ipea Instituto de Pesquisa Econômica Aplicada

Título do capítulo	CHAPTER 3 – METHODS: *POLICYSPACE2*
Autores(as)	Bernardo Alves Furtado
DOI	http://dx.doi.org/10.38116/978-65-5635-041-7/chapter3

Título do livro	*POLICYSPACE2*: MODELING THE REAL ESTATE MARKET AND PUBLIC POLICIES
Organizadores(as)	Bernardo Alves Furtado
Volume	-
Série	-
Cidade	Brasília
Editora	Instituto de Pesquisa Econômica Aplicada (Ipea)
Ano	2022
Edição	1
ISBN	978-65-5635-041-7
DOI	http://dx.doi.org/10.38116/978-65-5635-041-7

CHAPTER 3

METHODS: *POLICYSPACE2*

The description of the method follows the consolidation of the literature through the Overview, Design Concepts and Details (ODD) protocol (Grimm et al., 2020).[1] It assumes an understanding of the progressive and increasingly detailed model. The first step is the description of the purposes and objectives of the model. This purpose will serve to illustrate the adequacy of the model, regarding whether it effectively served the purpose or not. Next, the agents, their attributes and scales are described. Once the agents are known, the modeler describes the process execution program; that is, in what order events occur within the model. At this point the reader already has a good understanding of the model's specifications, and the ODD protocol recommends discussing the concepts underlying the proposed modeling: the design concepts. In this section, the basic principles of modeling and the conceptual aspects of complex systems are discussed. Finally, the details of the initialization of the model, the necessary data, the description, the rationality and the justification of the sub-processes, for example, the markets present in *PolicySpace2*, are described in sections 5, 6 and 7. The formulas and parameters for each core process are available in section 7.

1 ODD: PURPOSE

In this book, it was decided to use the ODD protocol in its entirety for the description and analysis of the model. Originating in the epidemiological literature, the ODD was adopted, adapted and extended by the community of scientists who use agent-based modeling (Grimm et al., 2006; 2020; Grimm and Railsback, 2012). The protocol embodies the most accurate description of the built model, its fundamental principles, purpose and details. The rigor of the protocol requires, for example, that when starting the description, the following passage is included: "the model description follows the ODD (Overview, Design Concepts and Details) protocol for describing individual and agent-based models" (Grimm et al., 2010, p. 2763).

The first item established in the protocol is the description of the purpose of the model. Given the comprehensive feature of the *PolicySpace2* and its intention to constitute itself as a platform, we established two central purposes, each one with its associated result, following Edmonds et al. (2019). As suggested, the first

1. This description is compatible with version number 1.1. Available at: <https://bit.ly/3LqCsir>.

purpose is linked to the public policy test carried out in chapter 6. The second is linked to the results of chapter 5.

The first purpose of *PolicySpace2* is to illustrate a potential explanation as to how different alternatives for investing municipal public resources in housing policies and income distribution among citizens can result in comparatively different behavior in local economic indicators. In the simulated case, the generation of families selected to receive the aid, as well as the amount of resources distributed and the way in which the benefits are offered are endogenous to the process. It is possible to understand the advantages of each of the alternatives: provision of housing, payment of rent or distribution of pecuniary assistance. With this, relevant comparative indications are obtained for better investment of public resources. The results are robust for a wide range of parameters, processes and different metropolitan regions (RMs).

Additionally, the *PolicySpace2* is a model whose purpose is descriptive and useful for making analogies (Edmonds and Meyer, 2017, p. 45; Grimm et al., 2020, p. 30).[2] In particular, *PolicySpace2* shows itself capable of articulating distant facets of analysis. For example, by what order of magnitude would improvements in labor market productivity affect house prices or corporate profits? Or, how does the presence of a larger stock of unoccupied properties affect household savings?

The *PolicySpace2* is descriptive in the sense that it makes it possible to analyze how different policy configurations affect the dynamics observed. The model allows the quantitative and formal representation of parameters and rules, as well as their associated results. Questions that could be included in this item are: i) how the percentage increase in construction firms affects macroeconomic indicators and income concentration; and ii) how the socioeconomic composition of families, given by the 2010 census, affects the intrinsic dynamics of inflation and unemployment.

The *PolicySpace2* also makes it possible to assess the compatibility of the results with the hypotheses raised. In this way, it serves to illustrate general principles.

Finally, it also serves as a methodological instrument for making analogies. It is easy to use the model to reason about the housing market in an endogenously integrated way with the rest of the economic system. In particular, this purpose is verifiable through the flexibility of the proposal and its relatively simple amendment process. For example, if a given rule for real estate trading is not satisfactory or adequate for a specific hypothesis, or for a set of evidence, other alternatives, such as auctioning, can be implemented.

2. Also, as recommended by the ODD protocol, the *PolicySpace2* is licensed under the MIT License. The complete code of the model is available at: <https://bit.ly/3wnnDZI>.

The explicit indication of the model's purpose is relevant because it indicates the patterns to be replicated that will serve as the model's evaluation criteria (Grimm and Railsback, 2012). As our purpose includes the possibility of "articulating distant aspects of analysis" with a view to understanding real estate market mechanisms, the patterns to be replicated will involve different aspects of the economic environment. With this, the intention is to verify if the behavior of the economic indicators of the model remains within reasonable margins, at the same time that the real estate market also presents a similar performance to the observed real market.

In this sense, the model would not serve its purpose if some meaningful change – for example, an increase in the influx of migrants, or growth of construction firms, or even a reasonable change in interest rates – resulted in unemployment of 90% or a Gini coefficient of 0.01, for example. Or, still, if the characterization of the real estate market resulting from the model was not in line with real data of the real estate market.

In general, the *PolicySpace2* replicates expected trajectories of stylized effects or facts in the behavior of markets. For example:

- the increase in dynamism in the real estate market, when a greater number of families participate in the market on a monthly basis, leads to an increase in economic performance, with greater price volatility and increased inequality;
- the increase in worker productivity leads to lower prices in the economy;
- the reduction in the number of firms consulted when making purchase decisions in the goods and services market – less competition – generates more inflation; and
- when there is an increase in population, there is an effect of scale and efficiency with greater inequality and greater production and consumption.

Other effects – for example, on household wealth or on the behavior of gross domestic product (GDP) and unemployment – are not so obvious and may suggest interesting connections:

- the fact that companies choose more workers with the criterion of distance, to the detriment of qualification, suggests an increase in prices; and
- the increase in taxes on firms leads to a reduction in their profits, but it also leads to an increase in government investments, with an increase in real estate prices and, therefore, greater savings for families.

As a result, the specific patterns that the *PolicySpace2* hopes to replicate are:

- for economic behavior in general, reasonable behavior of the indicators for GDP, inflation, unemployment and the Gini coefficient; and
- for the real estate market, similarity in the distribution of real property prices.

2 TRACE – PROBLEM FORMULATION

According to the modeling follow-up procedure suggested by the Transparent and Comprehensive Model Evaludation (TRACE) methodology (Grimm et al., 2014), we specify here the formulation of the problem. The model must specify the customers, the specification of the question to be answered, and the products needed to do so. Additionally, the applicability of the model and its extension possibilities must be explained.

The target audience of *PolicySpace2* is formed by scientists and public managers interested in understanding the broader mechanisms of the real estate market. As described, the model's intention is to provide a means of reasoning and making analogies about the real estate market, allowing the observation of the scale and comparative analysis of the effects of implementing changes in public policies and behaviors. This reasoning is possible because the model produces adequate macroeconomic indicators and is similar to real estate market prices. Additionally, the *PolicySpace2* tests alternative public housing policies. In terms of extension and extrapolation, some possibilities are listed by way of conclusion in the final considerations.

We do not exclude the possibility of using *PolicySpace2* for specific forecasting of the housing market, provided that more building data from the initial housing base is introduced. We also understand that it would be possible to include additional mechanisms or analyses that contribute to a broader understanding of phenomena already represented here, such as the municipal fiscal issue; the qualification of workers; inequality and social mobility; urban mobility and its CO emission effects; and the analysis of intersectoriality or innovation in firms. Any of these additional analyses should maintain adequate results overall and be validated for each case specifically.

3 ODD: ENTITIES, STATUS VARIABLES (ATTRIBUTES) AND SCALES

The *PolicySpace2* contains six entities that interact throughout the simulation: *individuals*, who are always organized into *families* and that inhabit *households*, with fixed locations; firms, which hire individuals, participate in the labor market and offer goods on the property and real estate markets; banks, which collect deposits and offer loans; and the spatial entities, the *municipalities*, which can be subdivided into weighted areas (APs) or not. A lot of information is collected monthly, so variable attributes are also recorded each month.

3.1 Individuals, workers

Individuals in *PolicySpace2* have their own identification (character, size: 36); gender (character: male, masculine; female, feminine); years of study (integer: 1, 2, 4, 6, 8, 9 to 15); and birthday month (integer: 1 to 12), which are invariable.

Additionally, individuals register the family to which they belong (*family_id, character*) and, perhaps the company at which they work, when employed (character). When not used, the variable assumes the null value (*none*). Belonging to the family can be updated through marriage, as can the link with the employing firm.

The age of individuals is obtained in proportion to official data, according to the smallest spatial area (municipality or AP) at the time of creation (integer: from 0 to 100) and is updated annually, according to the anniversary demographic processes.

Individuals also record a numeric variable (*float*) *money*, which represents the individual's current portfolio of financial resources. Note that the worker individually receives resources from the firm, but purchase decisions are made within the family, with the sum of resources from other family members. Employed individuals also record monthly, in the variable *distance* the distance traveled between home and work. When not used, the numeric variable returns to zero.

3.2 Families (collective)

Families are formed by one or more individuals and are, in practice, the environment for decision-making and financial sharing. Only the family identification variable remains the same throughout the simulation. All others are updated throughout each process.

The variable *members* registers the family member agents. As used in Object-Oriented Programming (OOP) modeling, the variable contains the individual that is a member of the family. Thus, the individual's variables, such as his financial portfolio, for example, remain accessible from within the family, as long as the individual remains a member.

