

Optimal policy: which, where, and why

Bernardo Alves Furtado^{1,3} and Gustavo Onofre Andreão^{1,2}

¹Department of Sectorial, Infrastructure and Innovation Studies,
Institute for Applied Economic Research.

²Institute of Economics, UNICAMP.

³National Council for Scientific and Technological Development.

Abstract

The paper exploits a simulation environment to compare "ex-ante" policy instruments. We create a progressive score to contrast six single and mixed instruments. The multiple simulation runs produce results for distinct instruments, cities, and goals. The outputs of the simulation provides a counterfactual arena where we explore public investment trade-offs quantitatively and empirically—which constitutes a rare (usually impossible) policy practice. Results suggest that a mixed policy instrument evaluated by a comprehensive indicator performs better. However, optimal policy classification changes when considering different places or goals.

Keywords: Public Policy, Policy Mix, Policy Design, Policy Instruments, Agent-based Model

JEL Classification: R23 , R31 , C63

1 Introduction

Policies should be evaluated. Contemporary societies demand that the multitude of data, knowledge, and scientific methods inform alternative paths of policy implementation and their comparative impact. What is the best policy? What will change in people's lives as governments invest? How much improvement shall one see as a result of this policy choice when compared with any other?

However, success refuses to be defined. Does a successful application of a policy instrument refer to increasing GDP? Is it enough? Or, does it refer to reducing inequalities? When? Where? Judged by what rationale?

Consider housing policy as the domain that needs improvement. Municipalities act as principal agents that decide how to invest public funds. Success would then refer to how many households received new properties in a given year? Or would a (continuous) rent voucher program for a larger number of households suffice? Would direct transfer of cash to households be proper? The increase in demand to construction companies is part of the success of the policy program?

Grasping this interplay of policy instruments requires defined goals, counterfactual analysis, and description of both instruments and recipients. Preferably done "ex-ante". Typically, however, policy evaluation cannot count on all of these requirements. They are not feasible.

We adapt an existing, open-source agent-based model simulation to apply different policy instruments and capture results after a 10-year period. We analyze both single and mixed policy designs—totaling 7 policy instruments—in an environment with 24 different cities and 1680 distinct simulation runs.

This setup enables the objective comparison among instruments and cities, within a context of 20 quantitative dimensions of impact. The comparison departs from a multidimensional evaluation indicator based on Principal Component Analysis (PCA). However, we modify the construction of the score so that it leads to a socially better—policymakers' preferred—results.

The development of the paper follows the recent literature, connecting its taxonomy and concepts to the procedures made. We also provide a test of robustness. The paper contributes as a quantitative, counterfactual, empirical environment where comparisons take place. Moreover, we propose a modified dimension reduction procedure that enables the synthesis of the data, following transparent decisions on policy impact.

Hence, the present paper promotes a picture of trade-offs of public investments, otherwise unavailable. Such picture is distinct from typical policy evaluation: occasionally done, "ex-post", for selected, immediate results; without considering alternative policy instruments.

The second section defines policy mixes, its instruments, and the taxonomy of evaluation. The section ends considering the uncertainties of the connection between instruments, indicators, and outcomes. Section 3 describes the simulation environment, the policy instruments proposed, and the construction of our score. We list the score step-by-step procedures, including the cities' and robustness' analysis. The paper concludes with the results, discussions and final considerations.

2 Literature

Policy mixes¹ involve "... complex arrangements of multiple goals and means which, in many cases, have developed incrementally over many years" [2, p. 395]. Policy mix serves to analyze synergies and tensions from different policies in a system [3]. In economics, policy mix refers back to post-great depression

¹[1] provides different definitions and comparisons.

macroeconomics² to describe the interplay between fiscal and monetary policies [6, 7]. Evolutionary economists in turn focused on innovation policy mix analyses.

Three generations of scholars and policymakers analyzed the connection between policy regime effectiveness and instrument choice [8]. The first generation (1950s-1980s) studied the effects of policy and regulation on business efficiency. Policy goals coincided with political goals and split into ideological, partisan or blame-avoidance. Analyses were either pro-market solutions or against them. The second generation (1980s-2000s) divided instruments between substantive (those that affects goods and services) and procedural (those that alter policy processes). Nevertheless, the authors focused on single rather than multiple instruments. Mainstream policy studies defined policymaker as a "... single, [perfectly] rational or bundled rational actor..." [4, p. 705]. Such concept of policymaker as a rational player then started to fade out from such policy studies. Institutional economics also focused on broadening the analysis of public governance [9].

Policy studies afterwards shifted towards the interaction among evolving agents that analyze, learn and act in complex systems. Contemporary third generation (2000s-) suggested modeling as a methodological tool, and proposed four steps for policy mixes [8]:

1. Assess a wide range of available policy instruments;
2. Focus on instruments that have synergies;
3. Use different instruments, such as self-regulation and incentives;
4. Consider information instruments.

In short, policy mixes assess (a) which instruments can be mixed; and (b) how they integrate [8].

Policy mixes research focus on avoiding inconsistency of instruments and incoherence of goals [1–4, 10]. Policy mixes that are both consistent and coherent are congruent: the goal for any policy mix and its processes.

Consistency "... captures how well the elements of the policy mix are aligned with each other, thereby contributing to the achievement of policy objectives" [1, p. 1626] and break into: consistency between strategies, between instruments, and between strategies and instruments.

Coherence³ refers to the "... synergistic and systematic policy making and implementation processes contributing – either directly or indirectly – towards the achievement of policy objectives" [1, p. 1626]. Coherence applies to different policy fields, the capabilities of policymakers, and the effects of policies.

[1] adds credibility, comprehensiveness, and different dimensions (policy field, governance and geography) to the concept of congruity. Perfect congruity is less tangible as analysis broadens, and optimal policy mix is harder to pinpoint⁴. [5] corroborates the difficulty to evaluate policy mixes and describes how it requires large amounts of quantitative and qualitative data. [3] emphasizes the dynamics of policymaking and states that tensions between different

²[4, 5] attribute the origin to 1960s Keynesians.

³see also [3].

⁴In line with [11, 12]

implementation processes may cause incongruity. Clearly, the more dimensions a policymaker has in mind during policy elaboration, implementation and assessment, the harder it is for the analyst to compare different policy mixes.

[13] shows that the selection in a principal-agent game⁵ becomes harder the more elements one has in mind during its decision-making process. Moreover, the decision process becomes more difficult with many dimensions, as the weights attributed to each element are subjective. Decision problems are reduced to one-dimension in which each option depends on different aspects. Decisions that consider multiple elements are either intractable or need simplifying to become feasible⁶.

[8] provides a taxonomy for situations with multiple goals and instruments (table 1). An optimal arrangement occurs when there are no conflicts among goals and among instruments. Otherwise, it fails. An optimal arrangement have low tensions and high synergies among the policy mix [3].

Goal mixes	Instrument mixes	
	Consistent	Inconsistent
Coherent	optimal	ineffective
Incoherent	misdirected	failed

Table 1: Typology of policy mix arrangements based on the relationship between goals and means. Source: [8, p. 8]

Intermediate cases happen when either goals or instruments have conflicts among them. Conflicting goals lead to misdirected arrangements, whereas conflicting instruments lead to ineffectiveness. Misdirected arrangements work properly but fail objectives. Moreover, synergies between instruments are insufficient given incoherent goals. Ineffective arrangements fail with coherent objectives. Inconsistent policy mix hardly reach coherent goals.

[3] argues that inconsistency and incoherence may be unknown during implementation and become problematic when received by other agents of the system. It is not trivial to know beforehand which policy mixes may be congruous. Uncertainty is the reason why *ex-ante* best policies are unknown. [12] provides a taxonomy of uncertainty in economic systems. Systems are uncertain when the distribution of risk is varied, additive and reliable.

Furthermore, when the decision making of agents is limited, the system has either procedural or fundamental uncertainty, the latter being a specific case of the former. Fundamental uncertainty contains novel actions⁷, whereas procedural uncertainty agents know the possible actions, but cannot assess their probabilities. Both cases include simultaneous actions and responses that depend on others [12]. Such actions do not need to be taken at the same

⁵The context is the election of representatives.

