

Statistical model checking of the Island Model: an established economic agent-based model of endogenous growth

Stefano Blando, Giorgio Fagiolo, Daniele Giachini, Andrea Vandin*

Institute of Economics and L'EMbeDS
Sant'Anna School of Advanced Studies
Pisa, Italy

n.surname@santannapisa.it

Ernest Ivanaj

Swiss Finance Institute
University of Geneva
Switzerland

ernest.ivanaj@etu.unige.ch

Agent-based models (ABMs) are increasingly used to study complex economic phenomena such as endogenous growth, but their analysis typically relies on ad-hoc Monte Carlo exercises without formal statistical guarantees. We show how statistical model checking (SMC), and in particular MultiVeStA, can automate and enrich the analysis of a seminal ABM: the Island Model of Fagiolo and Dosi, which captures the exploration-exploitation trade-off in technological search. We reproduce key stylized facts from the original model with formal confidence intervals, confirm the optimality of moderate exploration rates, and perform a counterfactual sensitivity analysis across returns to scale, skill transfer, and knowledge locality. Using MultiVeStA's built-in Welch's t-test, 6 out of 7 pairwise parameter comparisons yield statistically different growth trajectories, while the exception reveals a saturation effect in knowledge locality. Our results demonstrate that SMC offers a principled, reproducible methodology for the quantitative analysis of agent-based economic models.

1 Introduction

One of the key challenges in economics is understanding how economic systems evolve over time, and in particular identifying the sources of long-run economic growth. Traditional growth models, such as the Solow-Swan model [61], treat technological progress as exogenous, leaving unexplained the very engine of sustained growth. *Endogenous growth theory* emerged as a response, with contributions by Romer [54], Lucas [39], and Aghion and Howitt [2], explaining growth through internal mechanisms such as innovation, human capital, and knowledge spillovers (see Section 3). However, these models usually rely on simplifying assumptions such as representative agents, rational behavior, and equilibrium, which limit their adherence to real economic dynamics characterized by heterogeneity and interaction at the micro level [42].

More recently, agent-based computational economics (ACE) has been proposed as a solution to these problems. ACE offers a complementary approach [63, 24] where economic dynamics are modeled as collections of simple interacting agents with heterogeneous characteristics and adaptive behaviors, i.e. agent-based models (ABMs). These models, also known as multi-agent systems [70], can capture emergent macroeconomic phenomena such as growth, business cycles, or crises from interaction at the individual level (see Section 2). However, ABMs do not admit analytical solutions and their analysis relies on numerical simulations, where coming up with the *right* experiment design is difficult but crucial to obtain meaningful insights [59, 66].

In a recent line of research [66, 64, 51], approaches from computer science known as *statistical model checking* (SMC) [46, 1] have been proposed as a solution to the problem of rigorous analysis of ABMs. This research line builds on MultiVeStA [58, 35, 66], a statistical model checker designed for the quantitative analysis of discrete-event simulations. Compared to standard Monte Carlo approaches, SMC automatically determines the minimum number of simulations needed to achieve user-specified confidence levels, provides built-in formal hypothesis testing to compare different model configurations, and can target temporal properties through expressive query languages.

In this paper, we show how SMC, and in particular MultiVeStA, can be used to automate the analysis of a seminal ABM, the *Island Model* of Fagiolo and Dosi [29]. This model captures the exploration-exploitation tradeoff in technological search. Our goal is methodological rather than model-theoretic: we treat the Island Model as a black-box stochastic transition system observed through aggregate outputs, and show how MultiVeStA can equip its analysis with formal statistical guarantees—without re-encoding it in a formal modelling language. Our contributions are: (1) a detailed description of the Island Model accessible to non-experts; (2) software engineering aspects connected to the integration of the model with MultiVeStA; (3) a discussion of the obtained results, including a counterfactual analysis with formal pairwise hypothesis testing across different system parametrizations.

The paper is organized as follows. Section 2 provides background on agent-based modeling. Section 3 reviews endogenous growth theory. Section 4 introduces statistical model checking and MultiVeStA. Section 5 presents the Island Model. Section 6 describes our integration. Section 7 presents results. Section 8 concludes.

2 Agent-Based Modeling in Economics

Agent-based computational economics (ACE) models economies as collections of interacting autonomous agents [63]. Unlike traditional economic modeling based on representative agents and equilibrium analysis, ACE studies how macroeconomic patterns emerge from microeconomic interactions of heterogeneous individuals with bounded rationality.

Foundations. The foundations of ACE draw from multiple sources. Herbert Simon’s work on bounded rationality [60] challenged perfect optimization assumptions, arguing that real decision-makers use heuristics and satisficing strategies. The Santa Fe Institute’s artificial stock market [5] demonstrated how complex market dynamics emerge from adaptive trading rules. Thomas Schelling’s segregation model [57] showed how aggregate patterns arise from mild individual preferences.

Se

Several characteristics distinguish ACE from traditional macroeconomic modeling [24]: *Heterogeneity*, Agents differ in preferences, beliefs, strategies, and constraints, contrasting with representative-agent models; *Bounded rationality*, Agents use heuristics and learning rules rather than solving optimization problems; *Local interactions*, Agents interact with subsets of others rather than through centralized markets; *Out-of-equilibrium dynamics*, ABMs model transition paths and crises, not just steady states; *Emergence*, Macroeconomic phenomena arise from microeconomic interactions without being programmed.

Applications and Challenges. ABMs have been applied across economics: financial markets and bubbles [12, 37, 48, 3, 13], systemic risk [8], macroeconomic policy [25], and innovation dynamics [50, 26, 30, 15]. The Island Model belongs to this tradition of studying growth through technological search.

Despite flexibility, ABMs pose analytical challenges [28, 69]. Stochasticity means outcomes are distributions requiring many runs. Parameter sensitivity demands systematic exploration. Validation against empirical data is difficult. Statistical model checking addresses the first two challenges through rigorous, automated methods with quantifiable confidence.

3 Endogenous Growth Theory

Endogenous growth theory explains long-run growth from within the economic system, rather than treating technological progress as exogenous [55].

From Exogenous to Endogenous Growth. The Solow-Swan model [61] demonstrated that capital accumulation alone cannot sustain growth due to diminishing returns. The economy converges to a steady state where output grows only with exogenous technological progress. This explains convergence but leaves growth’s ultimate source unexplained.

Endogenous growth theory filled this gap. The key insight was that knowledge differs from physical capital: ideas are non-rival (one person’s use doesn’t diminish another’s) and potentially non-excludable. These properties create increasing returns at aggregate level [54]. Romer’s model formalized how firms invest in R&D to develop ideas that create spillovers benefiting others. Aghion and Howitt [2] developed a Schumpeterian approach emphasizing creative destruction.

The key mechanisms in endogenous growth include: *Learning by doing*, productivity improves as byproduct of experience [4]; *Human capital*, education increases productivity and innovation capacity [39]; *R&D*, intentional research generates new technologies [54]; *Knowledge spillovers*, ideas spread between firms and regions [38]; *Increasing returns*, scale economies create positive feedback [44].

The Evolutionary Tradition and the Exploration-Exploitation Tradeoff. A foundational contribution to the agent-based study of economic growth is the evolutionary theory of Nelson and Winter [50], which models firms as boundedly-rational entities that search for new technologies through stochastic routines rather than optimizing behavior. In their framework, innovation is an evolutionary process driven by variation (search for new techniques), selection (market competition), and retention (organizational routines). This perspective—where growth emerges from the collective dynamics of heterogeneous firms rather than from a representative agent’s optimization—provides the intellectual foundation for the Island Model and, more broadly, for the Schumpeterian tradition in agent-based economics [26, 27].

A central theme within this tradition is the exploration-exploitation tradeoff, formalized by March [49] in the context of organizational learning. March showed that organizations face a fundamental tension: *exploitation* of existing competencies yields reliable but incremental returns, while *exploration* of new alternatives is uncertain but potentially transformative. Crucially, the two strategies compete for scarce resources, and an excess of either leads to suboptimal outcomes—pure exploitation causes lock-in and stagnation, while pure exploration dissipates resources on unproven alternatives. This insight, further developed by Levinthal [47] in terms of the “myopia of learning,” has been influential across organizational theory, evolutionary economics, and machine learning (multi-armed bandits) [62].

The Island Model [29] directly operationalizes March’s tradeoff: agents choose between mining known islands (exploitation) and searching for new ones (exploration), with the aggregate balance determining the economy’s growth trajectory. The model further connects to fitness landscape theory [41] and the literature on technological search [33], where the structure of the search space shapes the returns to exploration.