The composition of family members can change through marriage or death; and by birth of the child of a female member of the family.

Another object that is accessed as a family variable is *house*. This variable refers to the household in which the family currently resides and may be the owner or tenant. At the same time, the identification of the region in which the current residence is located is recorded (*region_id*). There is also a list of possible properties of real estate, called *owned_houses*, which changes when there is a sale or purchase of a property by the family.

Other numerical variables, updated monthly and recorded within the family include: balance sheet (*balance*); savings (*savings*); wealth – in which financial assets and real estate are accounted for (*wealth*); payment of any loan installments; and calculation of permanent income (*last_permanent_income*).

3.3 Firms

Firms have a fixed location (*address*), which consists of a geographic object of the type *point* from class *shapely*, referring to their coordinates. All other variables are updated monthly and include the available account balance (*total_balance*), quantity produced (*amount_produced*), current price of the product[3] (*prices*), wages paid (*wages_paid*), profits earned and taxes paid (*taxes_paid*). The quantity sold (amount_sold), in number of units of the good (*float*) and billing (*revenue*) are reset to zero at the beginning of each month. Finally, firms also record the price assigned to the product.

Additionally, construction firms specifically contain information on whether or not they are building (Boolean, *building*), the inventory of properties built and not yet sold (*houses_inventory*), the cost, size, quality and region of the home under construction (*building_region*, *building_size*, *building_cost*, *building_quality*). Finally, there is an organization of the firm's cash flow, so that the sale of the property is not accounted for in cash, but in installments. To allow the payment flow of employees, the data is recorded in the variable (*cash_flow*), which is a dictionary containing the payment month and installment.

3.4 Households

Households also maintain a unique identification, address, quality and region of location, all of which are invariable attributes. The price (*float*, *price*) is adjusted based on changing regional prices and other factors.[4] The household, specifically, contains the identification of the family (*family_id*) referring to the occupant (owner or tenant) and the owning family (*owner_id*). When in lease, the residence maintains information regarding the payment of rent (*rent_data*); and, when vacant, records the number of months on sale (*on_market*).

3.5 Bank

A bank concentrates customer deposits - identified separately, records available cash resources (*float*, *balance*), referring to income from loans and expenditures and total assets (*outstanding_loans*). It also maintains a portfolio with identified

3. The *PolicySpace2* only has one homogeneous product per firm – differentiated by the location of the sale – and by the price. However, the program has an inventory that could include new products, also endogenously generated, but which remain with only one product.

4. See details in subsection 7.7.

deposits and maturities. For each loan granted, an instance, called loan, is created and records the age of the contract, the principal amount, the outstanding balance (principal plus interest), the payments made and the number of arrears, as well as two pieces of information regarding the loan settlement (*paid_off*) or the existence of late payments (*delinquent*).

3.6 RMs, municipalities and APs

Central in the spacing of the *PolicySpace2*, the regions are APs constructed by the Brazilian Institute of Geography and Statistics (IBGE). In some cases, the municipality has only one AP, which spatially coincides with the municipalities. From the point of view of location and generation of spatial data, APs are used as a spatial element whenever there is information available at that level (gender, age, population). Also in relation to the real estate market, each AP has its Quality of Life Index (QLI) updated individually. That is, the value is the same for all APs in the municipality at the beginning, but they are updated according to proportional population gains (or losses) throughout the simulation. From the point of view of tax collection and recording of macroeconomic statistical information (GDP, inflation, budget), the municipality is the reference unit. It works as a collection center, receiving the budget and distributing it proportionally to the populations of each AP. With this, the QLI maintains as attributes (*index*), the GDP, the population (pop), in addition to dictionaries that record taxes with monthly and accumulated values (*treasure*, *cumulative_treasure*, *applied_treasure*). Finally, each region contains an exogenous number of permits to build new homes (licenses).

3.7 Scales

In terms of time, the scale of operation of the *PolicySpace2* is monthly. The processes described, according to the sequence of subsection 4.2, take place every month. By default, the simulation runs from January 2010 to December 2020, so for 120 cycles. However, the model can also be configured to start in 2000, from the 2000 census data and the 2000 APs, and last for as many months as the modeler chooses, up to a maximum of 2030. From 2030 onwards, the IBGE mortality data are not configured, so it is not possible to continue to later years.

In spatial terms, as described in the previous subsection, the *PolicySpace2* operates at the intra-urban level, with the official limits of the IBGE APs, which are aggregations of census sectors and maintain statistically weighted sampled data. The APs aggregate in municipalities that are entities of the model and collect and distribute taxes. Finally, each simulation is independent for an IBGE (2016) Population Concentration Area (ACP), which is equivalent to the most dynamic and central part of the 46 RMs used.

Therefore, there is no migration or travel between RMs. For the case of São Paulo, for example, which has several RMs that are very integrated and close together, it is possible to simulate those of Campinas, Santos and São Paulo together. However, in this case, there will be daily trips between any point of the RMs as a whole, although mediated by the cost of transport.

4 ODD: OVERVIEW OF PROCESSES AND SEQUENCE OF EXECUTED STEPS

The general intuition of the model and the overview of the processes are presented initially in chapter 2. Formulas and rationality based on literature are made for each process in the description of the sub-models (section 7 of chapter 3). The relevant thing at this point, according to the prescription of the ODD protocol, is the order of execution of each step and each phase and the rationale for each ordering. At the same time, the variables that are updated at each step are detailed.

The *PolicySpace2* is a simulation with monthly periodicity which, in the default configuration, occurs 120 times (2010-2020). The central call of the model occurs through the program main.py, executed by the Python interpreter (version 3.7). This module determines the number of simulations that will be executed and makes the parameters compatible for each simulation. In default mode (*python main.py run*),[5] the simulation is performed only once, according to the parameters established in the file params.py and information on start and end dates, RM name, percentage of population to be considered, rescue location, among other details of each execution.

When the call is made in sensitivity mode, which involves simulation with parameter variation and production of comparative graphs, the module *main.py* organizes each individual run for each parameter group.

4.1 Agent generation

Three processes run sequentially. First, the agents are generated, determined by the chosen parameters. As this process is time consuming, it is possible to save the agents already created and just read the corresponding files in the next simulation. It is also possible to adjust so that new agents are created at each simulation.

The saved files are specific to a given combination of parameters that influence the generation of agents. So any change in any of these parameters necessarily leads to the generation of a new set of agents. The parameters that influence the generation of agents are described next.

5. See section 8, which describes the operationalization of the simulation.

1) *Members_per_family*: used only when the simulation start date is the year 2000. For 2010, the number of family members is read directly from the AP data.

2) *House_vacancy*: determines the vacancy in the number of output properties of the model in the first month.

3) *Simplify_pop_evolution*: is a parameter to simplify the process of generating agents by age. If determined to be false, this parameter probabilistically generates the population for each year of age. When true, which is the default choice, agent ages are drawn into six aggregated age groups.

4) *Percentage_current_pop*: is the percentage of the RM population to be considered. By default, we use 1% of the population. It is possible to run with 100% of the population, but the computational demand is relatively large and we only simulate in this case for the RM of Ipatinga, in Minas Gerais.

5) *T_licenses_per_region*: refers to the number of new building permits per AP each month. The default value is random, which, in practice, leads to the generation of one or no licenses per AP per month, which seems more than enough to supply the simulated civil construction market, since there are many licenses left over at the end of the process.

6) *Percent_construction_firms*: determines the percentage among the firms in the model that will be civil construction firms. The number of firms itself is proportional to the percentage of the simulated population and is based on real data on the number of firms per APs (in the standard year of 2010).

7) *Starting_day*: can be set to January 2000 or January 2010, the latter being the default parameter.

For details on the agent generation process, see subsection 4.1. Here, the important thing is to specify the generation order of the agents. The first agents to be created are the regions; that is, the smallest spatial units of analysis from the IBGE's APs. An iteration is made over the regions, in no specific order, since they are independent of each other, and the generation depends only on the data. From then on, agents, families and properties for that region are successively created for each region.

The number of agents created for each region is given by the percentage of the real population in that particular region (AP), according to the number of agents of each gender and of each age. Once the agents and families for each region are created, the agents are allocated as members of each family separately for adults (over 21 years old) and children. Agents are randomly shuffled before being allocated to families.

Allocation of adults is done as equitably as possible so that households have similar proportions of adults. The allocation of children is done for all children with the choice of families at random, so that there can be families with more children and families with no children.

Next, the properties are shuffled and distributed to a percentage of the families, according to the exogenous factor. In this first part of the process, the intention is that each family of this percentage has its own property. The remaining properties are randomly allocated to families and can be rented by families that do not own properties. This is repeated for each region of the current simulation. In practice, some families receive more than one property and others no property.

At the end of the process, each agent belongs to a family that has a residence address, which can be owned or rented and all properties have a family that owns it, although not all properties are occupied.

A second general process just controls the passing of days and each month processes all the steps described below. Although no process runs daily, quarterly or annually, this possibility exists in the model's programming.

4.2 Monthly processes

After the initialization of the agents – through generation or reading –, the simulation runs through the same activities, month by month, until the end of the established period. Commands are invoked in the following order.

1) Monthly interest and mortgage interest are read from exogenous data.
2) Each region (AP) provides new licenses for civil construction (exogenous). As the process is independent, the order does not change the execution.
3) New firms are incorporated into the model. While the absolute number of firms follows the empirically observed exogenous pattern, its allocation in each AP is probabilistically according to the number of employees and average profit.
4) Firms update their production. This process is also independent for each company and depends only on the number of employees and on their qualifications.
5) *PolicySpace*'s processes also include stochastic decisions for mortality and fertility, drawn from official data. The demographic processes of mortality, fertility and aging, with exogenous probability parameters and official data, are performed. The process takes place annually, in the month of each agent's birthday. The cohorts are operated by Federation Unit (UF), which is the origin of the probabilities, and in ascending order of age.

The first process that occurs is age advancement, and then the probability of marriage (for those over 21) and the probability of death are updated. For women, additionally for those between 14 and 50, the probability of giving birth is checked. If so, a new agent is created and incorporated into the mother's family.