⁶It is possible to decide multidimensional problems, nevertheless decision makers must be robust and it will take longer

⁷The author cite innovation and institutional change as sources of novelty. Technical and institutional changes are connected through co-evolutionary processes, which emphasizes the uncertainty [14–16]

time (especially in computational simulations) but at the same time frame. This dependency generates a procedural uncertainty of the consequences of one's actions. The payoff of an action depends on others under fundamental uncertainty⁸.

[5] provides a comparative taxonomy of a policy mix: rationales (or goals); domains; instruments; and actors⁹. Rationales support and shape the policymaking process (forward-looking), and determine how to assess results (backward-looking). These backward and forward looking reasoning justify policy implementation. Whereas policies have theories behind them, reasoning is provided retrospectively [4].

Actors are the principals of policy (in game theory terms) [4, 5]. The authors coined the term "policy subsystem" as the collective of agencies, regulators and other actors that shape policy. From an institutional standpoint, agents and rules' recipients shape the rules. Actors implement, assess and review such rules [18, 19].

Domains reflect types of policy, such as housing or energy policies. Different instruments may arise from domains. Energy policy, for example, may combine auctions and a public financing scheme to achieve a certain goal [20, 21]. In sum, policy mixes are the emergence of the Rationale-Domain-Instrument-Actor combination, through the dynamics of the policymaking process [1].

Instruments are the means that implement the policies. There is fundamental uncertainty about which aspect of a certain instrument is responsible for which outcome [4]. Especially so with policy mixes which include interactions, dynamic development, and learning [11]. As such, it is difficult to pinpoint which instrument within each domain of policy generates an outcome, given the interplay between the policy itself and the system characteristics'.

Following [11], policy elaboration and assessment are subjective. However, they should be evidence-based. As such, policymakers should use a wide range of indicators to assess the impacts of their policies, especially if they want to achieve transparency¹⁰.

[23] analyzes congruity beyond national borders in OECD countries. The authors find that similitude between policy mixes of neighboring or cooperating countries leads to better achievement of policy goals. In terms of applicability, [5] analyses a wide range of policy mixes. [2, 3, 8, 10, 23] provide single case studies.

3 Methods

This section describes the procedures taken to specify and construct the scores to evaluate an optimal policy mix. We first explain the baseline simulation model used to generate the policy mixes. Then, we specify the ensemble of single and mixed policies tested. We detail the construction of our score by

⁸We are aware that individual rationalities may clash under simpler terms, e.g. moral hazard problems under principal-agent games [17]

⁹We use the taxonomy of [10] due to its simplicity.

¹⁰See [22].

referring to the literature on Principal Component Analysis. Then, we make explicit our decision on what we expect policymakers prefer, and describe the step-by-step of our proposed evaluation score criteria. Finally, we describe two additional applications of results: a test on different cities, and another on robustness.

3.1 Baseline model: PolicySpace2

PolicySpace2 (PS2) is an agent-based simulation model that highlights interactions between agents and their environment, with both being updated according to explicit rules [24]¹¹ The PS2 model is "... a primarily endogenous computational agent-based model (ABM) that includes mortgage loans, housing construction, tax collection and investments, with firms and households interacting in real estate, goods and services, and labor markets" [24, par 1.5]. The model's main goal is to compare three policy instruments for Brasília, Brazil, between 2010 and 2020. The authors collect simulation results of macroeconomic indicators and geographical data. PS2 is open source and readily available.

We modified the original PolicySpace2 simulation to include the policy mix alternatives we propose. The controlled environment of the simulation enables the comparison of counterfactual alternative policy instrument applications, which would otherwise be impossible [26], and thus, produce the evaluation score.

3.2 Policy mix design

We associate the results, domains, policy instruments and agents of [24] to the taxonomy of [10] (see figure 1). Rationales reflect the preferences of policymakers, and take the form of simulation output indicators. When policymakers favor reduction of inequality, they focus on lower Gini coefficient. When they aim for production increase, the preferred indicator is the expansion of GDP. We detail those preferences on the "socially better" section 3.3.2, and describe their values on table 3.

PolicySpace2 [24] considers housing and social welfare as their policy domains. The policy instruments that implement the domains are:

1. Purchase of properties by municipalities and their transfer to lowest income families.
2. Rent voucher, in which municipalities pay the rent of lowest income families for 24 months, with possible renew when meeting the necessary criteria.
3. Monetary Aid transferred to families on a given lowest percentile for as long as they fit the necessary criteria. We will refer to Monetary Aid, simply as Cash.

¹¹For fuller documentation and procedures details, please see [25].

Actors include construction firms that provide new houses to the municipalities; the municipalities themselves, which allocate part of their funds towards the policies; and the recipients of the benefits, the lowest income households.

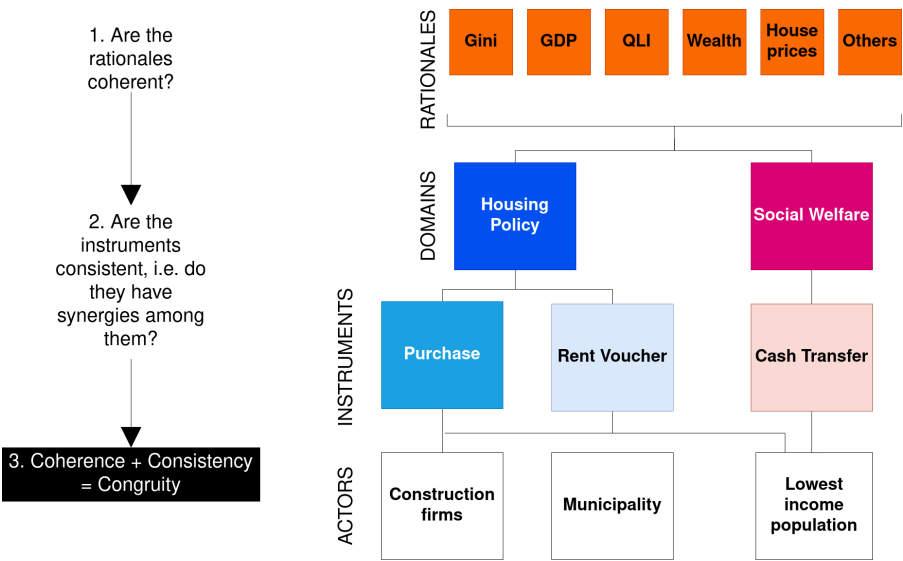


Fig. 1: Application of Magro and Wilson [10] taxonomy to the simulation environment of PolicySpace2. The rationales represent the preferences of policymakers. Domains constitute the space of policy, and instruments contain the single policy implementation procedures. Actors include acting agents, such as the municipalities and construction firms, but also the lowest income families, the beneficiaries. Hence, following Magro and Wilson, we identify two distinct domains, direct beneficiaries, but also indirect ones, the construction firms. Moreover, we link back to [8] and their definition of coherence and consistency (see table 1). Congruity is only achieved when rationales and instruments present synergism.

Figure 1 shows how to achieve congruity, from rationales to instruments. It also depicts the domains of policy and the actors. However, the figure considers single policy instruments only. We adapt the work of [24] to test policy mix instruments. Moreover, we propose the score evaluation procedure to evaluate whether they reach the rationales.

The PS2 model applies each policy instrument individually and exclusively so that each one is tested separately in a different simulation run. We propose a policy mix design to test a combination of policies occurring simultaneously. That is, a single simulation run may include portions of public investment to all three instruments at the same time. We modified the original code so that the modeler can alter the parameter "policy coefficient". This parameter

establishes how much of the public funds the municipalities invest in each policy instrument (see table 2).

Policy instrument/ Policy Regime		Purchase	Rent	Cash
Single Policy	Purchase	1	0	0
	Rent	0	1	0
	Cash	0	0	1
Policy Mix	Equal	.33	.33	.33
	Focused	.5	.25	.25
	Random	a*	b**	1 - a - b
Baseline		0	0	0

*: $a \sim U(0,1)$, **: $b \sim U(0,a)$

Table 2: Percentage of the policy available budget invested in each instrument for each policy regime. In the single policy regime, each instrument receives the total available fund. Under policy mix, each instrument receives a percentage. The random a and b coefficients refer to each simulation run.

