AUTOMATED ANALYSIS OF ECONOMIC AGENT-BASED MODELS ***BY (STATISTICAL) MODEL CHECKING***

Andrea Vandin

Associate Professor in Computer Science

Sant'Anna School of Advanced Studies – Pisa

Institute of Economics

Department of Excellence 2018 - 2022

2023-2027!

EMbeDS Economics and Management in the era of Data Science

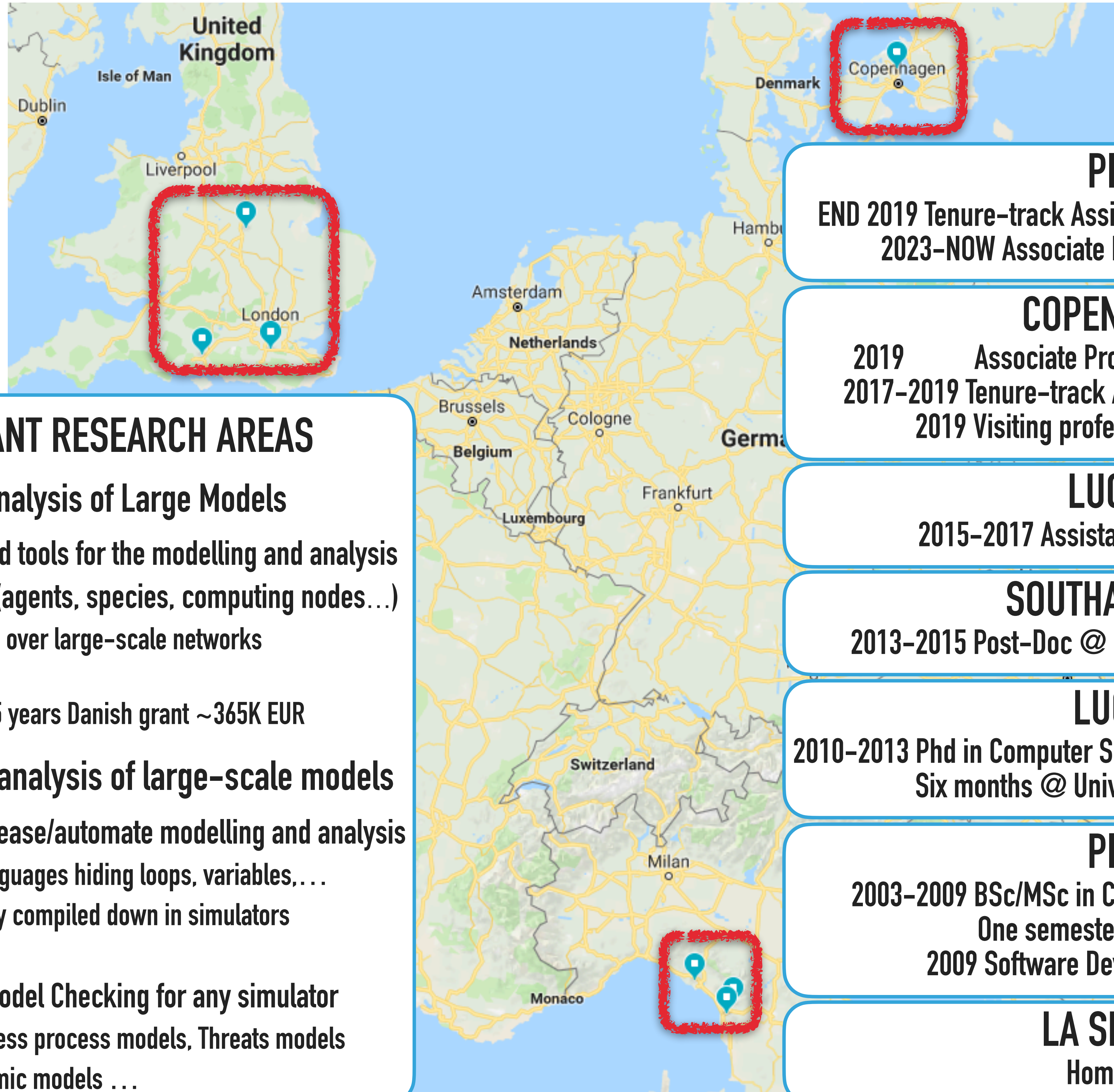
Adjoint Associate Professor

DTU Danmarks Tekniske Universitet

Classes 21t-22t, Software Validation and Verification, Unipi, 04-05/12/2023

Class 21t 04/12/2023

ABOUT ME



MOST RELEVANT RESEARCH AREAS

Quantitative Analysis of Large Models

- Scalable techniques and tools for the modelling and analysis
- Massive many entities (agents, species, computing nodes…)
  - Possibly interacting over large-scale networks
- Boolean networks
  - PI of prestigious 3.5 years Danish grant ~365K EUR

Simulation/statistical analysis of large-scale models

- High level languages to ease/automate modelling and analysis
  - Domain-specific languages hiding loops, variables,…
  - Models automatically compiled down in simulators
  - One-click analysis
- MultiVeStA: Statistical Model Checking for any simulator
  - Product lines, Business process models, Threats models
  - Agent-based economic models …

PISA

END 2019 Tenure-track Assistant Professor @ Sant'Anna
2023-NOW Associate Professor @ Sant'Anna

COPENHAGEN

2019 Associate Professor @ DTU
2017-2019 Tenure-track Assistant Professor @ DTU
2019 Visiting professor @ IMT

LUCCA

2015-2017 Assistant Professor @ IMT

SOUTHAMPTON

2013-2015 Post-Doc @ University of Southampton

LUCCA

2010-2013 Phd in Computer Science and Engineering @ IMT
Six months @ University of Leicester

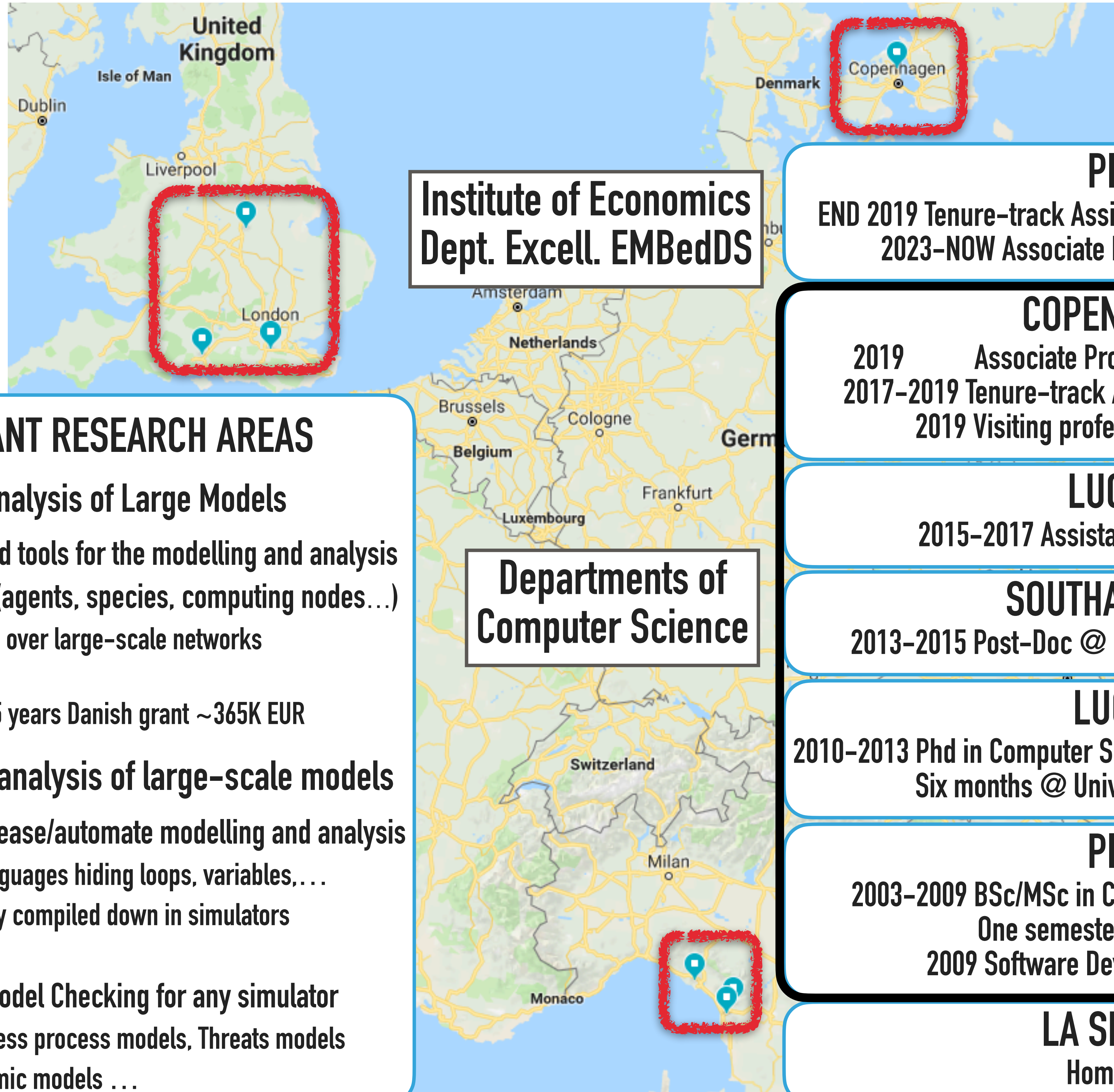
PISA

2003-2009 BSc/MSc in Computer Science @ UniPi
One semester@Queen Mary College London
2009 Software Developer @ ION Trading

LA SPEZIA

Hometown

ABOUT ME



Institute of Economics
Dept. Excell. EMBedDS

Departments of Computer Science

PISA

END 2019 Tenure-track Assistant Professor @ Sant'Anna
2023-NOW Associate Professor @ Sant'Anna

COPENHAGEN

2019 Associate Professor @ DTU
2017-2019 Tenure-track Assistant Professor @ DTU
2019 Visiting professor @ IMT

LUCCA

2015-2017 Assistant Professor @ IMT

SOUTHAMPTON

2013-2015 Post-Doc @ University of Southampton

LUCCA

2010-2013 Phd in Computer Science and Engineering @ IMT
Six months @ University of Leicester

PISA

2003-2009 BSc/MSc in Computer Science @ UniPi
One semester@Queen Mary College London
2009 Software Developer @ ION Trading

LA SPEZIA

Hometown

MOST RELEVANT RESEARCH AREAS

Quantitative Analysis of Large Models

- Scalable techniques and tools for the modelling and analysis
- Massive many entities (agents, species, computing nodes…)
  - Possibly interacting over large-scale networks
- Boolean networks
  - PI of prestigious 3.5 years Danish grant ~365K EUR

Simulation/statistical analysis of large-scale models

- High level languages to ease/automate modelling and analysis
  - Domain-specific languages hiding loops, variables,…
  - Models automatically compiled down in simulators
  - One-click analysis
- MultiVeStA: Statistical Model Checking for any simulator
  - Product lines, Business process models, Threats models
  - Agent-based economic models …

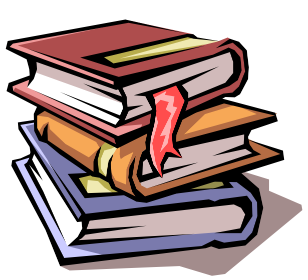
Large funding as 'Department of Excellence' 2018-2022 -> 2023-2027

- Led by Prof. Francesca Chiaromonte
  - Statistician, also working at PennState University, USA
- Involves a wide range of profiles: economics, management, law, statistics, computer science

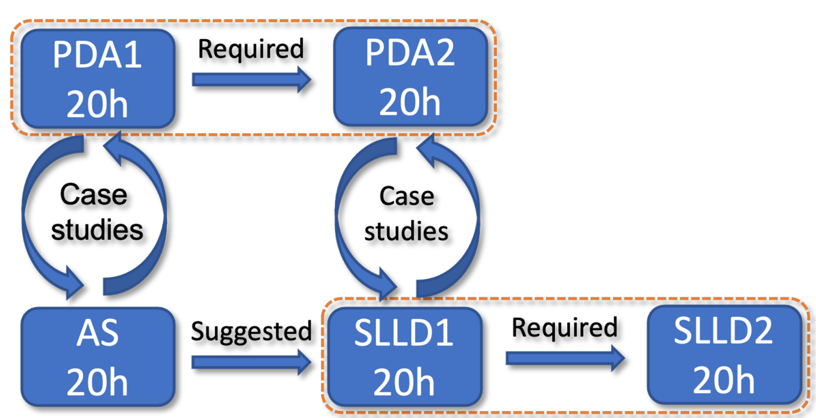
The mission of L'EMbeDS is to

- Foster data-driven statistical and computational approaches in the Social Sciences

Research



Teaching



2023-2027!

Sant'Anna School of Advanced Studies Pisa

Department of Excellence 2018 - 2022

EMbeDS

Economics and Management in the era of Data Science

ABOUT INSTITUTE OF ECONOMICS

Research Areas

Complexity Economics

The dynamics of industries and markets

Innovation, technical change and economic history

Data Science for economics and social science

Data Science for economics and social science

Statistical learning; analysis and methods for big & complex data; data mining; patterns of causality in economic data; statistical model checking; calibration and validation of economic models

Automated and Distributed Statistical Analysis of Economic Agent-Based Models

Andrea Vandin[a], Daniele Giachini[a], Francesco Lamperti[a,c], Francesca Chiaromonte[a,b]

[a] *Institute of Economics and EMbeDS, Sant'Anna School of Advanced Studies, Pisa, Italy.*
[b] *Dept. of Statistics and Huck Institutes of the Life Sciences, Penn State University, USA*
[c] *RFF-CMCC European Institute on Economics and the Environment, Milan, Italy.*

PUBLISHED AT HIGH-QUALITY ECONOMIC VENUE JEDC!

Abstract

We propose a novel approach to the statistical analysis of simulation models and, especially, agent-based models (ABMs). Our main goal is to provide a fully automated and model-independent tool-kit to inspect simulations and perform counter-factual analysis. Our approach: (i) is easy-to-use by the modeller, (ii) improves reproducibility of results, (iii) optimizes running time given the modeller's machine, (iv) automatically chooses the number of required simulations and simulation steps to reach user-specified statistical confidence, and (v) automatically performs a variety of statistical tests. In particular, our framework is designed to distinguish the transient dynamics of the model from its steady state behaviour (if any), estimate properties of the model in both "phases", and provide indications on the ergodic (or non-ergodic) nature of the simulated processes – which, in turns allows one to gauge the reliability of a steady state analysis. Estimates are equipped with statistical guarantees, allowing for robust comparisons across computational experiments. To demonstrate the effectiveness of our approach, we apply it to two models from the literature: a large scale macro-financial ABM and a small scale prediction market model. Compared to prior analyses of these models, we obtain new insights and we are able to identify and fix some erroneous conclusions.

Keywords: ABM, Statistical Model Checking, Ergodicity analysis, Steady state analysis, Transient analysis, Warmup estimation, T-test and power, Prediction markets, Macro ABM

AUTOMATED and DISTRIBUTED STATISTICAL ANALYSIS of ECONOMIC AGENT-BASED MODELS

Andrea Vandin

Sant'Anna School of Advanced Studies – Pisa

Institute of Economics

Department of Excellence 2018 - 2022

EMbeDS Economics and Management in the era of Data Science

Joint work with

Daniele Giachini, Francesco Lamperti, Francesca Chiaromonte

Paper, tool and models available at:
https://bit.ly/MultiVeStATool

Classes 21t 22t, Software Validation and Verification, Unipi, 04-05/12/2023

OUTLINE

1. Motivation, vision, and proposal
    1. Automated analysis with statistical guarantees for ABMs
    2. The MultiVeStA Statistical Model Checker
2. Transient Analysis of a large-scale financial macro ABM
    1. Estimation of expected outcome and Confidence Interval
    2. Counterfactual analysis for different model configurations
3. Steady-state analysis of a prediction market model
    1. Steady-state analysis by Replication and Deletion (RD)
    2. Warmup estimation
    3. Steady-state analysis by Batch Means (BM)
    4. A methodology for ergodicity analysis based on RD and BM
4. Conclusions & Future works

OUTLINE

1. **Motivation, vision, and proposal**
    1. Automated analysis with statistical guarantees for ABMs
    2. The MultiVeStA Statistical Model Checker
2. Transient Analysis of a large-scale financial macro ABM
    1. Estimation of expected outcome and Confidence Interval
    2. Counterfactual analysis for different model configurations
3. Steady-state analysis of a prediction market model
    1. Steady-state analysis by Replication and Deletion (RD)
    2. Warmup estimation
    3. Steady-state analysis by Batch Means (BM)
    4. A methodology for ergodicity analysis based on RD and BM
4. Conclusions & Future works

What is an Economic Agent-Based Model?

NON ABMs

- ‘Mainstream economists’ tend to reason in terms of models that
  - Are given as a unique monolithic model
    - Do not have focus on their single components, but on the overall dynamics of the model
    - *What* the system does, rather than *how* the system does
  - Have explicit representations of the laws governing the economic system
  - **Can be analysed analytically**

ABMs

- Some economists are getting interested in modeling an economic system in terms of its components
  - The **agents** that operate in it: firms, households, banks…
- The modeller does specify explicitly the laws governing the model.
- It describes explicitly
  - The behaviour of every agent
  - The interactions among the agents
  - **The laws governing the model then *emerge* from these behaviours and interactions**
- These types of models are often denoted as ABMs.
  - These are typically too difficult to be solved analytically
  - We need to do simulations
    - My message: **we need to do simulations well!**
    - A variant of model checking, statistical model checking, can help on this

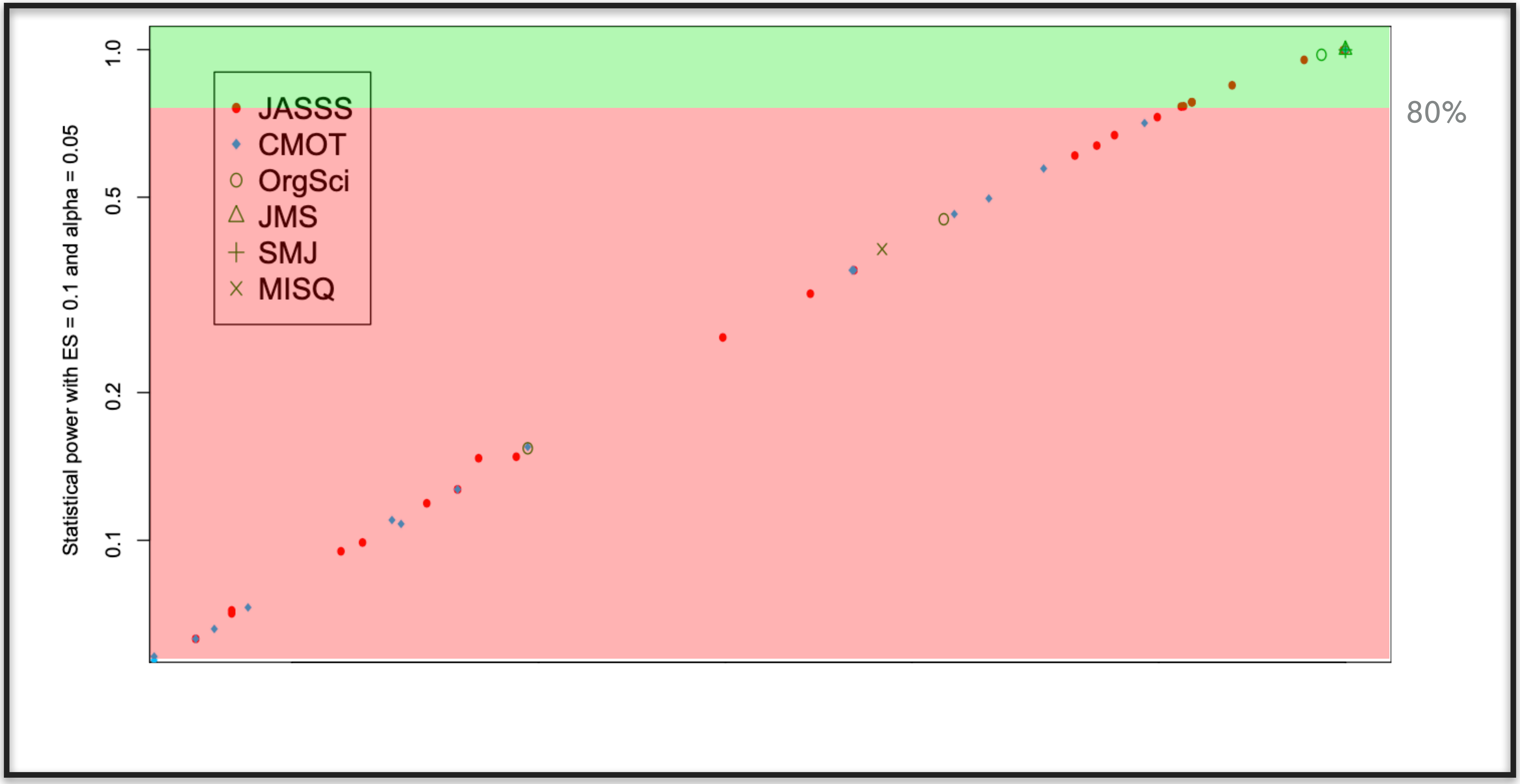
What is SMC?

The Model Checking problems

- **Model Checking (MC)**:
  - To decide whether a **non-deterministic** model satisfies a temporal logic property
- **Probabilistic MC (PMC)**:
  - To decide whether a **stochastic** model satisfies a temporal logic property with a probability greater than a certain threshold
- **Statistical MC (SMC):**
  - Simulation-based technique to **statistically approximate the PMC problem**
  - Only requires independent and identically distributed samplings (simulations)
    - Highly parallelizable
  - Many tools supporting it. E.g.
    - MultiVeStA, PRISM, UPPAAL, APMC, COSMOS, YMER, SAM, BIP,(P)VeStA…
  - Two main approaches: Probability estimation VS Hypothesis testing
    - Probability estimation → Real-valued property estimation

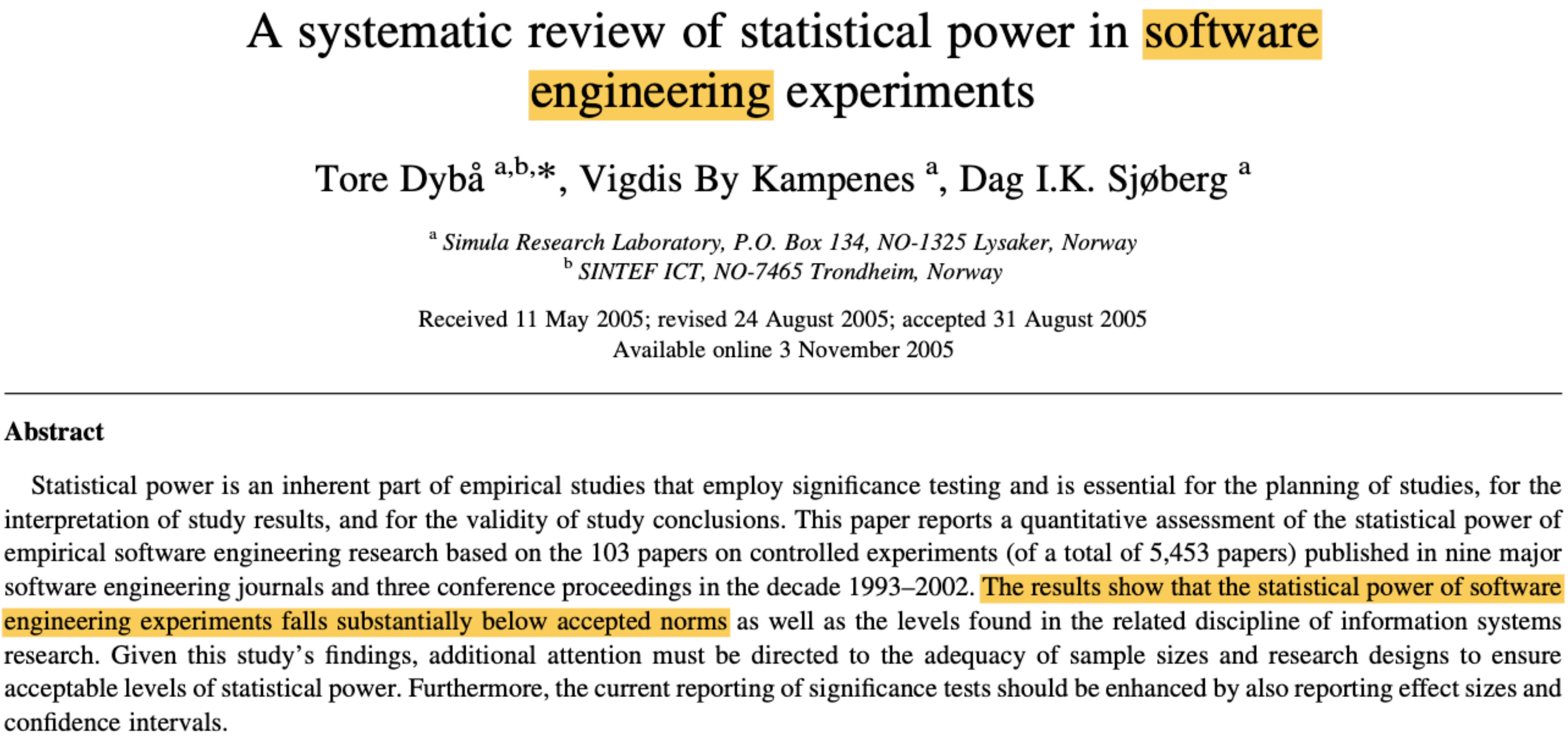
‘Quality’ of Statistical Analysis on 55 ABM from Management & Organisational Research

Adapted from Secchi, Seri, Computational and Mathematical Organization Theory, 2017



- The importance of designing well simulation-based analysis.
  - Power analysis on ‘are the expected outcomes of different configurations of parameters the same’?
- Power is 1 - P(Type II error)
  - Roughly, P(test=‘outcomes are different’ | outcomes are different)
  - “The value that seems to be more commonly accepted is 80%”
- “We need to encourage researchers to be more precise in the determination of the number of runs”

Similar studies can be found also in other communities



A systematic review of statistical power in software engineering experiments

Tore Dybå a,b,*, Vigdis By Kampenes a, Dag I.K. Sjøberg a

a Simula Research Laboratory, P.O. Box 134, NO-1325 Lysaker, Norway
b SINTEF ICT, NO-7465 Trondheim, Norway

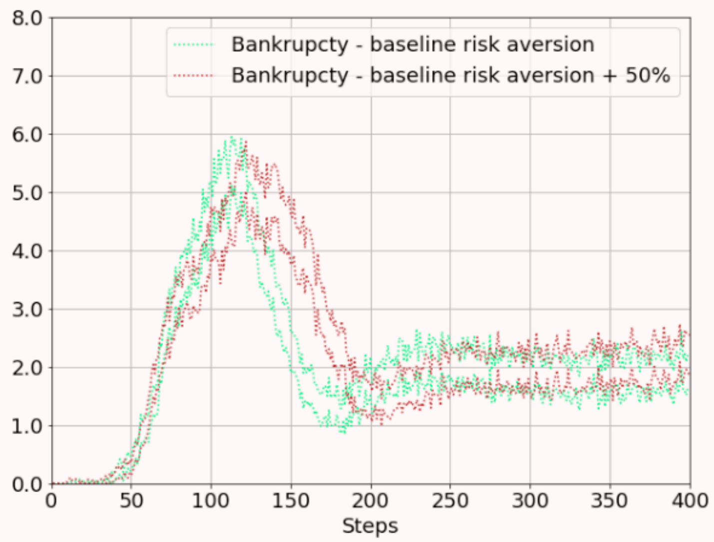
Received 11 May 2005; revised 24 August 2005; accepted 31 August 2005
Available online 3 November 2005

Abstract

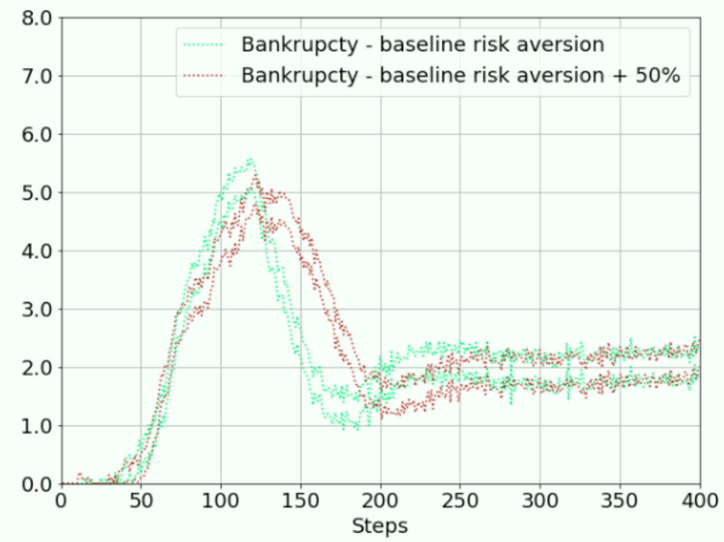
Statistical power is an inherent part of empirical studies that employ significance testing and is essential for the planning of studies, for the interpretation of study results, and for the validity of study conclusions. This paper reports a quantitative assessment of the statistical power of empirical software engineering research based on the 103 papers on controlled experiments (of a total of 5,453 papers) published in nine major software engineering journals and three conference proceedings in the decade 1993–2002. The results show that the statistical power of software engineering experiments falls substantially below accepted norms as well as the levels found in the related discipline of information systems research. Given this study's findings, additional attention must be directed to the adequacy of sample sizes and research designs to ensure acceptable levels of statistical power. Furthermore, the current reporting of significance tests should be enhanced by also reporting effect sizes and confidence intervals.