6) Then the immigration and marriage processes take place. By municipality, the annual number of migrants is calculated from exogenous data and linear estimation for the years without information, allocating it equally to each month of the year. The same procedure described in the generation process is performed. First there is the generation of agents, then the families and then the allocation of agents in the new families. However, when the property is allocated, the family is directed to the rental market. Families that are unable to rent a property are not included in the model.

7) In the marriage process, all agents – with their remarriage probability updated annually in the anniversary month – are probabilistically included in a list. The list is shuffled and pairs are randomly formed. The marriage and the formation of the new family only take effect if the new family manages to find a property to rent.

8) The goods market begins with the consumption of families. Families select a sample of firms at random that are exogenously fixed in size by a parameter.

9) The bank collects the mortgage payments due, according to the order in the portfolio, family by family.

10) Firms assess revenue, pay taxes and calculate profit or loss. They pay their employees and decide whether to update prices.

11) Prior to construction planning, the monthly global real estate vacancy is calculated. For each construction firm, the process of planning new properties takes place; next is verification of whether previously planned properties have been completed. If so, they enter the firm's sales portfolio.

12) 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. Subsequently, in a probabilistic way, according to an exogenous parameter, the firms assess whether they participate in the labor market in that month. If they choose to participate, they can fire an employee or open a new job.

13) Candidates and firms are shuffled. According to the last existing salary of the candidate and the corresponding income decile, it is verified whether there is car ownership for individual transport or not. The vacancies available are divided according to the exogenous parameter between those whose criterion is by proximity and those by qualification. Each of the lists is ordered so that the firms that pay the highest salaries choose first. The qualification process occurs first and is followed by the proximity process. For each firm that initiates the hiring process, a sample of candidates is selected.

14) After the selection process, however, firms and candidates are organized according to the value of the general indicator of offers; that is, each candidate who participated in each selection process ranks the firm according to salary and transportation cost. Additionally, the qualification is also included in the calculation of the indicator for firms that used this criterion. Thus, the final order is the score of the sum of the candidate's qualification (firm side), plus the salary deducted from the actual cost of transportation (candidate side). The firm of the candidate-firm pair with the highest score does the hiring and both exit the market. The next pair does the hiring, and so on, as long as the pair is still in the market.

15) Soon after, the real estate market transactions take place. A sample of families determined by exogenous parameter is chosen. All properties have their prices updated, and those that are vacant are included in the list of properties for sale and update the information regarding how long they are available on the market. Families are sorted by purchasing power, including potential mortgage loans. Properties are divided between the rental and sales market according to the proportion in the exogenous parameter.

16) The rental market comes first. In this case, families are ordered according to their permanent income variable. The rental properties make up a random sample. If there is a property whose rent is less than the family's permanent income, it chooses a random property. However, when the family is already settled – it is not a migrant family or one resulting from a new marriage – it will make the move only when the move is to a better (more expensive) residence. When the rent is not compatible with the family's income, it proposes a discount, proportional to the size of the vacant property supply, for the cheapest rental property.

17) In the buying and selling market, the family chooses a sample of properties available on the market and tries to buy the most expensive property in the sample. If the property chosen is within the limits of their savings,

they negotiate the price and make the purchase. If the property is within the limits of their savings plus financing, the family applies for housing finance. If financing is denied, the family exits the market. If the property is above the family's savings limits plus potential financing, the next cheapest property is looked at.

18) Families make investments when appropriate.

19) Municipalities invest the collected taxes in public improvements.

20) Monthly model statistics and information are collected.

5 ODD: MODEL DRAWING CONCEPTS

The design of the model sought to represent the processes from the point of view of families and firms, supported by the literature, when available. Following the proposal to build the description of the ODD model, the details of each submodel are shown in section 7.

5.1 ODD: basic principles – agents' decision process

The basic guiding principle of the *PolicySpace2* was the intention to describe the processes of the complex real estate market in an initially simple and intuitive way. From this basic model, constituent empirical data and a set of reasonable parameters were incorporated in order to observe the behavior of the economy after changes in processes and parameters. This was done to better understand the mechanisms of interaction and connection between the economy as a whole and the real estate market in particular, while including several dimensions of analysis of the problem. The *PolicySpace2* brings together in the same model the idiosyncrasies and location of the real estate, the asynchronous construction market, the relevance of production processes, the labor market and endogenous wages, real estate financing and changes in families in a specific empirical environment for each Brazilian RM.

Among the modeling traditions, the *PolicySpace2* manages to unite purely economic and non-spatial traditions, such as those reviewed by Dawid and Gatti (2018), with models originating from mobility analyses (Waddell, 2011), including endogenous wages, absent in typical real estate market analyses (Baptista et al., 2016; Guerrero, 2020; Hamill and Gilbert, 2016) and labor markets, as in the seminal model by Neugart and Richiardi (2012). In addition to including aspects of all these approaches, it is spatially more detailed than models of the traditional land use change stream (Parker et al., 2003), with spatial processes present in all markets.

In terms of design and modeling design, note that in the economic literature processes are traditionally instantaneous with equation solving and price adjustment in order to balance the market. In practice, however, the firm only knows demand after household consumption has taken place and does not have enough information to accurately determine prices and wages. Some central points of the simulation are based on modeling approaches present in the literature; others were included from our own concepts. Notably, we highlight below – as a complement to the description of the sub-models – processes that are difficult to design and how they were implemented.

1) Firm price decisions: Blinder (1994) reviews firms' practices based on a survey and identifies several distinct patterns in pricing decision making. We incorporated some of the behavior suggestions of firms as proposed by Blinder by including exogenous parameters that control the frequency with which firms update their prices. Firms also do not assess the labor market every month, but usually at intervals of three or four months. Additionally, the behavior was chosen in which the firm observes its own inventory to establish prices (Seppecher, Salle and Lavoie, 2017).

2) Salary decision: salaries are determined based on the firm's revenues, after discounting taxes and the size of overall unemployment. The higher the unemployment, the lower the volume of income to be distributed among workers. Distribution is made proportionately to the productivity of each.

3) Household consumption decision: the household consumption decision is based on the calculation of their permanent income (Dawid and Gatti, 2018), so that it is proportional to the monetary resources available, the family's savings and its assets. In practice, all amounts that exceed the calculation of permanent income are directed to savings, while values up to permanent income are directed to consumption. In some cases, when there are no immediate resources to pay loans, rents or consumption, resources are subtracted from savings.

4) Labor market and contracting decisions: the processes used generally follow those described in the literature (Neugart and Richiardi, 2012). They include, for example, the negative relationship between the supply of workers (unemployment) and the definition of wages. The searching by firms for more qualified workers is also followed. Additionally, *PolicySpace2* uses the proximity factor, more specifically the cost of transport conditioned on access to public and private transport services, as a criterion for the worker when choosing a firm. No interaction process or social networks are included in the sample that the firm uses for the selection processes.

5) Real estate negotiation process in the real estate market: the traditional price setting process follows a more abstract framework (Dipasquale and Wheaton, 1996) that is difficult to apply in practice or in ABM, as it is not micro-based and does not describe the path to equilibrium. Even so, the calculation of real estate prices from hedonic regressions is well established in the literature, according to Rosen's (1974) seminal text. The trading price, however, is difficult to measure (Glaeser and Nathanson, 2017). In the *PolicySpace2* we have designed a relatively ingenious process that allows for above and below market prices. Described in detail in subsection 7.7, the negotiation process itself involves two phases. On the demand side, the buyer makes a hedonic evaluation of the price, incorporating the intrinsic characteristics of the property, the municipal quality of life, which evolves according to the dynamism of its companies and is weighted by the weight of the population of each AP and the average income of the families in the neighborhood as a proxy for aspects of perception (Galster, 2001; Furtado, 2011). On the supply side, the seller correctly estimates the family's actual savings, including potential bank loans. The price adjustment, in turn, is made as an average of the two estimates, when savings are sufficient. Otherwise, the seller checks the size of the global property supply on the market (vacancies) and can probabilistically accept a certain level of discount.

6) Municipal management efficiency: an exogenous parameter linearly transforms the resources collected and transferred to the municipal treasury into changes to the municipality's QLI. This indicator is used in the formation of prices in the real estate market. However, note that this factor is only a referential component of the estimated price, and the actual transaction price will depend on the family's savings capacity.

7) Decision to participate in the real estate market and decision between renting or buying: we did not find enough elements to characterize when families decide to participate in the real estate market (Furtado and Souza, 2020), and we only have empirical indications of the frequency with which they do so. We also did not obtain any evidence of the decision process between living in own property or renting; the analysis of empirical data suggests that both families with greater and lesser economic power choose to rent or buy (Furtado and Galindo, 2010). Thus, these two processes are operated through exogenous parameters.

8) Decision to grant real estate financing: the process of the decision to grant a loan begins with the evaluation of three standard criteria: i) whether the bank has resources to lend; ii) if the requesting family no longer has

a loan; and iii) if the set of loans already offered does not exceed the exogenously fixed percentage of the bank's exposure; that is, if it has not exceeded the amount of deposits that must be compulsorily withheld.

9) Productivity of the firms: the productivity of firms varies in accordance with the qualification of their employees (Gaffeo et al., 2008). Two parameters make an overall adjustment of the productivity and consequential quantity of products that each firm produces.

5.2 ODD: emergency

Given the level of complexity to which the *PolicySpace2* evolved, with the number of mechanisms, parameters and empirical data, it is difficult to say which result is specifically emergent. However, sensitivity analysis – performed numerous times in the process of building and developing the model and analyzed in chapter 5 – suggests that worker productivity is the single mechanism with the greatest influence on the overall behavior of the economy. Regarding the real estate market, the efficiency of transforming funds raised into QLI improvement also seems to be especially influential on final prices transacted in the real estate market.

A third factor of relevance in the model is the initial empirical composition of the neighborhoods. Simulations that involve exactly the same set of initial parameters, but that use data for different RMs, result in behavior with a different trajectory for some central indicators, such as the evolution of unemployment, GDP or company profits, for example.

