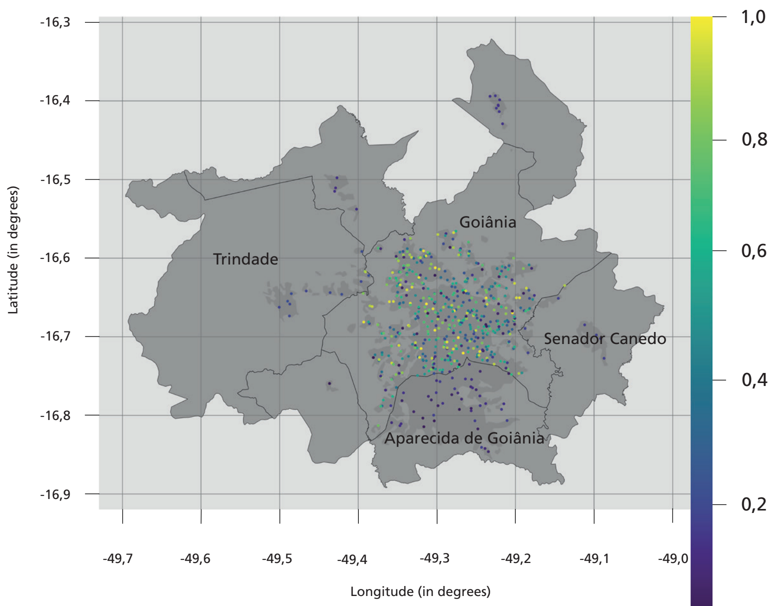
CHAPTER 7

GENERAL RESULTS: METROPOLITAN REGIONS OF BRAZIL

This chapter presents additional illustrative results for some other Population Concentration Areas (PCAs) and analyses in *PolicySpace* that have not been discussed in previous chapters.

Results suggest that spatiality is most striking at the level of the firms rather than at the families. These results – such as firms' profits, for instance – are not given directly by the data used as input at the beginning of the simulation.[1] They are developed as the simulation evolves. Thus, companies located closer to populous consumers' markets at the end of the period pay more taxes (figure 1), employ more workers (figure 2) and obtains higher profits (figure 3).

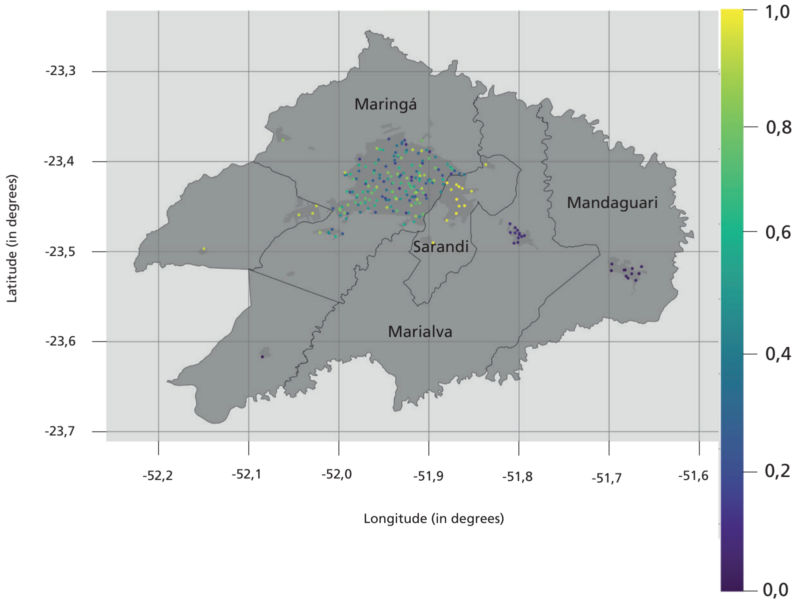
FIGURE 1
Taxes paid by firms at the end of the period for the PCA of Goiânia



Author's elaboration.