These ABMs of endogenous growth pose significant analytical challenges: their large state spaces and path-dependent dynamics preclude closed-form solutions, while the sensitivity of growth trajectories to parameter values demands systematic exploration backed by statistical guarantees. These features make them natural candidates for statistical model checking, which we introduce next.

4 Statistical Model Checking and MultiVeStA

The communities of verification and formal methods, part of computer science, proposed over the years several techniques for the analysis of systems. A notable example is the family of techniques known as model checking [21, 6].

Model checking. Traditional model checking verifies whether system M , given in some mathematical formalism, satisfies a property ϕ , typically given in a formal logic, written $M \models \phi$ [21]. For finite state spaces, this can be done exhaustively. However, ABMs have state spaces too large or infinite for exhaustive checking.

Probabilistic versions of model checking (PMC) [6, 40] can handle stochastic systems, but still require state space exploration. Instead of “does M satisfy ϕ ?”, these answer questions like “what is the probability that M satisfies ϕ ?”. Statistical model checking (SMC) [1, 46] answers this question by estimating such probabilities. It simulates M multiple times, evaluating ϕ on each trace, and using statistics to estimate the probability.

The key advantage is scalability: required simulations depend on desired confidence and precision, not state space size. This comes at the price of losing exactness in the analysis results. However, the statistical guarantees provided by SMC can be fine tuned. Two approaches exist [1]: hypothesis testing (using Sequential Probability Ratio Test [67]) and direct estimation with confidence intervals.

MultiVeStA. MultiVeStA [58, 66, 35] is a statistical analyzer that performs quantitative analysis on models of systems, with the only requirement being that of admitting probabilistic/stochastic (IID) simulations. Its architecture supports black-box integration with existing discrete-event simulators through a simple API. This makes SMC more accessible to external domains, allowing to directly analyze existing models written in non-formal general-purpose programming languages, or agent-based domain-specific languages. For example, MultiVeStA has been applied to several domains and simulators such as crowd steering scenarios [53], security threat analysis [10, 17], public transportation systems in smart cities [36, 20], product lines engineering [9, 65], business processes [22], decentralized finance [7], robotic systems [11, 14], collective adaptive systems [34, 19], agent-based models [66, 64, 51]. This included java-, c- and python-based simulators.

MultiVeStA alleviates the burden coming from performing reliable statistical analyses, as it automates all involved steps (e.g., triggering the minimum number of simulations required for obtaining user-specified confidence intervals, or comparing the results obtained for different model parameterizations). In particular, MultiVeStA enables the verification of properties defined in the MultiQuaTeX language, a practitioner-oriented domain-specific language allowing recursive and parametric queries corresponding to formal logics.

We complete this section with a brief introduction to the MultiQuaTeX language, which we use to express the properties of interest in our analysis. MultiQuaTeX uses recursive functions over simulation states. The basic building block is `s.rval("x")` returning observable x 's current value. For example, to compute expected $\log(\text{GDP})$ at time 101:

```

1 obsAtStep(x, obs) =
2   if (s.rval("steps") == x) then s.rval(obs)
3   else # obsAtStep(x, obs) fi;

```

The function recursively advances simulation (via #) until reaching the target step. Parametric queries enable systematic analysis:

```

1 eval parametric(E[ obsAtStep(t, "logGDP") ], t, 1, 10, 201);

```

For each simulation, this computes $\log(\text{GDP}(t))$ for $t \in \{1, 11, 21, \dots, 201\}$. By performing enough simulations, MultiVeStA uses these values to estimate the expected values of $\log(\text{GDP})$ in each time point, each equipped with a confidence interval of width specified by the user. Indeed, SMC suits ABMs well: the stochasticity and large state spaces of ABMs make exhaustive analysis particularly challenging, while SMC provides principled uncertainty quantification and systematic parameter exploration. In particular, given a desired statistical significance level α and interval width δ , MultiVeStA automatically determines and performs the minimum number of simulations needed to guarantee that each point estimate lies within a confidence interval of width at most δ , with statistical confidence $1 - \alpha$ [66].

Comparison with alternative approaches. The analysis of ABMs through simulation has a long tradition, and several methodologies exist beyond plain Monte Carlo. *Traditional sensitivity analysis* [56] varies parameters one at a time (or via factorial designs) and records the effect on output statistics, but typically relies on a fixed, predetermined number of simulations without formal guarantees on the precision of the estimates. *Parametric grid search* systematically covers the parameter space but faces the curse of dimensionality and, again, offers no principled criterion for the required sample size at each configuration. *Econometric meta-modeling* [43] fits surrogate models (e.g., regression or kriging) to simulation output, enabling efficient interpolation across the parameter space, but requires careful experimental design and may miss non-smooth or regime-switching behaviors common in ABMs. More recently, machine learning surrogates have been proposed to accelerate ABM calibration and parameter exploration [45], combining neural networks and gradient-boosted trees with intelligent sampling to efficiently navigate large parameter spaces. Comprehensive frameworks for the empirical validation of ABMs have also been developed [32, 31], addressing the interplay between calibration, sensitivity analysis, and comparison with empirical data.

SMC, as implemented in MultiVeStA, complements these approaches by offering *adaptive* sample sizes that are automatically determined to achieve user-specified statistical guarantees at every point of interest. Rather than fixing the number of runs ex ante, MultiVeStA adds simulation batches until the confidence interval width falls below the target δ , ensuring that regions of high variance receive proportionally more runs. Furthermore, MultiVeStA’s built-in hypothesis testing provides a rigorous framework for counterfactual comparisons that does not require fitting an intermediate model, operating directly on the simulation output with controlled Type I error and computable statistical power [66].

5 The Island model, an ABM to study endogenous growth

The Island Model [29] captures the exploration-exploitation tradeoff in technological search. Inspired by Phelps’ islands economy [52], it reinterprets “islands” as technologies in an abstract technology space.

5.1 Economic Motivation

The model translates core mechanisms of endogenous growth into an agent-based framework through precise analogies. *Exploration* corresponds to R&D investment: agents who leave a productive island to search for new ones bear a direct opportunity cost (foregone output) in exchange for the chance of discovering a superior technology—mirroring how firms allocate resources between current production and speculative research. *Imitation* captures technological diffusion: agents who observe a stronger productivity signal from a distant island migrate toward it, analogous to firms adopting proven innovations from competitors [50].

The spatial structure of the grid encodes the notion that more radical innovations—further from the known technological frontier (the center)—tend to be more productive but harder to reach, reflecting the empirical regularity that breakthrough technologies require longer search but yield higher returns [33]. The parameter ϕ governs *cumulative learning*: past skills carry over to newly discovered islands, capturing the learning-by-doing mechanism of Arrow [4] whereby a firm’s absorptive capacity grows with experience. The parameter ρ controls the spatial decay of productivity signals, modeling knowledge spillovers: low ρ corresponds to a regime of broad information diffusion (e.g., open science, strong patent disclosure), while high ρ restricts information to local clusters. Finally, α determines returns to scale in production—whether concentrating workers on a single technology yields increasing ($\alpha > 1$) or decreasing ($\alpha < 1$) marginal returns—directly connecting to the debate on agglomeration economies [44].

5.2 Model Components

Technology space: A $T \times T$ grid where each cell (x, y) may contain an island with probability π . The center $(T/2, T/2)$ always contains an island where all agents start, representing the initial technology.

Agents: N agents occupy grid positions in one of three states:

1. **Miners** (Type 1): Exploit known islands, producing output and broadcasting productivity signals
2. **Imitators** (Type 2): Target successful islands, modeling the adoption of established technologies
3. **Explorers** (Type 3): Search randomly for new islands, representing R&D

Island productivity: When discovered, the productivity of a new island/technology in position (x, y) is determined by:

$$s_{(x,y)} = (1 + \text{Poisson}(\lambda)) \cdot (|x - T/2| + |y - T/2| + \phi \cdot \text{skills}_i + \varepsilon) \quad (1)$$

where the Poisson term represents breakthroughs, distance from center captures the novelty premium (more distant technologies tend more productive), past skills reflect cumulative learning, and $\varepsilon \sim N(0, 1)$ adds noise.

Production: Miners at island (x, y) produce:

$$y_i = s_{(x,y)} \cdot m_{(x,y)}^{\alpha-1} \quad (2)$$

where $m_{(x,y)}$ is miner count and α controls returns to scale ($\alpha < 1$: decreasing returns/crowding; $\alpha = 1$: constant; $\alpha > 1$: increasing returns/agglomeration). GDP is total production across miners.