The Class in 3 Slides: Statistically Meaningful Counterfactual Analysis

97.5% CI
100 Simulations



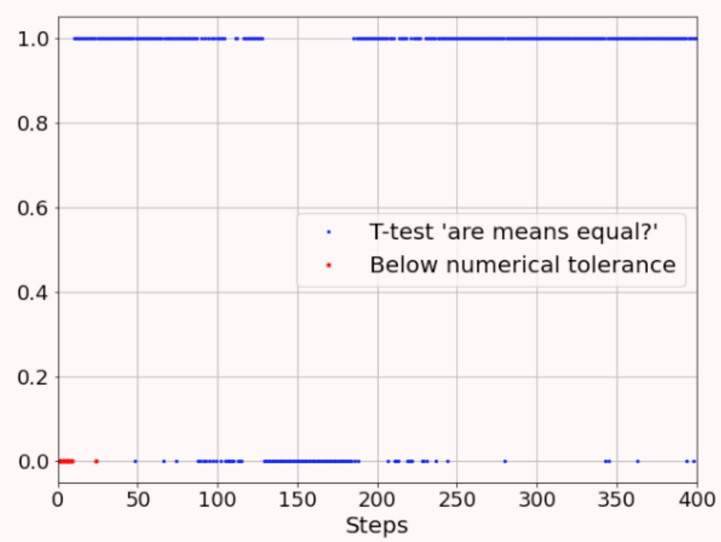
(a) CIs width for $\alpha = 0.025$ and $N = 100$ simulations



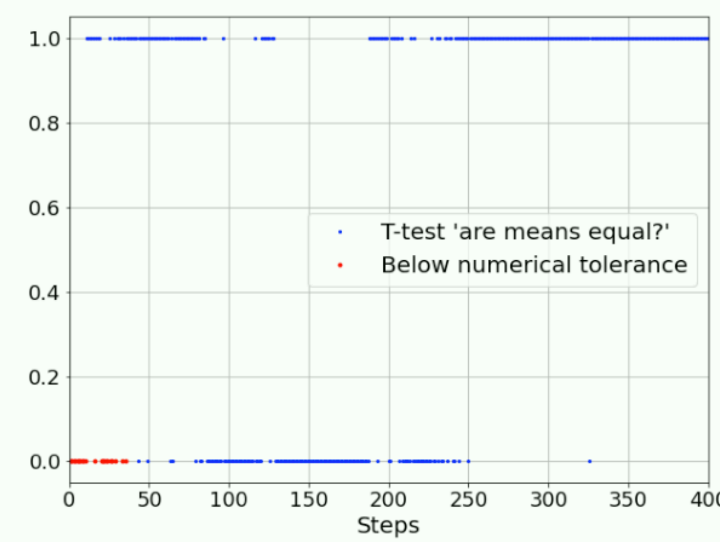
(b) CIs width for $\alpha = 0.025$ and $\delta = 0.5$

97.5% CI
MultiVeStA
'**Right**' number of simulations

Welch's t-test



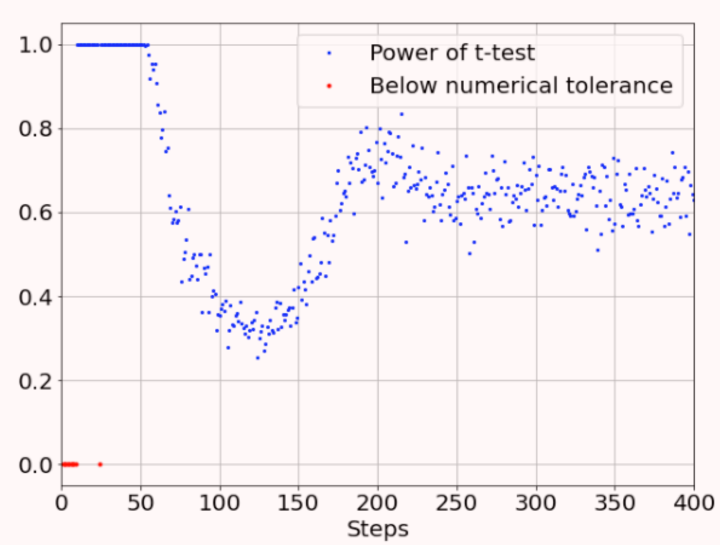
(c) T-test *are means in (a) point-wise equal* for significance $\alpha = 0.025$



(d) T-test *are means in (b) point-wise equal* for significance $\alpha = 0.025$

Welch's t-test

Power of the test
P(Test=0 | Real=0)



(e) Power of t-test in (c) for difference $\varepsilon = 0.5$

(f) Power of t-test in (d) for difference $\varepsilon = 0.5$

Power of the test
P(Test=0 | Real=0)

The Class in 3 Slides: Steady-State Analysis: Market Selection

Arbitrary choice of
- Number of sims
- Warmup period
- Time horizon

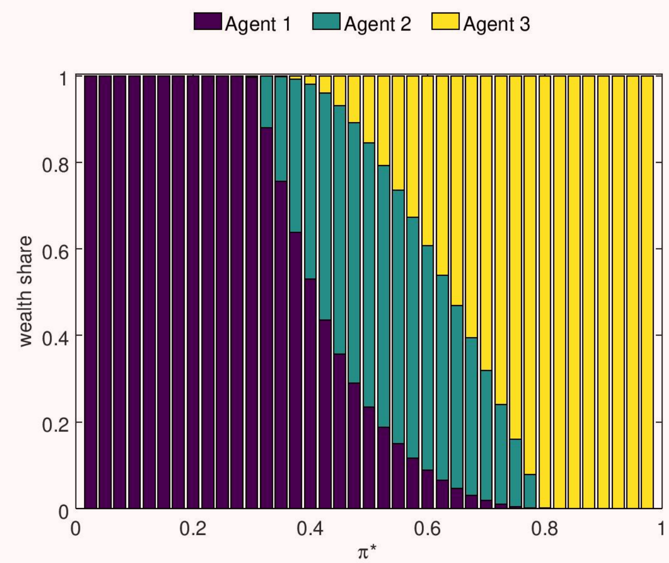
from [Kets et al2014]

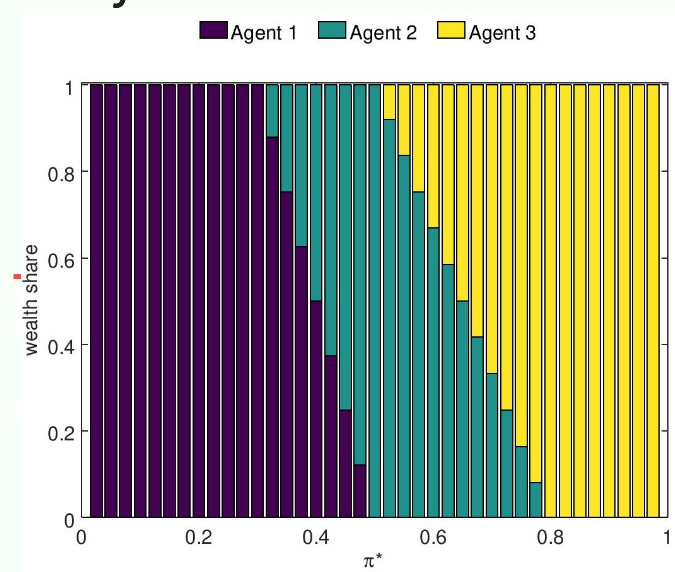
Automated choice of
- Number of sims
- Warmup period
- Time horizon

MultiVeStA

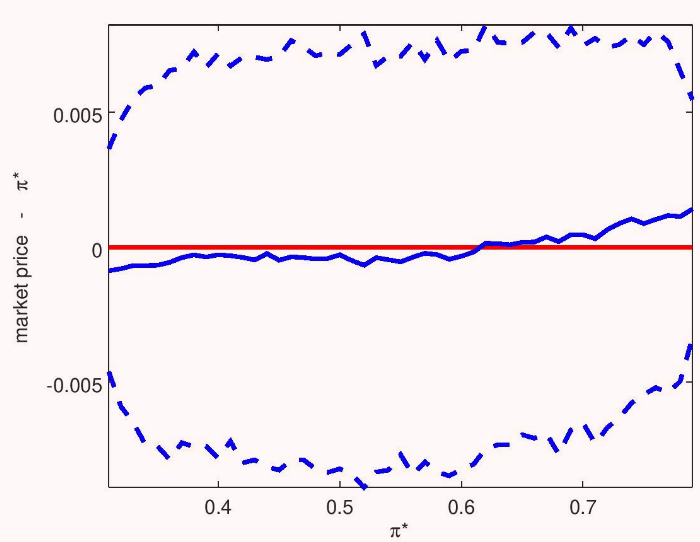
Same as analytical solution
from [Bottazzi,Giachini2019]

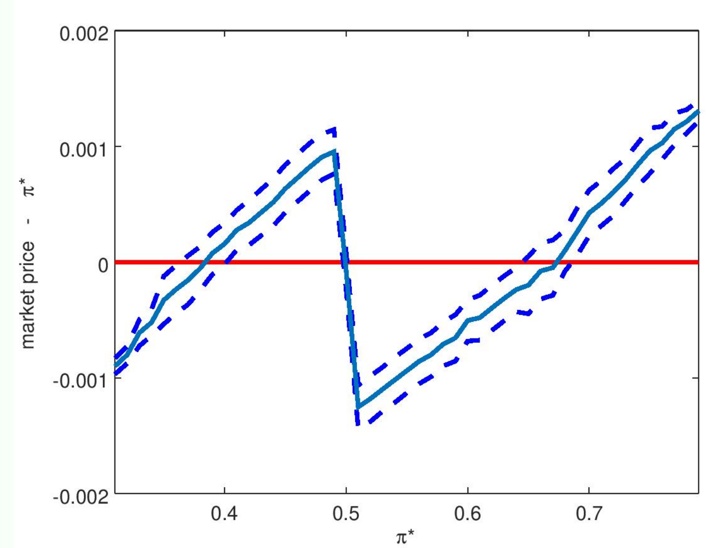
Agents wealth at steady state



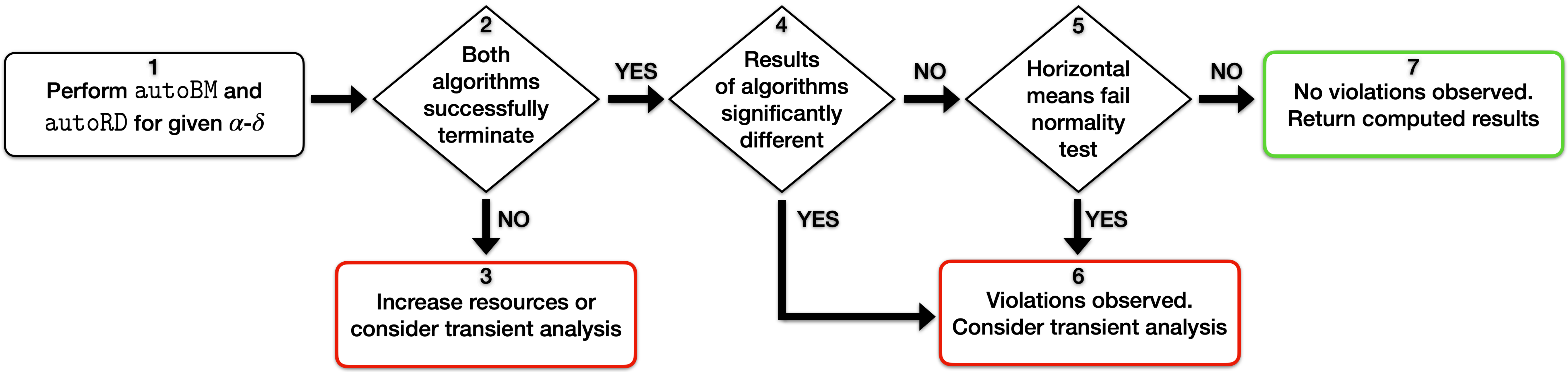


Does the market price match π^*?



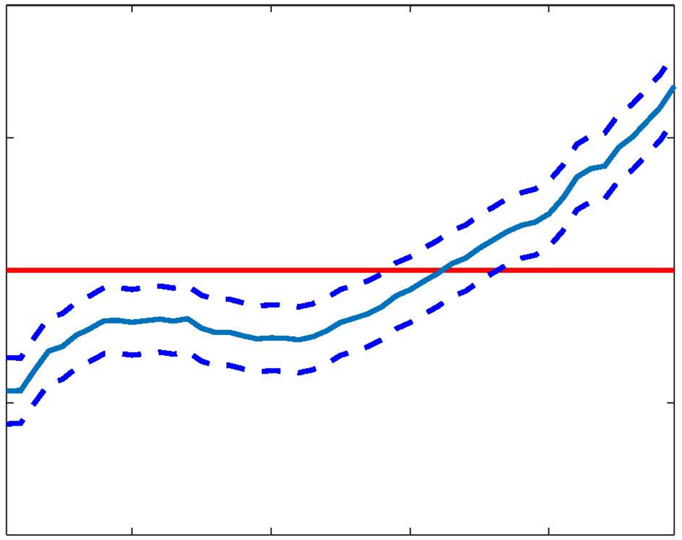


The Class in 3 Slides: a Methodology for Ergodicity Diagnostics



Our Proposed Approach to Simulation-Based Analysis

newstalkzb.co.nz/news/education/modern-lego-sets-more-complex-less-inspiring/



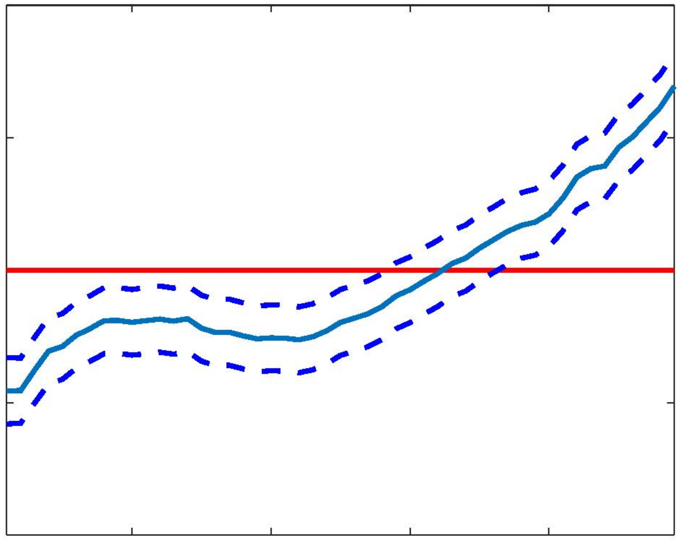
Our Proposed Approach to Simulation-Based Analysis

newstalkzb.co.nz/news/education/modern-lego-sets-more-complex-less-inspiring/

https://www.alamy.com/

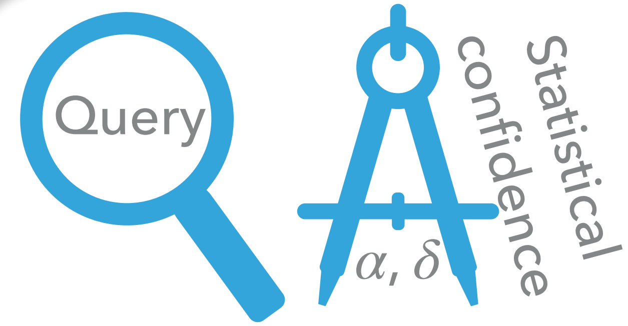
Handcrafted

- Mainly manual process
  - Time-consuming
  - Problems with replicability
  - Error-prone
    - Modify model, interpret CSV
- Ad-hoc implementations
  - Reliability? Efficiency?



Our Proposed Approach to Simulation-Based Analysis

newstalkzb.co.nz/news/education/modern-lego-sets-more-complex-less-inspiring/



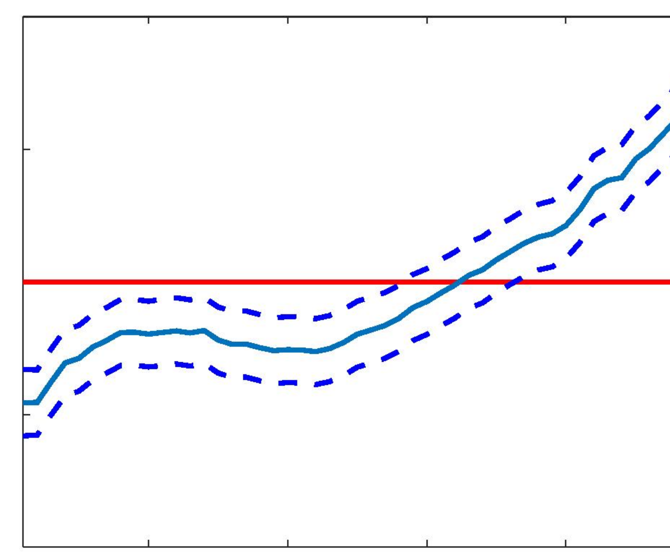
Statistical Model Checking

- Automatic
  - Time-saving and Reproducible
  - Promotes use of *standard* analysis
- Reference implementation
  - Reliable and Efficient

https://www.alamy.com/

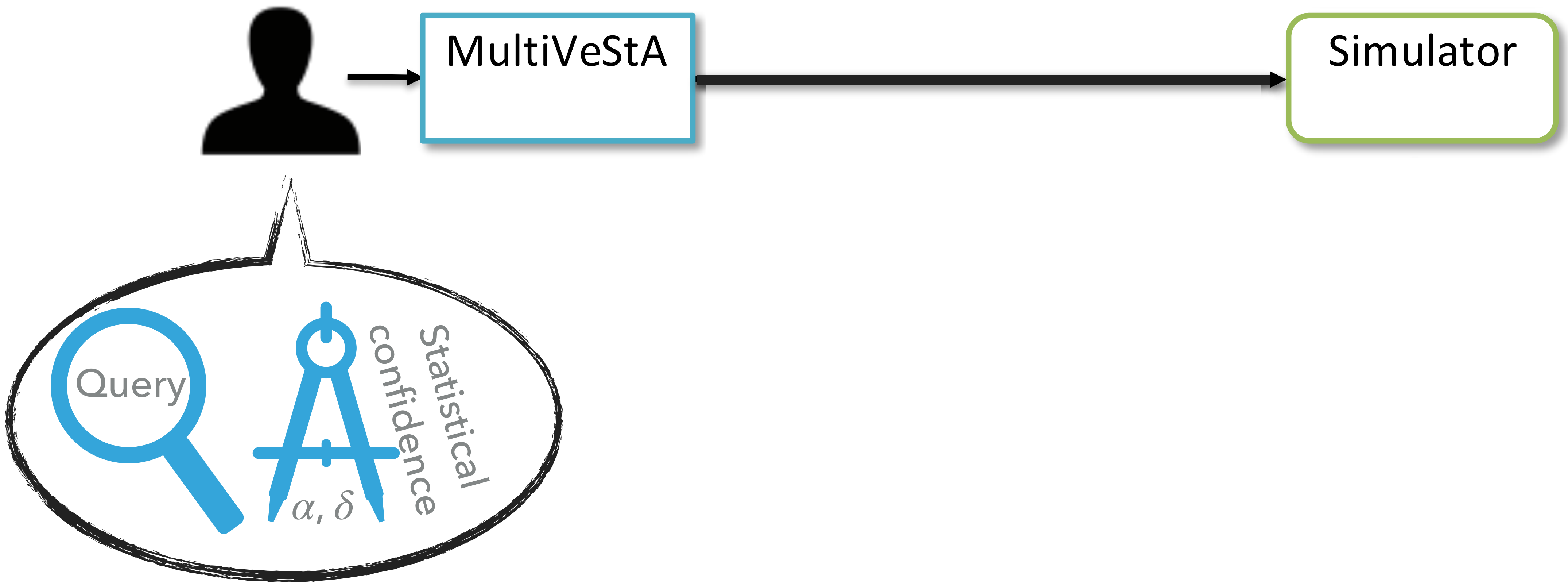
Handcrafted

- Mainly manual process
  - Time-consuming
  - Problems with replicability
  - Error-prone
    - Modify model, interpret CSV
- Ad-hoc implementations
  - Reliability? Efficiency?