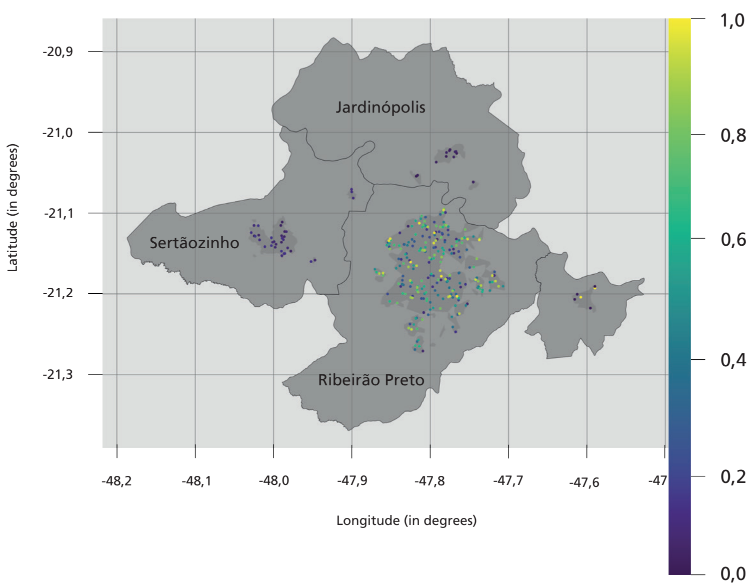
1. We use number of firms by municipality from official database (RAIS).

FIGURE 2
Employees by firms at the end of the period for the PCA of Maringá



Author's elaboration.

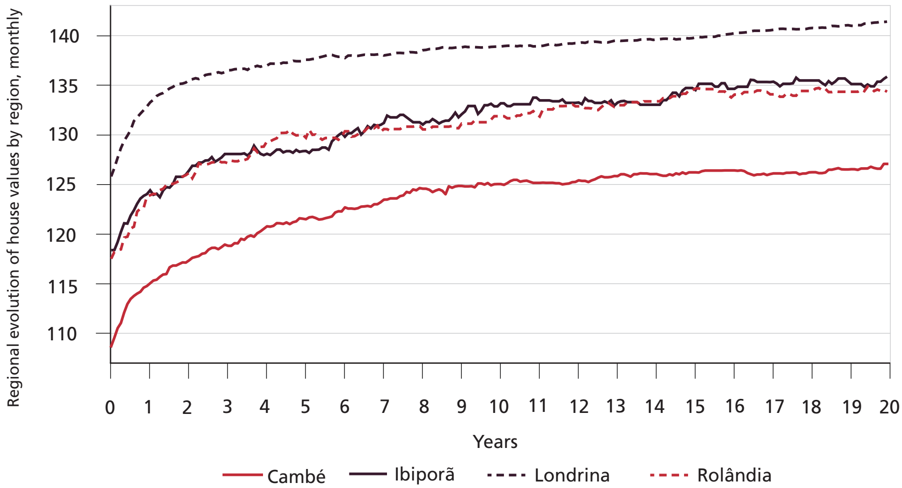
FIGURE 3
Firms' profits at the end of the period for the PCA of Ribeirão Preto



Author's elaboration.

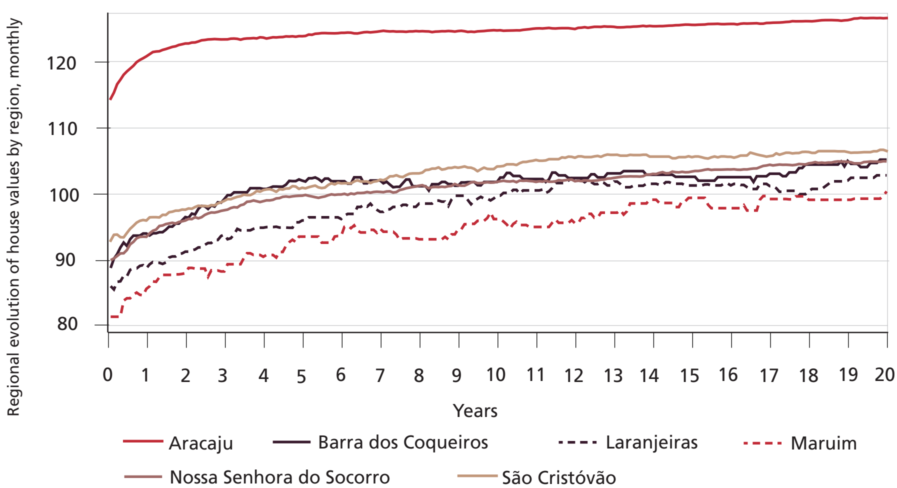
Consistently, real estate values are also higher in central municipalities. See the examples of the PCAs of Londrina (graph 1) and Aracaju (graph 2).

GRAPH 1
Evolution of average municipal real estate prices for the PCA of Londrina



Author's elaboration.