5.3 Dynamics

At each time step:

Signal transmission: Miners broadcast signals decaying with distance. Agent i receives signal from miner j with probability:

$$w_{ij} = \frac{m_{(x_j, y_j)}}{\sum_k \mathbf{1}[\text{Type}_k = 1]} \cdot \exp(-\rho \cdot d_{ij}) \quad (3)$$

where d_{ij} is Manhattan distance. Parameter ρ controls knowledge locality: low ρ creates global information; high ρ creates local bubbles.

Type transitions: At each step, each miner may become an explorer with probability ε , representing the willingness to abandon a known technology in search of a better one. Alternatively, a miner who receives a productivity signal stronger than its current production becomes an imitator, heading toward the more productive island. Conversely, an explorer who lands on an undiscovered island becomes a miner, with the island's productivity determined at the moment of discovery. Similarly, an imitator who reaches its destination island reverts to mining. These transitions are illustrated in Fig. 1.

Movement: Explorers move randomly in cardinal directions; imitators move deterministically toward destinations.

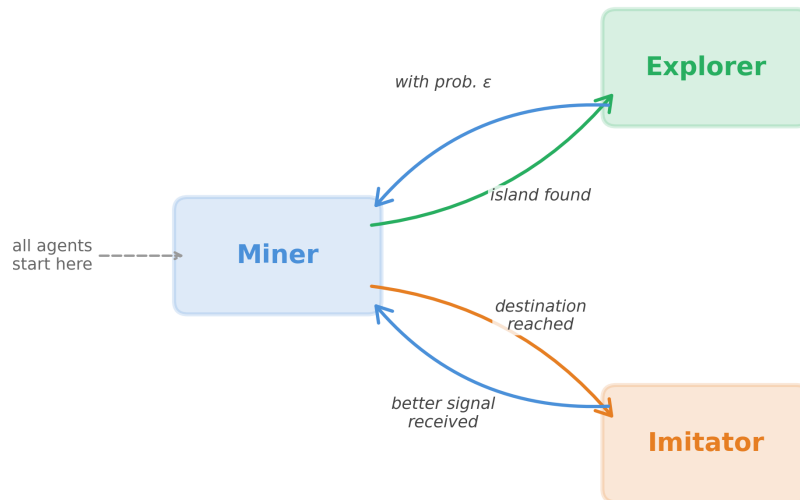


Figure 1: Agent type transitions. Miners are the productive state. Exploration (with probability ε) and imitation (upon receiving a stronger productivity signal) represent two distinct search strategies. Both explorers and imitators return to mining upon completing their search.

Table 1 summarizes parameters with economic interpretations.

Why Statistical Model Checking. The Island Model exhibits several features that make it particularly challenging to analyze with standard Monte Carlo methods. First, the dynamics are strongly *path-dependent*: early exploration successes or failures can lock the economy into qualitatively different growth regimes, leading to high inter-simulation variance. Second, the model can produce *multiple dynamic regimes*—sustained growth, stagnation, or lock-in on suboptimal technologies—depending on the stochastic sequence of discoveries and agent transitions. Third, the sensitivity to parameters such as α , ρ , and ε is non-trivial: small changes can shift the economy from one regime to another, but this is difficult to detect without controlled confidence intervals at each time point.

Table 1: Island Model Parameters

Parameter	Description	Default	Economic Interpretation
N	Number of agents	20	Labor force size
T	Simulation length	201	Time horizon (steps)
π	Island density	0.1	Technological opportunity
α	Returns to scale	1.5	Market structure
ε	Exploration probability	0.1	Innovation intensity
λ	Technology jump	1	Breakthrough frequency
ϕ	Past skills weight	0.5	Learning-by-doing strength
ρ	Knowledge locality	0.1	Information regime

A standard Monte Carlo approach computes sample means over a fixed number of runs, but provides no formal guarantee on the precision of those estimates, nor built-in tools to rigorously compare different parametrizations. SMC, and in particular MultiVeStA, addresses these limitations by automatically determining the number of simulations required to achieve a target confidence interval width at *every* time point of interest, and by providing built-in hypothesis testing to formally establish whether two parameter configurations produce statistically distinguishable trajectories. Our analysis focuses on aggregate observables (GDP, logGDP, AGR), expected trajectories over a finite horizon ($T = 201$), and one-parameter-at-a-time sweeps, choices that privilege interpretability and comparability with the original paper at the cost of leaving aside agent-level distributions and joint parameter effects

Original Implementation. The model was originally a monolithic MATLAB script combining parameters, state variables, and dynamics in a single file. While functional, this structure posed challenges: limited modularity (difficult to modify aspects independently), poor reusability (no stepping or intermediate queries), difficult external integration, and hard extensibility. These motivated the restructuring described next, which also required controlling random seeds for independent runs and preserving the order of random draws to maintain behavioral equivalence with the original.

6 Model’s implementation and integration details

In this section, we present code-specific implementation details of the model, as well as aspects connected to its integration with MultiVeStA. The original monolithic script was refactored into a modular, object-oriented architecture exposing the interfaces required by the statistical model checker.

6.1 Architecture Overview

We restructured into two main classes:

- **Model:** Main simulation class containing global state (grid, GDP, discoveries) and dynamics
- **Agent:** Individual agent class encapsulating state (position, type, production) and behaviors (movement, transitions)

This separation allows extending with new agent types or behaviors without modifying core logic.

6.2 Model Interface

Integrating a simulator with MultiVeStA requires exposing three basic actions [66]: (i) `reset(seed)`, which resets the simulator to its initial state and updates the random seed used for pseudo-random number generation, so that each simulation run is independent; (ii) `next`, which advances the simulation by one step; (iii) `eval(obs)`, which evaluates an observation on the current simulation state, where an observation can be any feature of the aggregate model or of any group of agents. In our MATLAB implementation, these correspond to methods of the `Model` class:

```

1 classdef Model
2     methods
3         function obj = setParams(obj, pi, alpha, eps, phi, rho, lambda, ...)
4         function obj = reset(obj, seed)           % action (i)
5         function obj = next(obj)                 % action (ii)
6         function value = evalObs(obj, variable)  % action (iii)
7     end
8 end

```

Listing 1: MultiVeStA integration interface (`Model` class).

An additional method, `setParams`, receives model parameters as strings from the command line and is called once at startup, enabling parameter sweeps via the `-otherParams` flag. The observable interface provides queryable quantities at each step:

```

1 function value = evalObs(obj, variable)
2     switch variable
3         case "GDP"
4             value = obj.GDP(obj.CurrentStep);
5         case "logGDP"
6             value = log(obj.GDP(obj.CurrentStep));
7         case "AGR"
8             value = obj.GDP(obj.CurrentStep) ...
9                 - obj.GDP(obj.CurrentStep-1);
10        case "AGR_total"
11            gdp_start = obj.GDP(t_start);
12            gdp_end   = obj.GDP(obj.CurrentStep);
13            value = (log(gdp_end) - log(gdp_start)) ...
14                / (obj.CurrentStep - t_start + 1);
15        end
16 end

```

Listing 2: Observable interface.

Observables include GDP (raw aggregate output), logGDP (logarithm of aggregate GDP, used for growth analysis), AGR (absolute growth rate between consecutive steps), and `AGR_total` (average growth rate over the entire simulation, computed as $(\log \text{GDP}_T - \log \text{GDP}_{t_0}) / (T - t_0)$, where t_0 is the first step with positive GDP).

The `-otherParams` flag specifies the class name, the evaluation method, the parameter method, and the parameter values (π , α , ε , φ , ρ , λ). MultiVeStA then calls `setParams` once, and iterates `reset/next/evalObs` for each simulation run until convergence.

6.3 Agent Class

The `Agent` class encapsulates individual state and behavior:

```

1 classdef Agent
2     properties
3         Type           % 1: Miner, 2: Imitator, 3: Explorer

```

```

4      X; Y          % grid position
5      Productivity; Production; Past_skills
6      end
7      methods
8          function obj = Agent(type, x, y)
9              obj.Type = type; obj.X = x; obj.Y = y;
10             obj.Productivity = 1; obj.Production = 0; obj.Past_skills = 1;
11         end
12         function obj = move(obj, d)
13             if d=="right", obj.X=obj.X+1; elseif d=="left", obj.X=obj.X-1;
14             elseif d=="up", obj.Y=obj.Y+1; elseif d=="down", obj.Y=obj.Y-1; end
15         end

```

Additional methods handle type transitions (becomeMiner, becomeExplorer, becomeImitator) and production.

This encapsulation enables adding new agent types or modifying behaviors without affecting the simulation loop.