This adaptation implies the possibility of mixed policy instruments in which the amount of funds invested does not need to be to a single instrument. Thus, the modeler may now distribute the investments to different instruments simultaneously. We tested three policy mixes combinations: (1) Equal, (2) Focused and (3) Random, along with simulation runs with a single instrument, and the no-policy baseline.

1. Single Policy Instrument

We first test a single policy mix instrument. Just one instrument consumes the whole municipality policy investment. In terms of the model itself, it would be a 0-0-1 combination of investment weights for all three possible policies. The investment reflects the total funds that the municipality allocates for policy investment. The municipalities also invest to improve the city’s infrastructure in general. The Quality of Life Index captures the effects of this investment (see [24] for details). This single policy is called ‘unitary’.

2. Equal Policy mix instrument

Equal policy mix instrument takes the municipal budget and divide it equally among all three policy instruments (Purchase, Rent, Cash). Each instrument receives one third of the policy municipal budget available to policy investment.

3. Focused Policy mix instrument

Focused policy mix instrument invests half of the available municipal budget towards the Purchase instrument, with the other half being equally divided into the two other instruments. The model implements the Purchase instrument first, thus it captures the lowest income households

present in the model in a given month. Next, the model applies Rent, the next instrument, to the immediately next lowest income households. Finally, the next ensemble of households receives the Cash investment.

4. Random Policy mix instrument

The model assigns random values for each of the three instruments in each simulation so that coefficients sum up to the total policy available funds. The model consumes each portion of the budget in the same order: Purchase, Rent, Cash.

3.3 Molding a summary score: from one to many, progressively

Every policy instrument generate effects. Even policies that at first produce no effect may affect at least agents' expectations (see for example [22, 27–30]). Moreover, money invested on a given policy could have otherwise provided alterations elsewhere. These effects or lack thereof produce different results depending on the recipients, their location, their previous and current state, to name a few [18, 31, 32]. Considering policy as a rule [14], and rules as institutions [19], context-specificity is expected [31] due to relevance of past decisions and past events over how future events will unfold [32].

Multiple indicators capture these changes. How actors perceive change, however, depends on which indicators they favor, how they decide on priorities and to what they compare the indicators with. [9, 33] characterize these design, preferences and choices as an evaluative criteria, whereas [34] provide an application.

When policymakers have multiple goals, such as to increase production and decrease inequality simultaneously, goals may have some degree of incoherence. A given policy instrument may increase families' consumption and firms' profits on one hand, but it may also reduce employment in a different cluster of firms, due to lower demand. Policymakers themselves may want to make larger public investments and reduce taxes at the same time. The more dimensions policymakers vow for when they devise, implement, and assess policies, the more likely they observe goal incoherence [1, 5].

This dispersion of multiple results—at times contradictory—suggests that the decision policymakers oblige to is nothing but trivial. They decide on which instrument is (supposedly) optimal, either individually or along with other instruments, and how much to invest, if at all, on each available instrument. Hence, we propose the construction of a score that captures the variance of the results, weighting their contribution on a previously decided socially better direction that should reflect policymakers preferences.

We use a weighted version of Principal Component Analysis (PCA) scores to decide on an optimal policy, as long as we know which dimension of results to include, and in which order. In the next sections we discuss the basics of PCAs, the preferences that reflect policymakers choices and the procedures used to construct the scores.

3.3.1 Basics of Principal Component Analysis

Principal Component Analysis (PCA) is a dimensionality reduction technique that is popular and easy-to-use [35]. PCA captures the maximum variance of the system where the data is the most spread out, based on the matrix of observations (rows) and variables (columns). The mathematical procedure chooses the "eigenvector with the highest eigenvalue" [35, p.5].

Once the method has identified the variance and direction of the system, the aim is to construct a successive linear combination of new variables—called components—that represent the original data in a condensed form [36]. By construction, the first component summarizes the most variance possible, given the original data. The second component comprises less variance than the first and more than the third component, and so on and so forth.

The components represent weights that multiple the original data to produce indexes (each component) and scores on those indexes (original data multiplied by the weights). These weights are called loadings. "Loadings are interpreted as the coefficients of the linear combination of the initial variables[dimensions] from which the principal components are constructed" [37, p.27]. Analyzing the loadings, the user may observe which original variables compose which component, their magnitude and direction.

3.3.2 Policymakers' preferences: what dimension would likely be socially better?

We define socially better in table 3 as the direction that would most likely bring larger benefits for most citizens. In general, which direction of an indicator would society prefer? We assume that policymakers agree with these *ad-hoc* social preferences. Furthermore, we define an additive relevance that reflects how each dimension of analysis is progressively added to the composed final score.

Dimension is each rationale (or goal), the result of the simulation. Dimensions vary from strictly economic indicators (such as GDP index and Inflation), to household information (Household consumption and median wealth), to firms' profits, and to urban (or metropolitan) indicators, such as the level of house vacancy, commuting, or rent and house prices

These choices mean we consider inequality as the the first and most relevant policy output dimension. However, inequality alone may not be enough to judge the quality of policy results. Hence, we progressively add all the other dimensions to build a more comprehensive view of policy results. We relax these choices of relevance and order in the robustness test.

3.3.3 Step-by-step procedures to construct the Policy Instrument Progressive (PIP) scores

This section describes the procedures to make the Policy Instrument Progressive (PIP) score, which reflects what we defined as socially better (see also *pseudocode*, in Appendix 1):

Dimensions: original simulation results	Progressive Relevance	Additive	Socially better	Category
Gini index	0		Lower	Inequality
GDP index	1		Higher	Production
Avg. Quality Life Index	2		Higher	Production
Inflation	3		Lower	Inequality
Unemployment	4		Lower	Inequality
Houses prices	5		Lower	Inequality
Perc. HH affordable rent	6		Higher	Inequality
Avg. HH consumption	7		Higher	Production
HH defaulting on rent	8		Lower	Inequality
HH commuting indicator	9		Lower	Inequality
Perc. HH zero consumption	10		Lower	Inequality
Perc. HH mortgage delinquency	11		Lower	Inequality
Firms profit	12		Higher	Production
HH median wealth	13		Higher	Production
Taxes by FPM (+ progressive)	14		Higher	Production
Taxes equally (+ progressive)	15		Higher	Production
Amount subsidized by policy	16		Lower	Inequality
Houses rents	17		Lower	Inequality
Houses vacancy	18		Lower	Production
Taxes locally (+ regressive)	19		Lower	Inequality

Table 3: The table include the dimensions—which are the indicators that result from the simulation runs –, the choice of additive relevance, the likely direction of the dimension that favors society, and two general categories of the dimension. We relax the relevance choice in a robust test. The category serves to enhance the analysis of results. Most choices are easy to make. Policymakers and society prefer lower inequality, more affordable rent, higher consumption and profits. Comparatively, we prefer taxes that are progressive to be higher than regressive taxes, based on the analysis of [38].

1. A Principal Component Analysis (PCA) is run successively for each added dimension. Thus, we run analyses with wider and wider matrices (more columns) each time, adding simulation results one by one. The final run is made for the full matrix of 1680 observations—individual simulation runs—and 20 dimensions (see table 3). These 1680 observations include 240 runs for each one of seven policy instruments proposed (single, mixed, and baseline—section 3.2). The observations also include 24 different metropolitan regions.
2. For each run, the PCA produces the weighted influence of each dimension to each of the components. These loadings represent the contribution of each dimension to each component.
3. We evaluate the first two components, which carry most of explained variance of the data. As such, we condense the information of the original indicators into the two components that carry the most variance of the data.