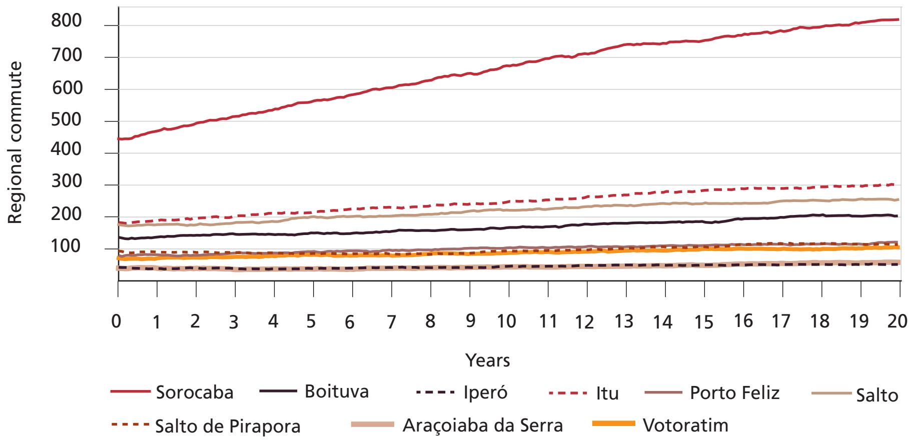
GRAPH 2
Evolution of average municipal real estate prices for PCA of Aracaju



Author's elaboration.

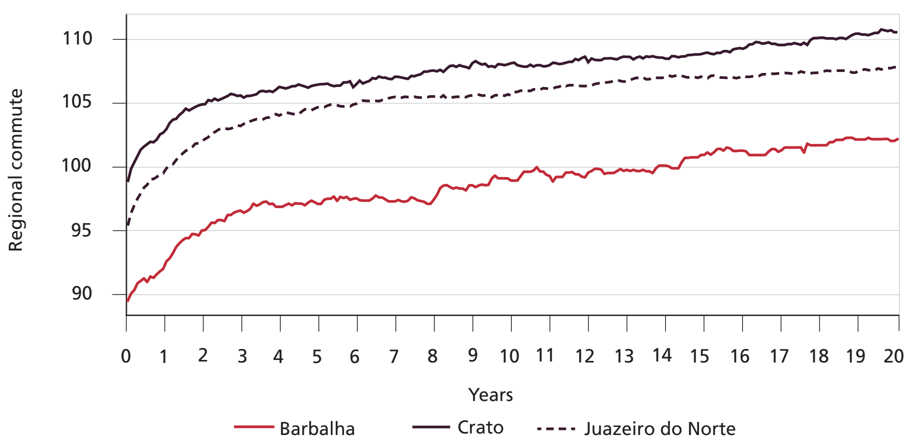
Finally, it is possible to observe for each variation of parameter, rule or PCA, the change in demand for municipal commuting, given that the location (and mobility) of families and employees are embedded in the model. Graph 3 presents the case of the PCA of Sorocaba, and graph 4 of Crato-Barbalha-Juazeiro do Norte.

GRAPH 3
Evolution of the demand for municipal transport commuting for the PCA of Sorocaba



Author's elaboration.

GRAPH 4
Evolution of the demand for municipal transport commuting for the PCA of Crato-Barbalha-Juazeiro do Norte



Author's elaboration.

CHAPTER 8

FINAL CONSIDERATIONS

This book presents in detail a computational platform and a spatial simulation model of public policies using an agent-based model.

In addition, it critically reviews the literature, conceptualizes *agent-based modeling* (ABM) and explores arguments and criticisms regarding the methodology. Next, the book details the model (*PolicySpace*) to make it useful, reproducible besides enabling changes and additions specific to readers' research questions and interests. In order to accomplish that, the description of the model follows the ODD protocol. Thus, it defines its purpose; describes its entities, states and scales; and details its processes and temporal execution. It also reports on the necessary input data and the model's initial configuration. Moreover, the book explains some sub-models, specifically: the operationalization of the markets; the functioning of firms; the parameters and the taxes. Finally, we provide instructions on the operation of PolicySpace and the various alternatives for running simulations and testing.

Validation is achieved successively. Firstly, we illustrate the adequacy of the economy through the reasonability of macroeconomic indicators. Secondly, we demonstrate the ability to replicate the collecting of taxes both in terms of percentage of GDP and percentage of the total amount collected. Thirdly, we compare the performance of the set of Population Concentration Areas (PCAs) using Kolmogorov-Smirnov test for the standard distribution for the taxes and their total. Finally, a series of sensitivity analyzes is presented in order to give the reader a better understanding of the model's mechanisms and behaviors.