6.4 MultiQuaTEx Queries

We developed two queries for different analyses:

Transient analysis of log(GDP) studies the evolution of output over time:

```

1  obsAtStep(x, obs) =
2  if (s.rval("my_time") == x) then s.rval(obs)
3  else # obsAtStep(x, obs) fi;
4  eval parametric(E[ obsAtStep(x, "logGDP") ], x, 1, 10, 201);

```

This computes the average $\log(\text{GDP}(t))$ for $t \in \{1, 11, 21, \dots, 201\}$, producing a time series of average log-output with confidence intervals. It is used for all parameter sweeps.

Average Growth Rate (AGR) computes the overall growth rate at the end of the simulation, used for the exploration-exploitation analysis:

```

1  obsAtStep(x, obs) =
2  if ( s.rval("my_time") == x )
3  then s.rval(obs)
4  else # obsAtStep(x, obs) fi ;
5  eval E[ obsAtStep(201, "AGR_total") ];

```

The execution pipeline consists of: configuration specification, query selection, MultiVeStA execution with block-based convergence checking, CSV output with means and confidence intervals, and post-processing visualization.

7 Analysis

We now present the results of applying MultiVeStA to the Island Model. We first reproduce two stylized facts from the original paper, then perform a counterfactual sensitivity analysis with formal hypothesis testing across different parameter configurations.

7.1 Experimental Setup

All experiments use a 95% confidence level ($\alpha_{\text{conf}} = 0.05$), block size of 30 simulations, and a precision threshold $\delta = 1$ for the width of the confidence interval. MultiVeStA automatically determines the number of simulation batches required to achieve convergence. The baseline configuration uses the

defaults from Table 1. Each simulation runs for up to $T = 201$ time steps, and the primary observable is the average $\log(\text{GDP}(t))$ evaluated at $t \in \{1, 11, 21, \dots, 201\}$.

7.2 Stylized Facts: The Role of Innovation

A fundamental prediction of endogenous growth theory is that sustained growth requires ongoing innovation [54, 2]. We verify this stylized fact by contrasting two scenarios under the baseline parameterization. This is shown in Fig. 2.

In the *stagnation* scenario, the exploration probability is set to zero at $t = 50$ ($\varepsilon \rightarrow 0$), removing all incentive for technological search. As shown in Fig. 2, average $\log(\text{GDP})$ grows during the initial phase when exploration is active, but plateaus after $t = 50$, stabilizing at around 10.4. Without exploration, agents exhaust the productivity of known islands and no new technologies are discovered.

In the *sustained growth* scenario, $\varepsilon = 0.1$ throughout. The economy exhibits approximately linear growth in log-output, reaching average $\log(\text{GDP}) \approx 22.7$ at $t = 201$ —more than double the stagnation level. The two trajectories diverge sharply after $t = 50$, confirming that continuous innovation is necessary and sufficient for sustained growth.

This result reproduces the stylized fact from Fagiolo and Dosi [29] (their Fig. 1a) and aligns with the core prediction of endogenous growth theory: economies that cease innovating converge to a stationary state.

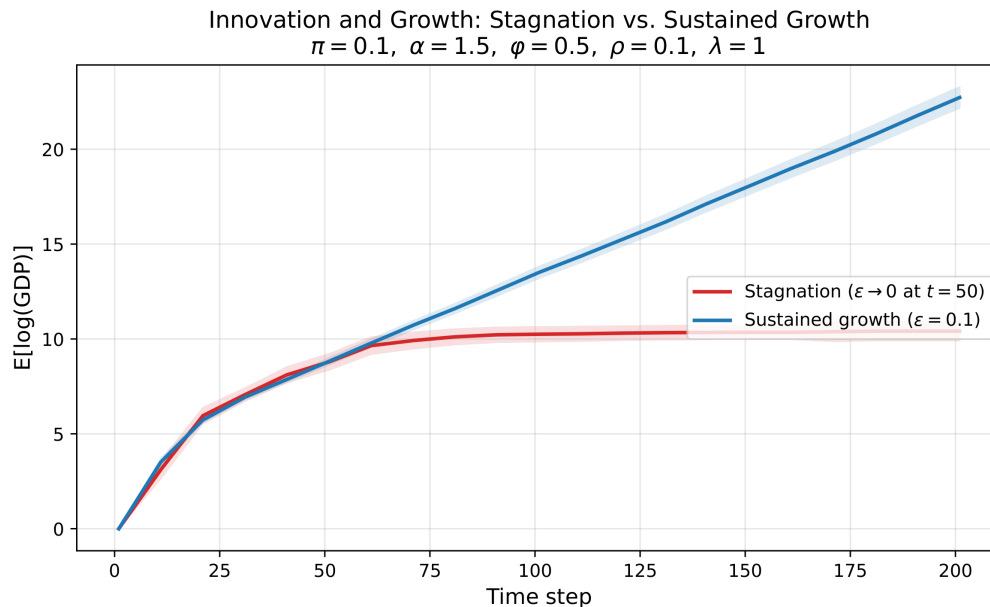


Figure 2: Stagnation vs. sustained growth. When exploration ceases at $t = 50$, $\log(\text{GDP})$ plateaus. With continuous exploration ($\varepsilon = 0.1$), growth is sustained. Shaded bands show 95% confidence intervals.

7.3 Exploration-Exploitation Trade-off

The exploration probability ε governs the fraction of miners who leave their current island to search for new technologies. While exploration drives long-run growth, it has an immediate cost: explorers do not

produce output while searching, creating a trade-off analogous to the multi-armed bandit problem [62].

Fig. 3 plots the Average Growth Rate (AGR) across 11 values of $\epsilon \in [0, 1]$. AGR increases sharply from $\epsilon = 0$ (no growth) to $\epsilon \approx 0.1$, where it peaks, and then declines monotonically. At high exploration rates ($\epsilon > 0.5$), most agents are searching rather than producing, and growth stabilizes at a lower level.

The peak at $\epsilon = 0.1$ reproduces the finding of Fagiolo and Dosi [29] (their Fig. 6d): a moderate exploration rate optimally balances the discovery of new technologies against the exploitation of existing ones. This is consistent with March’s [49] theoretical prediction that organizations perform best with a balanced exploration-exploitation strategy.

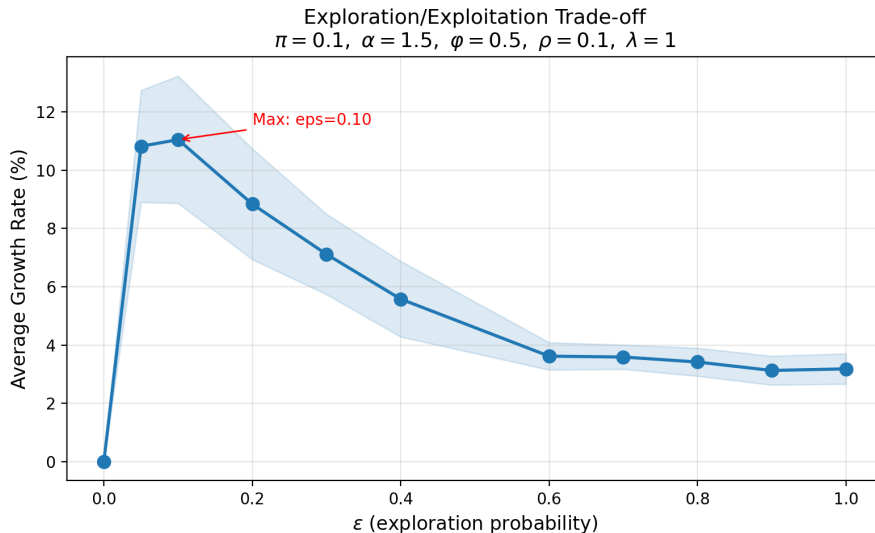


Figure 3: Average Growth Rate vs. exploration probability ϵ . Growth peaks at $\epsilon \approx 0.1$, demonstrating the exploration-exploitation trade-off. Error bars show 95% confidence intervals.

7.4 Counterfactual Analysis

We perform a counterfactual sensitivity analysis by varying individual parameters while holding the others at their baseline values, following the methodology of Fagiolo and Dosi [29] (their Fig. 8). For each parameter configuration, MultiVeStA runs simulation batches (block size 30) until the confidence interval width converges below $\delta = 1$.