4. For each one of the two components (c) and each dimension (dim), we evaluate if the loading ($L_{dim,c}$) is in the direction that is socially better (SB_d), as defined in table 3.
5. Then, we start implementing what makes a regular PCA score into a desired score that maximizes previously chosen directions (d) of what is defined as socially better (SB) (see equation 1). In practical terms, when the direction of the loading is the opposite of what is defined as socially better, we invert the direction by multiplying the loading by -1 .
6. These procedures guarantee that the information and magnitude contained in the loading adds to the final PIP indicator (PIP), contributing positively or negatively. In doing so, the final score of each observation is not the typical linear combination of loadings, but actually, the desired summation towards the "ex-ante" definition of socially better. As a typical score, the numbers refer to the context of the data set, thus, are comparable between instruments, cities, or domains.

$$L_{ind,c} = \begin{cases} L_{ind,c}, & \text{if } L_d \mapsto SB_d \\ L_{ind,c} * -1, & \text{otherwise} \end{cases} \quad (1)$$

7. Next, we multiply each observation value by the loading for both components (c) 1 and 2. The score (S) for each observation (i) is the average of the score of each component, weighted by its explained variance (w) (see equation 2).

$$S_i = \frac{\sum_{c=1}^2 x_i * L_{ind,c} * w_c}{w_1 + w_2} \quad (2)$$

8. The PIP Indicator for each policy is the summation of all observations of that policy instrument.

$$PIP_{policy} = \sum_{policy} S_{policy} \quad (3)$$

3.4 Cities

Effects of policy implementation depend on the region where they are applied and when they are applied [18, 20, 31, 39, 40]. In fact, cities' intrinsic spatial structure, density and location of citizens, along with their characteristics and those of cities' firms configure heterogeneous recipient for policies. In essence, the same policy may produce opposing results solely because their region of application is distinct.

Policy results are presented for the ensemble of all 24 cities considered in the study. However, in a separate results' section we present cities' average results for all policies, along with results per city per policy.

3.5 Robustness

As expected, the results are dependent on the dimensions chosen and their order of relevance. We run some alternative configurations to relax these constraints. First, we run the procedures for two sets of (generally speaking) similar indicators. We divide the 20 indicators into two categories: production and inequality, specified in table 3. Secondly, we run the exact same procedures one thousand times, however, with random order of dimensions. In doing so, we remove the choice of which indicator should be considered first, second or last and all of them stochastically rotate for all positions of relevance. Thus we remove the selection bias of relevance order.

We then calculate a distribution of mean results by policy. For each individual run, we average results for each step of additional dimension. Therefore, the score represents the average value across all dimensions for each run. This procedure generates 1,000 values that represent means of means.¹²

4 Results

4.1 Illustration of the score construction

Figure 2 shows how we organized the information extracted from the data towards positive or negative contributions to the final score. The top panel represents what the method produces as a result of the variance of the original data. How the applied math brings the observed variance into the loadings. The bottom panel shows the loadings organized into contributions to the score. Notice that the loadings may flip along each one of the components axis, or not flip at all.

4.2 Policy instruments' results

Results show that the score of each policy regime varies significantly depending on the successive inclusion of each dimension (figure 3). Were inequality the single dimension to be considered, the Purchase policy instrument would score much worse than the other instruments. This is expected given that each property bought and passed on to a specific household is rather expensive. Thus, in a sense, few families receive a proportionally large share of policy investments.

However, when GDP is included in the score, the previously worst instrument, Property, becomes the best. That probably results from the increase in demand for firms' production provided that municipal government enters the market as a buyer.

Furthermore, Firms' profits seem to be an influential dimension. When included in the analysis, Firms' profits change the overall results, placing the no-policy Baseline on top, and reducing all other scores. That illustrates the idea that policy implementation affects society in contradictory ways. It might be on the interest of firms to have higher demand and profit. It might

¹²In total, we calculate scores for 20,000 PCAs, that is, 20 dimensions for each run times 1,000 runs.

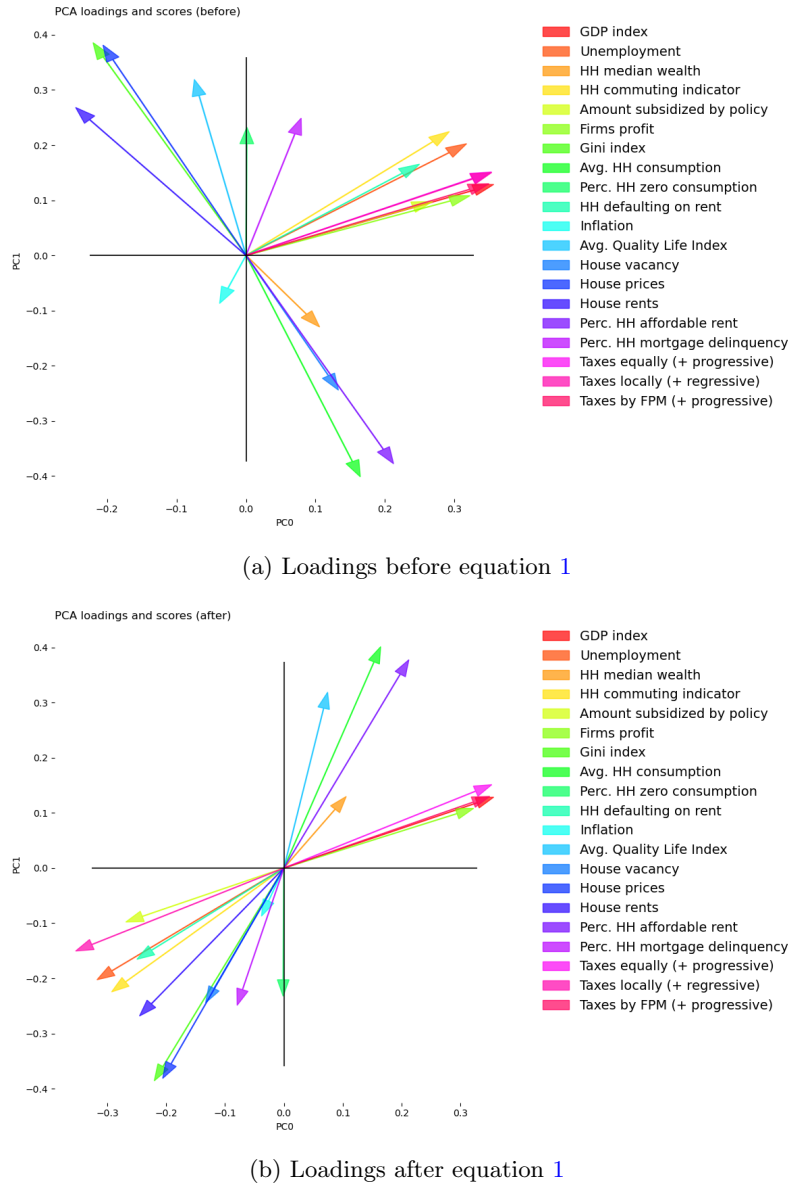


Fig. 2: Biplot illustration of the loadings for both principal component 1 and 2 before and after changes applied by the procedures' steps. Figure (a) reflects the loadings for a regular PCA analysis, whose results derive from the variance in the data. Figure (b) shows how each loading contributes to components 1 and 2, after the implementation of the proposed score. It shows how we organize the data towards favorable or unfavorable, flipping along the components axis, maintaining the magnitude. See that the purple arrow—affordable—belongs to the top-right quadrant in (a) and lower-left in (b).

even increase employment and salaries on these specific construction firms. Many households, however, may also go unattended by the instrument as the municipality focus on a fewer number of families.

In the end, if policymakers consider all dimensions, then the mixed policies (Focused, Equal and Random) perform better (see table 4 for the final rightmost PIP scores that include all dimensions). As expected, when policy-makers consider a more comprehensive indicator, policy mix instruments seem to perform better.