In any case, for a large set of analyzed variations, there is growth or maintenance of GDP and jobs, with relatively low inflation, with few combinations in which it exceeds 20%. The Gini coefficient is also stable for all configurations with most final values between 0.4 and 0.55. Tax payments and general household consumption are also preserved for the vast majority of parameter and mechanism configurations.

5.3 ODD: adaptation

Several mechanisms include decision-making based on the situation observed locally and its consequent adaptation.

1) Families decide to apply for real estate financing if the most expensive property in their sample is above their available savings.

2) Candidates choose the firm according to a ranking criterion, which includes their own transport cost, with their characteristic of being a public or individual transport user.

3) New couples abandon plans to create new families if they are not successful in the rental market. The same is true of new immigrant families.
4) Firms hire or fire employees according to their performance sales made in the consumer market.
5) Construction firms weigh projected profitability and the size of the current housing supply when planning to build new properties.

5.4 ODD: objectives

The *PolicySpace2* does not contain explicit utility functions. In any event, we can list some objectives pursued by the agents throughout the simulation:

- candidates of working age are always looking for jobs;
- unoccupied properties are always available for rent or sale;
- families always seek to buy the most expensive property in their sample, although they only move to the most expensive (better) when at least one family member is employed;
- firms seek to hire workers with the highest possible qualifications or who live closer, in order to reduce their own cost of subsidized transport;
- sellers seek to extract the greatest savings from purchasing families;
- banks seek to make loans to families that have proportionally enough equity to make the payments; and
- families seek to keep their consumption in line with their calculated permanent income.

5.5 ODD: learning

The *PolicySpace2* does not contain endogenous methods of altering behavior from past experience. We anticipate the possibility of implementing an endogenous change in the worker's qualifications.

5.6 ODD: forecast

There is also no explicit provision in the *PolicySpace2*. At three moments, agents consider future implications of present actions.

1) In civil construction, planning for the construction of new properties involves present profitability, as well as the size of the present offer to decide whether to start the construction process of properties that will be ready in the future.

2) Families, when carrying out the calculation of permanent income, consider current interest in the economy and the impact on their future wealth to decide on the share of present consumption.

3) Firms look at past demand and its consequent effect on their finances to decide in the present on whether to increase or decrease the size of the workforce and their future production capacity.

5.7 ODD: perception

Globally, only the indicators referring to unemployment and the general vacancy of properties are known to all agents. The rest of the information is restricted to agents at the time of their interaction, as follows.

1) The dynamics of the entry of families in the real estate market and construction firms. In the case of real estate market negotiations, families know the calculated prices of properties for a small sample. Sellers estimate (correctly) household savings when joint pricing takes place. Additionally, families are able to estimate the average income component of neighborhood families by calculating the real estate prices on offer in their sample.

2) Firms look only at their revenue, the skills of their workers, and overall unemployment to determine wages. Prices are assigned using the firm's own information.

3) Civil construction firms are able to calculate the expected profitability in the planning of new properties from the average size, quality and price of some properties for the intended regions.

4) Workers know the salary offered by the firm when they participate in a selection process and also calculate the distance and cost of transport from their current residence to the firm.

5) Banks know the assets of their clients, as well as their presence or absence in the portfolio when deciding whether to grant real estate financing.

6) Families know the prices and distances of a sample of firms at the time of the consumption decision.

7) Municipalities know the amount of their population on a monthly basis. They are also effective in collecting all taxes; therefore, they know the owners and tenants of real estate, those who have transacted real estate, family consumption, payment of workers and the profit of firms.

5.8 ODD: interaction

Agents interact with each other in the three markets; and with the bank, when they need financing or wish to make investments, as described below.

1) In the real estate market, interaction is mediated by competition, and the family seeks to buy the most expensive, best-quality property from its sample and rent the cheapest of the sample that its budget allows.
2) In the goods market, interaction is also mediated by the prices and location of the firms, among those contained in the sample – which varies every month – of the family.
3) Also on the labor market, hiring depends on the competition among agents, considering the criteria of greatest qualification and cost of proximity to the firm.

There are two other implicit interaction processes. The first refers to the influence of the income of the set of families in a given neighborhood on the prices of real estate in that location. The second refers to the influence of the collection dynamics of firms in a given municipality and its impact, via tax collection, on the municipal QLI.

5.9 ODD: stochasticity

Random processes are used numerous times throughout a simulation as in *PolicySpace2*. In general, every process described as probabilistic – such as participation in the process of marriage, fertility or mortality – involves drawing a number between zero and one, for example, and verifying whether the number is greater or smaller than the probability being tested. They are also random processes every time there is a sample selection among the agents in the model. By way of example, the input data suggest that a 79-year-old male residing in the Federal District in 2020 has a 0.0438 probability of dying over the next year. In the birthday month, the model processes a random number. If the value drawn is greater than the probability, the agent remains in the simulation.

The seed that determines the random number generating process in the simulation is controlled so that it is possible to replicate exactly the same simulation, using the same seed, regardless of the numerous existing random processes. Randomness is counterbalanced by simulating the model over and over again, with different seeds, and the result is presented as the average of the various simulations.

The following processes involve stochasticity.[6]

6. The list is not guaranteed to be exhaustive. Search the referenced code available on GitHub for an even more detailed analysis. Available at: <https://bit.ly/3wHq8Fs>.

1) Agent generation:
   a) in the pairing between qualification and years of study;
   b) in age attribution, within the age group;
   c) in the allocation of initial financial resources;
   d) in the birthday month;
   e) in the entry of immigrants, from the choice of an existing agent to replicate the characteristics – except for financial resources;
   f) at the time of shuffling and allocating the agents in the families;
   g) in choosing the specific address of the property within the AP;
   h) the size and quality of the property;
   i) in the allocation of urban or rural to municipalities with only one AP;
   j) in the process of allocating properties to families; and
   k) in the initial balance sheet of the firms.

2) In the real estate market:
   a) in the composition of the sample of properties that each family verifies;
   b) in the probability of granting real estate financing by the bank, to transpose the criterion of the amount of borrowed resources in proportion to the family's assets;
   c) in the seller's evaluation, if a discount proposal on the calculated price is accepted, in proportion to the vacancy of overall properties;
   d) in choosing the property to be leased, among those that fit the budget in the selected sample;
   e) in the evaluation of the construction company, if it decides to start building a new property;
   f) in productivity, in addition to markup of the construction firm; and
   g) in the monthly increment of licenses for the construction of the APs.

3) On the labor market:
   a) when verifying whether the firm participates in the market in a given month;
   b) in choosing which worker to fire, once the decision is made;
   c) in building the sample of candidates that the firm evaluates;
   d) in the shuffling of available jobs; and
   e) in the probability, according to the income decile of the worker's last salary, for owning a private car.

4) On the market for goods:
   a) in the composition of the sample of firms to be evaluated by the family;
   b) in choosing the decision criterion based on prices or proximity; and
   c) in the firm's decision, prices are updated.

5) Demographics:
   a) in the mortality decision, linked to the input data;
   b) in the fertility decision and consequent process of generating a new agent;
   c) in the distribution of assets and debts of agents who die and are the last in a given family, but who have generated descendants in other families; and
   d) in the composition of the agents that participate in the marriage process, in the shuffling of the list.

5.10 ODD: collectives

The great collective of *PolicySpace2* is the family. The family behaves as a decision-making unit in the consumption and real estate processes, but it is composed of its members who act individually in the labor market.

5.11 ODD: observation – collected data

The data collection and storage process can be configured in the module *run.py*. Data can be recorded monthly, quarterly or annually. Data from firms, banks, construction companies, by municipalities and general are always saved. Optionally, data on individual agents, on agents who died during the simulation, on properties and on families can also be saved.

The call to save the data is made as the last monthly process. The module *output.py* then calculates and saves the series of statistics about each chosen agent group. The file with general data, for example, calculates and saves the following information for the RM set:

- the current month;
- price index;
- GDP index;
- GDP growth;
- unemployment;
- average number of workers per firm;
- household wealth;
- household savings;
- total balance sheet of firms;
- profits of firms;
- Gini coefficient;
- accumulated household consumption;
- inflation;
- QLI of municipalities;
- real estate vacancy;
- average real estate prices;
- families who are renting properties;
- proportion of families whose rent is less than 30% of their monthly budget;
- investment by municipalities from equitably divided resources;
- locally, Property Tax (IPTU), Tax on Real Estate Transactions (ITBI) and segment on consumption;
- or through the Municipal Participation Fund (FPM), portions of the Tax on Industrialized Products (IPI), Income Tax (IR); and
- taxes paid by the bank.

All these statistics are reported in the form of evolution graphs over the simulation period.

In the default simulation (*run*), the graphs only show the evolution of the indicators for the standard configuration of the parameters and the chosen RM – the Federal District and surroundings.

The model results contribute to the purpose of the simulation, especially when performing simulation analysis of the *sensitivity* type, in which you can compare results for ranges of parameter changes.

6 ODD: TEMPLATE INITIALIZATION AND REQUIRED INPUT DATA

The model can be initialized for each of the RMs under analysis.[7] Model initialization can occur with the same seed; that is, with the same random number generator process, so as to repeat exactly the same results, or with a different *seed*, in order to produce different results for each simulation.

In addition to the chosen RM, it is necessary to identify the elements below.

1) The percentage of the population to be used as a sample. We usually run the simulations with approximately 1% to 2% of the population. The model saves the initial generation of agents, so that another simulation, with the same chosen RM and the same percentage of population, only needs to read the agents saved previously. Other factors should remain the same for reading agents.
2) Average number of members per family – for 2000.
3) Percentage of vacancy of residences.
4) Decision to simplify the population by age groups.[8]
5) The start year of the simulation (2000 or 2010). Housing vacancy is estimated from the analysis of the 2000 census (Nadalin, Furtado and Rabetti, 2018).