To formally compare the resulting trajectories across different parameter values, we use MultiVeStA’s statistical hypothesis testing module, following the methodology described in [66]. Given two sets of simulation traces obtained under different parameterizations, MultiVeStA applies a Welch’s t-test [68] at each time step to test the null hypothesis H_0 that the two configurations produce equal expected values of the observable. The test computes the t-statistic from the sample means and variances of the two groups, and rejects H_0 at significance level $\alpha_{\text{conf}} = 0.05$ when the statistic falls outside the acceptance region. MultiVeStA also reports the statistical power of the test, quantifying the probability of correctly detecting a difference when one exists. The t-test results are shown at the bottom of each sweep figure: a filled dot (\bullet) indicates that the null hypothesis is not rejected (means are equal), while a cross (\times) indicates rejection (means are statistically different).

Out of 7 pairwise comparisons across three parameters, 6 reject the null hypothesis of equal means at $t = 201$ with power above 0.85, confirming that the observed differences in growth trajectories are statistically significant. The only exception is $\rho = 3.0$ vs. 5.0, where the test does not reject equality at any time step, suggesting that the effect of knowledge locality saturates beyond a certain threshold.

7.4.1 Returns to Scale (α)

The parameter α controls whether production exhibits decreasing ($\alpha < 1$), constant ($\alpha = 1$), or increasing ($\alpha > 1$) returns to the number of miners on an island. We tested values $\alpha \in \{0.9, 1.0, 1.1\}$, all of which achieved convergence with $\delta = 1$. Overall, the maximum number of required simulations was 60 for $\alpha = 1.1$ (at multiple time steps from $t = 131$ onward, due to higher variance in the super-linear regime), while $\alpha = 0.9$ and $\alpha = 1.0$ converged with 30 simulations at most steps.

Fig. 4 shows the resulting trajectories. Growth increases monotonically with α . Even a mild degree of increasing returns ($\alpha = 1.1$) produces markedly higher growth than constant returns ($\alpha = 1.0$), consistent with the agglomeration effects predicted by the model.

The pairwise t-tests (Fig. 4, bottom) confirm that all three pairs are statistically distinguishable at $t = 201$ ($\alpha_{\text{conf}} = 0.05$), with power above 0.85 in all cases. The $\alpha = 1.0$ vs. 1.1 comparison shows equal means at early time steps ($t \in [11, 71]$), with the test rejecting equality from $t = 81$ onward as the trajectories diverge.

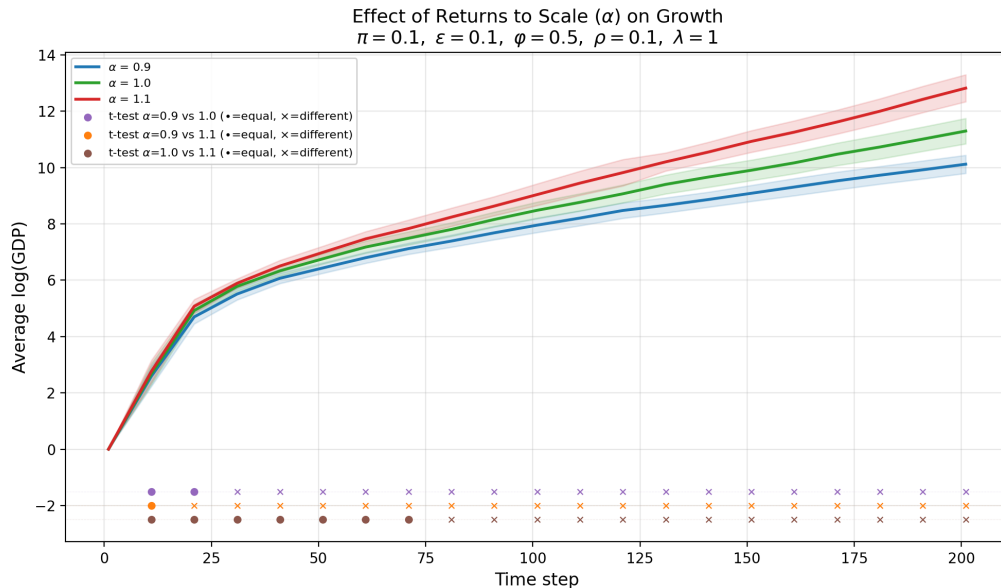


Figure 4: Effect of returns to scale α on average $\log(\text{GDP})$. Shaded bands show 95% confidence intervals. Bottom: pairwise t-test results (\bullet = equal means, \times = different means).

7.4.2 Skill Transfer (φ)

The parameter φ weights the contribution of an agent's past skills when determining the productivity of a newly discovered island. Only $\varphi \in \{0, 0.1\}$ achieved convergence; higher values produced variance too large for $\delta = 1$.

Fig. 5 shows that even a small amount of skill transfer ($\varphi = 0.1$) produces visibly higher growth than no transfer ($\varphi = 0$), with average $\log(\text{GDP})$ reaching 8.1 vs. 7.4 at $t = 201$. This captures the learning-by-doing mechanism emphasized by Arrow [4]: productivity gains are cumulative and embodied in workers' experience.

The pairwise t-test (Fig. 5, bottom) confirms that the two means are statistically different at $t = 201$ (power > 0.99). At early time steps ($t \in [11, 41]$) the test does not reject equality, as the trajectories have not yet diverged sufficiently.

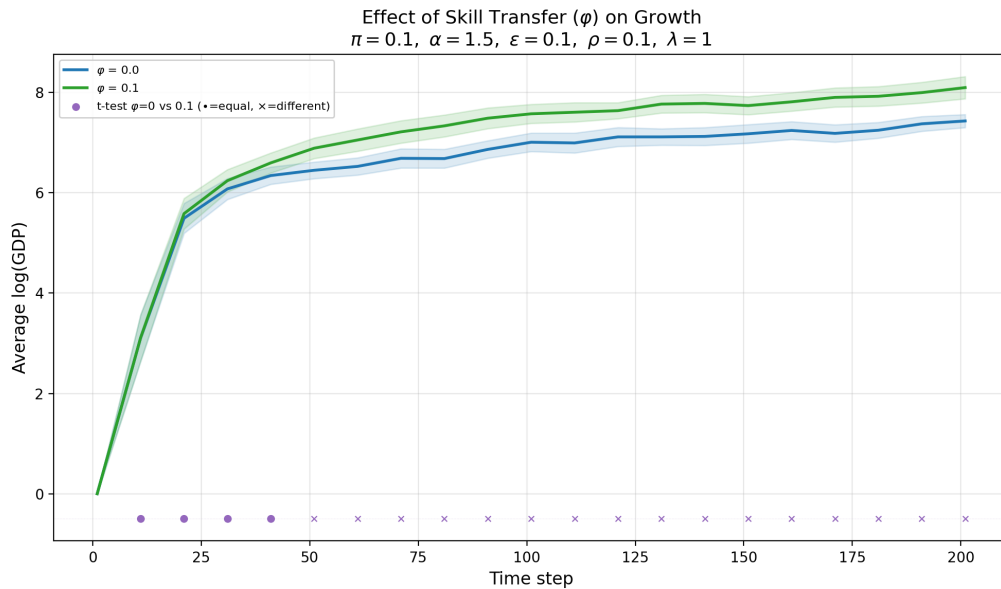


Figure 5: Effect of skill transfer φ on average $\log(\text{GDP})$. Even low skill transfer ($\varphi = 0.1$) produces higher growth than none. Bottom: pairwise t-test results (• = equal means, × = different means).

7.4.3 Knowledge Locality (ρ)

The parameter ρ controls how rapidly productivity signals decay with distance. Low ρ creates global knowledge diffusion; high ρ restricts information to nearby islands. We tested $\rho \in \{1.0, 3.0, 5.0\}$.

Fig. 6 shows that lower ρ promotes growth: average $\log(\text{GDP})$ reaches 9.0 at $\rho = 1.0$ vs. 8.0 at $\rho = 5.0$. When knowledge diffuses more broadly, agents can identify and imitate productive technologies regardless of distance.

The t-test results (Fig. 6, bottom) reveal an interesting asymmetry: the pairs $\rho = 1.0$ vs. 3.0 and $\rho = 1.0$ vs. 5.0 show statistically significant differences (power > 0.99), while $\rho = 3.0$ vs. 5.0 does not reject equality at any time step (power = 1.0). This suggests a non-linear relationship where the transition from broad to moderate knowledge diffusion has a measurable effect on growth, but further restricting diffusion beyond $\rho = 3.0$ does not produce additional distinguishable changes, indicating a saturation effect.