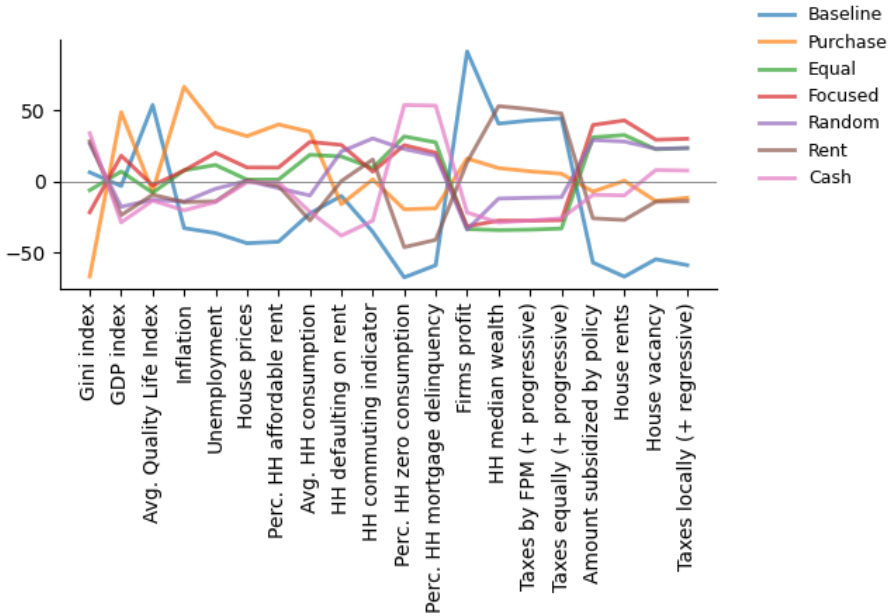


Fig. 3: Evolution of policies' scores as each dimension, from left to right, is included in the calculus. In the end, at the rightmost position when all dimensions are present, The Focused policy instrument scores the highest. Considering all dimensions, any policy is better than the no-policy baseline scenario.

4.3 Cities' results

The cities' results show that each city responds differently from one another. Their response varies intrinsically, as the PIP score includes each dimension. We illustrate the results for the medium-sized city of Belo Horizonte (see figure 4). As the first dimensions are added (leftmost part of the graph), the PIP

score moves closely to each other for all policies. As more and more dimensions are included, the scores start to disaggregate.

When all dimensions have been added (at the rightmost position of the graph), we can see that all policies' implementation for Belo Horizonte perform worse when compared to other cities, hence the negative values. However, among themselves, the no-policy Baseline is the most negative, with Random, Equal and Focused performing slightly better (see table 4 for all results).

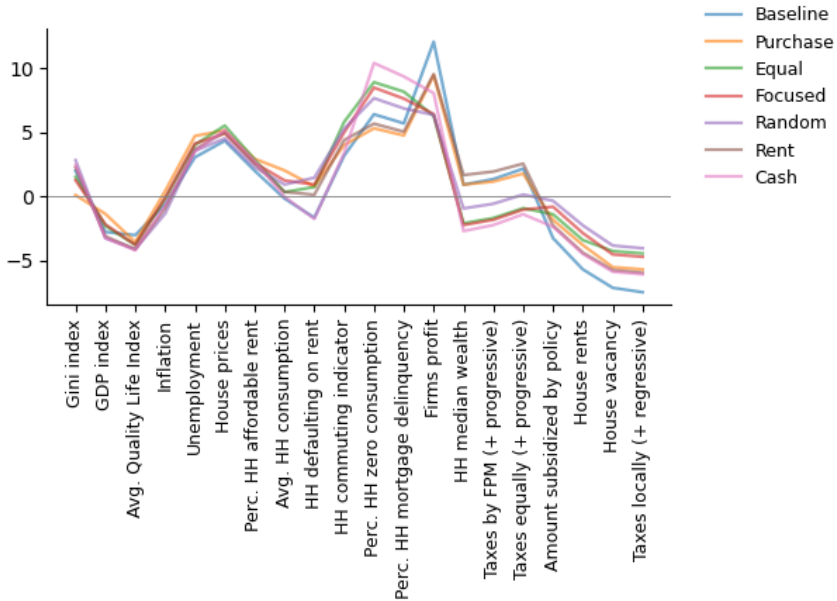


Fig. 4: PIP scores for the city of Belo Horizonte.

Figure 5 highlights the unique response from each city to the dimensions. The figure represents the summation of all policies' PIP scores for each city. Note that the same indicator affects each city in a different way. However, as figure 6 shows, policies have similar behavior depending on the city and results differentiate themselves as more and more indicators are considered.

4.4 Robustness results

We present the robustness results in two steps. First, we show the two category scores: Inequality and Production, and then the distribution of the average results of 1,000 runs.

Figure 7 shows the results for the Inequality category. This category includes dimensions that tend to favor households across the population in

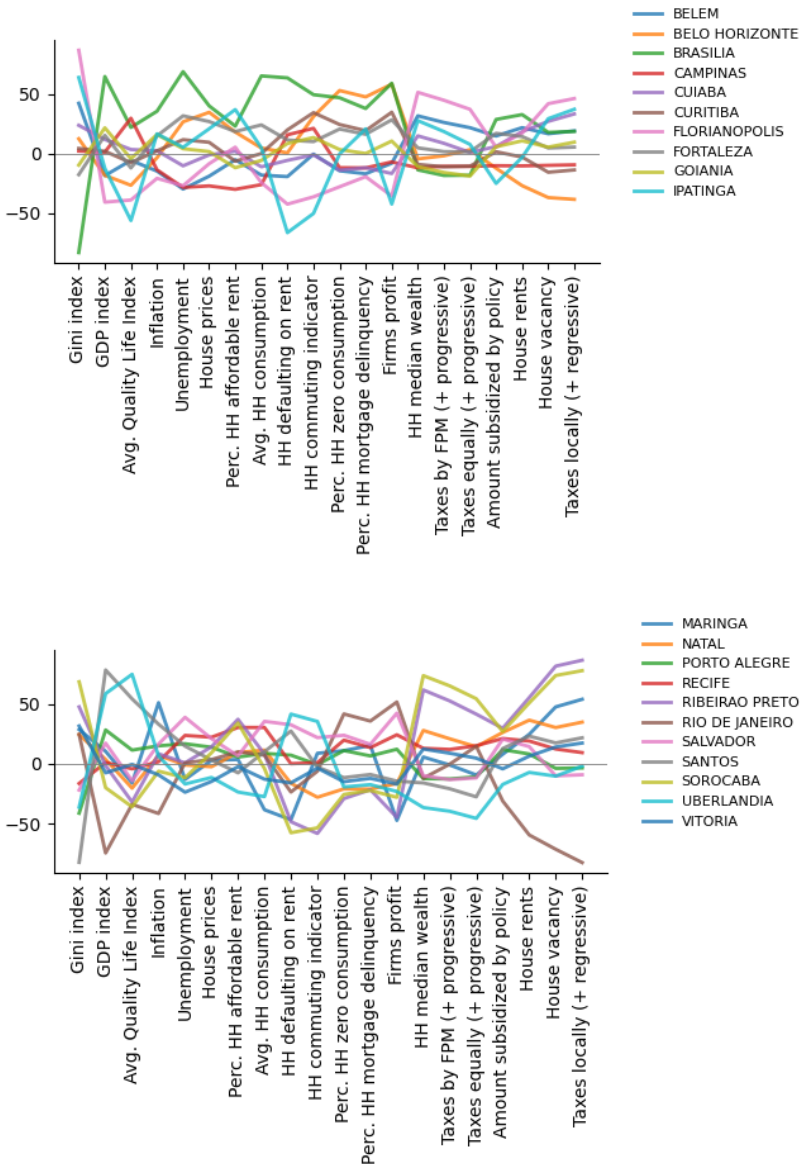


Fig. 5: Sum of all policies' implementation per city per indicator. Each city responds differently to policies in general and to each additional indicator in particular. As expected, a city's given configuration of citizens and firms, along with their own characteristics make the same policy produce results that are distinct in magnitude and direction. Obs.: We excluded São Paulo, Londrina and Campo Grande to keep similar Y-axis bounds. All patterns are included in figure 6. Full results table per city is available as supplemental material.