The data needed to run the *PolicySpace2* include the ones described next.[9]

1) Proportion of urban population per municipality.[10]

7. Manaus, Belém, Macapá, São Luís, Teresina, Fortaleza, Crajubar (current RM of Cariri: urban area of Juazeiro do Norte, Crato and Barbalha), Natal, João Pessoa, Campina Grande, Recife, Maceió, Aracaju, Salvador, Feira de Santana, Ilhéus-Itabuna, Petrolina-Juazeiro, Belo Horizonte, Juiz de Fora, Ipatinga, Uberlândia, Vitória, Volta Redonda-Barra Mansa, Rio de Janeiro, Campos dos Goytacazes, São Paulo, Campinas, Sorocaba, São José do Rio Preto, Santos, Jundiaí, São José dos Campos, Ribeirão Preto, Curitiba, Londrina, Maringá, Joinville, Florianópolis, Porto Alegre, Novo Hamburgo-São Leopoldo, Caxias do Sul, Pelotas-Rio Grande, Campo Grande, Cuiabá, Goiânia, Brasília.

8. Named parameters: *members_per_family*, *house_vacancy*, *simplify_pop_evolution* in the model.

9. Since the data used is required at the time of model initialization, we chose to describe it here, although the ODD protocol recommends doing so in the section immediately following.

10. Resident population, by sex and household status. Available at: <https://bit.ly/3MyTk81>.

2) The Municipal Human Development Index (IDHM), calculated by Fundação João Pinheiro (FJP) and by Ipea.[11]

3) The shapefiles (geo-referenced files) of the minimum statistical units considered, and in the case of the *PolicySpace2*, these are the IBGE's APs, for municipalities with more than one PA, and the municipality itself for those without statistical subdivisions. Additionally, urban areas are also used, as defined by the IBGE, according to the standard of census sectors.[12]

When starting the model in 2000, the weighting areas designed for the 2000 census are used. When the model start is specified for 2010, then the APs from the 2010 census are used. Note that the APs have a different design, geographic spatialization, quantity and code between the two censuses. Both APs used were built from the shapefiles of IBGE census sectors and the list of sectors that make up each AP.

The number of firms per IBGE APs, in fact the number of establishments, is the result of the effort of researcher Vanessa Nadalin, who processed the original bases of the Annual Social Information List (Rais) of the former Ministry of Labor and georeferenced, in the proprietary environment of the ESRI/Galileo system, the geographic coordinates of each company's location through its address and Postal Address Code (CEP).

From the georeferenced Rais processed, we calculated the number of establishments per AP, for each set of APs (2000 and 2010). The simulation, when it starts in 2000, uses data from the 2002 and 2012 Rais. When it starts in 2010, it uses data from the 2010 and 2017 Rais. Note that for 2010 all municipalities were incorporated and not just those belonging to the ACPs of interest. For reasons of confidentiality, the APs with fewer than three establishments were all updated with information from three establishments.[13]

Data by gender and age for 2010 were extracted directly from table 1,378 of the IBGE Automatic Recovery System (Sidra).

The population estimates come from data prepared by the IBGE for the Federal Audit Court (TCU) and published annually.[14] The model uses the estimates to infer population growth, in addition to the endogenous process of fertility and mortality existing in the model.

11. Available at: <https://bit.ly/3lspLJB>.

12. Available at: <https://bit.ly/39CcP10>; <https://bit.ly/3yV3OLf>.

13. Seven APs were imputed for 2002 and five for 2012, both with APs based on the 2000 census. For 2010, since the APs of all municipalities were used, 64 APs contained fewer than three establishments, with a minimum of three being imputed. For 2017, there were 38 imputations.

14. Available at: <https://bit.ly/3LwmyTN>.

For 2000, the years of study by weighting area (V4300) were extracted. For 2010, the education level information (V6400) was available, with the following structure:

- uneducated and incomplete elementary school;
- complete elementary and incomplete high school;
- complete high school and incomplete higher education;
- college degree; and
- not determined.

Based on this information, the model randomly transforms (uniform distribution) into years of study.[15]

The Central Bank series (in 433, 4390, 25497) described in section 6, contain the data required for initialization.

Finally, for 2010, the number of members per family is no longer exogenous – a parameter chosen by the modeler, and is replaced by the average size of families, within the scope of the APs, as determined by the IBGE in the 2010 census.

The process of generating the model families occurs once, before the start of the simulation, according to the choices made by the modeler. To save time, it is possible to use agents previously created in subsequent simulations, as long as the creation parameters remain the same.

6.1 Regions

The process of creating all entities in the *PolicySpace2* starts by importing the geographic limits of the minimum areas used, that is, the IBGE's APs. Thus, depending on the start year of the simulation (2000 or 2010), the corresponding shapefiles will be used. The information needed to create each region, as referenced in the model, is just its unique code – and its geographic boundaries.

Throughout the simulation, the region maintains a series of aggregated information regarding the families and firms located in its territory and updated monthly. They are: i) population; ii) IDHM; iii) GDP; iv) number of permits available for new construction; v) budget record, referring to transfers and taxes received; and vi) sum of displacements carried out by families.

These all start with zero and are endogenously calculated by the model, except for the IDHM index, whose value for the first month is read from data provided

15. Map between available information (educational level) and transformation in years of study: one for one or two; two for four or six or eight; three for nine or ten or eleven; four for twelve or thirteen or fourteen; and five for one or two or four or six or eight or nine.

by Ipea and FJP in the Municipal Development Atlas. From each region, using its AP code at IBGE, the other entities are created within the scope of each region.

6.2 Agents

The number of agents, or individuals, in each region is determined from the gender and age information for each AP. Additionally, years of study for each agent are derived from the results of the census sample. Thus, only the month of birth (uniform distribution) and a stipend of financial resources (also with uniform distribution, between 1 and 34) do not follow registration information.

During the simulation, other relevant factors for the agents are readjusted endogenously: i) belonging to the family (mortality, marriage); ii) employment and salary; and iii) commuting to work.

6.3 Families

The families will gather the agents' budget and carry out the purchase and sale or rental of properties. Participation in the goods market also occurs within the family. At the time of creation, the number of families depends on the number of individuals, and the exogenous parameter depends on the average number of members per family. For 2010, this information is read directly from the average number of members per family per AP from IBGE.

Once the agents and families are created, a process of allocation of individuals in each family is carried out. The generated agents are initially divided between adults (over 21 years old) and children. The first step seeks, insofar as agents are available, to allocate at least one adult to each family. The distribution seeks to maintain a similar number of adults per family, according to the number of adults and families in the region. Subsequently, in the same way, children are distributed among the existing families.

6.4 Residences

The number of homes created is always some percentage higher than the number of families, so there is always a number of empty homes (Nadalin and Igliori, 2016). This percentage is determined as a model parameter, suggesting values between 9% and 11% of the number of families. The residences created have a size between 20 m² and 120 m² and a quality level between 1 and 4. The two parameters are chosen uniformly.

The initial price of the residence is the product of its size, level of quality and the IDHM of the region in which the residence is located. As a result, throughout the *PolicySpace2*, the price composition depends on intrinsic characteristics of the residence (size and quality), but also on a location factor, which, as will be seen, varies as the model develops.

As for the address, in the municipalities where there are no subdivisions by APs, the location of the residences is chosen according to the proportion between urban and rural municipal, using the shapefiles of the IBGE for urbanized areas (2000 and 2010).[16] In these cases, given the probabilistic decision to locate in an urban or rural area, geographic coordinates that are contained in the corresponding urban or rural municipal polygons are selected. When the municipality has more than one AP, the location of the residence occurs in any location that is within the AP. Note that municipalities that contain more than one AP are more populous and contain APs with smaller territorial extensions.

A portion of the households created, determined by an exogenous parameter, is distributed linearly (one household per family on the list). The remaining portion is distributed randomly among all families. When the family receives a residence, it registers the property and, if it still doesn't have an address, it moves. In this way, some families may be left without a residence, while others may receive more than one. Families that at the end of the process remain without a property turn to the rental market. At the end of the process, all properties are registered in the name of a family.

6.5 Firms

The number of firms per APs is determined from the number of firms initially existing in the AP, according to data processed by Rais for 2000 and 2010. Additionally, a percentage of these firms, according to an exogenous parameter, is created as a construction firm, which will operate in the real estate construction market. In addition to the random location within the region, firms also receive an initial equity that follows a value taken from a beta distribution (with alpha parameter = 1.5 and beta = 10), multiplied by 10 thousand. The firms participate on the labor market, hiring and firing employees. The size and skill of their workforce determines the output that is sold on the goods market. Firms' locally conditioned decisions involve setting prices and wages and the timing of hiring and downsizing. Building firms additionally decide in which region to build new homes.

6.6 Banks

In this version, the *PolicySpace2* has only one bank that receives and remunerates customer deposits and makes loans. Its creation considers only an exogenous parameter, that is, the basic interest rate of the economy. Throughout the simulation, the series of the Central Bank related to the following indexes are used:

- the Broad National Consumer Price Index (IPCA) – series 433;

16. Available at: <https://bit.ly/3LyFJfC>.

- the average monthly interest rate on credit operations with earmarked resources (individuals, real estate financing at market rates) – series 25,497; and
- the interest rate (Selic accumulated in the month) – series 4,390.

7 ODD: SUBMODELS

This section describes in detail, substantiates and justifies sub-processes of the *PolicySpace2*, including the purpose and rationale of each sub-model, as well as the parameters and formulas. With this, we will detail the markets for goods, work and real estate, the rationality of firms, bank agents and real estate loans, in addition to demographic processes – mortality, fertility, immigration and marriage. Default parameter values are listed in the appendix of this book.

7.1 ODD: rationality of firms

The firms (i) participate in the labor market, hiring and firing workers (l), in the goods market, with sales to families (h) and, for construction-type firms, in the real estate market, producing and selling homes (H). The sequence involving the firms is:

- exogenous growth of new firms, according to observed population growth;
- the quantity produced varies according to the number of employees and their qualifications;
- when making sales, firms update balance sheet and monthly invoicing and pay consumption tax at the time of sale (details in the goods market);
- calculation and payment of salaries, proportional to qualification/productivity;
- payment of company tax and profit calculation;
- decision making on prices and adjustments;
- construction-type firms plan and build homes;
- decision making on activity on the labor market and participation; and
- participation in the real estate market for construction firms.