These results demonstrate that MultiVeStA's counterfactual analysis can effectively detect meaningful differences across parameter configurations. The formal statistical guarantees complement the original analysis of Fagiolo and Dosi [29], confirming that variations in returns to scale, skill transfer,

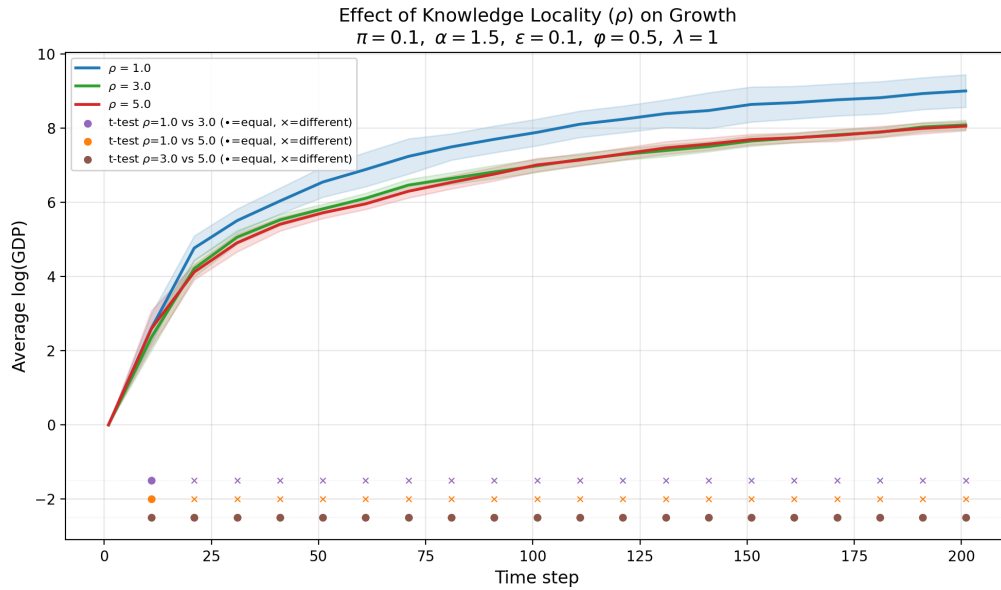


Figure 6: Effect of knowledge locality ρ on average $\log(\text{GDP})$. Lower ρ (broader knowledge diffusion) promotes growth. Bottom: pairwise t-test results (\bullet = equal means, \times = different means).

and knowledge locality produce genuinely distinct growth dynamics.

8 Conclusions

We have shown how MultiVeStA [58] can automate and enrich the analysis of the Island Model [29], a seminal agent-based model of endogenous growth. Our experiments reproduced key stylized facts with formal confidence intervals, confirmed the optimality of moderate exploration rates ($\varepsilon \approx 0.1$), and established through counterfactual analysis that 6 out of 7 pairwise parameter comparisons yield statistically different growth trajectories, with the single exception ($\rho = 3.0$ vs. 5.0) revealing a saturation effect in knowledge locality. By automating convergence checking and hypothesis testing, MultiVeStA provides a principled alternative to ad-hoc Monte Carlo approaches, and a reusable template for the rigorous analysis of agent-based models across economics and beyond; the approach scales with simulation cost and variance rather than state-space size, though higher-dimensional parameter sweeps would require more selective experimental designs.

Future Work. Several directions warrant investigation. Recently, MultiVeStA has been integrated with process mining techniques to *explain* analysis results [17, 19, 16, 18]; applying these to agent traces could reveal decision patterns invisible at the aggregate level, taking inspiration from techniques to discover process collaborations [23]. We also plan to explore steady-state analyses, already supported by MultiVeStA, and to apply the framework to richer model variants incorporating financial sectors [30] and large-scale macro-financial ABMs [26, 27, 25], where SMC could enable rigorous policy analysis—for instance, assessing the impact of prudential regulation on long-run growth or formally testing whether financial frictions shift the economy between growth regimes. Finally, data-driven calibration against historical data and the introduction of extended agent types represent natural extensions of this work.

References

- [1] Gul Agha & Karl Palmkog (2018): *A Survey of Statistical Model Checking*. *ACM Transactions on Modeling and Computer Simulation* 28(1), pp. 1–39, doi:10.1145/3158668.
- [2] Philippe Aghion & Peter Howitt (1992): *A Model of Growth Through Creative Destruction*. *Econometrica* 60(2), pp. 323–351, doi:10.2307/2951599.
- [3] Mikhail Anufriev & Giulio Bottazzi (2012): *Asset Pricing with Heterogeneous Investment Horizons*. *Studies in Nonlinear Dynamics & Econometrics* 16(4), doi:10.1515/1558-3708.1903.
- [4] Kenneth J. Arrow (1962): *The Economic Implications of Learning by Doing*. *The Review of Economic Studies* 29(3), pp. 155–173, doi:10.2307/2295952.
- [5] W. Brian Arthur, John H. Holland, Blake LeBaron, Richard Palmer & Paul Taylor (1996): *Asset Pricing Under Endogenous Expectations in an Artificial Stock Market*. In: *The Economy as an Evolving Complex System II*, Addison-Wesley, pp. 15–44, doi:10.1201/9780429496639-2.
- [6] Christel Baier & Joost-Pieter Katoen (2008): *Principles of model checking*. MIT Press, doi:10.1093/comjnl/bxp025.
- [7] Massimo Bartoletti, James Hsin-yu Chiang, Tommi A. Junttila, Alberto Lluch-Lafuente, Massimiliano Mirelli & Andrea Vandin (2022): *Formal Analysis of Lending Pools in Decentralized Finance*. In Tiziana Margaria & Bernhard Steffen, editors: *Leveraging Applications of Formal Methods, Verification and Validation. Adaptation and Learning - 11th International Symposium, ISoLA 2022, Rhodes, Greece, October 22-30, 2022, Proceedings, Part III, Lecture Notes in Computer Science* 13703, Springer, pp. 335–355, doi:10.1007/978-3-031-19759-8_21. Available at https://doi.org/10.1007/978-3-031-19759-8_21.
- [8] Stefano Battiston, J. Doyne Farmer, Andreas Flache, Diego Garlaschelli, Andrew G. Haldane, Hans Heesterbeek, Cars Hommes, Carlo Jaeger, Robert May & Marten Scheffer (2016): *Complexity Theory and Financial Regulation*. *Science* 351(6275), pp. 818–819, doi:10.1126/science.aad0299.
- [9] Maurice H. ter Beek, Axel Legay, Alberto Lluch-Lafuente & Andrea Vandin (2020): *A Framework for Quantitative Modeling and Analysis of Highly (Re)configurable Systems*. *IEEE Trans. Software Eng.* 46(3), pp. 321–345, doi:10.1109/TSE.2018.2853726. Available at <https://doi.org/10.1109/TSE.2018.2853726>.
- [10] Maurice H. ter Beek, Axel Legay, Alberto Lluch-Lafuente & Andrea Vandin (2021): *Quantitative Security Risk Modeling and Analysis with RisQFLan*. CoRR abs/2101.08677. arXiv:2101.08677.
- [11] Lenz Belzner, Rocco De Nicola, Andrea Vandin & Martin Wirsing (2014): *Reasoning (on) Service Component Ensembles in Rewriting Logic*. In Shusaku Iida, José Meseguer & Kazuhiro Ogata, editors: *Specification, Algebra, and Software - Essays Dedicated to Kokichi Futatsugi, Lecture Notes in Computer Science* 8373, Springer, pp. 188–211, doi:10.1007/978-3-642-54624-2_10. Available at https://doi.org/10.1007/978-3-642-54624-2_10.
- [12] Giulio Bottazzi, Giovanni Dosi & Igor Rebesco (2005): *Institutional architectures and behavioral ecologies in the dynamics of financial markets*. *Journal of Mathematical Economics* 41(1-2), pp. 197–228, doi:10.1016/j.jmateco.2004.02.006.
- [13] Giulio Bottazzi & Daniele Giachini (2017): *Wealth and price distribution by diffusive approximation in a repeated prediction market*. *Physica A: Statistical Mechanics and its Applications* 471, pp. 473–479, doi:10.1016/j.physa.2016.12.012.
- [14] Roberto Bruni, Andrea Corradini, Fabio Gadducci, Alberto Lluch-Lafuente & Andrea Vandin (2015): *Modelling and analyzing adaptive self-assembly strategies with Maude*. *Sci. Comput. Program.* 99, pp. 75–94, doi:10.1016/J.SCICO.2013.11.043. Available at <https://doi.org/10.1016/j.scico.2013.11.043>.
- [15] Gianluca Capone, Franco Malerba, Richard R. Nelson, Luigi Orsenigo & Sidney G. Winter (2019): *History friendly models: retrospective and future perspectives*. *Eurasian Business Review* 9(1), pp. 1–23, doi:10.1007/s40821-019-00121-0.