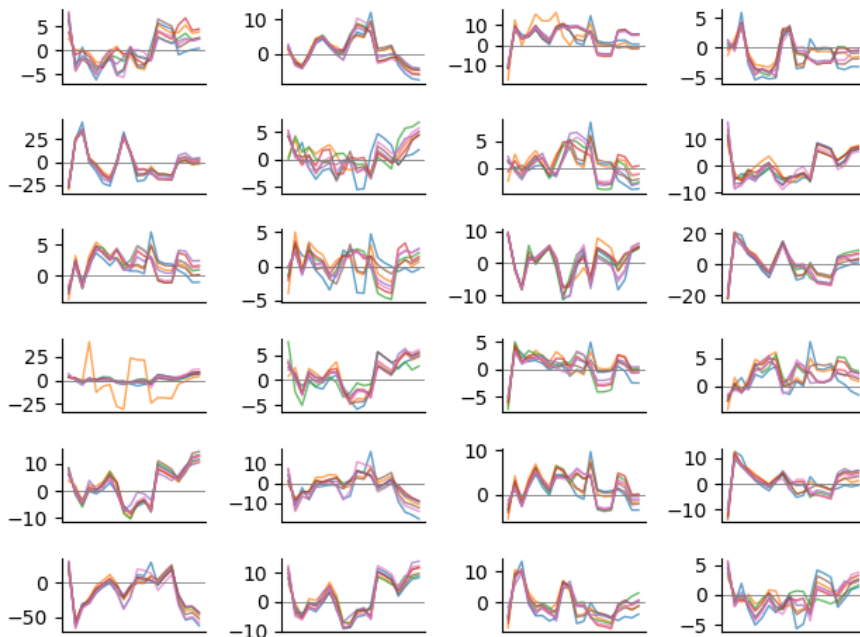


Fig. 6: Summary of policies for all available cities. Policies labels are the same of figure 4, above. Also, same progression of parameters in figure 6. Cities from left to right, top to bottom: Belém, Belo Horizonte, Brasília, Campinas–Campo Grande, Cuiabá, Curitiba, Florianópolis–Fortaleza, Goiânia, Ipatinga, Londrina–Maringá, Natal, Porto Alegre, Recife–Ribeirão Preto, Rio de Janeiro, Salvador, Santos–São Paulo, Sorocaba, Uberlândia, Vitória. Most cities have similar patterns for all of the policies. When all indicators have been added—at the rightmost position of each graph—then the effects of each individual policy is highlighted.

a more homogeneous way. That includes, obviously, the Gini index, but also inflation and unemployment, and some dimensions related to poverty, such as how affordable is rent relative to wages, and how many families go a month without consuming in the goods and services market.

The behavior of the PIP score as more dimensions are added varies much less when compared to the full array of dimensions. In the end (the rightmost part of the graph), the single Cash instrument performs much better than the others, followed by the three mixed policy instruments. The Purchase instrument—seemingly a concentrating one—performs worse than the baseline, whereas Rent is slightly above it. This suggests that direct transfer of funds to households would be an adequate instrument when policymakers favor these rationales.

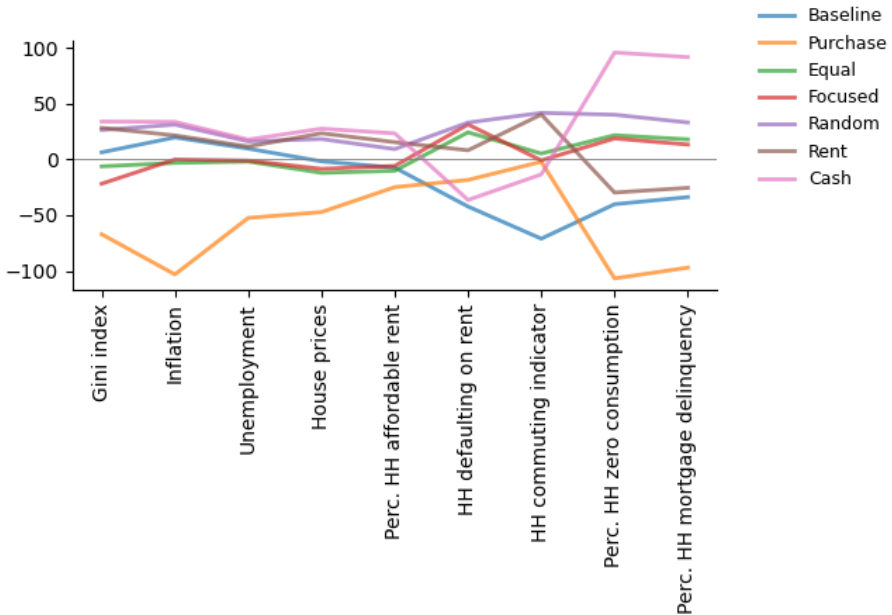


Fig. 7: Results for PIP score considering selected indicators that fall into a general category of inequality information. Cash policy instrument performs the better, whereas Purchase, the worst. Mixed policy instruments show as second best.

The so-called production category also present a more stable evolution of the PIP score as it includes more dimensions. The category considers the GDP index, but also consumption and wealth of families and firms, along with taxes collected, which is correlated with production. Note also that the funds reserved to invest in the policy instrument (amount subsidize by policy) and the average Quality of Life Index reflect the amount of resources invested by municipalities. For this category, single policies do better—Rent and Cash—and worse, Property. All of the three mixed policies and the Baseline perform around the average PIP score.

The robust test demonstrate that running the procedures with random order of dimensions favors the mixed policy instruments. All of the three mixed policies have higher overall scores (see figure 9 and table 4). The no-policy Baseline policy is the most volatile, with average scores that is worse than its counterparts.

Considering the single policy instruments, the ranking of the results suggest that Cash is the best one, followed by Rent and then Property. Mixed instruments seem to do well in most cases, except for the PIP Production, in

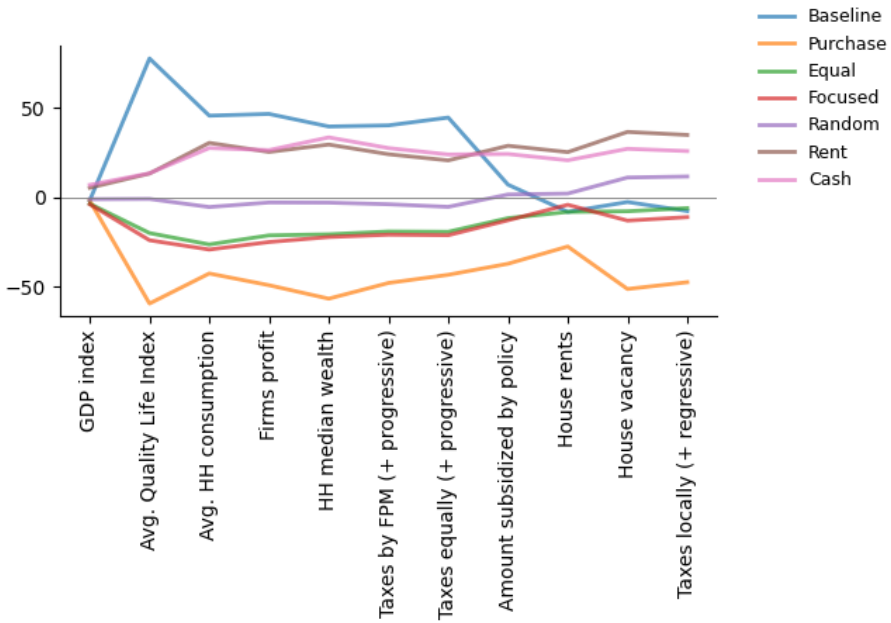


Fig. 8: Results for PIP score considering selected indicators that fall into a general category of production information. Baseline performs well until the score includes the amount of funds invested in the instrument, house rents and regressive taxes. Single policy instruments of Rent and Cash perform similarly through the composition of the dimensions and scores the best in the end.

which Rent and Cash do the best, and Property, the worst. Property also performs worse than the no-policy Baseline for both categories of inequality and production, having the second to worst in the robust score.

5 Discussion

Results show that scores—or success of a specific single or mixed policy instrument—depend on the rationales used to evaluate this success. Clearly, policy effects depend on what society considers their driven goals, their rationales. Such results suggest that the more factors or dimensions one takes into consideration, the harder it becomes to assess any object of study [13]. This is in accordance with [1], who makes a similar analysis regarding policies: that the more dimensions one takes into consideration the harder it is to pinpoint an optimal policy mix.

The analysis also demonstrate that the more inclusive in terms of results the policymaker is, the interests of more groups of citizens need to be included in the results. That is, all parties are contemplated when evaluation includes indicators from different areas.