7.2 ODD: new firms

The positive number[17] of new firms is determined exogenously, maintaining the base of firms empirically observed in a given period (growth in the number of establishments between 2002 and 2012 for simulations starting in 2000 and

17. In this version of *PolicySpace2* no firms die. However, it may happen that there are firms that do not have employees and, therefore, do not generate new production.

between 2010 and 2017 for 2010 simulations). The new firms for each municipality will be located probabilistically in the most dynamic APs. Specifically, the firm is more likely to be located in APs whose average earnings and employees are also proportionately higher. Once the AP is chosen, the firm is created the same way as the firms created in the first month of the simulation.

7.3 ODD: production

When created, each firm (i) generates a product of its own. Although the simulation supports the creation of several products per firm, only one item is produced per firm. The firm's production (Q_i) is proportional to the qualification of each employee (q_l), raised to an alpha factor of exogenous productivity and adjusted by a beta premium.

$$Q_{i,t} = \sum_{l}^{L} \frac{q_l^{\alpha}}{\beta} . \quad (1)$$

Therefore, the quantity produced per firm ($Q_{i,t}$) is the sum, for all employees, of the years of study (qualification) adjusted by two productivity parameters, raised to alpha (α) and divided by beta (β).

7.4 ODD: decision making on prices – via inventory

The firm checks whether the quantity sold in the month is greater than the present inventory. If so, it increases the price by a percentage of mark-up (π) defined exogenously by the modeler. In this sense, it compares its reality (the production carried out) with the behavior of demand for decision-making whose order of magnitude is external. The frequency with which the firm checks whether prices change is determined by the parameter ζ.

7.5 ODD: decision making on salaries

The firms (i) decide the salaries ($\omega_{l,t}$)of each worker (l) according to total sales ($TR_{i,t}$) and current global unemployment (U_t) and weight it by the productivity of each worker (q_l^{α}), withholding the worker's tax at source (tax_l).

For construction firms, proceeds from the sale of real estate are accounted for over a period of 24 months (no), in order to maintain a more constant cash flow for salary payments, given the more concentrated sales and production.

$$\omega_{l,t} = TR_{i,t} * (1 - U_t) * \frac{q_l^{\alpha}}{\sum_l^L q^{\alpha}} * (1 - tax_l). \quad (2)$$

7.6 ODD: market for goods

Sales. Decision making on consumption and savings. Each family chooses the firm in which it will consume and the amount to be spent, deciding with equal probability ($P(.5)$) whether by price or by distance, given a sample of firms (ς). The amount of product purchased will depend on the price and the family's consumption decision. The goods and services are homogenous, and technology is fixed (Lengnick, 2013). Taxes on consumption are collected at the time of sale. The firm sells all the production requested by the family, up to the limit of its availability.

The decision of households between how much to allocate for immediate consumption and how much for savings is based on current macroeconomic modeling practice. Dawid and Gatti summarize the calculation of permanent income in this way ($PI_{h,t}$): "it is a linear function of current and expected future incomes and of financial wealth. (…) All income in excess of permanent income will be saved and added to financial wealth" (Dawid and Gatti, 2018, p. 78).

$$PI_{h,t} = i_t * \overline{Y_{h,t0-t}} + i_t * \frac{\overline{Y_{h,t0-t}}}{r_t} + w_t * r_t. \quad (3)$$

Using r_t for current interest (t), $i_t = r_t/(1 + r_t)$, $\overline{Y_{h,t0-t}}$ the average salary income of the family for all previous periods of the simulation and w_t is the family's asset wealth, including real estate and bank deposits.

7.7 ODD: real estate market

Every month of the simulation, the call to the real estate market takes place after all the previous sub-processes. The only process that takes place afterwards is the distribution of taxes and data collection. Market share is initially based on the selection of homes available for sale or rent and families interested in one or the other. On the supply side, housing availability is endogenous, with all vacant properties included in the market. On the demand side, there is an endogenous portion, generated from marriage and the formation of new families. Additionally, migrant families seek the market, and enter the simulation, based on the growth estimated by the IBGE annually. Finally, existing families are randomly chosen to participate in the market (σ), reflecting empirical statistics that suggest that, on average, about 6% of families move each year (Causa, Woloszko and Leite, 2019).

The rental market is defined by an exogenous portion of families and properties, given that the decision to buy or rent is difficult to rationalize, considering the uncertainties about future macroeconomic behavior (Malmendier and Steiny, 2017; Furtado and Souza, 2020). Households selected to participate in the market whose budget – including potential bank loans – does not contain sufficient funds to buy the cheapest property available on the market will participate only in the rental market. As a result, the real estate market starts with the construction of

four groups of agents: families looking for properties to buy or rent and vacant properties available for sale or rent.

In any case, the sale value of the properties is updated monthly (P_{ask}). The price is a direct result of the characteristics of the property ($H_{s,q}$) (size and quality) and its location ($N_{m,t}$). The location, in turn, is also updated monthly by the neighborhood's QLI (AP do IBGE) (m). The indicator depends on the taxes collected (tax_t) and passed on to the municipality, weighted by population variation ($\frac{pop_{m,t-1}}{pop_{m,t}}$) and by a linear magnitude adjustment parameter (ψ).

$$N_{m,t} += \sum_m tax_t * \psi * \frac{pop_{m,t-1}}{pop_{m,t}}. \quad (4)$$

Additionally, the price may (or may not) be discounted by the size of the global offer of properties (V) for sale at RM, depending on a decision parameter; and also by the proportion of family income in the neighborhood normalized between zero and one for the set of RM (τ) (Ge, 2017). Finally, the number of months the property has been for sale (T) can interfere with the price, limited in its fall (γ) and with adjusted intensity (κ).

$$P_{ask} = (1 + V) * H_{s,q} * N_{m,t} * (1 + \tau * N_q) * ((1 - \gamma) * e^{\kappa * T} + \gamma). \quad (5)$$

On the family side, the supply (P_{offer}) includes a home loan estimate (L_h) when savings are not enough. Families seek to buy the best property, equivalent to the most expensive (Goldstein, 2017). The final price (P) traded is the simple average of the asking and offer price.

$$P_{offer} = S_h \vee S_h + L_h \ if \ P_{ask} < P_{offer}$$

$$P = \frac{(P_{ask} + P_{offer})}{2} \quad (6)$$

The bank loan is limited by the factor *loan-to-value* (LTV), so that the ratio between the loan amount and the price of the property cannot exceed a certain proportion. Additionally, there is an upper limit (ρ^+) and a lower one (ρ^-) for the final price. When the selling price is above the total offer, but within the limits of the lower limit, there is a probability that the seller will accept the discount proportionally to the current vacancy in the market $P(\sum Listed / \sum h)$.

$$\frac{L_h}{P} \leq LTV$$

$$if \ \frac{P_{ask}}{P_{offer}} > \rho_+ \longrightarrow P = P_{offer} * \frac{\rho_+}{2}$$

$$if P_{ask} > S_h > \rho_- \longrightarrow P = S_h \mid P(\sum Listed / \sum h) \quad (7)$$

Box 1 presents the sequential listing of events that occur in the real estate market.

BOX 1
Real estate market: sequence of steps

1) Random selection of families interested in buying or renting. Families created through marriage and newly arrived immigrant families are added.

2) Inclusion of all unoccupied residences.

3) Updating of prices for all real estate (sales and rent).

4) Families interested in purchasing consult the bank for the possibility of potential loans. They check what their credit limit is, given the age of the oldest member, long-term family budget, and interest.

5) Entry into the real estate market takes place in order of purchasing power, with families with greater resources and access to credit choosing first.

6) Among the unoccupied households whose owners are families, and not builders, a percentage is selected[1] for the rental market. Therefore, the size of the housing supply for rent is exogenous. On the supply side, endogenously, those interested in rents are families without financial resources to purchase.

7) Next, families randomly choose a subsample of properties.[2] They order the properties in the group starting with the most expensive, which indicates the highest quality (Gilbert, Hawksworth, and Swinney, 2009; Goldstein, 2017), and check whether their savings, or savings plus credit, are enough to buy the home. If there are no properties in their group or if the bank does not provide the loan, and the existing savings are not enough to make the purchase, the family leaves the market.

8) The transaction price is the average of the household's savings and the calculated house price, limited to a higher proportion by an exogenous parameter. In this way, there is a match between supply and demand, with the seller correctly estimating the buyer's ability to pay, and the buyer performing an accurate analysis of the property's value.

9) In the event that the family's savings are less than the estimated selling price, the family can make an offer with a value lower than the seller's request. The probability that the seller will accept is equal to the size of the offer. Therefore, if there is a 10% vacancy for properties in a given month, there is a 10% chance that the seller will accept the offer.

10) Once the transaction is completed, the property transfer tax is collected; and the money transferred from the buying family to the seller or construction firm.

11) The purchasing family then has at least two residences and decides whether to move or not.

12) Families and residences for rent are sent to the rental market. The rental market also welcomes immigrant or newlywed families.

13) The family, in the rental market, is also looking for a subsample of real estate.[3] They randomly chooses a property whose rent value is compatible with their monthly expenses (through the calculation of permanent income).

14) If there are no properties in these conditions, the family will seek to negotiate a discount that may vary according to the size of the properties available. The lower the number of properties on offer, the lower the requested discount.

15) Once the property, family and price have been chosen, the family that already has a fixed residence (rented or owned) checks if the property to be rented is of better quality (more expensive), in relation to the current one. The family that is not allocated (immigrant or newly married) moves.

16) The rent payment is collected.

Author's elaboration.
Notes: [1] Parameter: *RENTAL_SHARE*: default value = 0.4.
[2] Parameter: *SIZE_MARKET*: default value = 10.
[3] The same parameter of the real estate market is used, *SIZE_MARKET*, multiplied by three, in order to guarantee a larger subsample in the rental market, since it is more competitive.