- [16] Roberto Casaluca, Andrea Burattin, Francesca Chiaromonte, Alberto Lluch Lafuente & Andrea Vandin (2024): *White-box validation of quantitative product lines by statistical model checking and process mining*. *Journal of Systems and Software* 210, doi:10.1016/j.jss.2024.111983.
- [17] Roberto Casaluca, Andrea Burattin, Francesca Chiaromonte & Andrea Vandin (2023): *Process Mining Meets Statistical Model Checking: Towards a Novel Approach to Model Validation and Enhancement*. In Cristina Cabanillas, Niels Frederik Garmann-Johnsen & Agnes Koschmider, editors: *Business Process Management Workshops*, Springer International Publishing, Cham, pp. 243–256, doi:10.1007/978-3-031-25383-6_18.
- [18] Roberto Casaluca, Andrea Burratin, Francesca Chiaromonte, Alberto Lluch-Lafuente & Andrea Vandin (2024): *Enhancing Threat Model Validation: A White-Box Approach based on Statistical Model Checking and Process Mining*. In Bernardo Breve, Giuseppe Desolda, Vincenzo Deufemia & Lucio Davide Spano, editors: *Proceedings of the First International Workshop on Detection And Mitigation Of Cyber attacks that exploit human vulnerabilities (DAMOCLES 2024) co-located with 17th International Conference on Advanced Visual Interfaces (AVI 2024), Arenzano (Genoa), Italy, Arenzano, Italy, June 4th, 2024, CEUR Workshop Proceedings 3713, CEUR-WS.org*, pp. 9–20. Available at https://ceur-ws.org/Vol-3713/paper_2.pdf.
- [19] Roberto Casaluca, Max Tschaikowski & Andrea Vandin (2024): *White-Box Validation of Collective Adaptive Systems by Statistical Model Checking and Process Mining*. In Tiziana Margaria & Bernhard Steffen, editors: *Leveraging Applications of Formal Methods, Verification and Validation. REoCAS Colloquium in Honor of Rocco De Nicola - 12th International Symposium, ISoLA 2024, Crete, Greece, October 27-31, 2024, Proceedings, Part I, Lecture Notes in Computer Science 15219, Springer*, pp. 204–222, doi:10.1007/978-3-031-73709-1_13. Available at https://doi.org/10.1007/978-3-031-73709-1_13.
- [20] Vincenzo Ciancia, Diego Latella, Mieke Massink, Rytis Paskauskas & Andrea Vandin (2016): *A Tool-Chain for Statistical Spatio-Temporal Model Checking of Bike Sharing Systems*. In Tiziana Margaria & Bernhard Steffen, editors: *Leveraging Applications of Formal Methods, Verification and Validation: Foundational Techniques - 7th International Symposium, ISoLA 2016, Imperial, Corfu, Greece, October 10-14, 2016, Proceedings, Part I, Lecture Notes in Computer Science 9952*, pp. 657–673, doi:10.1007/978-3-319-47166-2_46. Available at https://doi.org/10.1007/978-3-319-47166-2_46.
- [21] Edmund M. Clarke, Orna Grumberg, Daniel Kroening, Doron A. Peled & Helmut Veith (2018): *Model checking, 2nd Edition*. MIT Press. Available at <https://mitpress.mit.edu/books/model-checking-second-edition>.
- [22] Flavio Corradini, Fabrizio Fornari, Andrea Polini, Barbara Re, Francesco Tiezzi & Andrea Vandin (2021): *A formal approach for the analysis of BPMN collaboration models*. *J. Syst. Softw.* 180, p. 111007, doi:10.1016/J.JSS.2021.111007. Available at <https://doi.org/10.1016/j.jss.2021.111007>.
- [23] Flavio Corradini, Sara Pettinari, Barbara Re, Lorenzo Rossi & Francesco Tiezzi (2024): *A technique for discovering BPMN collaboration diagrams*. *Softw. Syst. Model.* 23(6), pp. 1323–1343, doi:10.1007/S10270-024-01153-5. Available at <https://doi.org/10.1007/s10270-024-01153-5>.
- [24] Herbert Dawid & Domenico Delli Gatti (2018): *Agent-Based Macroeconomics*. In: *Handbook of Computational Economics*, 4, Elsevier, pp. 63–156, doi:10.2139/ssrn.3112074.
- [25] Herbert Dawid, Simon Gemkow, Philipp Harting, Sander van der Hoog & Michael Neugart (2012): *The EURACE@Unibi Model: An Agent-Based Macroeconomic Model for Economic Policy Analysis*. Technical Report 05-2012, Bielefeld University, doi:10.2139/ssrn.2408969.
- [26] Giovanni Dosi, Giorgio Fagiolo & Andrea Roventini (2010): *Schumpeter Meeting Keynes: A Policy-Friendly Model of Endogenous Growth and Business Cycles*. *Journal of Economic Dynamics and Control* 34(9), pp. 1748–1767, doi:10.1016/j.jedc.2010.06.018.
- [27] Giovanni Dosi, Marcelo C. Pereira, Andrea Roventini & Maria Enrica Virgillito (2017): *Micro and Macro Policies in the Keynes+Schumpeter Evolutionary Models*. *Journal of Evolutionary Economics* 27(1), pp. 63–90, doi:10.1007/s00191-016-0466-4.
- [28] Annalisa Fabretti (2013): *On the Problem of Calibrating an Agent Based Model for Financial Markets*. *Journal of Economic Interaction and Coordination* 8(2), pp. 277–293, doi:10.1007/s11403-012-0096-3.

- [29] Giorgio Fagiolo & Giovanni Dosi (2003): *Exploitation, Exploration and Innovation in a Model of Endogenous Growth with Locally Interacting Agents*. *Structural Change and Economic Dynamics* 14(3), pp. 237–273, doi:10.1016/s0954-349x(03)00022-5.
- [30] Giorgio Fagiolo, Daniele Giachini & Andrea Roventini (2020): *Innovation, finance, and economic growth: an agent-based approach*. *Journal of Economic Interaction and Coordination* 15(3), pp. 703–736, doi:10.1007/s11403-019-00258-1. Available at <https://doi.org/10.1007/s11403-019-00258-1>.
- [31] Giorgio Fagiolo, Mattia Guerini, Francesco Lamperti, Alessio Moneta & Andrea Roventini (2019): *Validation of Agent-Based Models in Economics and Finance*. In: *Computer Simulation Validation*, Springer, pp. 763–787, doi:10.1007/978-3-319-70766-2_31.
- [32] Giorgio Fagiolo, Alessio Moneta & Paul Windrum (2007): *A Critical Guide to Empirical Validation of Agent-Based Models in Economics: Methodologies, Procedures, and Open Problems*. *Computational Economics* 30(3), pp. 195–226, doi:10.1007/s10614-007-9104-4.
- [33] Lee Fleming (2001): *Recombinant Uncertainty in Technological Search*. *Management Science* 47(1), pp. 117–132, doi:10.1287/mnsc.47.1.117.10671.
- [34] Vashti Galpin, Anastasis Georgoulas, Michele Loreti & Andrea Vandin (2018): *Statistical Analysis of CARMA Models: an Advanced Tutorial*. In Björn Johansson & Sanjay Jain, editors: *2018 Winter Simulation Conference, WSC 2018, Gothenburg, Sweden, December 9-12, 2018*, IEEE, pp. 395–409, doi:10.1109/WSC.2018.8632456. Available at <https://doi.org/10.1109/WSC.2018.8632456>.
- [35] Stephen Gilmore, Daniel Reijnsbergen & Andrea Vandin (2017): *Transient and Steady-State Statistical Analysis for Discrete Event Simulators*. In: *Integrated Formal Methods - 13th International Conference, IFM 2017, Turin, Italy, September 20-22, 2017, Proceedings*, pp. 145–160, doi:10.1007/978-3-319-66845-1_10. Available at https://doi.org/10.1007/978-3-319-66845-1_10.
- [36] Stephen Gilmore, Mirco Tribastone & Andrea Vandin (2014): *An Analysis Pathway for the Quantitative Evaluation of Public Transport Systems*. In Elvira Albert & Emil Sekerinski, editors: *Integrated Formal Methods - 11th International Conference, IFM 2014, Bertinoro, Italy, September 9-11, 2014, Proceedings, Lecture Notes in Computer Science 8739*, Springer, pp. 71–86, doi:10.1007/978-3-319-10181-1_5. Available at https://doi.org/10.1007/978-3-319-10181-1_5.
- [37] Cars H. Hommes (2006): *Heterogeneous Agent Models in Economics and Finance*. In: *Handbook of Computational Economics*, 2, Elsevier, pp. 1109–1186, doi:10.1016/s1574-0021(05)02023-x.
- [38] Adam B. Jaffe, Manuel Trajtenberg & Rebecca Henderson (1993): *Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations*. *The Quarterly Journal of Economics* 108(3), pp. 577–598, doi:10.2307/2118401.
- [39] Robert E. Lucas Jr. (1988): *On the Mechanics of Economic Development*. *Journal of Monetary Economics* 22(1), pp. 3–42, doi:10.1016/0304-3932(88)90168-7.
- [40] Joost-Pieter Katoen (2016): *The Probabilistic Model Checking Landscape*. In: *Proceedings of the 31st Annual ACM/IEEE Symposium on Logic in Computer Science, LICS '16*, Association for Computing Machinery, New York, NY, USA, p. 31–45, doi:10.1145/2933575.2934574. Available at <https://doi.org/10.1145/2933575.2934574>.
- [41] Stuart A. Kauffman (1993): *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, New York, doi:10.1126/science.260.5113.1531.
- [42] Alan Kirman (2011): *Complex Economics: Individual and Collective Rationality*. Routledge, London, doi:10.23941/ejpe.v4i2.81.
- [43] Jack P. C. Kleijnen (2015): *Design and Analysis of Simulation Experiments*, 2nd edition. Springer, doi:10.1007/978-3-319-18087-8.
- [44] Paul Krugman (1991): *Increasing Returns and Economic Geography*. *Journal of Political Economy* 99(3), pp. 483–499, doi:10.1086/261763.