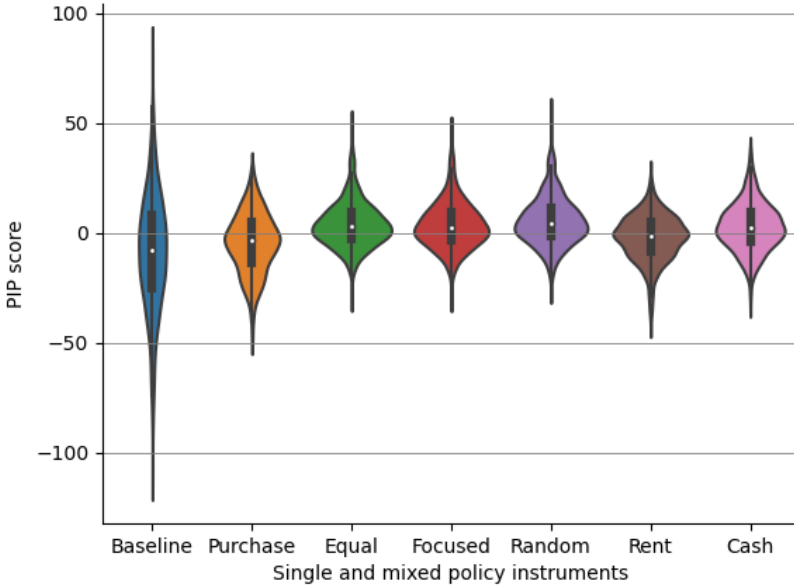


Fig. 9: Distribution of PIP score values for 1,000 runs. Each single value is the average of all PIP scores, from 1 to 20 dimensions. The no-policy Baseline and the single Property policy perform worse on average than all of the other ones, although with more volatility. All of the mixed policy instruments have higher averages when compared with single ones. Cash is the only single policy instrument that averages positive scores overall. For average values and a comparison of the other scores, see table 4.

The analysis further indicates that policymakers that favor a more comprehensive evaluation should consider mixed policy instruments. They are more likely to be balanced and take into account a wider range of perspectives. The use of policy mixes rather than single instruments becomes a better solution for policymakers who want to encompass more dimensions into its policy assessment.

All in all, it is not trivial to decide which policy instrument is the best, as it depends on the metric used to evaluate. Some patterns seem clear, though. When policymakers are interested in general results, mixed instruments perform better. When having clear goals, such as to reduce inequality or to boost production, single instruments prove more adequate.

Despite the agreement of our results with the literature, the distinct pattern of results for the cities highlight that policy recipients react consistently distinctly to the added dimensions. This idiosyncratic behavior probably derives from their intrinsic nature, historical construction, and present characteristics.

Instrument	PIP	PIP BH	PIP Inequality	PIP Production	Robust PIP
Random	23.43	-4.02	33.13	11.62	5.46
Equal	23.51	-4.42	17.99	-6.06	3.81
Focused	30.05	-4.69	13.37	-11.09	3.54
Cash	7.66	-6.05	91.73	25.83	2.80
Rent	-13.97	-5.93	-25.45	34.85	-2.11
Property	-11.59	-5.67	-97.01	-47.47	-4.54
Baseline	-59.09	-7.45	-33.77	-7.69	-8.97

Table 4: Summary results of PIP scores. PIP refers to the score following the relevance of indicators shown in table 3. PIP BH illustrates the case of a single city (Belo Horizonte). PIP Inequality and Production show the results of each individual domain, and Robust PIP refers to the average of 1,000 simulation. Considering all results, mixed policies fair very well overall, whereas single policies have their best results at specific domains. Property and Baseline have generally low results.

The point here is that the same policy instrument performs differently giving the attributes of the location where it is applied. This message contains practical information for policymakers. Finding a best instrument for a specific city, by no means guarantee that the same instrument fits any other city. Such conclusion is in line with [18, 31, 32, 41] understandings that path dependence plays an important role on determining how a policy will fit into the pre-existent institutional environmental. That is, what will be the interplay of a policy with the other institutions, be them formal or informal.

Referring back to the literature, our analysis confirm that congruity becomes harder to be achieved the more dimensions the policymaker has in mind while evaluating its policy mix [1]. Our analysis also confirms that not only the number of dimensions matter, but also which dimensions. And above all, our analysis also suggests that which and how many dimensions matter as much as the city. In [8] terms, inconsistency between instruments and incoherence between goals may entirely depend on what are the dimensions taken into account when evaluating the policy mix itself, depending on the municipality and its pre-intervention characteristics.

It may be at best impractical to pinpoint a one-size-fits-all policy mix (which) that disregards both the recipient place (where) and the dimensions (why) on which the policy will be assessed. To determine an optimal policy mix is both a question of which instrument, where to apply, aiming at which goal. By asking these questions, policymakers may avoid policy incongruities in an easier fashion.

6 Final considerations

We take advantage of a simulation environment to test the effects of single and mixed policy instruments in 20 output indicators. Following the literature, we evaluate whether policymakers rationale—the goals achieved as shown by

the indicators—favor one or another policy instrument. In order to subsidize policymakers, we create a process to reduce dimensionality whilst modifying it towards socially better results. The application of the evaluation process to the simulation runs produces non-trivial results that reveal nuances of the instruments, their combination, and the idiosyncrasies of each recipient city.

We find that it is impractical to pinpoint one policy mix that is optimal for all analyzed cities at the same time. Such problem is only highlighted by the fact that defining what is optimal depends on the dimensions that the policymaker uses to assess its own policies. Moreover, the difficult to decide on an optimal policy instrument is not only a matter of how many dimensions, but also of which dimensions, and which relevance. Yet, mixed policy instruments seemed better for comprehensiveness intentions. Cash policy performed better to reduce inequality, whereas Rent vouchers enhanced production. No-policy baseline and Property achieved generally lower comparative results.

Hence, the message of the paper is that policymakers should make sure to know which dimension of society their policy instrument is trying to improve. Moreover, it is necessary to know which instrument and dimension is effective for which region. Computational simulation seems to be a relatively cheap and fast means to respond these questions before making policy decisions.

References

- [1] Rogge, K.S., Reichardt, K.: Policy mixes for sustainability transitions: An extended concept and framework for analysis. *Research Policy* **45**(8), 1620–1635 (2016). <https://doi.org/10.1016/j.respol.2016.04.004>. Accessed 2022-04-14
- [2] Kern, F., Howlett, M.: Implementing transition management as policy reforms: A case study of the Dutch energy sector. *Policy Sciences* **42**(4), 391–408 (2009). <https://doi.org/10.1007/s11077-009-9099-x>. Accessed 2022-04-14
- [3] Nykamp, H.: Policy Mix for a Transition to Sustainability: Green Buildings in Norway. *Sustainability* **12**(2), 446 (2020). <https://doi.org/10.3390/su12020446>. Accessed 2022-04-14
- [4] Flanagan, K., Uyarra, E., Laranja, M.: Reconceptualising the ‘policy mix’ for innovation. *Research Policy* **40**(5), 702–713 (2011). <https://doi.org/10.1016/j.respol.2011.02.005>. Accessed 2022-04-14
- [5] Cunningham, P., Edler, J., Flanagan, K., Laredo, P.: Innovation policy mix and instrument interaction: A review. NESTA Working paper **13** (2020)
- [6] Tobin, J.: *Fiscal Policy: Its Macroeconomics in Perspective*, Rochester, NY (2001). <https://papers.ssrn.com/abstract=275514>