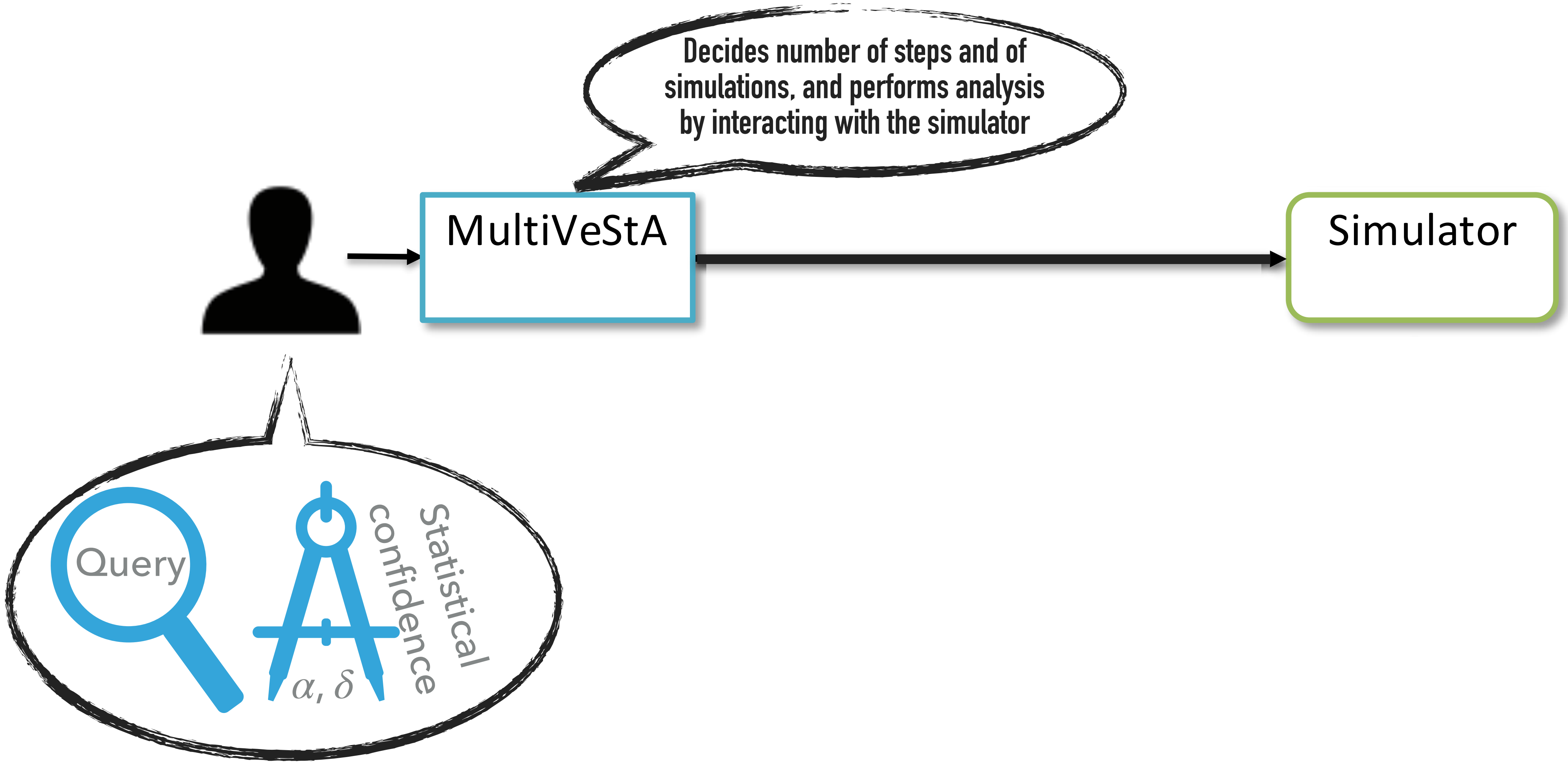


MultiVeStA: SMC For Discrete-Event Simulators

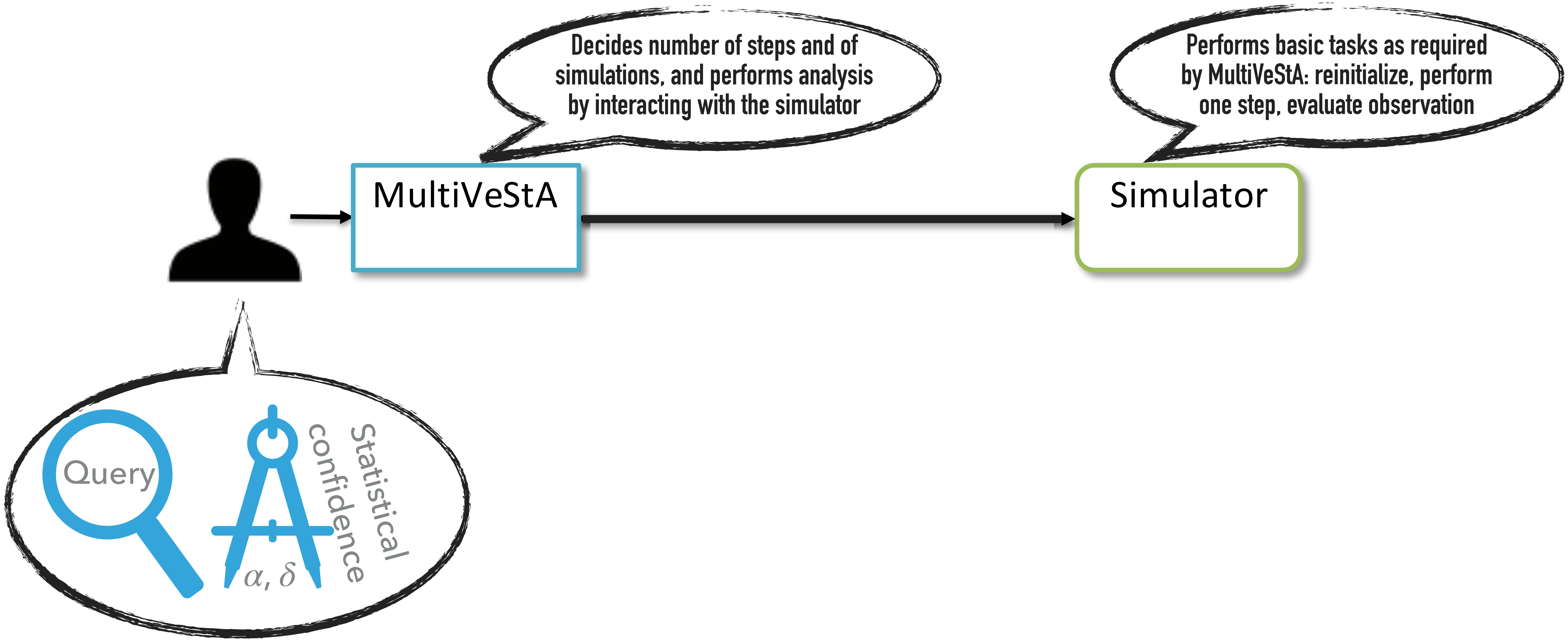
MultiVeStA: SMC For Discrete-Event Simulators



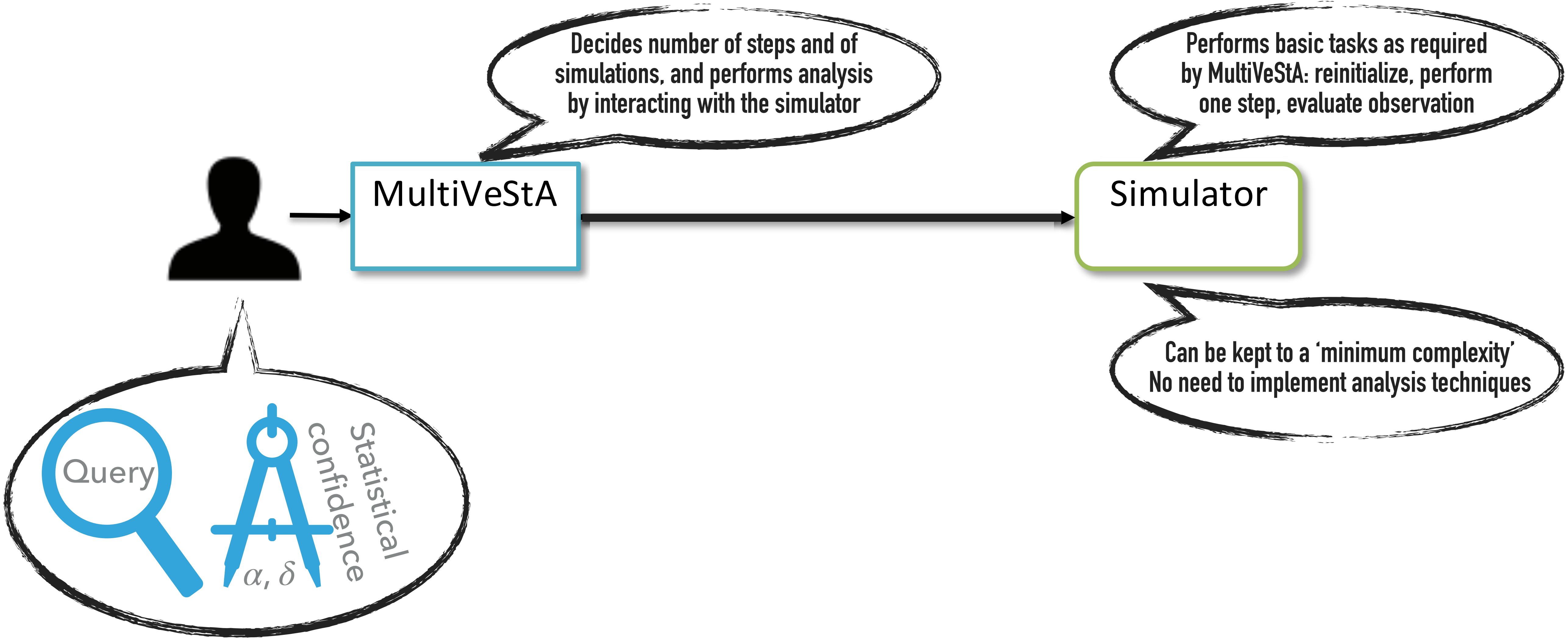
MultiVeStA: SMC For Discrete-Event Simulators



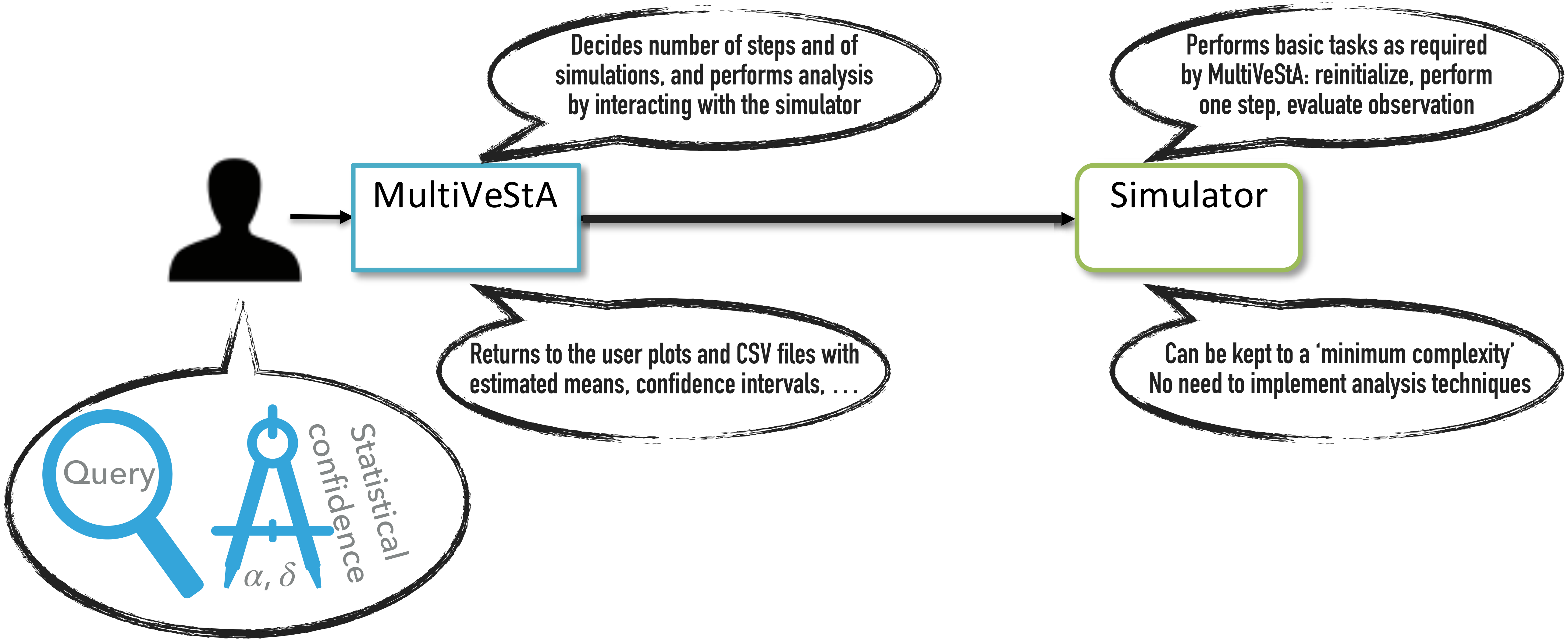
MultiVeStA: SMC For Discrete-Event Simulators



MultiVeStA: SMC For Discrete-Event Simulators

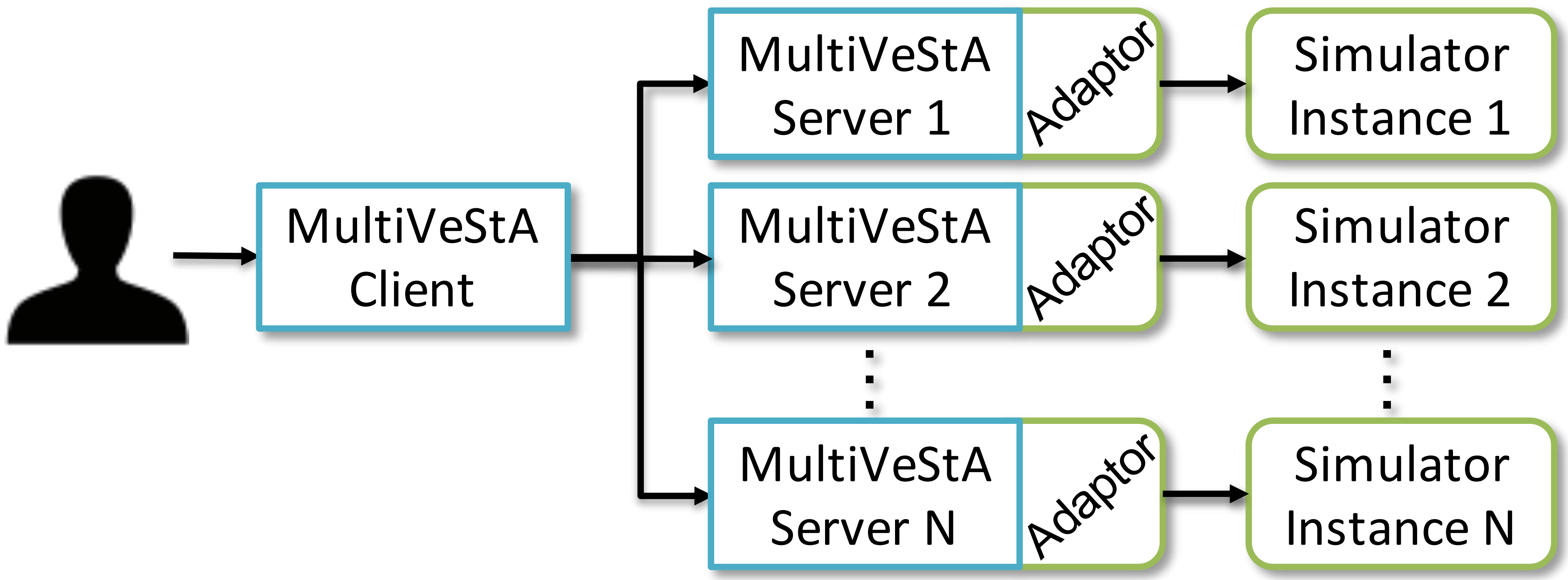


MultiVeStA: SMC For Discrete-Event Simulators

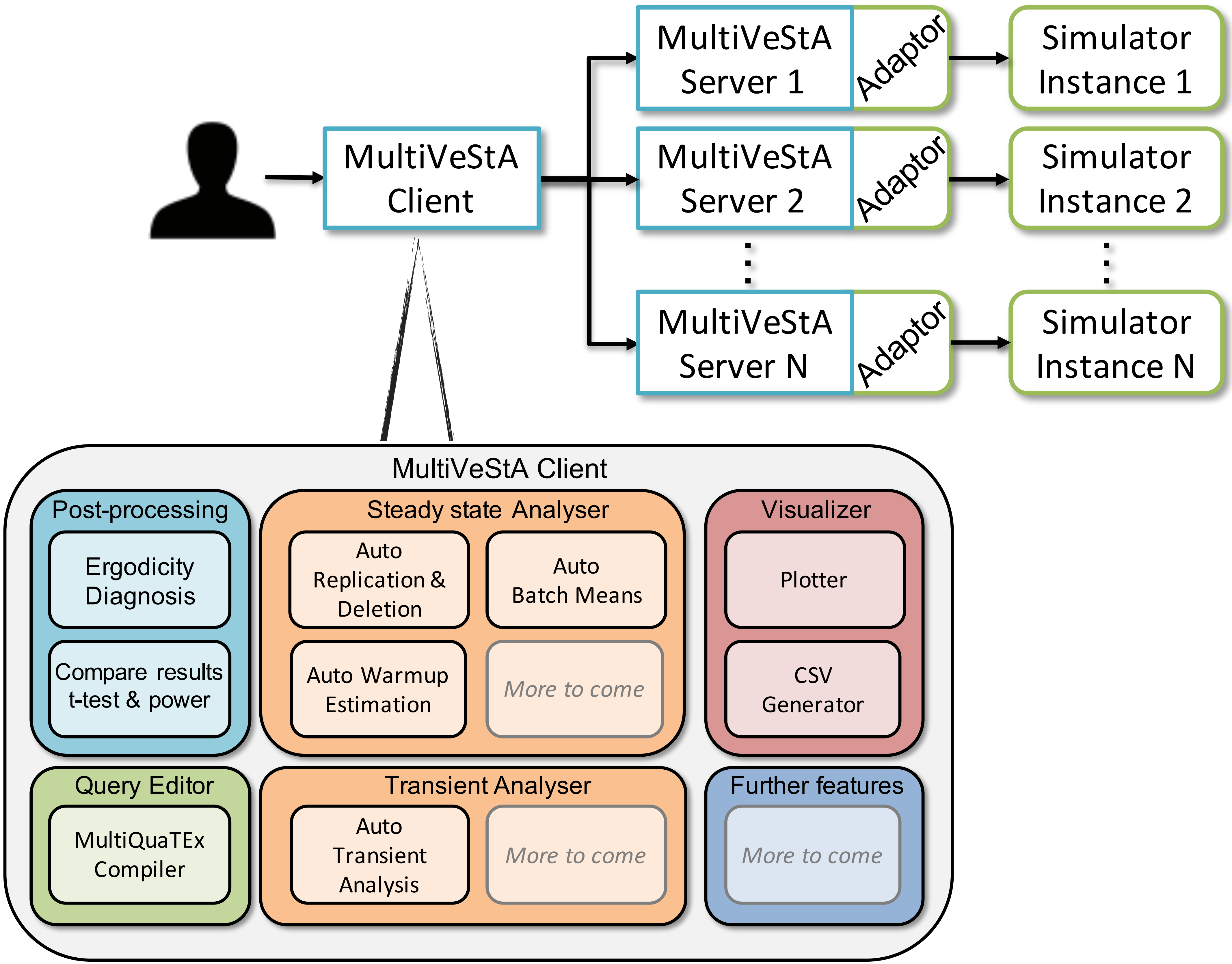


MultiVeStA: SMC For Discrete-Event Simulators

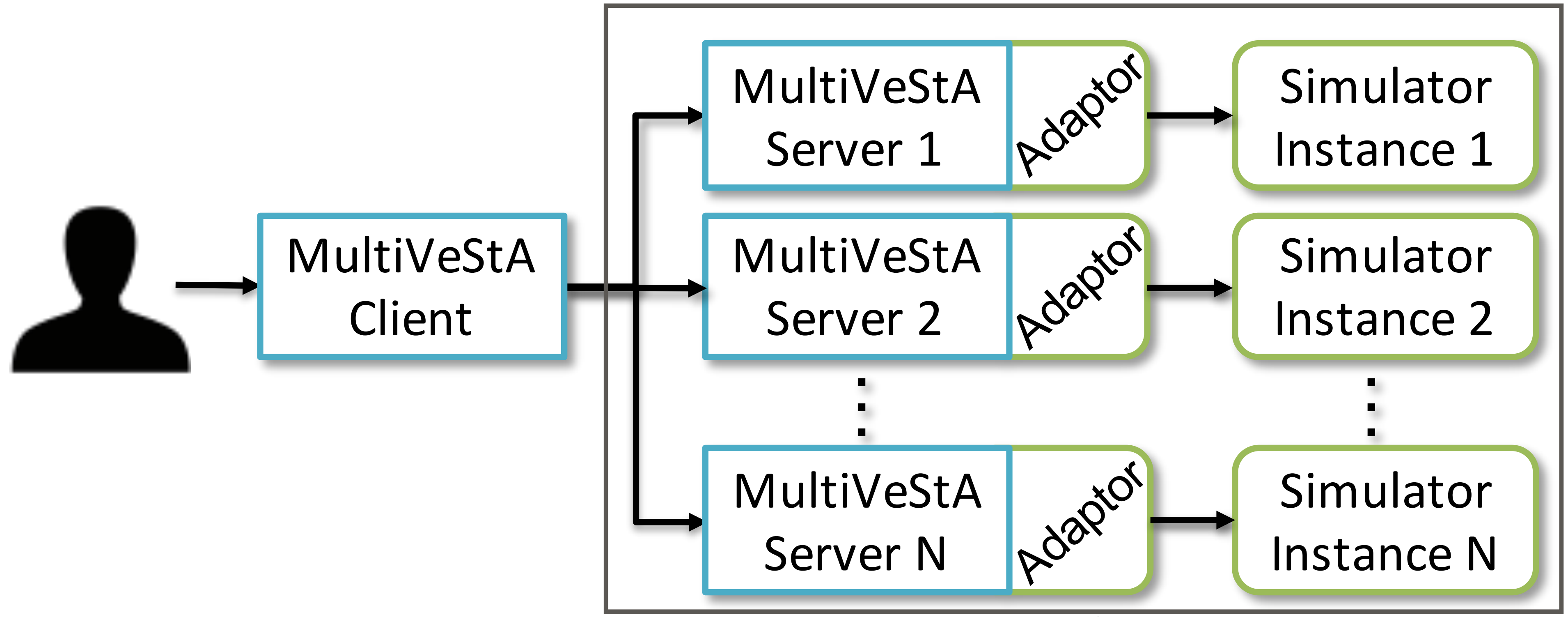
MultiVeStA: SMC For Discrete-Event Simulators