7.8 ODD: sales

In the case of sales, families looking for a property to buy are ordered, and those with greater financial resources – including possible bank loans – choose first. Each family then "visits" (assembles a group of possibilities) that is three times the number of firms that the family consults in the goods market. Typically, with the default values, the family initially examines thirty properties among those for sale; removes from the sample the properties that are more expensive than its ability to pay, with the possibility of borrowing; and then tries to buy the best property; that is, the most expensive (Goldstein, 2017).

When applying for a loan, the success of the transaction is dependent on credit approval by the bank. If the savings balance is sufficient, or if they get the loan, the family buys the property. The value of real estate financing (L_h) is at most a percentage (χ) of the permanent income of the family (PI_h) and the maximum between the number of months (m) until the oldest family member reaches 75 years of age or the loan reaches 360 months. If the loan is not granted by the bank, the family tries to purchase the second property on their list.

$$L_h = PI_h * \chi * m. \tag{8}$$

The property price is based, on the one hand, on the family's ability to pay and, on the other, on the calculated price of the property. The logic is that the seller estimates the maximum value the family has available, and the buyer makes the offer based on the calculated price of the property. The final price is established as the simple average of the two. If the previous step is met; that is, if the savings are greater than the estimated price of the property or if the savings plus the successful loan is greater than the price, then the transaction is signed. The property is transferred and registered to the purchasing family. Resources change hands and ITBI is paid for the region where the property is located.

Every time the family buys a property, it evaluates the possibility of moving to the new property. If the family does not have another residence, they will move to the property purchased. If they have another home, they will move to the lowest-priced property – thus reflecting its location, size and quality – if all adults in the family are unemployed. If at least one adult in the family is employed, the family will move to the property with the highest value.

The rental and sales market is described in the next subsection.

7.9 ODD: rent

As in the sales market, families looking for rent examine a sample of rental properties. The rental price is given by a fixed percentage of the property price, which is variable. The family initially seeks a property that is within their monthly payment possibilities, according to the calculation of their permanent income. If they

do not find one, they then make an offer for the cheapest property among those initially selected. The percentage reduction in the offer made is equivalent to the size of properties available on the market. If the standard maximum number of rental properties is available, this reduction is 0.03 percentage point (pp), or close to 10% of the standard index (0.29%) on the property price. If the list of available properties is restricted, this reduction request is smaller. Finally, after negotiation, in the case of new families (migrants, marriage), the contract is signed and the family moves. If the family already has a residence in the RM, it checks if the quality (price) of the new residence is better. Otherwise, it gives up the rental. In cases where families are unable to honor rent payments, landlords pay the loss. In any case, default numbers remain low throughout the simulation.

7.10 ODD: civil construction market

Construction firms are derived from the main entity firms, with some specific changes. There are two main methods: i) plan and operationalize construction of a new residence; and ii) and actually build one. The first planning stage is to choose the region for construction, according to those that have lots available. The lots (called in the simulation *licenses*) are offered by the city as new subdivisions and cost the index price of the region ($N_{m,t}$) times the percentage of the cost of the lot in relation to the cost of the project. Determining the size of the new household (H_s) and the quality (H_q) chosen are random. However, they impact cost, which is calculated as the product of size and quality, and the firm's productivity, which is a function of its profit margin $f(\pi)$. The builder, then, based on the planning of the house, chooses similar houses – that is, with a difference in size greater or less than ten and with greater or lesser quality than one – where lots are available. Then the average prices are calculated , limited to 100 households per region (APs) under analysis. The construction company chooses to build in the region (N) with the highest profit $N_{\pi,m,t}$.

$$N_{\pi,m,t} = \overline{P_{ask,m,t}} - \left(H_s * H_q * f(\pi) * N_{m,t} * (1 + v)\right). \quad (9)$$

The process of building the residence itself occurs through the accumulation of products that the traditional firm produces. When the amount of products accumulated is greater than or equal to the cost of construction, the firm makes the financial adjustment, registers the new property and makes it available for sale on the real estate market. The size of the offer (property vacancy V) makes up the prices at the time of negotiation.

7.11 ODD: banks and real estate financing

The financial system of *PolicySpace2* is simplified and aims only to remunerate households' deposits and provide resources for financing on the real estate market. A bank only controls the loan bureaucracy and household deposits. Firstly, during the simulated month, the bank collects the payments of the families' real estate financing; then, it opens up the possibility of credit for financing families in the real estate market; and, finally, it remunerates household investments.

The bank has restrictions that are observed when applying for loans to families to finance the purchase of real estate. The bank must have cash in excess of the loan application. In addition, it is necessary that the total of resources already made available to society does not exceed the percentage defined by the modeler (the mandatory regulation fit). Finally, the bank does not grant a new loan to families that have not completed the payment of previous transactions. Each family can therefore only have one active loan.

The maximum available credit can be consulted by the family and is obtained by calculating its permanent income, limited to the maximum percentage of indebtedness, given by an exogenous parameter. Additionally, the number of years converted into months is considered for the oldest member of the family to reach the maximum age. Finally, the amount possibly to be borrowed is defined by the maximum monthly commitment times the number of months until the monthly term, divided by the interest. Additionally, the loan is limited to a property price ratio, stipulated by the LTV parameter.

Interest on deposits is fixed, exogenously established, fixed at the time of the contract, according to the Central Bank series, and credited monthly. The cost of financing also follows the market value observed for individuals. The installments are charged from the deposits made by the customer and available in the account. If the resources are sufficient for the monthly payment owed, it is paid. Any overdue installments are accumulated and can be paid in subsequent periods, immediately after there is a balance.

7.12 ODD: demography – mortality, fertility, aging and inheritance

Three processes of the agents' life cycle – given their dynamics – are inserted in the *PolicySpace2*: mortality, fertility and aging.

Aging is operated directly. Given the birthday month that each agent receives in the act of creation (*sic*), when the current month is the birthday month, the agent's age is increased by one year. The birthday month is also used as the annual month of check-up of the agent, checking the official probability, according to the state mortality table by gender and calendar year, and the possibility of fertility. Women aged between 14 and 50 years are also evaluated in terms of

their likelihood of becoming pregnant, also in accordance with state fertility data, by age and calendar year. The birth process is like the initial agent creation process. The child is added to the parent's family.

When a death occurs, the death certificate is registered (the data are available in the list named *grave*). If the deceased agent is not the last member of the family, the real estate, savings and debt remain with the family. The agent is dropped from the firm's records if he or she worked. However, if the deceased agent is the last member of the family, the existence of relatives from the original family is verified, if the agent has married from the beginning of the simulation (2010).[18] In this case, an inventory is needed and a series of checks are carried out.

1) The residential property is vacant.
2) If the deceased agent owes a debt to the bank and relatives, the relative who receives the most expensive property also receives the debt. If there are still other relatives and properties, they are randomly distributed (given the immediate indivisibility of the property). If there is only debt, it is distributed randomly among the relatives.
3) The savings balance is withdrawn from the bank and distributed equally among relatives.

7.13 ODD: immigration

The immigration process takes place at the beginning of the monthly processes, right after the demographic check (aging, mortality, fertility) and before consumption and interaction in the markets. The main purpose of this sub-module is to keep the total population compatible with the growth levels observed in the metropolises over the period. Thus, for each year, the population equivalent each month to 1/12 of the migrants is created and added to the model. To this end, estimated municipal data for the 2001-2017 period are used. Missing data are estimated by a simple linear function.

The generation process is similar to the beginning of the simulation of the *PolicySpace2*. By municipalities, the new estimated agents are created and then allocated to families, in the same way as the initial generator module. The new families then go through the process of acquiring or renting properties. Unlike the initial moment, in which all families obtain a property, in the migratory process, respecting the logic that the creation of properties is endogenous to the model,

18. Note that, in this case, the model's genealogical record of relatives only operates for marriages from the start year of the simulation. It is not feasible to estimate previous family ties.

families that cannot obtain a property – due to lack of resources or lack of supply in the rental market – are not added to the simulation. When they obtain residency, new agents and families are registered in the model.

7.14 ODD: weddings

The possibility of marriage occurs shortly after immigration. The insertion of marriage into *PolicySpace2* brings dynamism to families. It is possible to create new families from single adults or with the separation of existing couples and the formation of new unions. This family dynamic affects the economic support of the model in several ways:

- first, the family is the financial decision-making collective, participating in the property and real estate markets together;
- the family is also the repository of wealth, both in terms of real estate held in the portfolio and in bank deposits; and
- additionally, family changes lead to the endogenous configuration of participation in the real estate market, since, when married, family members with other adults need to acquire or rent a new home.

The process starts exogenously, according to the parameter that determines the proportion of agents that will participate in the process. In the standard model, this initial proportion is 3.4% of people per month. From this contingent of people, a second probability check by age is applied. Finally, an additional constraint with truncation is considered, with zero probability of marriages for agents up to 20 years old. Gender restrictions are not considered. The selected people are shuffled and matched two by two. The following options are considered for each engaged pair.

1) If both original families have at least one other adult (age 21 or older), the new spouses leave the family and form a new family if they are successful in the real estate market.
2) If only one of the families has another adult, then the new spouse of the family that has an adult leaves the family and moves into the family with the other new spouse. In this sense, he or she leaves behind the family goods, but starts to share the goods of the new family.
3) If both families have no other adults, but possibly have children, then they merge into a new single family, with all the assets of the component families.

7.15 ODD: labor market

The labor market processes are the penultimate occurrence of the month, immediately preceding the real estate market. The first step is the offer of the agents, followed by the interest of the firms and the adjustment between candidates and companies. The choices of interested workers are weighted by their mobility and the salaries offered. Firms do not evaluate the labor market each month. They enter the job market following the exogenous frequency parameter (ι); of course, this can be adjusted to every month, for example. When in the market, firms participate as a contractor if they make a profit – or financial stability, zero profit. When there is a loss, they send an employee away, who is chosen at random.