- [7] Reynolds, A.: The Fiscal-Monetary Policy Mix. *Cato Journal* **21**, 263 (2001)
- [8] Howlett, M., Rayner, J.: Design Principles for Policy Mixes: Cohesion and Coherence in ‘New Governance Arrangements’. *Policy and Society* **26**(4), 1–18 (2007). [https://doi.org/10.1016/S1449-4035\(07\)70118-2](https://doi.org/10.1016/S1449-4035(07)70118-2). Accessed 2022-04-14
- [9] Ostrom, E.: *Understanding Institutional Diversity*. Princeton University Press, Princeton (2005)
- [10] Magro, E., Wilson, J.R.: Complex innovation policy systems: Towards an evaluation mix. *Research Policy* **42**(9), 1647–1656 (2013). <https://doi.org/10.1016/j.respol.2013.06.005>. Accessed 2022-04-14
- [11] Arthur, W.B.: Inductive reasoning and bounded rationality. *The American Economic Review* **84**(2), 406–411 (1994)
- [12] Dequech, D.: Uncertainty: A Typology and Refinements of Existing Concepts. *Journal of Economic Issues* **0**(3), 621–640 (2011). <https://doi.org/10.2753/JEI0021-3624450306>. Accessed 2018-04-17
- [13] North, D.C.: A Transaction Cost Theory of Politics. *Journal of Theoretical Politics* **2**(4), 355–367 (1990). <https://doi.org/10.1177/0951692890002004001>
- [14] Vazquez, M., Hallack, M.: The role of regulatory learning in energy transition: The case of solar PV in Brazil. *Energy Policy* **114**, 465–481 (2018). <https://doi.org/10.1016/j.enpol.2017.11.066>. Accessed 2018-04-25
- [15] Vazquez, M.: Institutional dynamics in an economy seen as a complex adaptive system. *Bocconi IEF Working Paper* **104**, 33 (2018)
- [16] Nelson, R.R.: The co-evolution of technology, industrial structure, and supporting institutions. *Industrial and Corporate Change* **3**(1) (1994)
- [17] Kreps, D.M.: Part III: Noncooperative game theory. In: *A Course in Microeconomic Theory*, pp. 353–551. Princeton University Press, ??? (1990)
- [18] North, D.C.: Transaction costs, institutions, and economic performance. In: *Transaction Costs, Institutions, and Economic Performance. Occasional Papers / International Center for Economic Growth*, pp. 5–32. ICS Press, San Francisco, Calif (1992)
- [19] Crawford, S.E., Ostrom, E.: A grammar of institutions. *American Political Science Review* **89**(3), 582–600 (1995)

- [20] Held, A., Ragwitz, M., Gephart, M., De Visser, E., Klessmann, C.: Design features of support schemes for renewable electricity. Ecofys task report **2** (2014)
- [21] Hochstetler, K., Kostka, G.: Wind and Solar Power in Brazil and China: Interests, StateBusiness Relations, and Policy Outcomes. *Global Environmental Politics* **15**(3), 74–94 (2015). https://doi.org/10.1162/GLEP_a.00312
- [22] Cardim de Carvalho, F.J.: Economic policies for monetary economies: Keynes’s economic policy proposals for an unemployment-free economy. *Revista de Economia Política* **17**(4), 31–51 (1997)
- [23] Constantini, V., Crespi, F., Palma, A.: Characterizing the policy mix and its impact on eco-innovation in energy-efficient technologies. Sustainability environmental economics and dynamic studies working paper series **11** (2015)
- [24] Alves Furtado, B.: Polycspace2: Modeling markets and endogenous public policies. *Journal of Artificial Societies and Social Simulation* **25**(1), 8 (2022). <https://doi.org/10.18564/jasss.4742>
- [25] Furtado, B.A.: PolicySpace2 : Modelando Mercado Imobiliário e Políticas Públicas, 1st edn. Ipea, Rio de Janeiro (2022)
- [26] Arnold, K.F., Harrison, W.J., Heppenstall, A.J., Gilthorpe, M.S.: DAG-informed regression modelling, agent-based modelling and microsimulation modelling: a critical comparison of methods for causal inference. *International Journal of Epidemiology* **48**(1), 243–253 (2019). <https://doi.org/10.1093/ije/dyy260>. Accessed 2022-10-07
- [27] Cardim de Carvalho, F.J.: Alta inflação e hiperinflação: Uma visão pós-keynesiana. *Revista de Economia Política* **10**(4), 62–82 (1990)
- [28] Árida, P., Lara-resende, A.: Inertial inflation and monetary reform in Brazil. *Textos para Discussão Departamento de Economia PUC-Rio* **85** (1985)
- [29] Pereira, L.B.: Inertial inflation and the Cruzado plan. *World Development* **15**(8), 1035–1044 (1987). [https://doi.org/10.1016/0305-750X\(87\)90170-7](https://doi.org/10.1016/0305-750X(87)90170-7)
- [30] Kinzig, A.P., Ehrlich, P.R., Alston, L.J., Arrow, K., Barrett, S., Buchman, T.G., Daily, G.C., Levin, B., Levin, S., Oppenheimer, M., Ostrom, E., Saari, D.: Social Norms and Global Environmental Challenges: The Complex Interaction of Behaviors, Values, and Policy. *BioScience* **63**(3), 164–175 (2013). <https://doi.org/10.1525/bio.2013.63.3.5>

- [31] North, D.: Institutions and the Performance of Economies over Time. In: Ménard, C., Shirley, M.M. (eds.) *Handbook of New Institutional Economics*, pp. 21–30. Springer, Berlin (2008)
- [32] North, D.C.: The new institutional economics and third world development. In: Harriss, J., Hunter, J., Lewis, C.M. (eds.) *The New Institutional Economics and Third World Development*, pp. 17–26. Routledge, London (1997)
- [33] Ostrom, E.: Background on the Institutional Analysis and Development Framework. *Policy Studies Journal* **39**(1), 7–27 (2011). <https://doi.org/10.1111/j.1541-0072.2010.00394.x>
- [34] Nigussie, Z., Tsunekawa, A., Haregeweyn, N., Adgo, E., Cochrane, L., Floquet, A., Abele, S.: Applying Ostrom’s institutional analysis and development framework to soil and water conservation activities in north-western Ethiopia. *Land Use Policy* **71**, 1–10 (2018). <https://doi.org/10.1016/j.landusepol.2017.11.039>
- [35] Verma, A., Angelini, O., Di Matteo, T.: A new set of cluster driven composite development indicators. *EPJ Data Science* **9**(1), 1–21 (2020). <https://doi.org/10.1140/epjds/s13688-020-00225-y>. Number: 1 Publisher: SpringerOpen. Accessed 2022-07-26
- [36] Jolliffe, I.T., Cadima, J.: Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **374**(2065), 20150202 (2016). <https://doi.org/10.1098/rsta.2015.0202>. Publisher: Royal Society. Accessed 2022-07-26
- [37] Husson, F., Lê, S., Pagès, J.: *Exploratory Multivariate Analysis by Example Using R* vol. 15. CRC press Boca Raton, ??? (2011)
- [38] Furtado, B.A.: Modeling tax distribution in metropolitan regions with PolicySpace. *Journal on Policy and Complex Systems* **5**(1) (2019). <https://doi.org/10.18278/jpcs.5.1.6>
- [39] Mueller, B.: Why public policies fail: Policymaking under complexity. *Economia* **21**(2), 311–323 (2020). <https://doi.org/10.1016/j.econ.2019.11.002>. Accessed 2022-01-28
- [40] Adepetu, A., Keshav, S., Arya, V.: An agent-based electric vehicle ecosystem model: San Francisco case study. *Transport Policy* **46**, 109–122 (2016). <https://doi.org/10.1016/j.tranpol.2015.11.012>. Accessed 2019-05-02

- [41] North, D.C.: Institutions, Institutional Change, and Economic Performance. The Political Economy of Institutions and Decisions. Cambridge University Press, Cambridge ; New York (1990)

7 Appendix

This appendix shows the *pseudocode* applied to construct the PIP scores. Each variable in program 1 represents the dimensions mentioned in the text. For each additional dimension included, a new matrix of observations and dimensions is made. For each new matrix (variables), the program ran a PCA analysis. Then, loadings are evaluated to check if they are in accordance to table 3 socially better preferences. When not, the loading signal is inverted. A second loop for each component calculates the observation score by multiplying the original value by the modified loading. The score is also weighted by the variance explained of each component. Final PIP score is the summation of all observation scores per policy instrument. See equations and details in section 3.

Program 1 PIP score *pseudocode* implementation

```

for each [variable 0 ... variable 19]
  variables += variable
  run PCA(variables)
    for each component [PC0 PC1]
      for each loading
        evaluate loading direction versus socially better
        if not desired direction
          invert signal
      for each component [PC0 PC1]
        observation score = modified loading * original data value
        observation score weighted by variance explained
      observation score = weighted average components
  PIP = summation of all scores by policy instrument

```