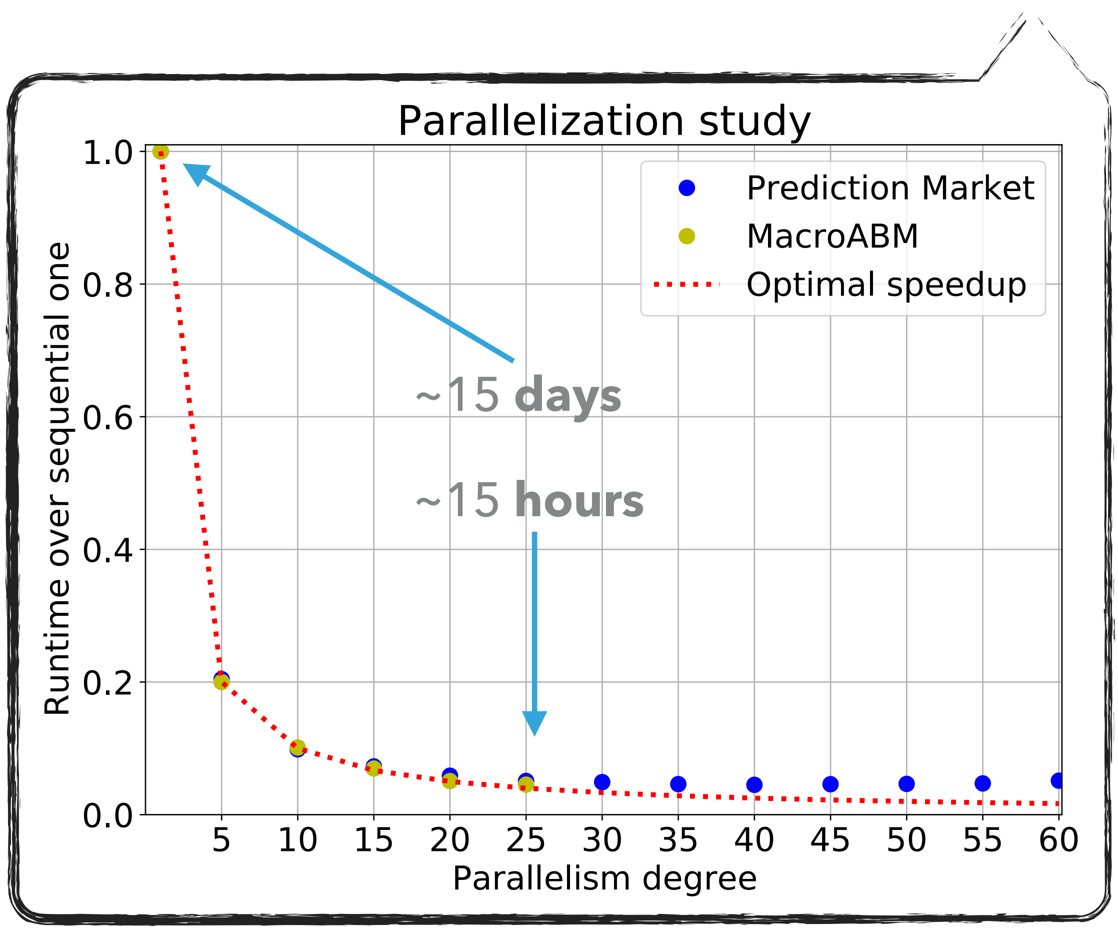


MultiVeStA: SMC For Discrete-Event Simulators

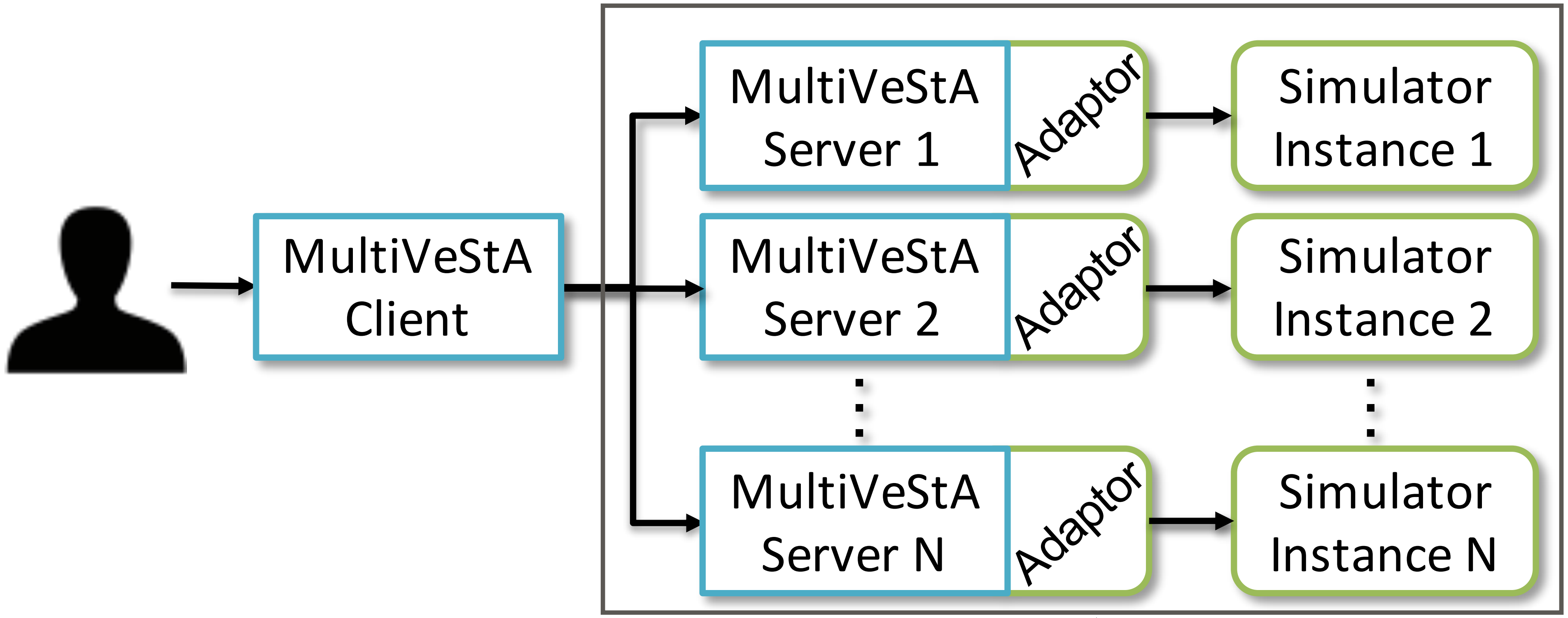


MultiVeStA: SMC For Discrete-Event Simulators

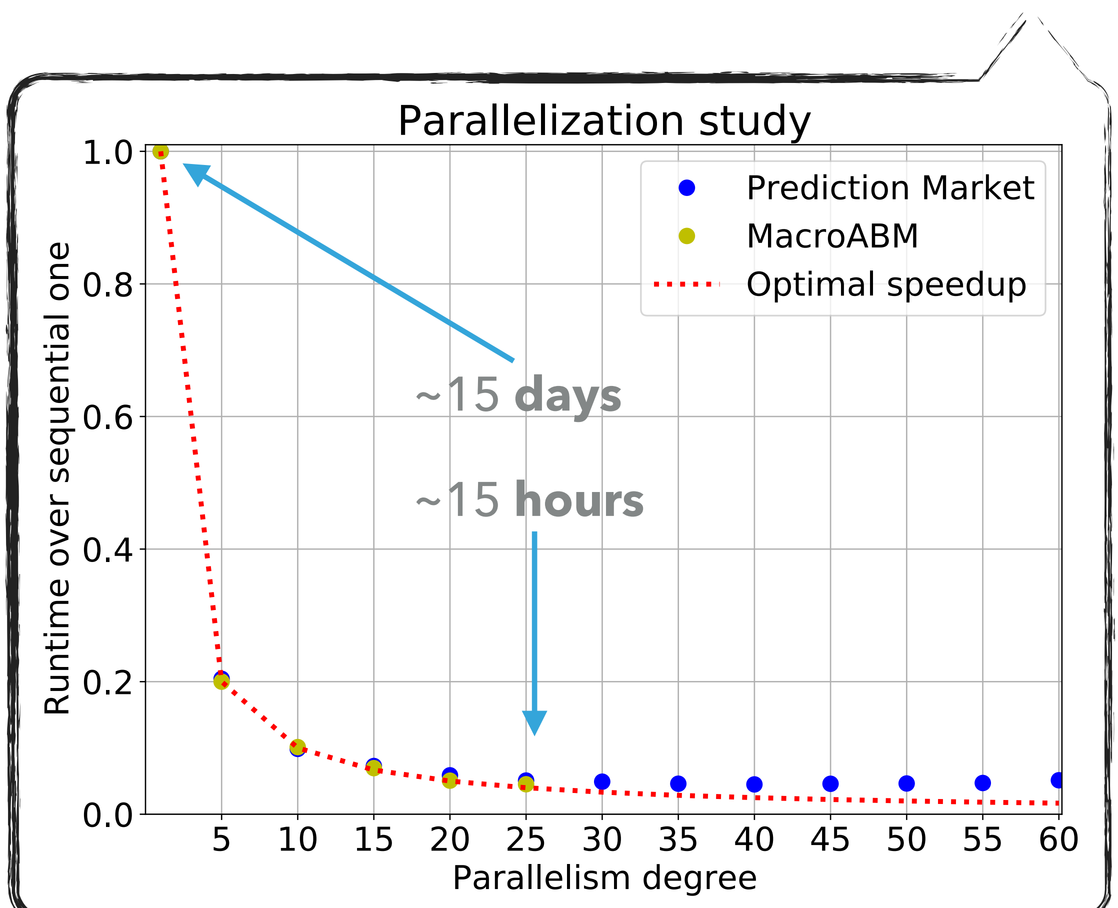




MultiVeStA: SMC For Discrete-Event Simulators



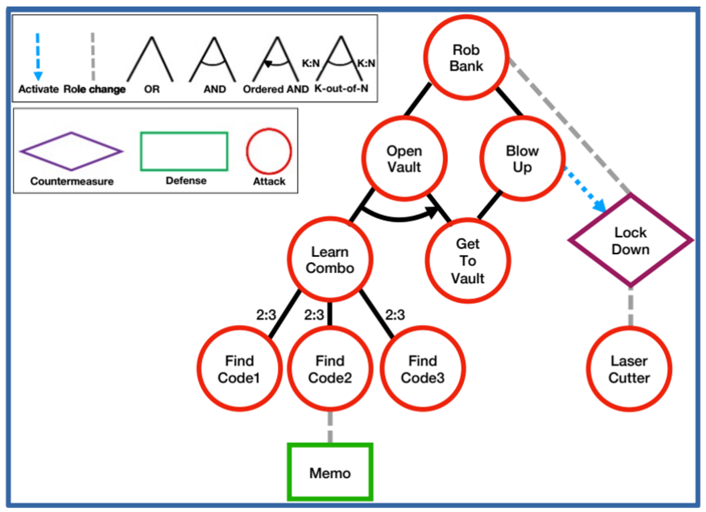
PROJECTS AVAILABLE!



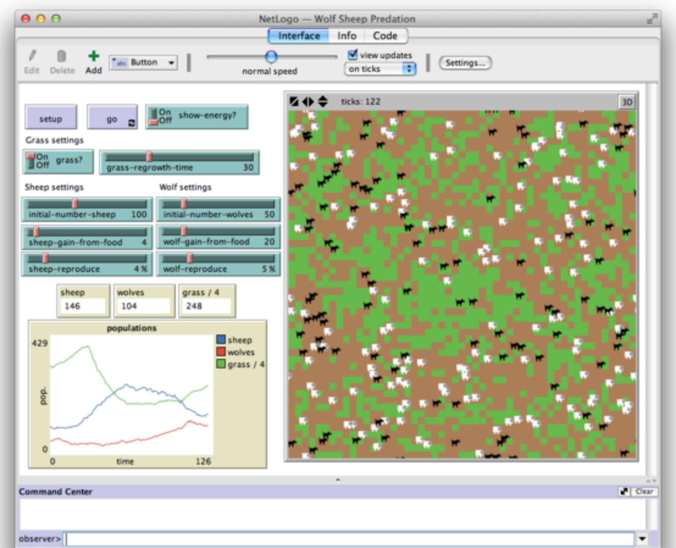
Java · R · NetLogo · python · C++ · JMAB JABM · MAUDE

Would you like to join the MultiVeStA family?

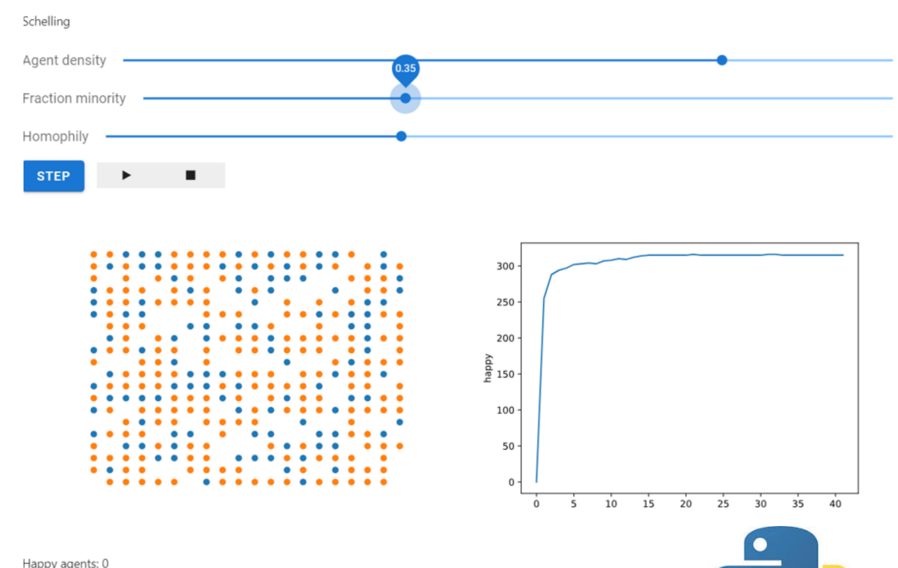
- **Projects available**
  - As an exam for this course
  - As starting points for Master projects?
  - As starting points for longer collaborations!?



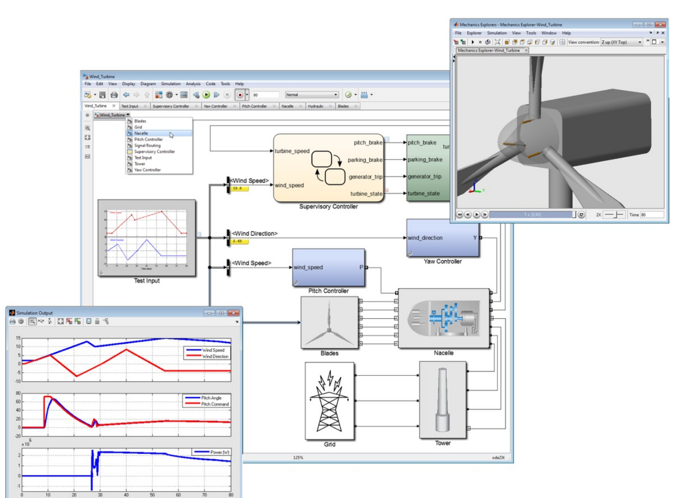
RisQFLan - Security



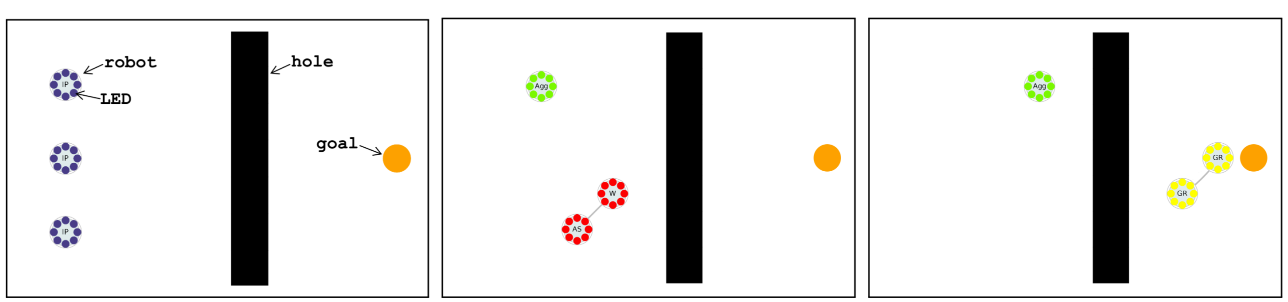
NetLogo multi-agent modeling
millions of students/teachers/researchers



Mesa: ABM in Python



Matlab Simulink



Maude - rewriting logic

More…

OUTLINE

1. Motivation, vision, and proposal
   1. Automated analysis with statistical guarantees for ABMs
   2. The MultiVeStA Statistical Model Checker
2. **Transient Analysis of a large-scale financial macro ABM**
   1. Estimation of expected outcome and Confidence Interval
   2. Counterfactual analysis for different model configurations
3. Steady-state analysis of a prediction market model
   1. Steady-state analysis by Replication and Deletion (RD)
   2. Warmup estimation
   3. Steady-state analysis by Batch Means (BM)
   4. A methodology for ergodicity analysis based on RD and BM
4. Conclusions & Future works

Transient Analysis by autoIR: Large Macro ABM

Large-scale macro financial ABM from Caiani et al, JEDC, 2016

- An economy with households, consumption/capital firms, commercial banks, government, central bank
- Thousands of agents
- Implemented in JMAB: Java framework for macro stock-flow consistent ABM models.
  - Side product: any model implemented in JMAB is now natively integrated with MultiVeStA

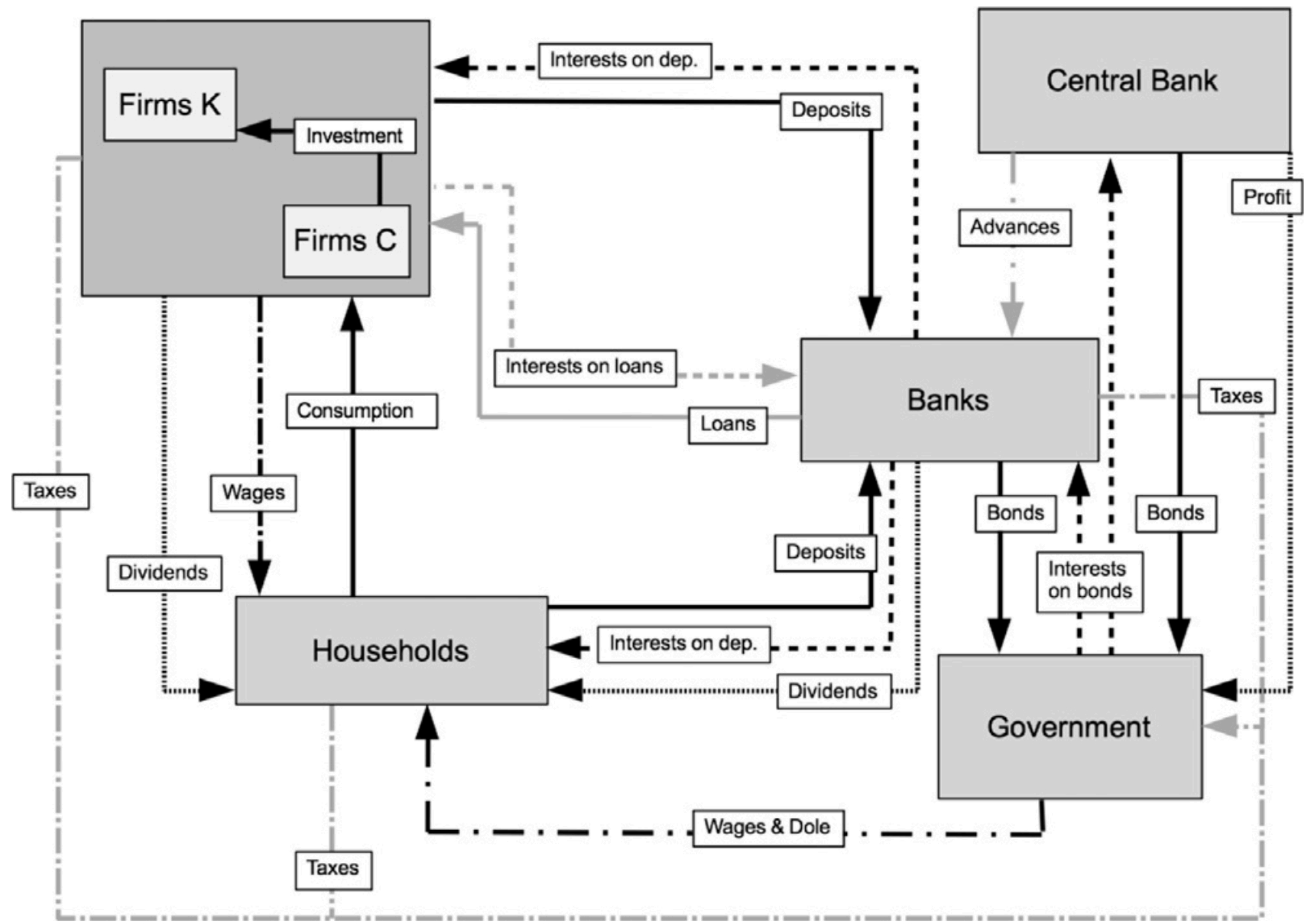


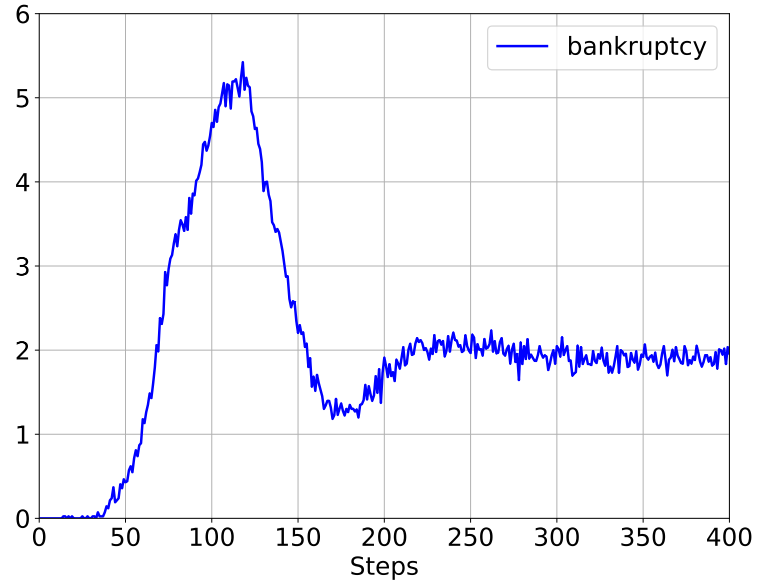
Fig. 1. Flow diagram of the model. Arrows point from paying sectors to receiving sectors.

Transient Analysis by autoIR: Large Macro ABM

Large-scale macro financial ABM from Caiani et al, JEDC, 2016

- An economy with households, consumption/capital firms, commercial banks, government, central bank
- Thousands of agents
- Implemented in JMAB: Java framework for macro stock-flow consistent ABM models.
  - Side product: any model implemented in JMAB is now natively integrated with MultiVeStA

Evolution of bankruptcies



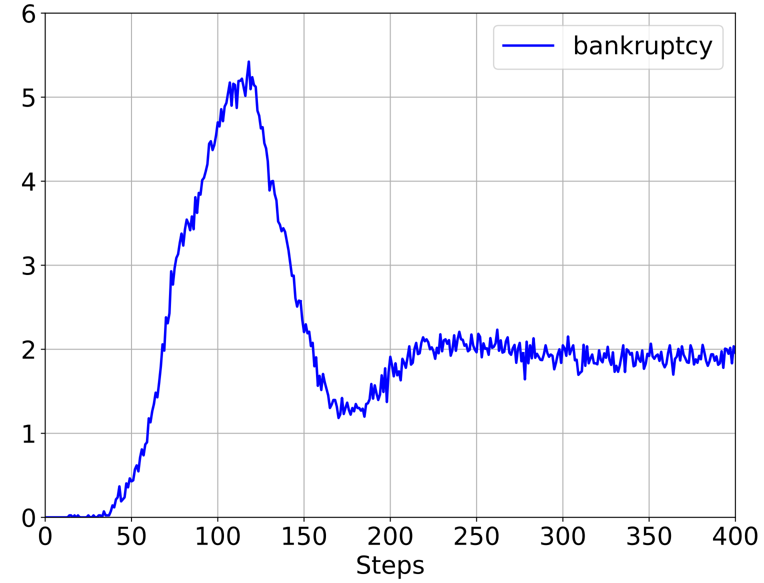
Transient Analysis by autoIR: Large Macro ABM

Large-scale macro financial ABM from Caiani et al, JEDC, 2016

- An economy with households, consumption/capital firms, commercial banks, government, central bank
- Thousands of agents
- Implemented in JMAB: Java framework for macro stock-flow consistent ABM models.
  - Side product: any model implemented in JMAB is now natively integrated with MultiVeStA

Evolution of bankruptcies

y_1

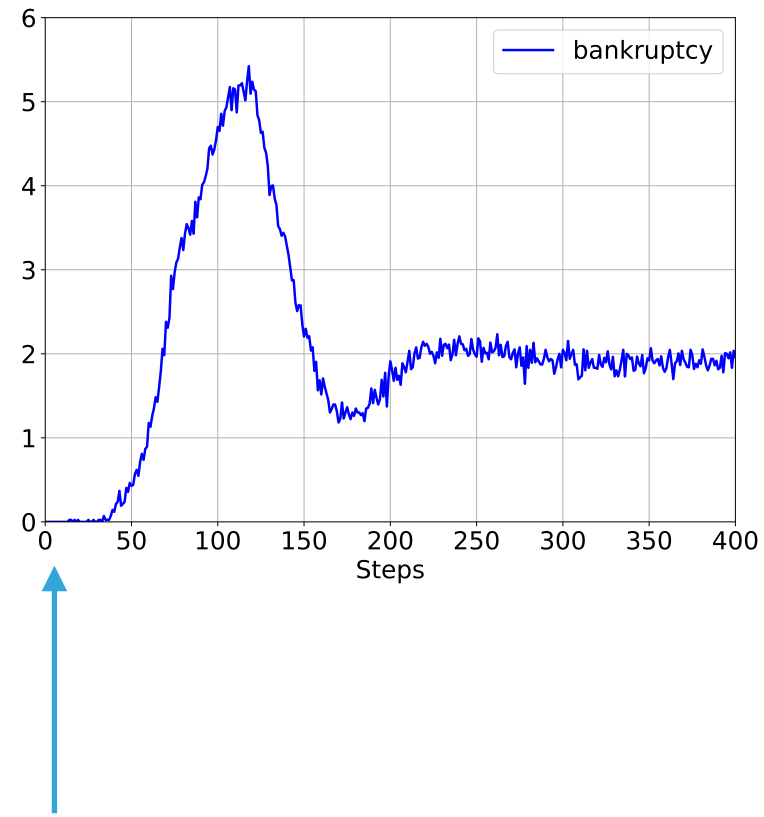


Transient Analysis by autoIR: Large Macro ABM

Large-scale macro financial ABM from Caiani et al, JEDC, 2016

- An economy with households, consumption/capital firms, commercial banks, government, central bank
- Thousands of agents
- Implemented in JMAB: Java framework for macro stock-flow consistent ABM models.
  - Side product: any model implemented in JMAB is now natively integrated with MultiVeStA

Evolution of bankruptcies



y_1

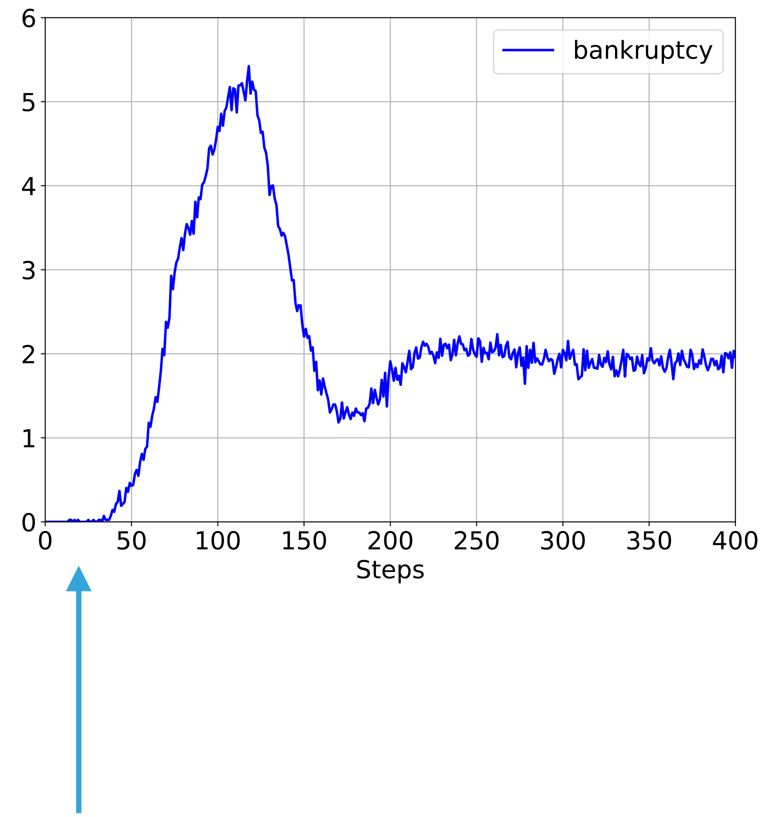
Transient Analysis by autoIR: Large Macro ABM

Large-scale macro financial ABM from Caiani et al, JEDC, 2016

- An economy with households, consumption/capital firms, commercial banks, government, central bank
- Thousands of agents
- Implemented in JMAB: Java framework for macro stock-flow consistent ABM models.
  - Side product: any model implemented in JMAB is now natively integrated with MultiVeStA

$y_1 \quad y_2$

Evolution of bankruptcies



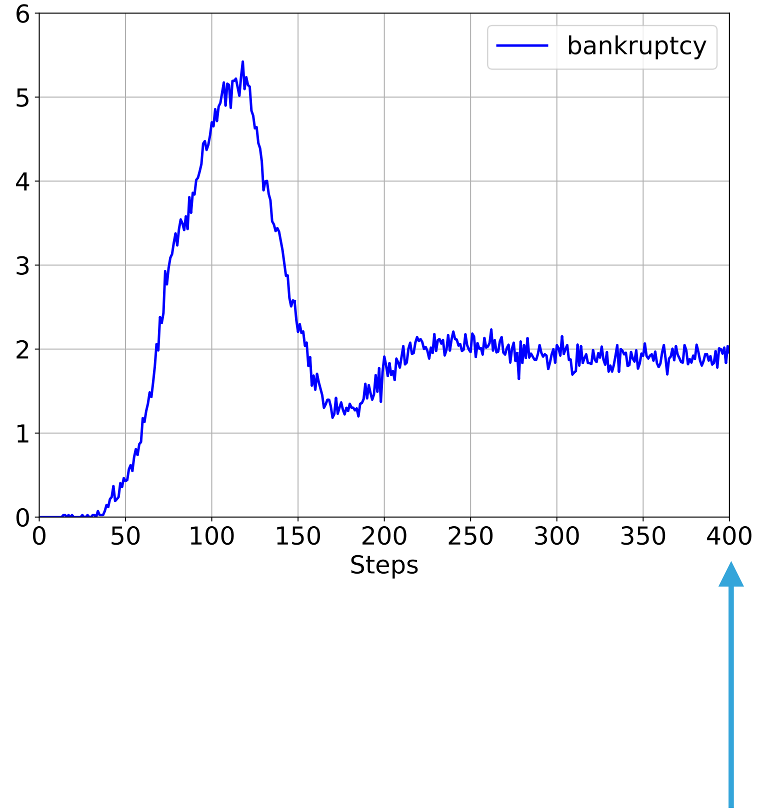
Transient Analysis by autoIR: Large Macro ABM