The group of candidates comprises all those who are not retired (less than 70 years old), are not minors (over 16) and are not currently working. The list of job applicants is randomly shuffled every month. Job seekers evaluate the (endogenous) distances between home and business, their financial capacity and the consequent use of the public or private transport system and the salaries potentially offered in the ranking of offers.

Pairing. First, firms are randomly divided between the group that will seek candidates according to the qualification criterion and the group that emphasizes proximity. Group sizes are defined by an exogenous parameter (η) which can even be zero (all by qualification) or one (all by proximity). Firms in each group are ordered in descending order by the base wages offered, considering current unemployment in the standard model. In practice, this means that companies are ranked in terms of a percentage of their recent turnover. In addition, the higher overall employment is, the lower the percentage of firms' sales that will be predicted for payment of salaries. Firms with high revenues choose first.

Each firm then selects a random pool of candidates from among those who remain through the process without receiving an acceptance from any firm. The size of this group of interviews is exogenously determined (σ). The ranking of candidates involves an adjustment of the interests of the firm and of the candidate's interests. On the firm's side, qualification is included in the calculation; from the candidate's point of view, the potential salary of the firm ($w_{l,t}$) and the distance between their residence and the firm ($d_{l,h\,-\,i,t}$), weighted by the transport system available to the candidate ($c^{tr}_{l,t}$), given their previous income level. In practice, candidates with access to a private vehicle penalize distance more heavily, both for the monetary cost and the cost of travel time.

$$s_{l,i,t} = q_l + \sum_i w_{l,t} - d_{l,h-i,t} * c^{tr}_{l,t}. \tag{10}$$

In pairing only by proximity, qualification is not included in the calculation, and the choice of the firm is based on proximity, weighted by the transportation

system available to the candidate. According to the level of the last salary deciles (endogenous) and empirical data, the model uses the probabilities associated with the income deciles to identify workers with and without a private vehicle. The process is started again in the following month, creating new lists of firms and candidates and conducting pairing.

7.16 ODD: collection and distribution of taxes

The process of collecting and distributing taxes maintained the process described in Furtado (2018b; 2018c).[19] For each triggering event, five types of taxes are collected. They are: on household consumption (proxy for the Tax on Circulation of Goods and Services – ICMS); on wages paid to workers (*proxy* for the Individual Income Tax – IRPF); on corporate profits (proxy for the Corporate Income Tax – IRPJ); about the property (*proxy* for property tax – IPTU); and on the transfer of real estate (proxy for the ITBI). Once collected, taxes are redistributed at the end of the month according to two rules, established through the following parameters:

- presence or absence of the FPM, with a default value of presence; and
- maintenance of municipal budgets as they are or, alternatively, merging the municipal funds into a single budget within the scope of the RM (ACP), with the status quo used as the default value.

7.16.1 ODD: FPM

The first form of distribution occurs according to the FPM (*FPM_DISTRIBUTION* = *true*). The collection is given from a percentage (23.5%)[20] of taxes on workers' wages and firms' profits. Distribution among the municipalities of an RM follows the same proportionality seen in the real distribution of FPM.

7.16.2 ODD: merging municipal budgets

The second form of distribution follows the alternative proposal of merging the municipal budget and, consequently, its redistribution in an egalitarian way, weighted by the population, among the municipalities. The status quo, still, *ALTERNATIVE0* = *true*, keeps the municipalities independent, so that the municipal collection (municipal taxes and ICMS share) generates benefits only for the municipality where the companies are located, regardless of the workers' residence.

All funds collected and transferred to the municipal treasury are converted linearly through parameter ($N_{m,t}$) and become an increase or decrease in the municipal QLI,

19. In that simulation, the extreme cases of resources distributed exclusively locally or equitably were also tested.
20. This percentage follows the FPM legislation on IRPF and IRPJ. The percentage can be changed via the parameter *TAXES_STRUCTURE*, with the FPM key. There is no IPI in the model. The exogenous proportionality is observed annually until 2016. Thereafter, the values for 2016 are used.

according to the formula presented in equation 9. Note that the parameters only change the way in which the taxes collected are spatially distributed; they do not change the amount of resources. However, redistribution alone is sufficient to generate distinct and robust effects on macroeconomic indicators.

7.17 ODD: designing policy experiments

This subsection describes the mechanisms implemented to carry out the test of experimental design of public policies at the municipal level. The discussion and results of the analysis are contained in chapter 6. Based on endogenous tax resources collected in each municipality, according to the five tax types (IRPF, IRPJ, ICMS, ITBI, IPTU), the experiment consists of retaining a percentage, defined by a parameter (δ), and applying it into three separate policies. The applications are compared with the simulation for which there is no retention of resources and there is no policy enforcement, which is considered the standard case.

In the application of policy alternatives, an endogenous registry of families is prepared based on the list of families with the lowest permanent income ($PI_{h,t}$) to the exogenously determined quantile (θ). When the policy is applied, the families that were included in the registry during the last year and still reside in the same municipality are ranked according to the permanent income calculation, so that the poorest family will be the first to receive any policy aid.

Public policies at the municipal level are described below.

1) Baseline: in the standard case, there is no retention of resources, and the investment normally occurs to increase the municipal QLI.

2) Acquisition and distribution of real estate: in the experiment in which *POLICIES* has the value *buy*, the available resource is used to purchase municipal properties and transfer ownership to families. In this case, the properties that are ready and still available in the portfolio of construction companies and located in the municipality implementing the policy are selected. The properties are then ordered so that the cheapest available will be the first to be purchased by the municipal government. Families who own property are excluded from this policy, and only those who are tenants are eligible. With the families and properties ordered and the amount to be invested available, the municipality buys the properties and transfers them to the families successively, until the resources are exhausted. Recipient families move immediately. The ITBI that applies to the transaction is collected normally and transferred to the municipal coffers, which in practice configures a discount to the municipality in relation to the market price offered by the construction company.

3) *Vouchers* for rent payment: in the experiment in which *POLICIES* is designated as *rent*, in the same manner, only families that do not have their own property are eligible. In this case, again, as long as available resources last, the municipality distributes vouchers that guarantee the payment of the rent of the current residence of the families for the next 24 months. If the family moves to another property, the amount related to the remaining payments, if any, is reverted to the same fund and will benefit other families. The voucher is also only provided once every two years to the same family. However, if the family is added to the registry again at the end of the period, it may possibly receive the aid again.

4) Monetary aid. The third experiment is selected when *POLICIES* is set to *wage*. In this case, the entire available monthly resource is distributed equally among the families of each municipal registry as pecuniary aid. Families remain on the register for a period of one year. In this case, there is no requirement for the absence of own property. That is; if the family's income is below the income quantile, as determined by the policy, it will be included in the register and receive the monthly aid from available resources, divided by the number of families qualified in that month.

8 OPERATIONALIZATION OF SIMULATION

The model simulation *PolicySpace2* requires the installation of the free and open-source program Python[21] and several of its libraries.[22]

The default simulation uses Python version 3.6, but we also simulated it in versions 3.7 and 3.8. Geographic libraries (*gdal*, *fiona*, *shapely and geopandas*) are not immediately installable, so we suggest installing via the Python library aggregator called Conda.[23] Our Conda version was 4.8.4. We also use, throughout production and testing, the educational license interface *PyCharm Professional*, made available by *JetBrains*. The simulation was performed concurrently and interchangeably on computers with Windows and Linux operating systems. We also suggest the creation of a specific environment for the simulation, in order to avoid conflicts between library versions. Use the following sequence of commands in the terminal (with Conda installed) to create, activate the environment and install the following libraries:

```
conda create --name ps2 python=3.6
conda activate ps2
```

21. Available at: <https://bit.ly/3wN29Eu>.
22. To access the complete list, see the GitHub repository. Available at: <https://bit.ly/3PEHhZ1>.
23. Available at: <https://bit.ly/3wCMybM>.

```
conda install shapely gdal -c conda-forge
conda install fiona pandas geopandas numba
descartes scipy seaborn pyproj matplotlib six
cycler statsmodels joblib scikit-learn flask
flask-wtf psutil -c conda-forge
```

With this, the modeler will have created an environment, activated the environment and installed the main libraries together, ensuring the compatibility of versions between them.

To run the simulation itself, it is necessary to clone the repository on GitHub, which can also be done in a computer terminal, with the Git program installed, using the command *git clone*.[24]

Next, set at least the option *output_path* of the file *conf/run.py*, in order to choose where to write the model's table and graph output. Other changes can be made in the parameters module, in the file *conf/params.py*.

The *PolicySpace2* contains automated execution of several parallel simulations, using more than one core from the computer. All commands accept the -n 3 parameters to specify the number of times – for example, three – that the simulation will be run, presenting individual results and the averages. Additionally, you can specify -c 4, to identify the number of *cores of* the computer that will be used simultaneously. Thus, the simplest command for a simulation is *run*:

```
python main.py run
```

To run a simulation with a given set of parameters ten times, using two computer cores, use the command:

```
python main.py -c 2 -n 10 run
```

The sensitivity analysis, also built into the simulation, requests that for each continuous parameter, the modeler inform, in the following order, separated by a colon: parameter name: minimum parameter value: maximum parameter value and number of intervals between the minimum and maximum. As a result:

```
python main.py sensitivity ALPHA:0:1:7
```

In the case of parameters of the true or false type, just the parameter name is enough, for example:

```
python main.py sensitivity WAGE_IGNORE_
UNEMPLOYMENT
```

24. Available at: <https://bit.ly/3sU1sZc>.

Especially in the case of comparison between RMs, it is enough for the modeler to include the names of the regions of interest, separated by a hyphen and enclosed in double quotes, as in the following example:

```
python main.py sensitivity "PROCESSING_ACPS-
BRASILIA-CAMPINAS-FORTALEZA-BELO HORIZONTE"
```

It is also possible to combine continuous, Boolean parameters or RM names, as long as the parameters are separated by spaces:

```
python main.py -n 4 -c 12 sensitivity
MARKUP:0.05:0.15:4 WAGE_IGNORE_UNEMPLOYMENT
"PROCESSING_ACPS-BRASILIA-VITORIA"
```