- [45] Francesco Lamperti, Andrea Roventini & Amir Sani (2018): *Agent-based model calibration using machine learning surrogates*. *Journal of Economic Dynamics and Control* 90, pp. 366–389, doi:10.1016/j.jedc.2018.03.011.
- [46] Axel Legay, Benoît Delahaye & Saddek Bensalem (2010): *Statistical Model Checking: An Overview*. In: *Runtime Verification (RV 2010)*, LNCS 6418, Springer, pp. 122–135, doi:10.1007/978-3-642-16612-9_11.
- [47] Daniel A. Levinthal & James G. March (1993): *The Myopia of Learning*. *Strategic Management Journal* 14(S2), pp. 95–112, doi:0.1002/smj.4250141009.
- [48] Thomas Lux & Frank Westerhoff (2009): *Economics Crisis*. *Nature Physics* 5(1), pp. 2–3, doi:10.1038/nphys1163.
- [49] James G. March (1991): *Exploration and Exploitation in Organizational Learning*. *Organization Science* 2(1), pp. 71–87, doi:10.1287/orsc.2.1.71.
- [50] Richard R. Nelson & Sidney G. Winter (1982): *An Evolutionary Theory of Economic Change*. Harvard University Press, Cambridge, MA, doi:10.1086/261177.
- [51] Marco Pangallo, Daniele Giachini & Andrea Vandin (2025): *Statistical Model Checking of NetLogo Models*. CoRR abs/2509.10977, doi:10.48550/ARXIV.2509.10977. arXiv:2509.10977.
- [52] Edmund S. Phelps (1969): *The New Microeconomics in Inflation and Employment Theory*. *American Economic Review* 59(2), pp. 147–160, doi:/10.1016/b978-0-12-554001-8.50009-0.
- [53] Danilo Pianini, Stefano Sebastio & Andrea Vandin (2014): *Distributed statistical analysis of complex systems modeled through a chemical metaphor*. In: *International Conference on High Performance Computing & Simulation, HPCS 2014, Bologna, Italy, 21-25 July, 2014*, IEEE, pp. 416–423, doi:10.1109/HPCSIM.2014.6903715. Available at <https://doi.org/10.1109/HPCSim.2014.6903715>.
- [54] Paul M. Romer (1990): *Endogenous Technological Change*. *Journal of Political Economy* 98(5), pp. S71–S102, doi:10.1086/261725.
- [55] Paul M. Romer (1994): *The Origins of Endogenous Growth*. *Journal of Economic Perspectives* 8(1), pp. 3–22, doi:10.1257/jep.8.1.3.
- [56] Andrea Saltelli, Marco Ratto, Terry Andres, Francesca Campolongo, Jessica Cariboni, Debora Gatelli, Michaela Saisana & Stefano Tarantola (2008): *Global Sensitivity Analysis: The Primer*. Wiley, doi:10.1111/j.1751-5823.2008.00062_17.x.
- [57] Thomas C. Schelling (1971): *Dynamic Models of Segregation*. *Journal of Mathematical Sociology* 1(2), pp. 143–186, doi:10.1080/0022250x.1971.9989794.
- [58] Stefano Sebastio & Andrea Vandin (2013): *MultiVeStA: Statistical model checking for discrete event simulators*. *Performance Evaluation* 70(6), pp. 457–475, doi:10.4108/icst.valuetools.2013.254377.
- [59] Davide Secchi & Raffaello Seri (2017): *Controlling for false negatives in agent-based models: a review of power analysis in organizational research*. *Computational and Mathematical Organization Theory* 23(1), pp. 94–121, doi:10.1007/s10588-016-9218-0. Available at <https://doi.org/10.1007/s10588-016-9218-0>.
- [60] Herbert A. Simon (1955): *A Behavioral Model of Rational Choice*. *The Quarterly Journal of Economics* 69(1), pp. 99–118, doi:10.2307/1884852.
- [61] Robert M. Solow (1956): *A Contribution to the Theory of Economic Growth*. *The Quarterly Journal of Economics* 70(1), pp. 65–94, doi:10.2307/1884513.
- [62] Richard S. Sutton & Andrew G. Barto (2018): *Reinforcement Learning: An Introduction*, second edition. MIT Press, Cambridge, MA, doi:10.1017/s0263574799211174.
- [63] Leigh Tesfatsion & Kenneth L. Judd (2006): *Handbook of Computational Economics: Agent-Based Computational Economics*. 2, Elsevier, Amsterdam, doi:10.1109/mci.2008.929849.
- [64] Andrea Vandin (2024): *Statistical Model Checking of Python Agent-Based Models: An Integration of MultiVeStA and Mesa*. In Bernhard Steffen, editor: *Bridging the Gap Between AI and Reality - Second International Conference, AISoLA 2024, Crete, Greece, October 30 - November 3, 2024, Proceedings, Lecture*

- Notes in Computer Science* 15217, Springer, pp. 398–419, doi:10.1007/978-3-031-75434-0_26. Available at https://doi.org/10.1007/978-3-031-75434-0_26.
- [65] Andrea Vandin, Maurice H. ter Beek, Axel Legay & Alberto Lluch-Lafuente (2018): *QFLan: A Tool for the Quantitative Analysis of Highly Reconfigurable Systems*. In Klaus Havelund, Jan Peleska, Bill Roscoe & Erik P. de Vink, editors: *Formal Methods - 22nd International Symposium, FM 2018, Held as Part of the Federated Logic Conference, FloC 2018, Oxford, UK, July 15-17, 2018, Proceedings, Lecture Notes in Computer Science* 10951, Springer, pp. 329–337, doi:10.1007/978-3-319-95582-7_19. Available at https://doi.org/10.1007/978-3-319-95582-7_19.
- [66] Andrea Vandin, Daniele Giachini, Francesco Lamperti & Francesca Chiaromonte (2022): *Automated and distributed statistical analysis of economic agent-based models*. *Journal of Economic Dynamics and Control* 143, doi:10.1016/j.jedc.2022.104458.
- [67] Abraham Wald (1945): *Sequential Tests of Statistical Hypotheses*. *The Annals of Mathematical Statistics* 16(2), pp. 117–186, doi:10.1214/aoms/1177731118.
- [68] Bernard L. Welch (1947): *The Generalization of ‘Student’s’ Problem when Several Different Population Variances are Involved*. *Biometrika* 34(1–2), pp. 28–35, doi:10.2307/2332510.
- [69] Paul Windrum, Giorgio Fagiolo & Alessio Moneta (2007): *Empirical Validation of Agent-Based Models: Alternatives and Prospects*. *Journal of Artificial Societies and Social Simulation* 10(2), doi:10.1007/s10614-007-9104-4.
- [70] M. Wooldridge (2009): *An Introduction to MultiAgent Systems*. Wiley. Available at <https://books.google.it/books?id=X3ZQ7yeDn2IC>.