Large-scale macro financial ABM from Caiani et al, JEDC, 2016

- An economy with households, consumption/capital firms, commercial banks, government, central bank
- Thousands of agents
- Implemented in JMAB: Java framework for macro stock-flow consistent ABM models.
  - Side product: any model implemented in JMAB is now natively integrated with MultiVeStA

$y_1 \quad y_2 \quad \dots \quad y_{400}$

Evolution of bankruptcies



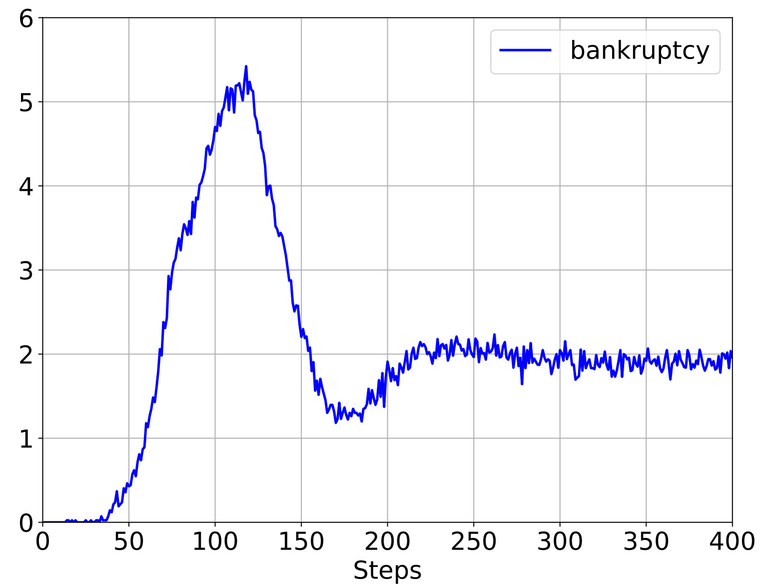
Transient Analysis by autoIR: Large Macro ABM

Large-scale macro financial ABM from Caiani et al, JEDC, 2016

- An economy with households, consumption/capital firms, commercial banks, government, central bank
- Thousands of agents
- Implemented in JMAB: Java framework for macro stock-flow consistent ABM models.
  - Side product: any model implemented in JMAB is now natively integrated with MultiVeStA

$$
\begin{matrix}
y_{1,1} & y_{1,2} & \cdots & y_{1,400} \\
y_{2,1} & y_{2,2} & \cdots & y_{2,400} \\
\vdots & \vdots & \vdots & \vdots \\
y_{n,1} & y_{n,2} & \cdots & y_{n,400}
\end{matrix}
$$

Evolution of bankruptcies

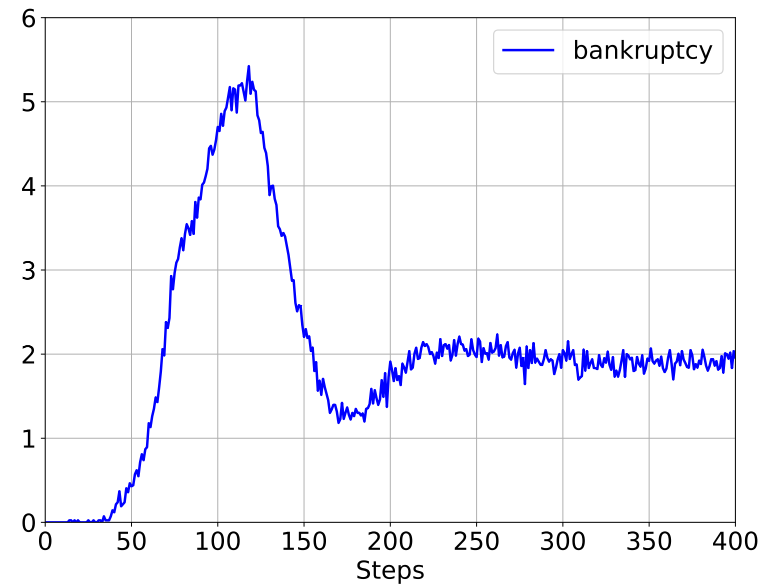


Transient Analysis by autoIR: Large Macro ABM

Large-scale macro financial ABM from Caiani et al, JEDC, 2016

- An economy with households, consumption/capital firms, commercial banks, government, central bank
- Thousands of agents
- Implemented in JMAB: Java framework for macro stock-flow consistent ABM models.
  - Side product: any model implemented in JMAB is now natively integrated with MultiVeStA

Evolution of bankruptcies



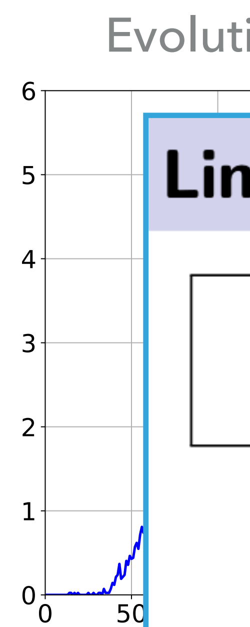
$$\begin{array}{cccc} y_{1,1} & y_{1,2} & \cdots & y_{1,400} \\ y_{2,1} & y_{2,2} & \cdots & y_{2,400} \\ \vdots & \vdots & \vdots & \vdots \\ y_{n,1} & y_{n,2} & \cdots & y_{n,400} \end{array}$$

$$\sum_{i=1}^{n} \frac{y_{i,2}}{n} = \overline{Y}_2 \approx E[Y_2]$$

Transient Analysis by autoIR: Large Macro ABM

Large-scale macro financial ABM from Caiani et al, JEDC, 2016

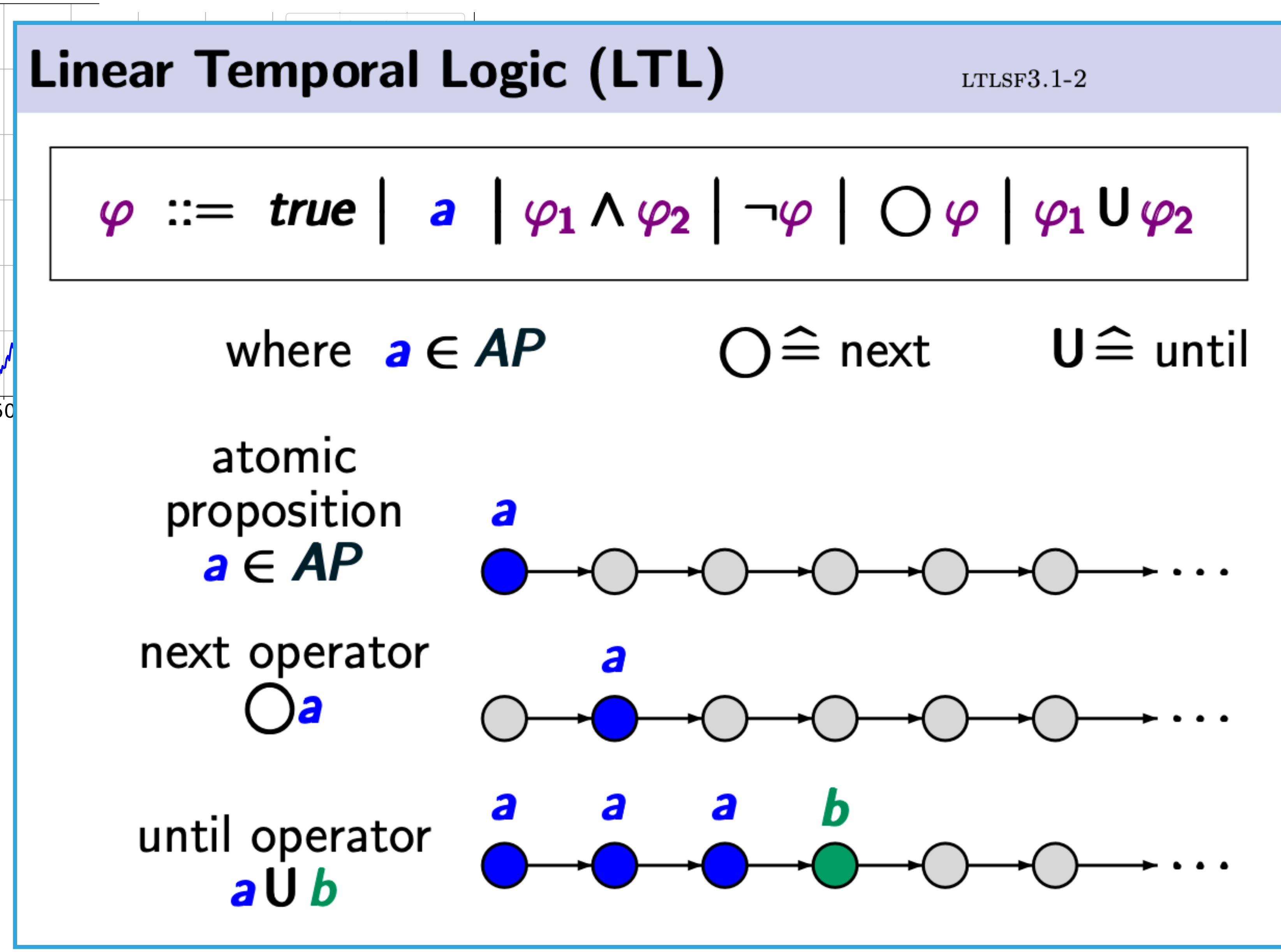
- An economy with households, consumption/capital firms, commercial banks, government, central bank
- Thousands of agents
- Implemented in JMAB: Java framework for macro stock-flow consistent ABM models.
  - Side product: any model implemented in JMAB is now natively integrated with MultiVeStA



Does This Remind You Anything?

$$\begin{array}{cccc} y_{1,1} & y_{1,2} & \cdots & y_{1,400} \\ y_{2,1} & y_{2,2} & \cdots & y_{2,400} \\ \vdots & \vdots & \vdots & \vdots \\ y_{n,1} & y_{n,2} & \cdots & y_{n,400} \end{array}$$

$$\sum_{i=1}^{n} \frac{y_{i,2}}{n} = \overline{Y}_2 \approx E[Y_2]$$



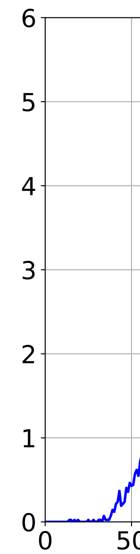
Transient Analysis by autoIR: Large Macro ABM

Large-scale macro financial ABM from Caiani et al, JEDC, 2016

- An economy with households, consumption/capital firms, commercial banks, government, central bank
- Thousands of agents
- Implemented in JMAB: Java framework for macro stock-flow consistent ABM models.
  - Side product: any model implemented in JMAB is now natively integrated with MultiVeStA

Does This Remind You Anything?

Evoluti



$y\ \ 1$	$y\ \ 2$	$\cdots$	$y\ \ 400$
$y_{2,1}$	$y_{2,2}$	$\cdots$	$y_{2,400}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$y_{n,1}$	$y_{n,2}$	$\cdots$	$y_{n,400}$

$$\sum_{i=1}^{n} \frac{y_{i,2}}{n} = \overline{Y}_2 \approx E[Y_2]$$

Interpretation of LTL formulas over TS LTLSF3.1-SEM-TS

given: TS $\mathcal{T} = (S, Act, \rightarrow, S_0, AP, L)$
without terminal states
LTL formula φ over AP

$\mathcal{T} \models \varphi$ iff $s_0 \models \varphi$ for all $s_0 \in S_0$
iff $trace(\pi) \models \varphi$ for all $\pi \in Paths(\mathcal{T})$
iff $Traces(\mathcal{T}) \subseteq Words(\varphi)$
iff $\mathcal{T} \models Words(\varphi)$

satisfaction relation for LT properties

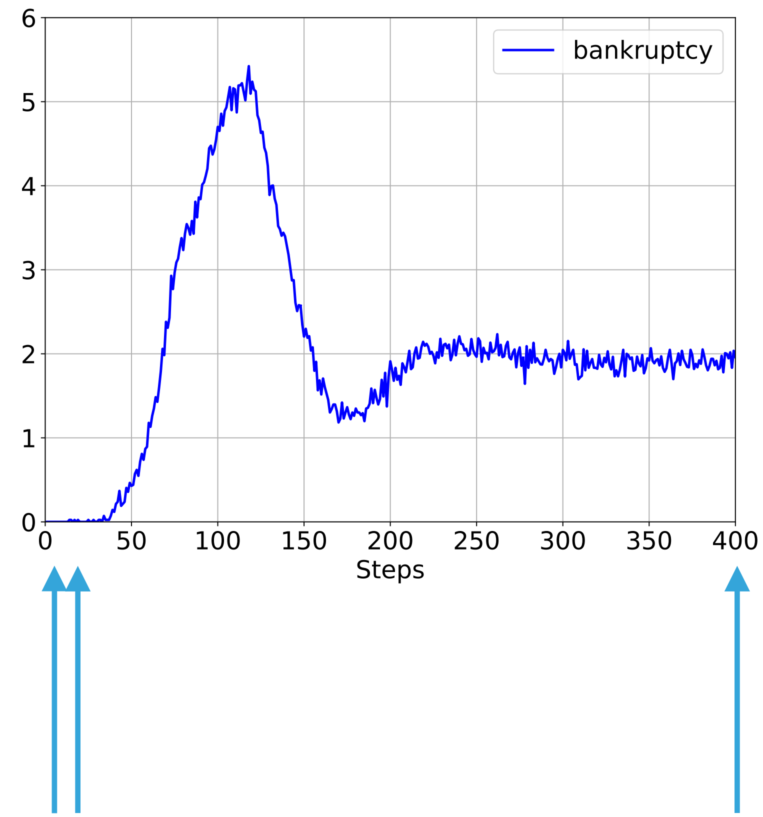
106 / 416

Transient Analysis by autoIR: Large Macro ABM

Large-scale macro financial ABM from Caiani et al, JEDC, 2016

- An economy with households, consumption/capital firms, commercial banks, government, central bank
- Thousands of agents
- Implemented in JMAB: Java framework for macro stock-flow consistent ABM models.
  - Side product: any model implemented in JMAB is now natively integrated with MultiVeStA

Evolution of bankruptcies



$$\begin{array}{cccc} y_1 & y_2 & \cdots & y_{400} \\ y_{2,1} & y_{2,2} & \cdots & y_{2,400} \\ \vdots & \vdots & \vdots & \vdots \\ y_{n,1} & y_{n,2} & \cdots & y_{n,400} \end{array}$$

$$\sum_{i=1}^{n} \frac{y_{i,2}}{n} = \overline{Y}_2 \approx E[Y_2]$$

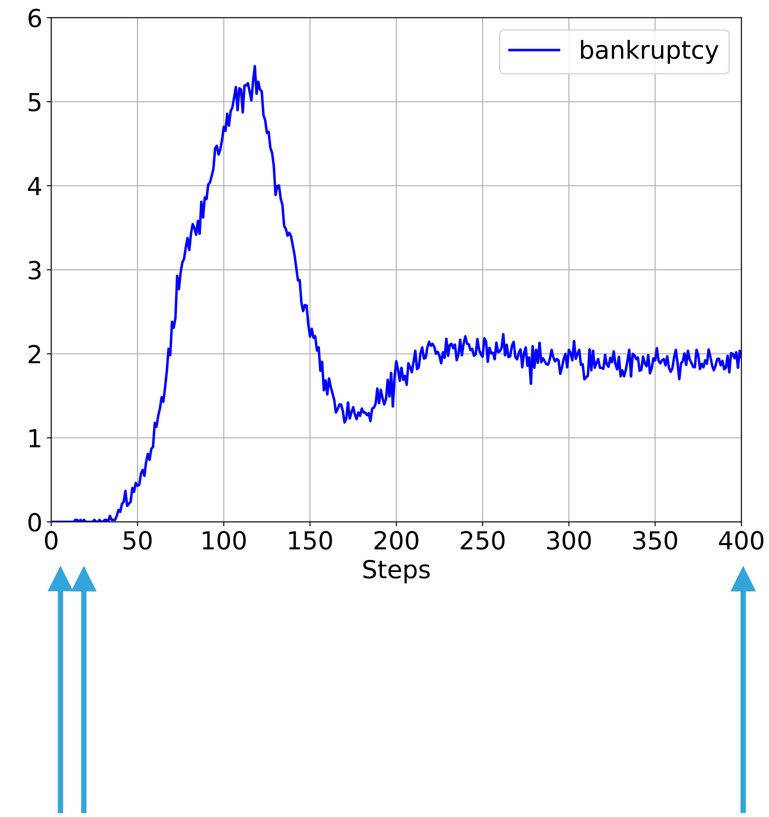
What is the correct value of n?

Transient Analysis by autoIR: Large Macro ABM

Large-scale macro financial ABM from Caiani et al, JEDC, 2016

- An economy with households, consumption/capital firms, commercial banks, government, central bank
- Thousands of agents
- Implemented in JMAB: Java framework for macro stock-flow consistent ABM models.
  - Side product: any model implemented in JMAB is now natively integrated with MultiVeStA

Evolution of bankruptcies



$$\begin{matrix} y_1 & y_2 & \cdots & y_{400} \\ y_{2,1} & y_{2,2} & \cdots & y_{2,400} \\ \vdots & \vdots & \vdots & \vdots \\ y_{n,1} & y_{n,2} & \cdots & y_{n,400} \end{matrix}$$

$$\sum_{i=1}^{n} \frac{y_{i,2}}{n} = \overline{Y}_2 \approx E[Y_2]$$

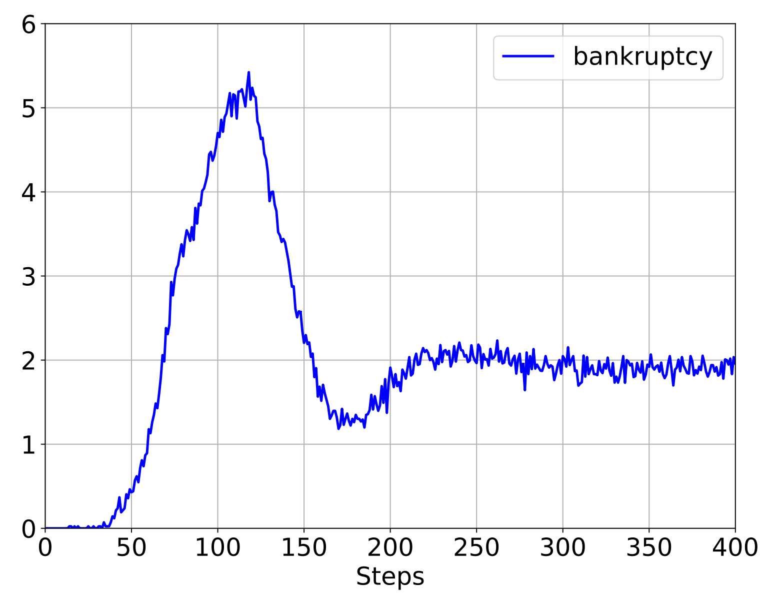
What is the correct value of n?
Typical answer: 100

Transient Analysis by autoIR: Large Macro ABM

Large-scale macro financial ABM from Caiani et al, JEDC, 2016

- An economy with households, consumption/capital firms, commercial banks, government, central bank
- Thousands of agents
- Implemented in JMAB: Java framework for macro stock-flow consistent ABM models.
  - Side product: any model implemented in JMAB is now natively integrated with MultiVeStA

Evolution of bankruptcies



$$\begin{matrix} y_{1,1} & y_{1,2} & \cdots & y_{1,400} \\ y_{2,1} & y_{2,2} & \cdots & y_{2,400} \\ \vdots & \vdots & \vdots & \vdots \\ y_{n,1} & y_{n,2} & \cdots & y_{n,400} \end{matrix}$$

$$\sum_{i=1}^{n} \frac{y_{i,2}}{n} = \overline{Y}_2 \approx E[Y_2]$$

What is the correct value of n?
Typical answer: 100

Our answer:
The question itself is ill-posed
Each property and time step might require a different n

Transient Analysis by autoIR: Large Macro ABM

Large-scale macro financial ABM from Caiani et al, JEDC, 2016

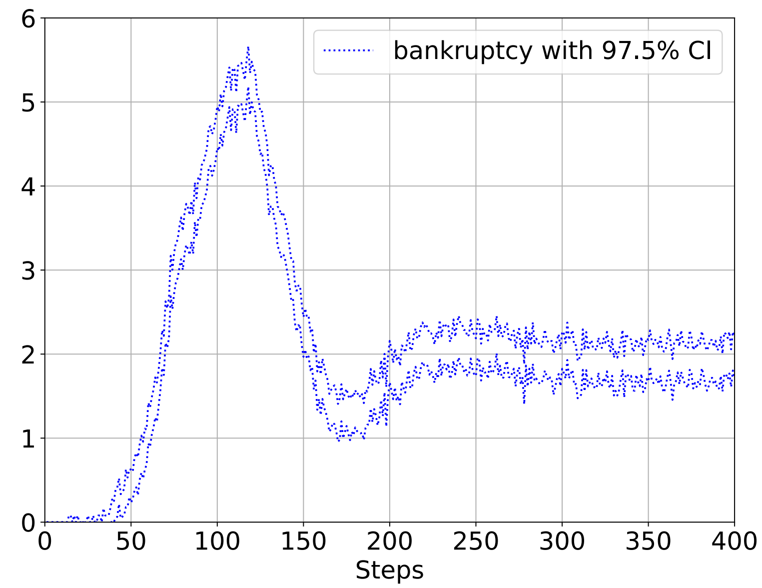
- An economy with households, consumption/capital firms, commercial banks, government, central bank
- Thousands of agents
- Implemented in JMAB: Java framework for macro stock-flow consistent ABM models.
  - Side product: any model implemented in JMAB is now natively integrated with MultiVeStA

$$\overline{Y}_t \pm \mathbf{t}_{n-1,1-\frac{\alpha}{2}} \cdot \sqrt{\frac{s_t^2}{n}}$$

$$\begin{matrix} y_{1,1} & y_{1,2} & \cdots & y_{1,400} \\ y_{2,1} & y_{2,2} & \cdots & y_{2,400} \\ \vdots & \vdots & \vdots & \vdots \\ y_{n,1} & y_{n,2} & \cdots & y_{n,400} \end{matrix}$$

$$\sum_{i=1}^{n} \frac{y_{i,2}}{n} = \overline{Y}_2 \approx E[Y_2]$$

Evolution of bankruptcies



What is the correct value of n?
Typical answer: 100

Our answer:
The question itself is ill-posed
Each property and time step might require a different n

Transient Analysis by autoIR: Large Macro ABM

Large-scale macro financial ABM from Caiani et al, JEDC, 2016

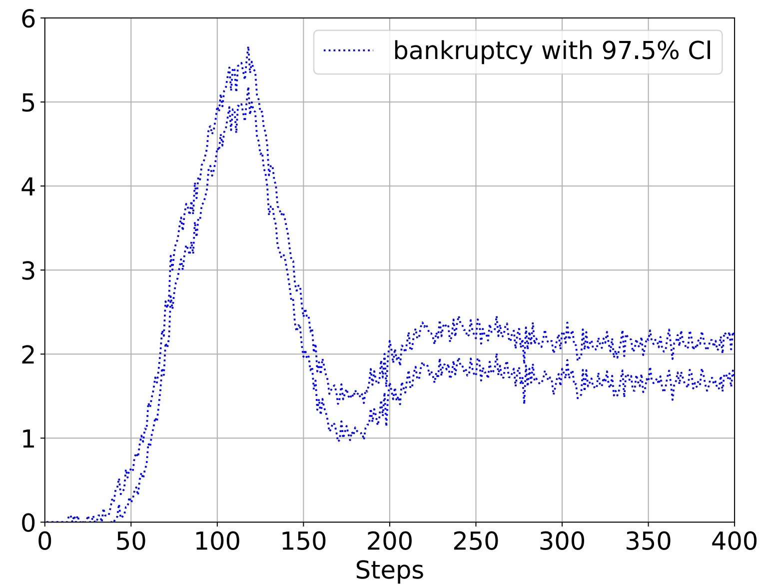
- An economy with households, consumption/capital firms, commercial banks, government, central bank
- Thousands of agents
- Implemented in JMAB: Java framework for macro stock-flow consistent ABM models.
  - Side product: any model implemented in JMAB is now natively integrated with MultiVeStA

$$\overline{Y}_t \pm \mathbf{t}_{n-1,1-\frac{\alpha}{2}} \cdot \sqrt{\frac{s_t^2}{n}}$$

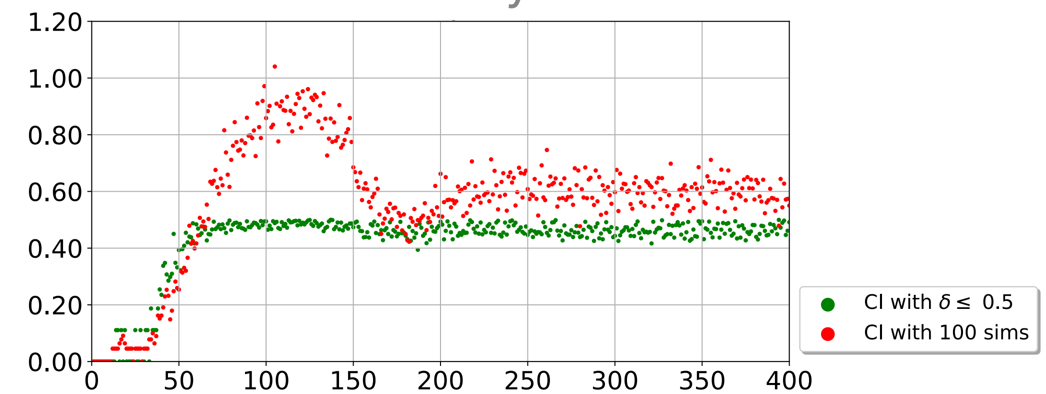
$$\begin{matrix} y_{1,1} & y_{1,2} & \cdots & y_{1,400} \\ y_{2,1} & y_{2,2} & \cdots & y_{2,400} \\ \vdots & \vdots & \vdots & \vdots \\ y_{n,1} & y_{n,2} & \cdots & y_{n,400} \end{matrix}$$

$$\sum_{i=1}^{n} \frac{y_{i,2}}{n} = \overline{Y}_2 \approx E[Y_2]$$

Evolution of bankruptcies



Confidence Intervals width
MultiVeStA VS by 100 sims



What is the correct value of n?
Typical answer: 100

Our answer:
The question itself is ill-posed
Each property and time step might require a different n