The markets in the model are simple and the analysis is only indicative. However, we believe the effective contribution of the book is to make available to the community a comprehensive empirical model for the 46 Brazilian PCAs, which allows quite easily cross-analysis of different elements that influence public policy analysis. Thus, the modeler may observe effects of firms' decisions (frequency of price adjustments or decisions on salaries) in a spatialized and dynamic way, with results in tax collection or demand for mobility, for example. Further, the analyst may observe effects of workers' productivity, of the rules of tax distribution or the behavior of the real estate market on numerous combinations of parameters and alternatives.

In addition to the contribution of the platform, the study made some considerations regarding the effects of tax distribution among municipalities of the same PCA. The results reinforce the relevance of the Municipal Participation Fund (MPF) as a distributive rule. The tested alternative of fusion of municipalities for the purposes of distribution showed only marginal gains. However, in the absence of MPF, the municipal fusion would result in effects similar to MPF presence, according to the simulations.

Anyway, this initial effort allows us to put forward simple expansion possibilities that take advantage of the effort already made, as described in the section on modularization of the platform (section 2 of chapter 2).

Illustratively, we detail them below.

1) The functioning of firms, with sophistication of decision rules or innovations in products, creating heterogeneous products or differentiated by sectors of activity.

2) Mobility – with fine-tuning of differentiated routes in the work-home path and inclusion of modals in the transportation system, taking advantage of the fact that the generation of demand is given and it is compatible with any other variation of the model as a whole.

3) Analysis of the qualification of the employees and the productivity of firms, changing qualification from fixed to variable. One alternative would be to use the already existing account of household savings as investment in family members' qualification.

4) Demographics – deepen family (or even gender analysis) as a number of parameters associated with the family and demography are already implemented. It could be easily done with the creation of a process of constitution (and dissolution) of new families.

5) The deepening analysis of the immense amount of data already generated, which has not been fully completed.

Further:

1) Analysis of the platform in other scales, such as intra-urban detailing, or regional analysis, micro-based.

2) Credit market, investment and financing analysis, in order to compose a more robust macroeconomic model and allow the execution of more specific validations, such as the one carried out by Guerini and Moneta (2017).[1]

1. Guerini, M.; Moneta, A. A method for agent-based models validation. *Journal of Economic Dynamics and Control*, v. 82, p. 125-141, 2017.

3) More detailed analysis of the tax question and alternative tax alternatives.
4) And finally – but again not comprehensively – detailing the specificities of the real estate market with a typical life cycle of real estate and its link to the system of financing the economy.

That is it.[2]

2. For collaborations, please get in touch: <bernardo.furtado@ipea.gov.br>.

Ipea – Institute for Applied Economic Research

Press and Communications Office

PUBLISHING DEPARTMENT

Coordination
Cláudio Passos de Oliveira

Supervision
Andrea Bossle de Abreu

Typesetting
Aline Cristine Torres da Silva Martins
Carlos Henrique Santos Vianna
Mayana Mendes de Mattos
Vinícius Arruda de Souza

Cover design
Aline Cristine Torres da Silva Martins
Mayana Mendes de Mattos (estagiária)
Vinícius Arruda de Souza (estagiário)

The manuscripts in languages other than Portuguese published herein have not been proofread.

Ipea Bookstore
SBS – Quadra 1 – Bloco J – Ed. BNDES, Térreo
70076-900 – Brasília – DF – Brazil
Fone: + 55 (61) 2026 5336
Email: livraria@ipea.gov.br

Composed in adobe garamond pro 11/13,2 (text)
Frutiger 67 bold condensed (headings, graphs and tables)
Rio de Janeiro - RJ - Brazil

Bernardo Alves Furtado is a researcher at Ipea and at the National Council for Scientific and Technological Development (CNPq). He holds a PhD in Geosciences made in co-tutorship with a doctorate in Economics. He has been a professor and coordinator of several courses, has published articles, chapters and books. Presently, he investigates urban public policies and is a professor at Ipea's master course on policy and development. From 2014 onwards, his research focus emphasizes on spatial analysis from a complex systems approach and agent-based modeling methodology.

Ipea's mission

Enhance public policies that are essential to Brazilian development by producing and disseminating knowledge and by advising the state in its strategic decisions.

ISBN 978-85-7811-326-1

9 788578 113261

ipea Institute for Applied Economic Research

MINISTRY OF
PLANNING, DEVELOPMENT AND MANAGEMENT

BRASIL
BRAZILIAN GOVERNMENT