Transient Analysis by autoIR: Large Macro ABM

Large-scale macro financial ABM from Caiani et al, JEDC, 2016

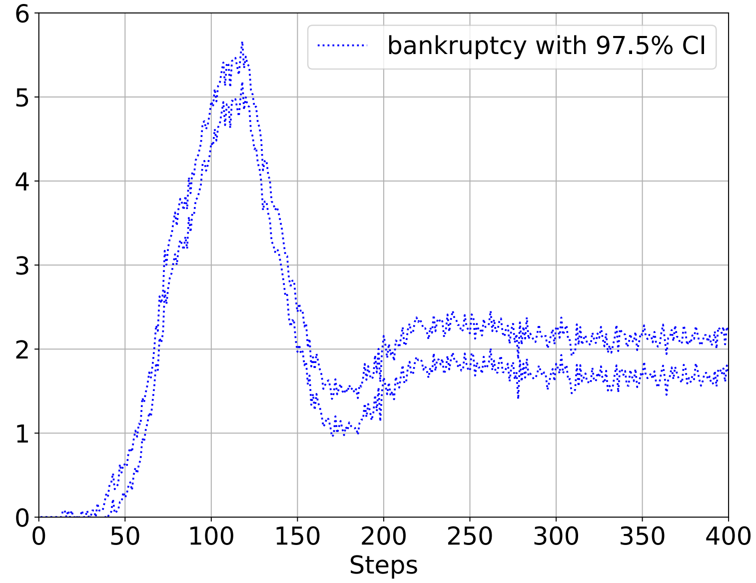
- An economy with households, consumption/capital firms, commercial banks, government, central bank
- Thousands of agents
- Implemented in JMAB: Java framework for macro stock-flow consistent ABM models.
  - Side product: any model implemented in JMAB is now natively integrated with MultiVeStA

$$\overline{Y}_t \pm \mathbf{t}_{n-1,1-\frac{\alpha}{2}} \cdot \sqrt{\frac{s_t^2}{n}}$$

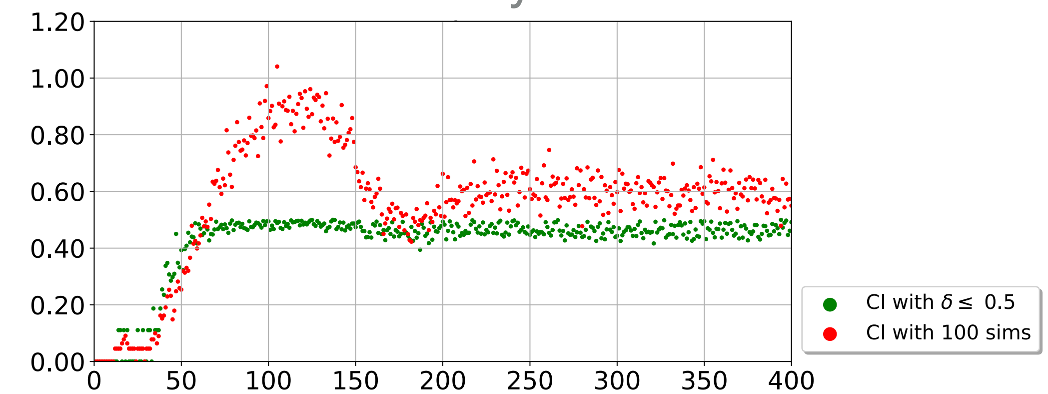
$y_{1,1}$	$y_{1,2}$	$\cdots$	$y_{1,400}$
$y_{2,1}$	$y_{2,2}$	$\cdots$	$y_{2,400}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$y_{n,1}$	$y_{n,2}$	$\cdots$	$y_{n,400}$

$$\sum_{i=1}^{n} \frac{y_{i,2}}{n} = \overline{Y}_2 \approx E[Y_2]$$

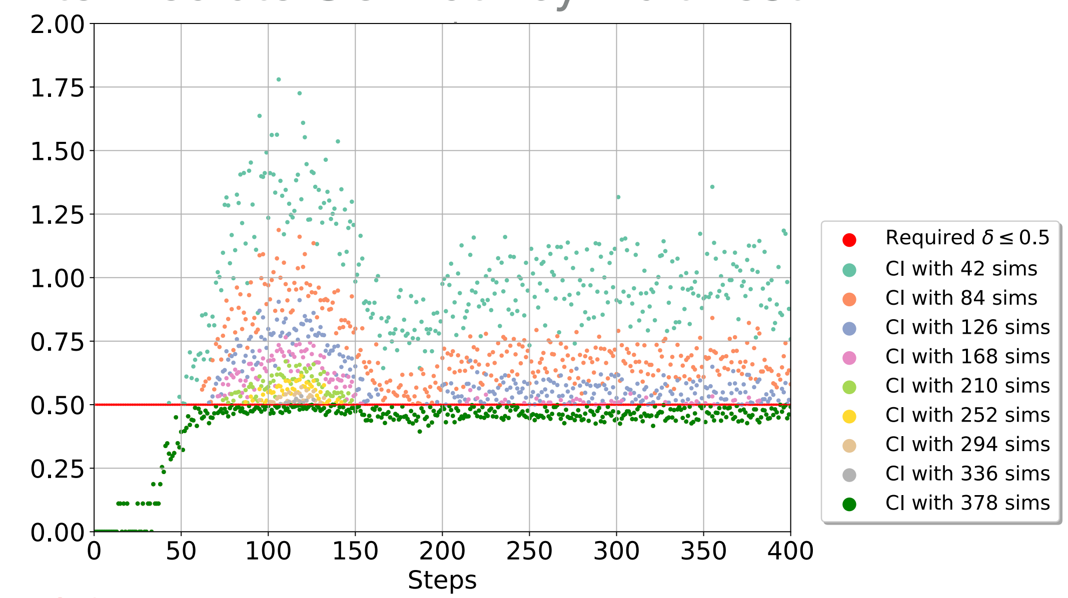
Evolution of bankruptcies



Confidence Intervals width MultiVeStA VS by 100 sims



Intermediate CIs width by MultiVeStA



What is the correct value of n?
Typical answer: 100

Our answer:
The question itself is ill-posed
Each property and time step might require a different n

Transient Analysis by autoIR: Large Macro ABM

Large-scale macro financial ABM from Caiani et al, JEDC, 2016

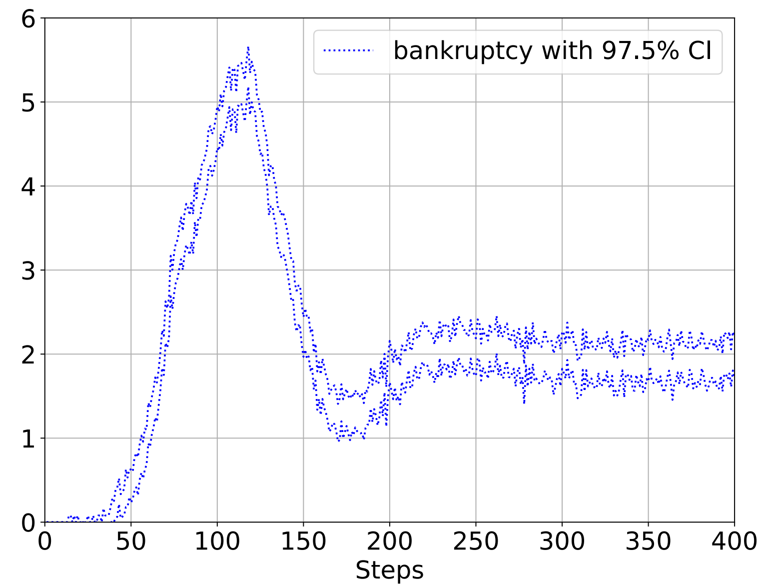
- An economy with households, consumption/capital firms, commercial banks, government, central bank
- Thousands of agents
- Implemented in JMAB: Java framework for macro stock-flow consistent ABM models.
  - Side product: any model implemented in JMAB is now natively integrated with MultiVeStA

$$\overline{Y}_t \pm \mathbf{t}_{n-1,1-\frac{\alpha}{2}} \cdot \sqrt{\frac{s_t^2}{n}}$$

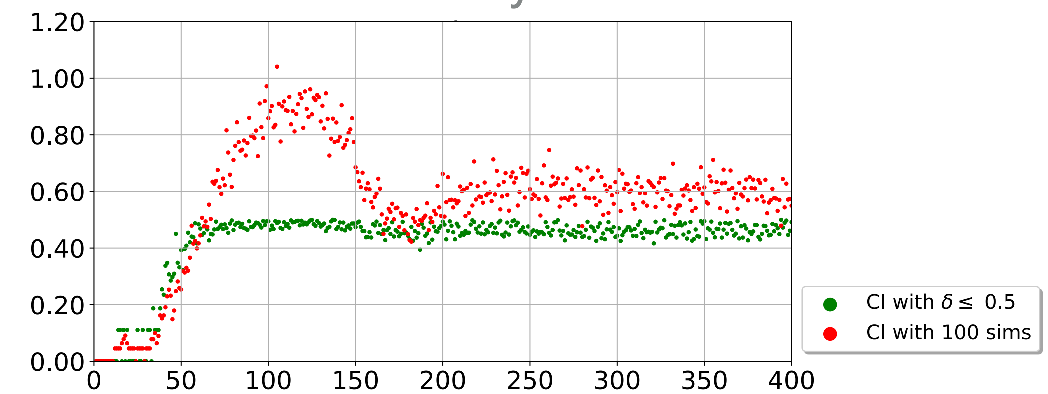
$$\begin{matrix} y_{1,1} & y_{1,2} & \cdots & y_{1,400} \\ y_{2,1} & y_{2,2} & \cdots & y_{2,400} \\ \vdots & \vdots & \vdots & \vdots \\ y_{n,1} & y_{n,2} & \cdots & y_{n,400} \end{matrix}$$

$$\sum_{i=1}^{n} \frac{y_{i,2}}{n} = \overline{Y}_2 \approx E[Y_2]$$

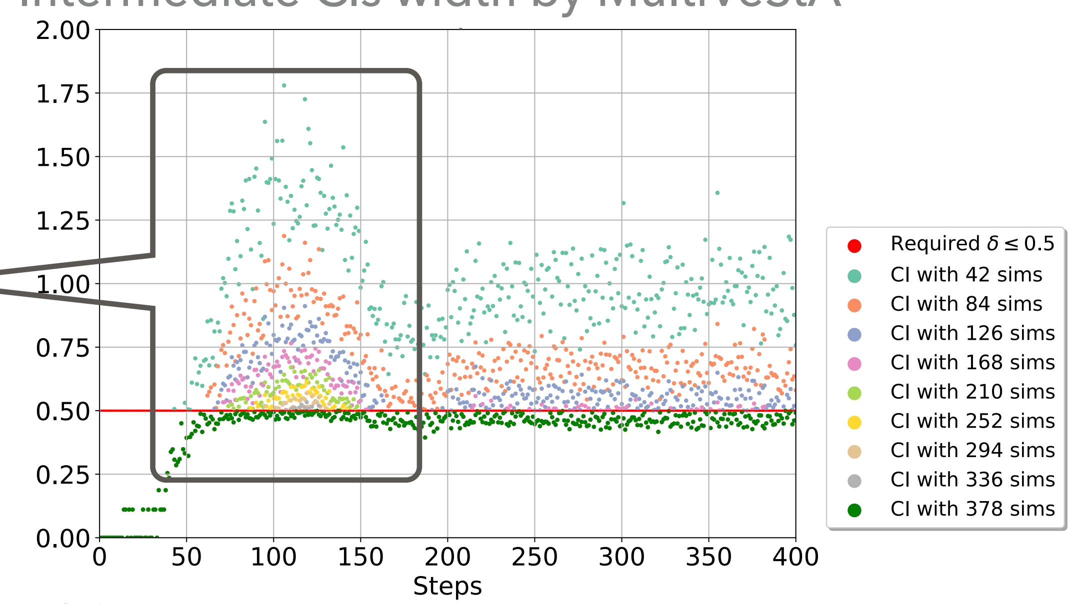
Evolution of bankruptcies



Confidence Intervals width MultiVeStA VS by 100 sims



Intermediate CIs width by MultiVeStA



More simulations required here
We save a lot of time halting the last simulations to 150 steps

What is the correct value of n?
Typical answer: 100

Our answer:
The question itself is ill-posed
Each property and time step might require a different n

Transient Analysis by autoIR: Large Macro ABM

Large-scale macro financial ABM from Caiani et al, JEDC, 2016

- An economy with households, consumption/capital firms, commercial banks, government, central bank
- Thousands of agents
- Implemented in JMAB: Java framework for macro stock-flow consistent ABM models.
  - Side product: any model implemented in JMAB is now natively integrated with MultiVeStA

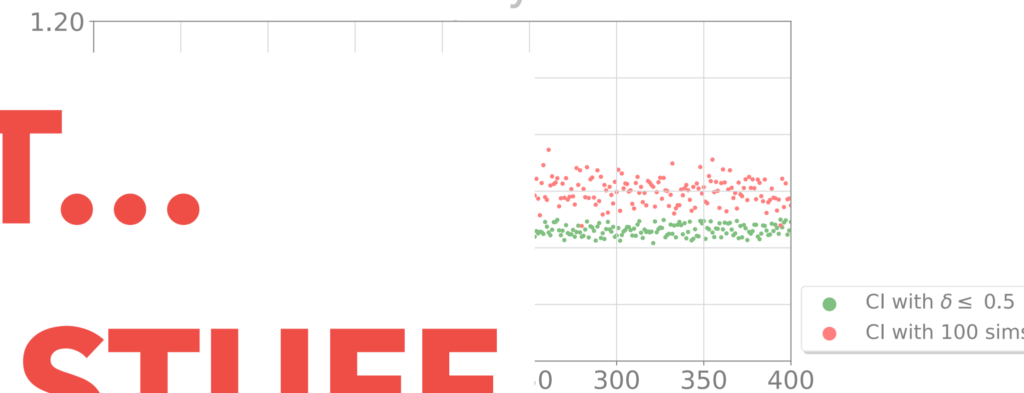
$$\overline{Y}_t \pm t_{n-1,1-\frac{\alpha}{2}} \cdot \sqrt{\frac{s_t^2}{n}}$$

$$\begin{matrix} y_{1,1} & y_{1,2} & \cdots & y_{1,400} \\ y_{2,1} & y_{2,2} & \cdots & y_{2,400} \\ \vdots & \vdots & \vdots & \vdots \\ y_{n,1} & y_{n,2} & \cdots & y_{n,400} \end{matrix}$$

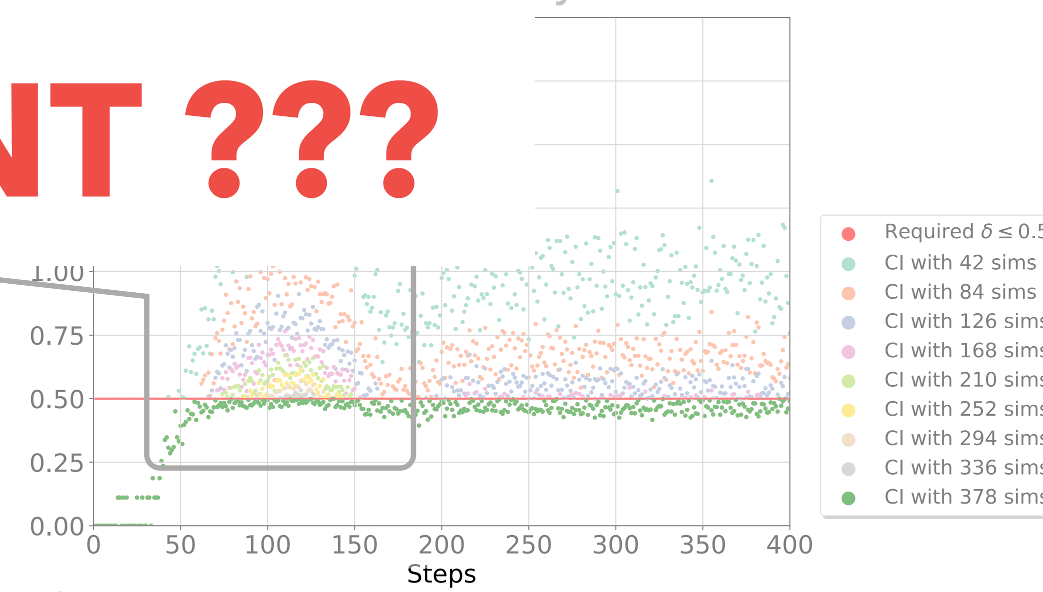
$$\sum_{i=1}^{n} \frac{y_{i,2}}{n} = \overline{Y}_2 \approx E[Y_2]$$

Evolution of bankruptcies

Confidence Intervals width
MultiVeStA VS by 100 sims



by MultiVeStA



We save a lot of time halting the last simulations to 150 steps

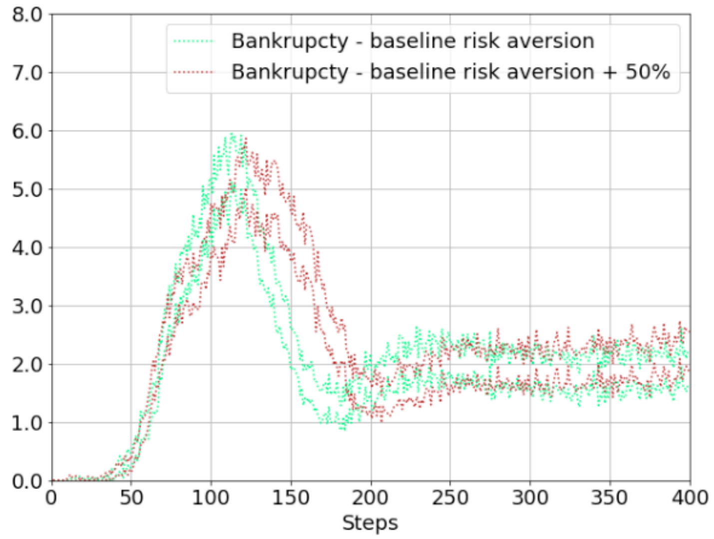
OK, BUT...
IS ALL THIS STUFF IMPORTANT ???

What is the correct value of n?
Typical answer: 100

Our answer:
The question itself is ill-posed
Each property and time step might require a different n

Statistically Meaningful Counterfactual Analysis

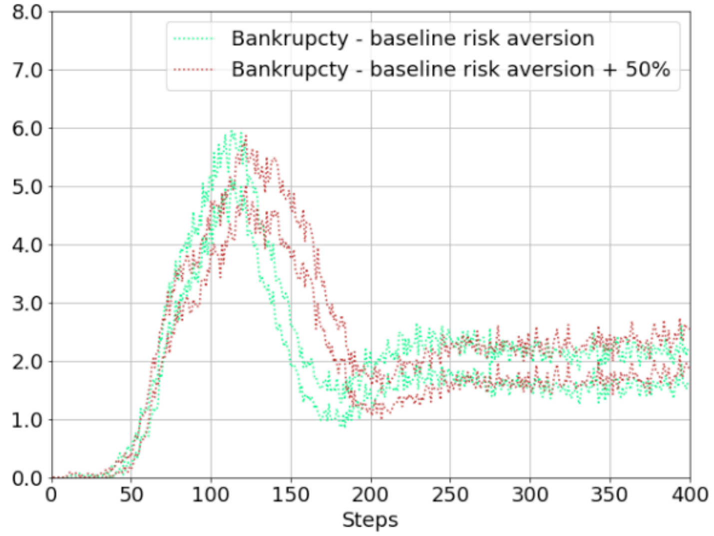
97.5% CI
100 Simulations



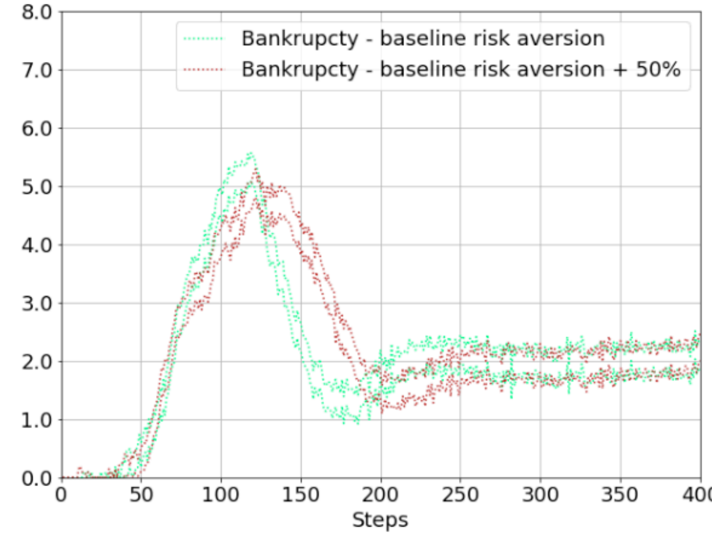
(a) CIs width for $\alpha = 0.025$ and $N = 100$ simulations

Statistically Meaningful Counterfactual Analysis

97.5% CI
100 Simulations



(a) CIs width for $\alpha = 0.025$ and $N = 100$ simulations

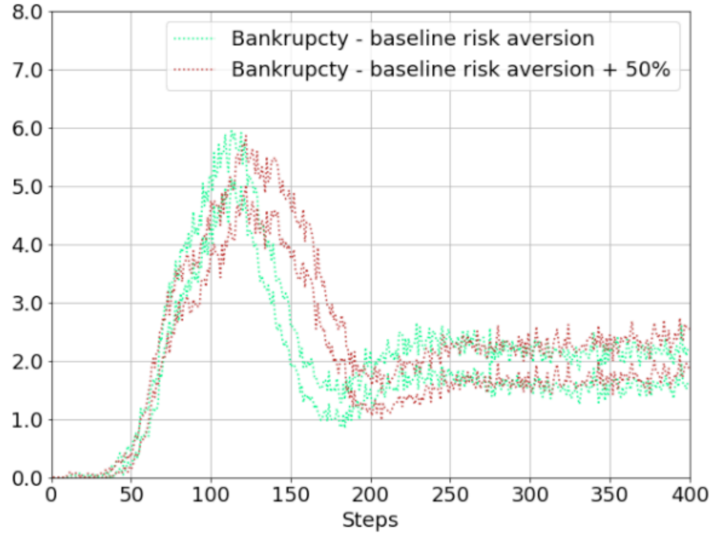


(b) CIs width for $\alpha = 0.025$ and $\delta = 0.5$

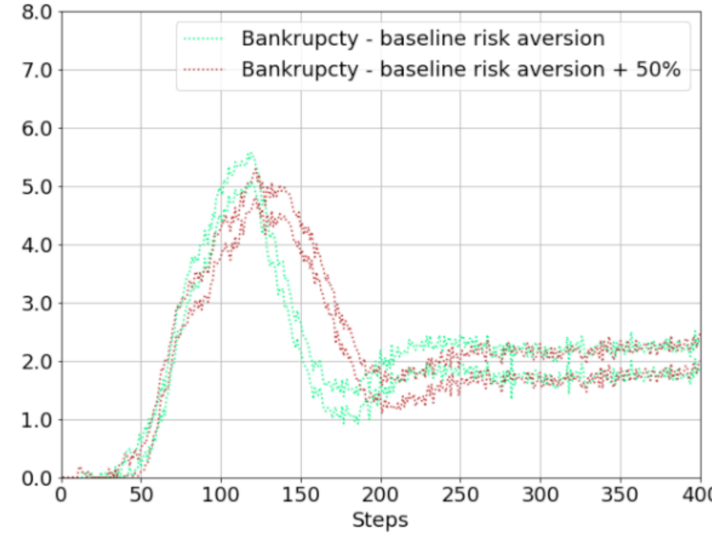
97.5% CI
MultiVeStA
$\delta = 0.5$

Statistically Meaningful Counterfactual Analysis

97.5% CI
100 Simulations



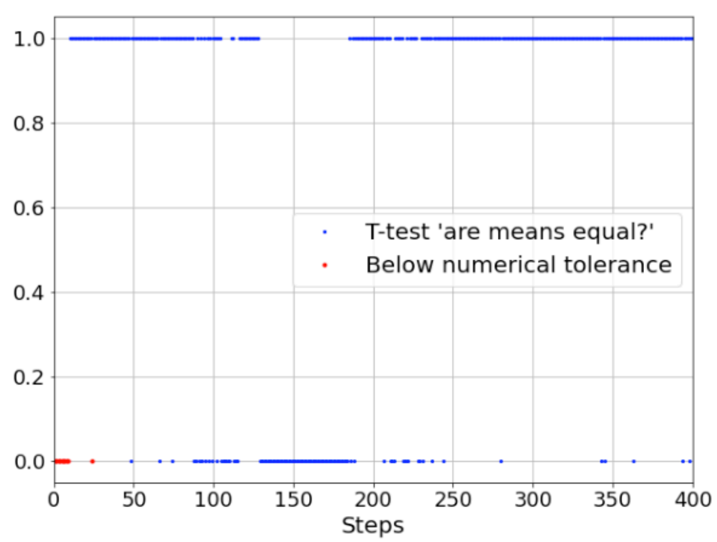
(a) CIs width for $\alpha = 0.025$ and $N = 100$ simulations



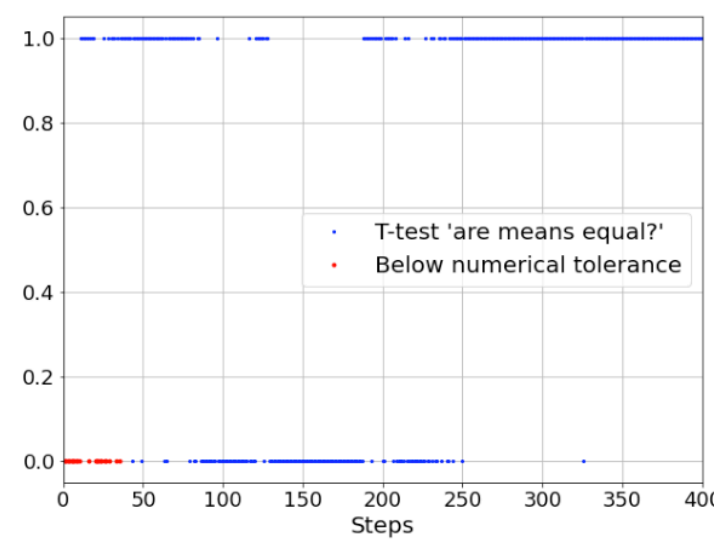
(b) CIs width for $\alpha = 0.025$ and $\delta = 0.5$

97.5% CI
MultiVeStA
$\delta = 0.5$

Welch's t-test with
significance α=2.5%
[Welch1947]



(c) T-test *are means in (a) point-wise equal* for significance $\alpha = 0.025$

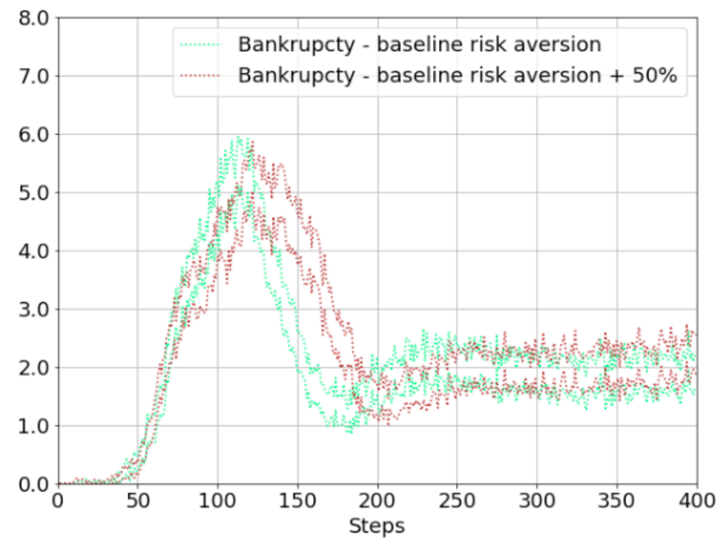


(d) T-test *are means in (b) point-wise equal* for significance $\alpha = 0.025$

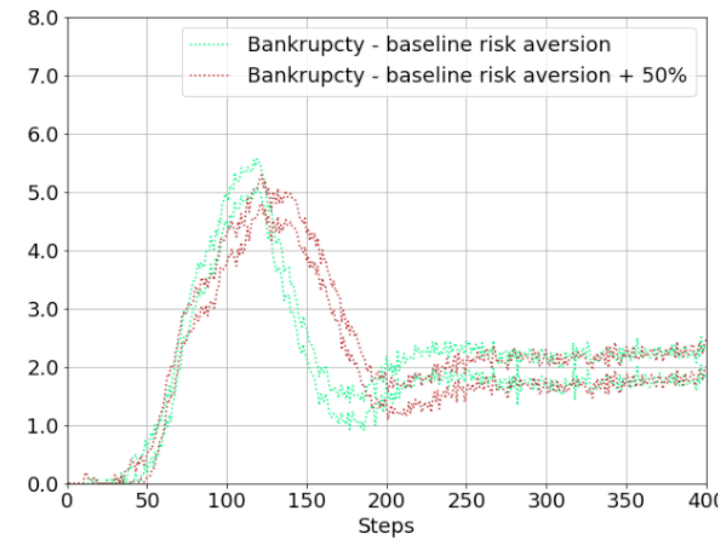
Welch's t-test with
significance α=2.5%
[Welch1947]

Statistically Meaningful Counterfactual Analysis

97.5% CI
100 Simulations



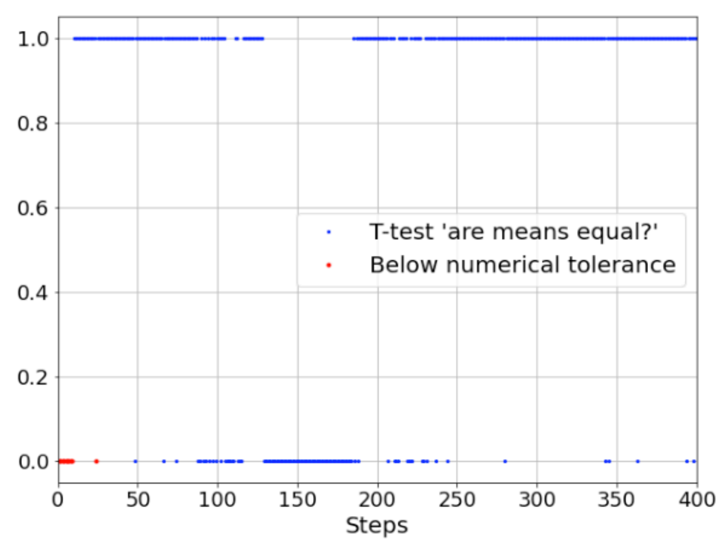
(a) CIs width for $\alpha = 0.025$ and $N = 100$ simulations



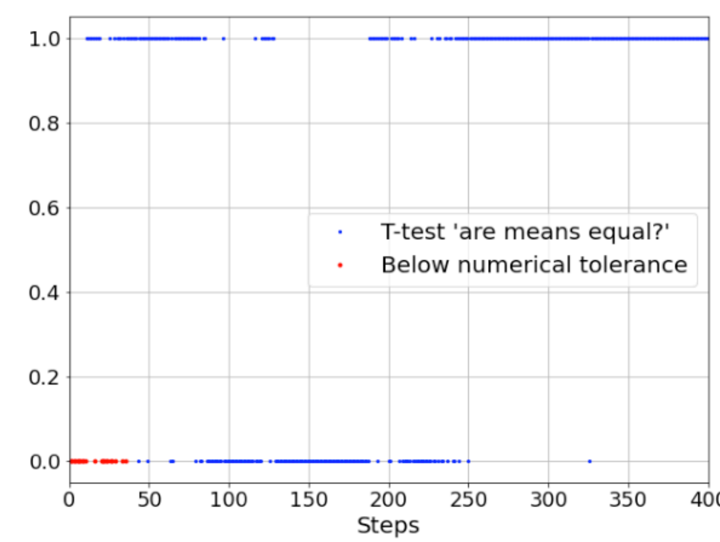
(b) CIs width for $\alpha = 0.025$ and $\delta = 0.5$

97.5% CI
MultiVeStA
$\delta = 0.5$

Welch's t-test with
significance α=2.5%
[Welch1947]



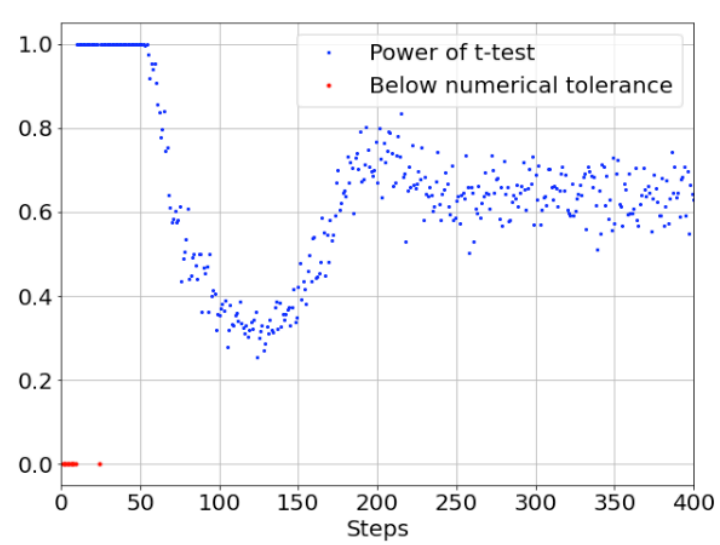
(c) T-test *are means in (a) point-wise equal* for significance $\alpha = 0.025$



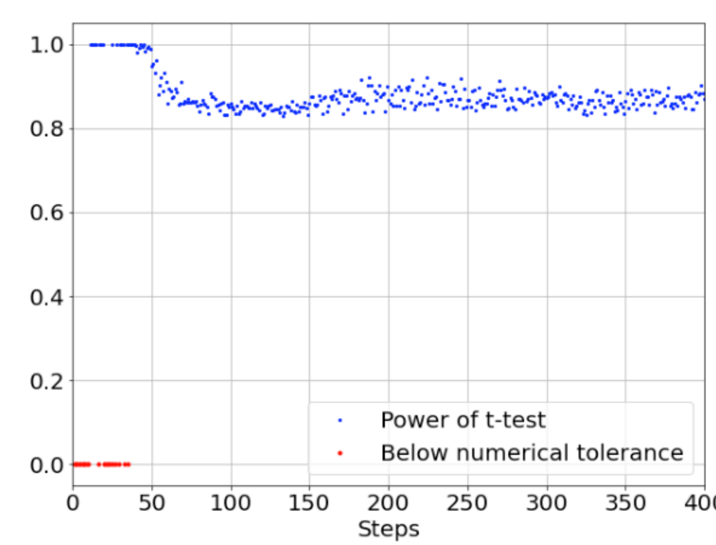
(d) T-test *are means in (b) point-wise equal* for significance $\alpha = 0.025$

Welch's t-test with
significance α=2.5%
[Welch1947]

Power of the test
P(Test=0 | Real=0)
1 - P(Type II error)
[Chow2002]



(e) Power of t-test in (c) for difference $\varepsilon = 0.5$

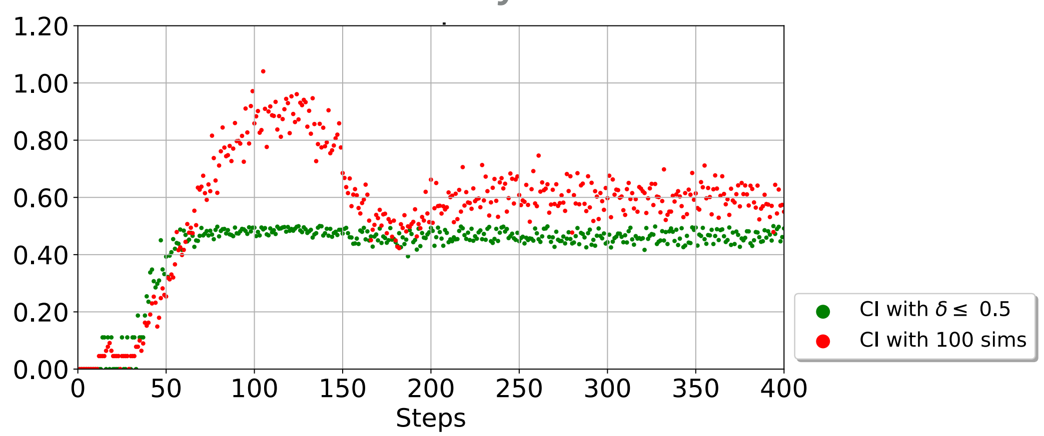


(f) Power of t-test in (d) for difference $\varepsilon = 0.5$

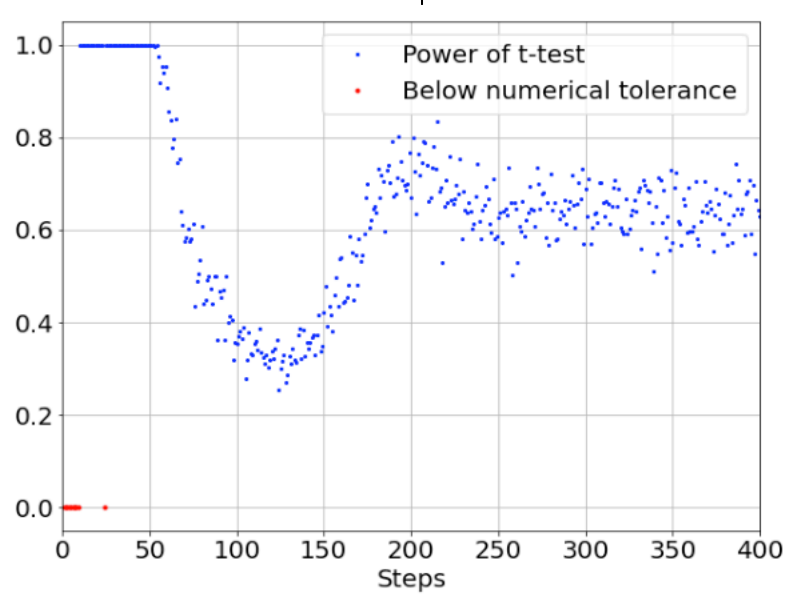
Power of the test
P(Test=0 | Real=0)
1 - P(Type II error)
[Chow2002]

Width of the Confidence Intervals and T-Test Power

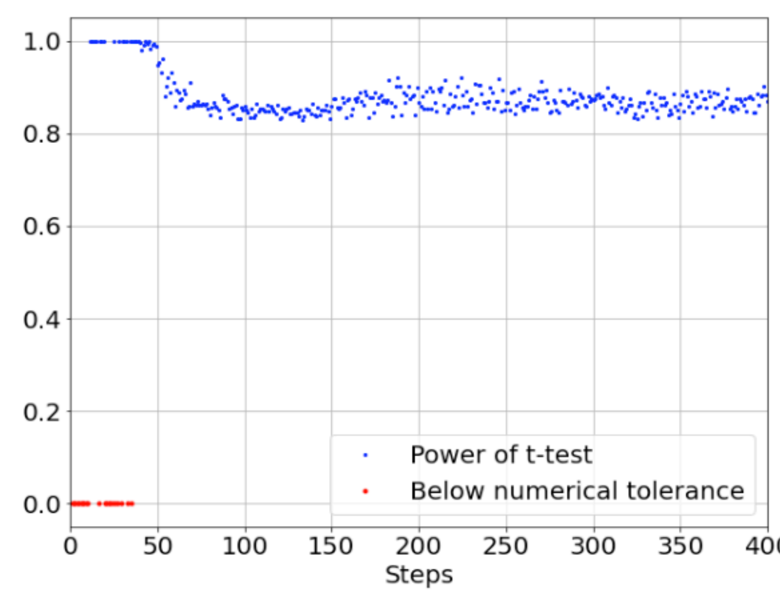
Confidence Intervals width
MultiVeStA VS by 100 sims



Power of the test
P(Test=0 | Real=0)
1 - P(Type II error)
[Chow2002]



(e) Power of t-test in (c) for difference $\varepsilon = 0.5$



(f) Power of t-test in (d) for difference $\varepsilon = 0.5$

Power of the test
P(Test=0 | Real=0)
1 - P(Type II error)
[Chow2002]

Transient Analysis by autoIR: How To Do It in MultiVeStA?

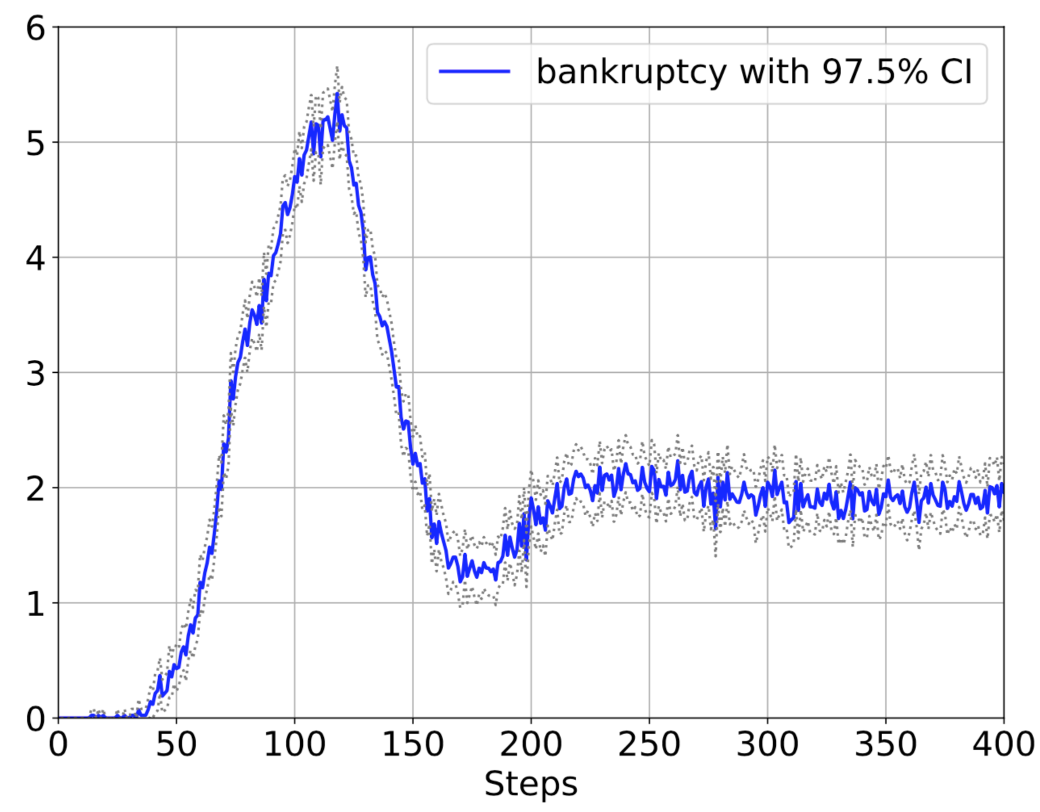
Large-scale macro financial ABM from Caiani et al, JEDC, 2016

- An economy with households, consumption/capital firms, commercial banks, government, central bank
- Thousands of agents
- Implemented in JMAB: Java framework for macro stock-flow consistent ABM models.
  - Side product: any model implemented in JMAB is now natively integrated with MultiVeStA

A query to study the evolution of bankruptcies

```
obsAtStep(t,obs) = if (s.eval("steps") == t)
                       then s.eval(obs)
                       else next(obsAtStep(t,obs))
                   fi ;
eval autoIR(E[ obsAtStep(t,"bankruptcy") ],                          ,t,1,1,400) ;
```

δ=0.5



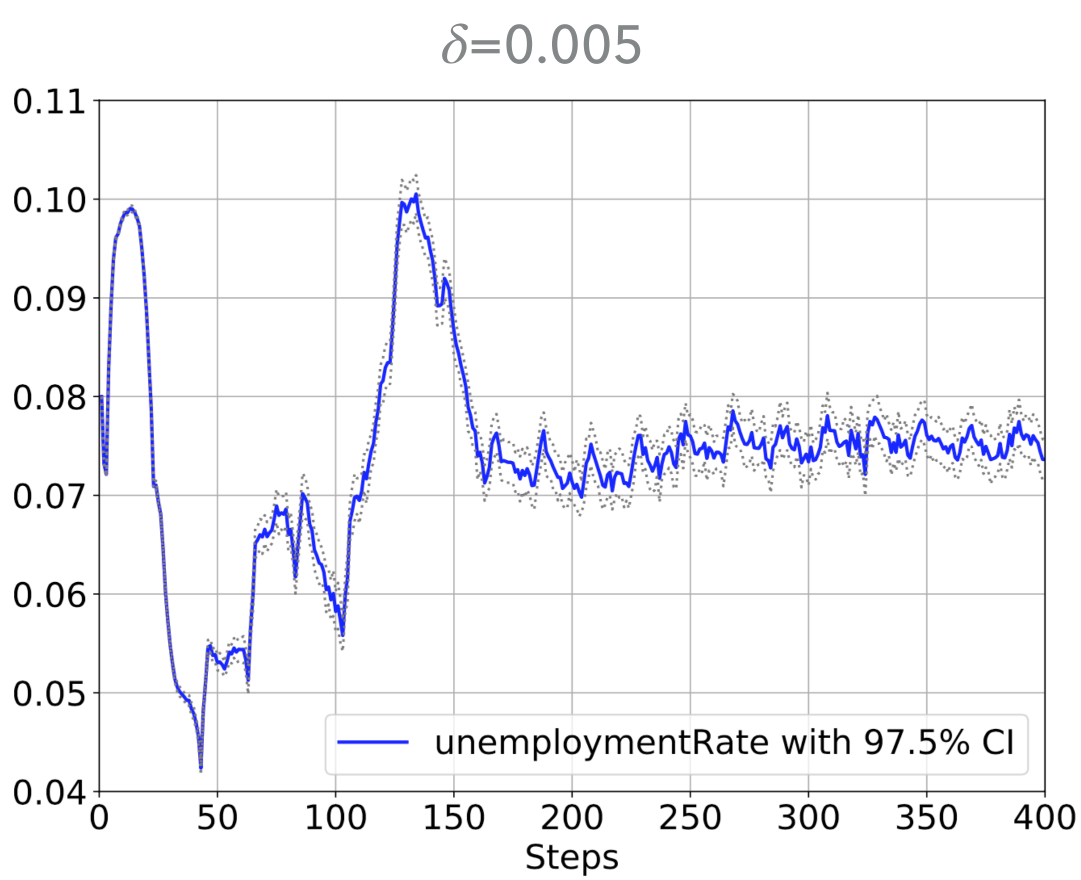
Transient Analysis by autoIR: How To Do It in MultiVeStA?

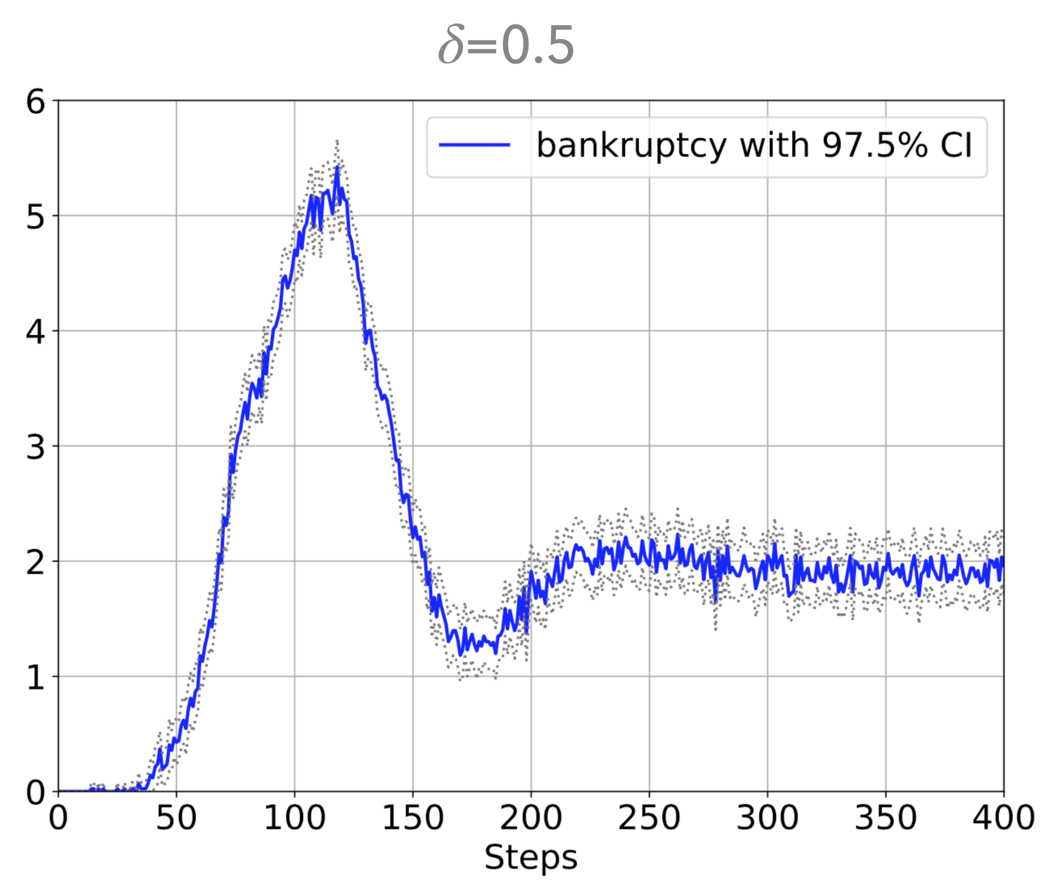
Large-scale macro financial ABM from Caiani et al, JEDC, 2016

- An economy with households, consumption/capital firms, commercial banks, government, central bank
- Thousands of agents
- Implemented in JMAB: Java framework for macro stock-flow consistent ABM models.
  - Side product: any model implemented in JMAB is now natively integrated with MultiVeStA

A query to study the evolution of bankruptcies **and** unemployment rate

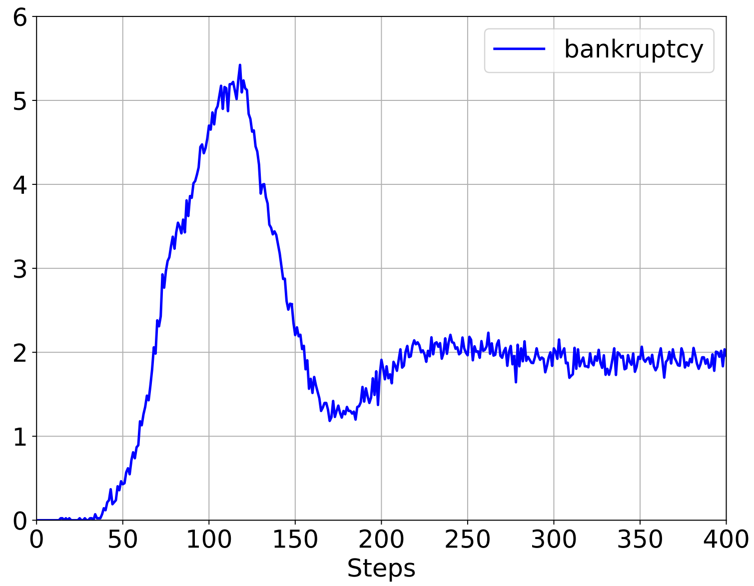
```
obsAtStep(t,obs) = if (s.eval("steps") == t)
                        then s.eval(obs)
                        else next(obsAtStep(t,obs))
                   fi ;
eval autoIR(E[ obsAtStep(t,"bankruptcy") ],E[ obsAtStep(t,"unemploymentRate") ],t,1,1,400) ;
```





Does This Remind You Anything?

Evolution of bankruptcies



Linear Temporal Logic (LTL)

LTLSF3.1-2

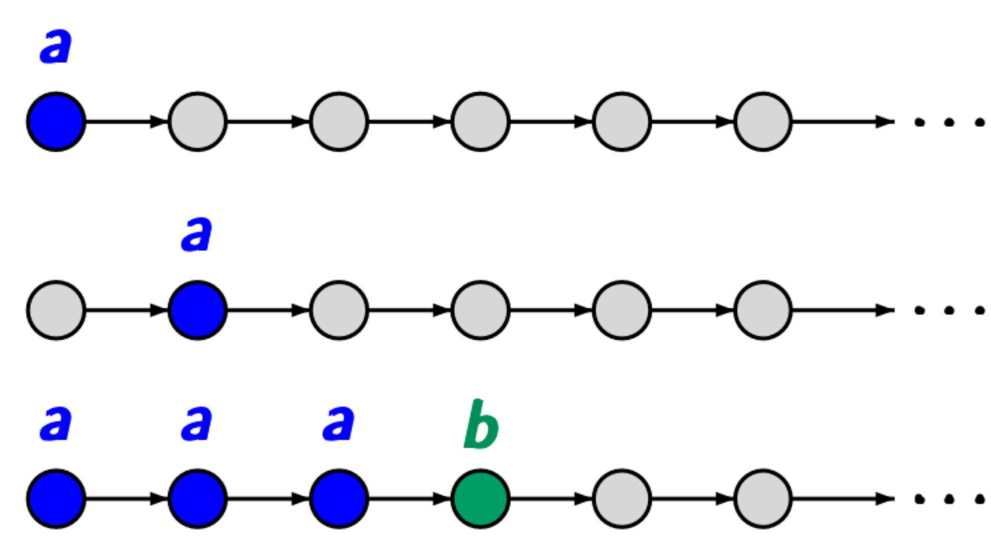
$$\varphi ::= true \mid a \mid \varphi_1 \wedge \varphi_2 \mid \neg\varphi \mid \bigcirc\varphi \mid \varphi_1 \mathsf{U} \varphi_2$$

where $a \in AP$ $\qquad \bigcirc \widehat{=}$ next $\qquad \mathsf{U} \widehat{=}$ until

atomic proposition $a \in AP$

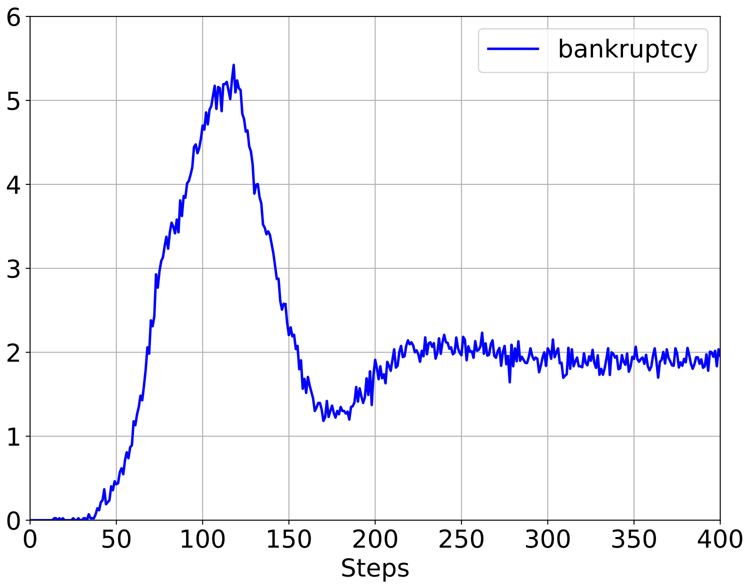
next operator $\bigcirc a$

until operator $a \mathsf{U} b$



Does This Remind You Anything?

Evolution of bankruptcies



y_1

Linear Temporal Logic (LTL)

LTLSF3.1-2

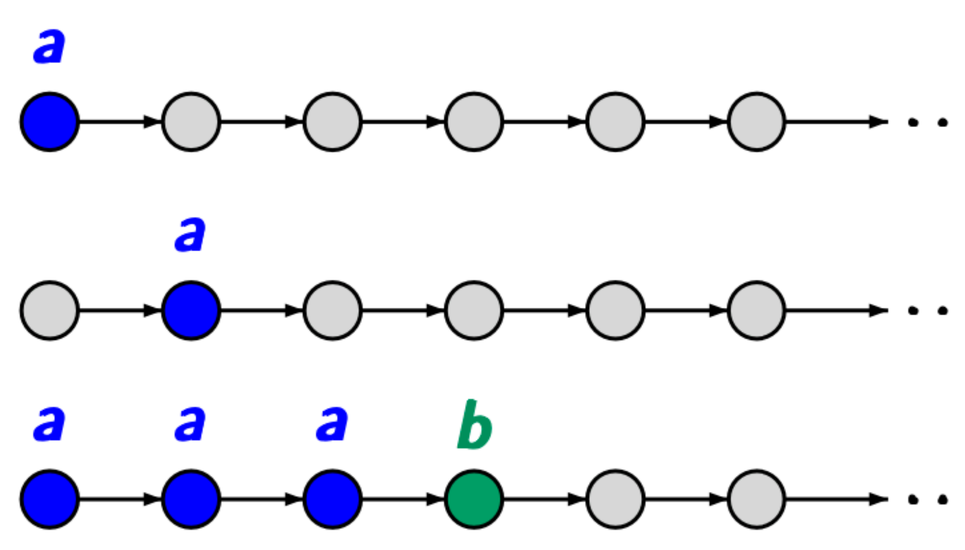
$$\varphi ::= \mathit{true} \mid a \mid \varphi_1 \wedge \varphi_2 \mid \neg\varphi \mid \bigcirc\varphi \mid \varphi_1 \,\mathsf{U}\, \varphi_2$$

where $a \in AP$ $\bigcirc \,\hat{=}\,$ next $\mathsf{U} \,\hat{=}\,$ until

atomic proposition $a \in AP$

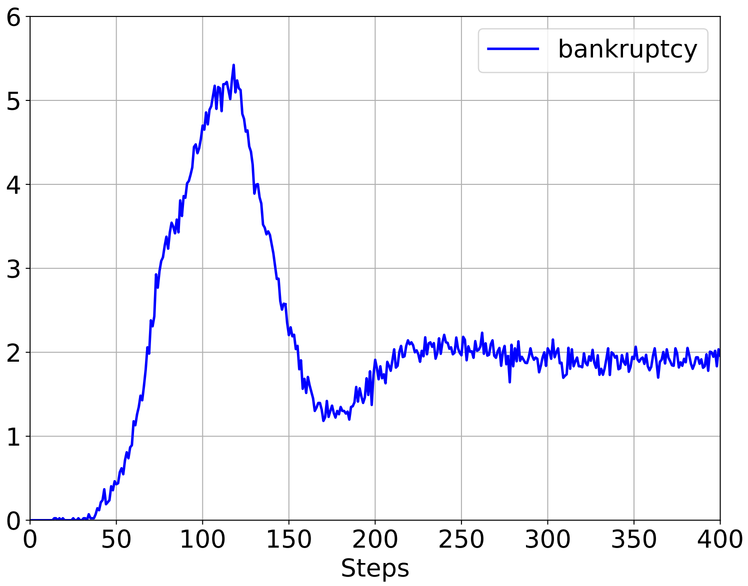
next operator $\bigcirc a$

until operator $a \,\mathsf{U}\, b$

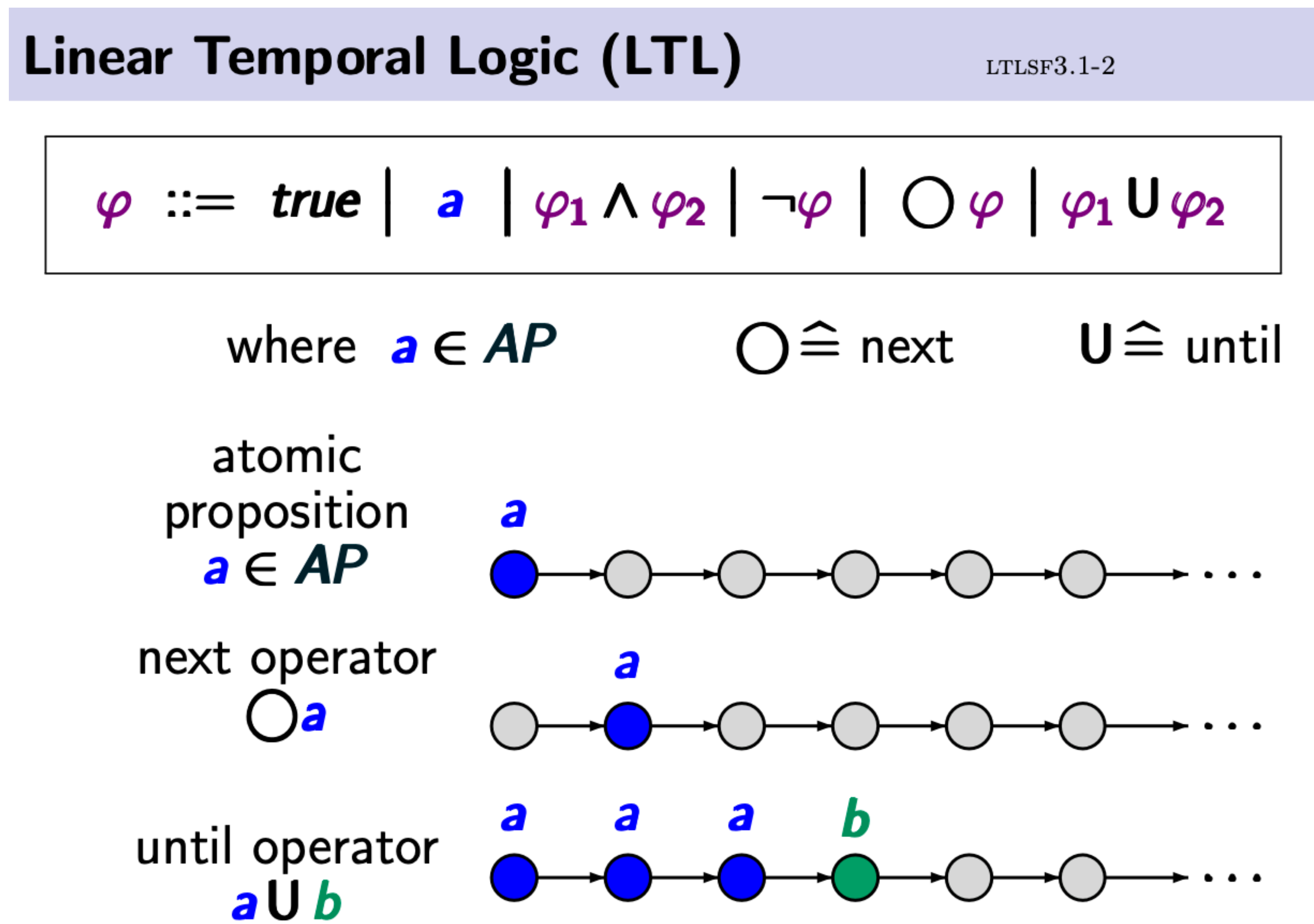


Does This Remind You Anything?

Evolution of bankruptcies

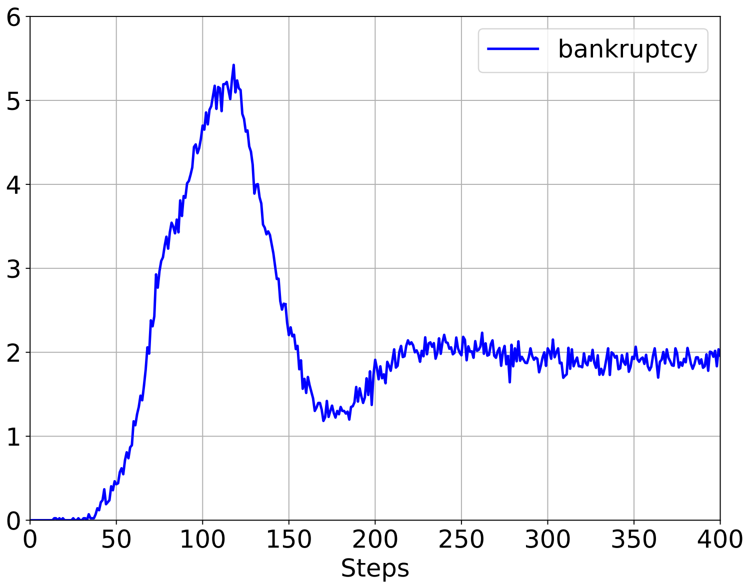


y_1 y_2



Does This Remind You Anything?

Evolution of bankruptcies



$y_1 \quad y_2 \quad \dots \quad y_{400}$

Linear Temporal Logic (LTL)

LTLSF3.1-2

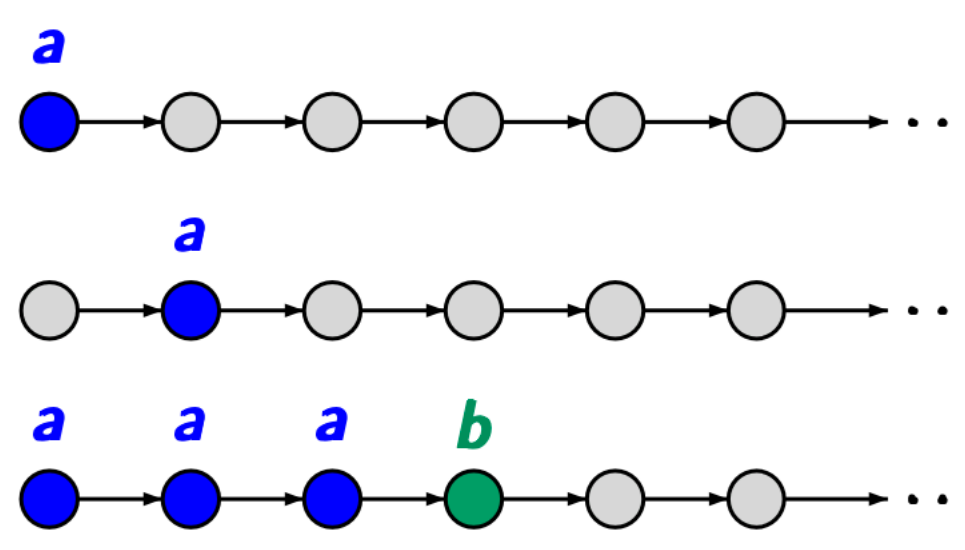
$$\varphi \ ::= \ \mathit{true} \ \mid \ a \ \mid \ \varphi_1 \wedge \varphi_2 \ \mid \ \neg\varphi \ \mid \ \bigcirc\varphi \ \mid \ \varphi_1 \,\mathsf{U}\, \varphi_2$$

where $a \in AP$ $\qquad \bigcirc \mathrel{\widehat{=}}$ next $\qquad \mathsf{U} \mathrel{\widehat{=}}$ until

atomic proposition $a \in AP$

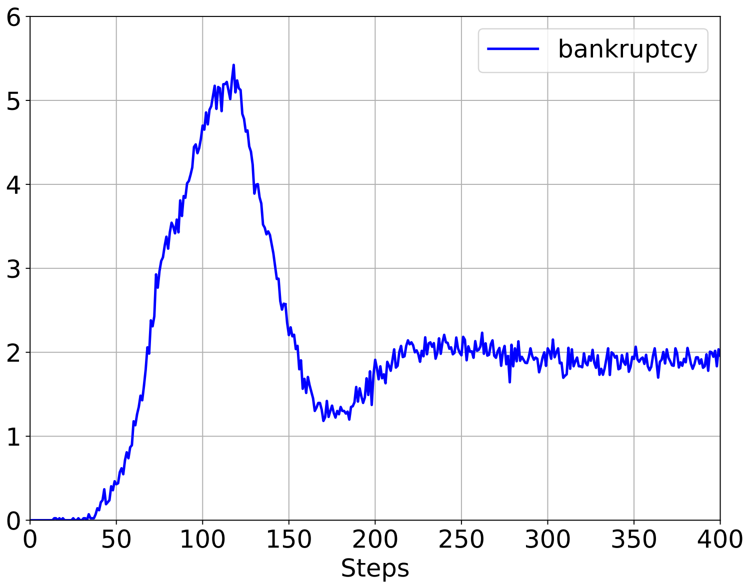
next operator $\bigcirc a$

until operator $a \,\mathsf{U}\, b$



Does This Remind You Anything?

Evolution of bankruptcies



$$y_{1,1} \quad y_{1,2} \quad \cdots \quad y_{1,400}$$

$$y_{2,1} \quad y_{2,2} \quad \cdots \quad y_{2,400}$$

$$\vdots \quad \vdots \quad \vdots \quad \vdots$$

$$y_{n,1} \quad y_{n,2} \quad \cdots \quad y_{n,400}$$

Linear Temporal Logic (LTL)

LTLSF3.1-2

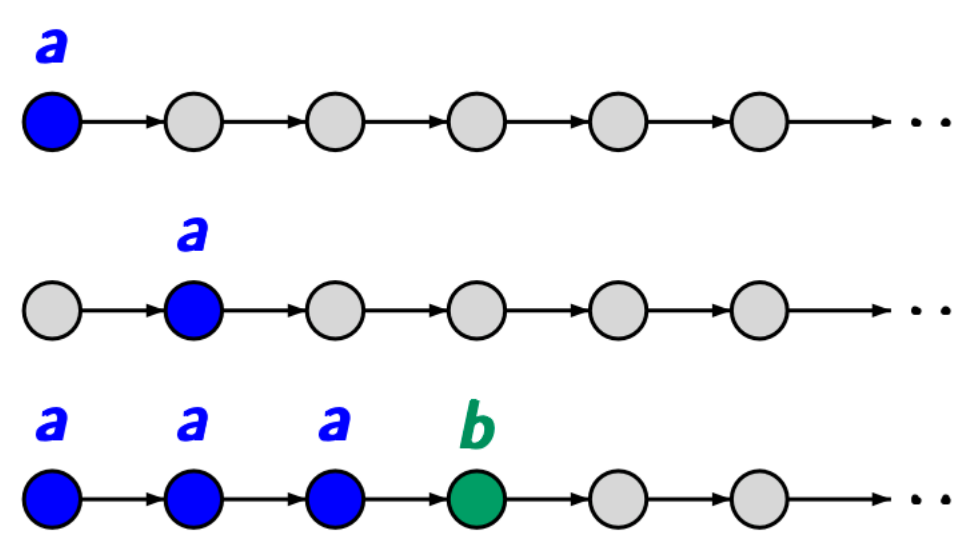
$$\varphi \ ::= \ \mathit{true} \ \big| \ a \ \big| \ \varphi_1 \wedge \varphi_2 \ \big| \ \neg\varphi \ \big| \ \bigcirc \varphi \ \big| \ \varphi_1 \, \mathsf{U} \, \varphi_2$$

where $a \in AP$ $\bigcirc \mathrel{\widehat{=}}$ next $\mathsf{U} \mathrel{\widehat{=}}$ until

atomic proposition $a \in AP$

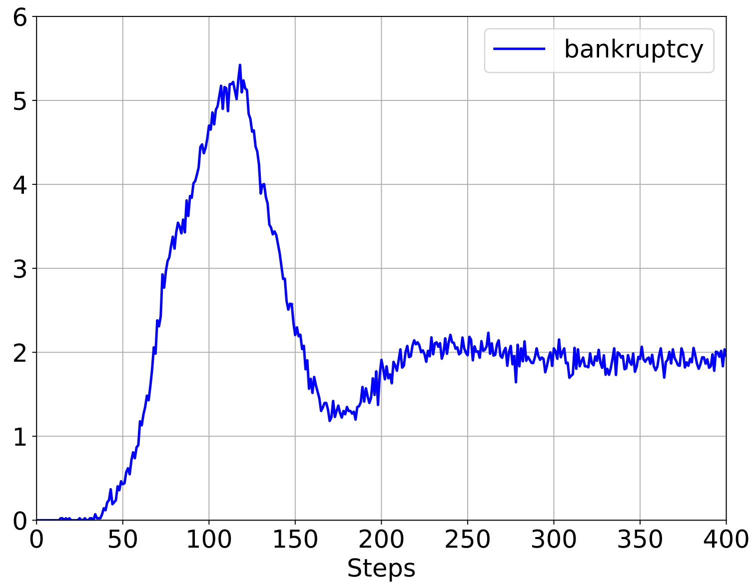
next operator $\bigcirc a$

until operator $a \, \mathsf{U} \, b$



Does This Remind You Anything?

Evolution of bankruptcies



$y_{1,1}$	$y_{1,2}$	$\cdots$	$y_{1,400}$
$y_{2,1}$	$y_{2,2}$	$\cdots$	$y_{2,400}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$y_{n,1}$	$y_{n,2}$	$\cdots$	$y_{n,400}$

$$\sum_{i=1}^{n} \frac{y_{i,2}}{n} = \overline{Y}_2 \approx E[Y_2]$$

Linear Temporal Logic (LTL)

LTLSF3.1-2

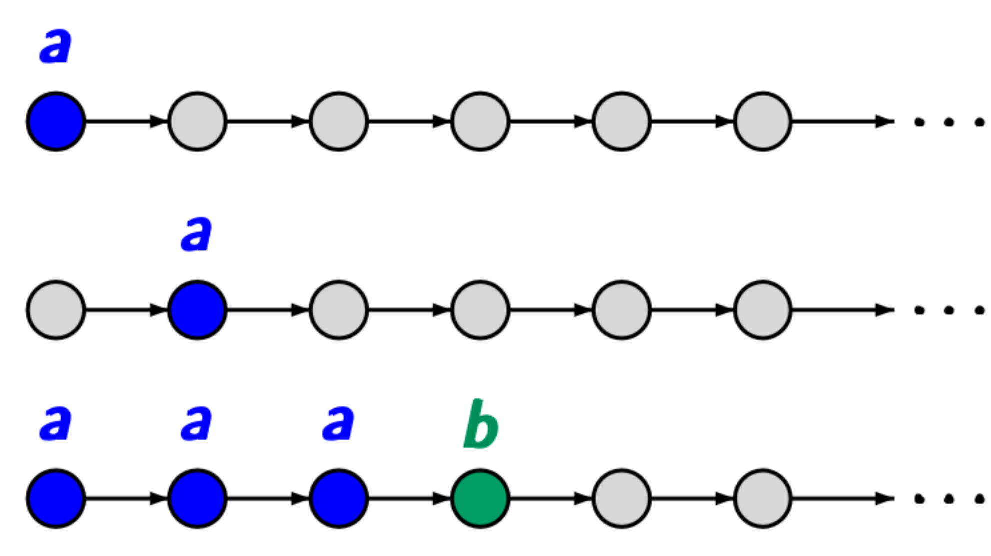
$$\varphi \ ::= \ true \ \mid \ a \ \mid \ \varphi_1 \wedge \varphi_2 \ \mid \ \neg\varphi \ \mid \ \bigcirc \varphi \ \mid \ \varphi_1 \, \mathsf{U} \, \varphi_2$$

where $a \in AP$ $\quad \bigcirc \hat{=}$ next $\quad \mathsf{U} \hat{=}$ until

atomic proposition $a \in AP$

next operator $\bigcirc a$

until operator $a \, \mathsf{U} \, b$



OUTLINE

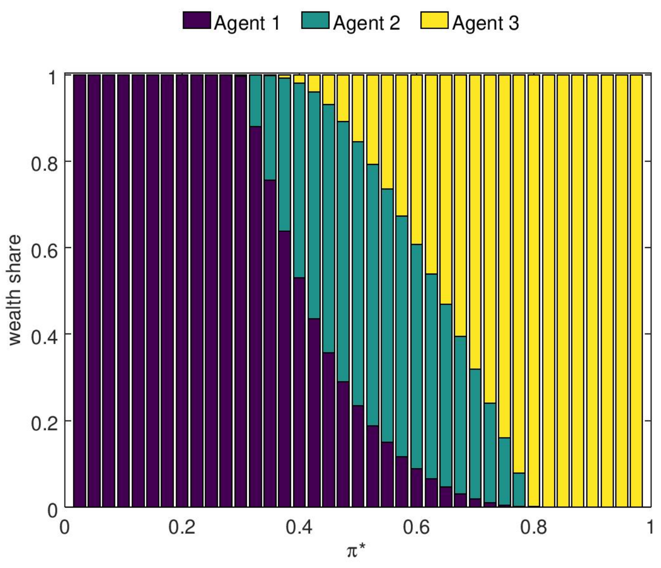
1. Motivation, vision, and proposal
    1. Automated analysis with statistical guarantees for ABMs
    2. The MultiVeStA Statistical Model Checker
2. Transient Analysis of a large-scale financial macro ABM
    1. Estimation of expected outcome and Confidence Interval
    2. Counterfactual analysis for different model configurations
3. **Steady-state analysis of a prediction market model**
    1. **Steady-state analysis by Replication and Deletion (RD)**
    2. Warmup estimation
    3. Steady-state analysis by Batch Means (BM)
    4. A methodology for ergodicity analysis based on RD and BM
4. Conclusions & Future works

Steady-State Analysis by autoRD: Market Selection

Simple repeated betting market from Kets et al, AAAI 2014

- 1 event realises at every step with a fixed probability $\pi*$
- 3 Fractional Kelly bettors with a belief on $\pi*$ and place bets accordingly

Agents wealth at steady state



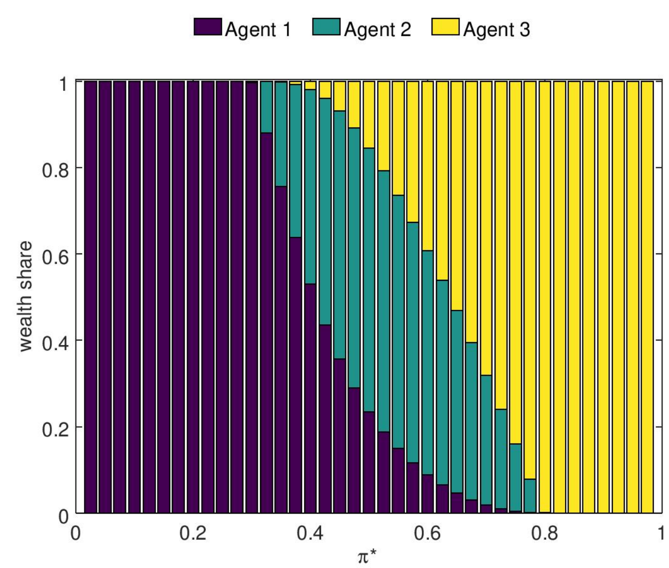
Steady-State Analysis by autoRD: Market Selection

Simple repeated betting market from Kets et al, AAAI 2014

- 1 event realises at every step with a fixed probability π^*
- 3 Fractional Kelly bettors with a belief on π^* and place bets accordingly

$$\begin{array}{cccc} y_{1,1} & y_{1,2} & \cdots & y_{1,m} \\ y_{2,1} & y_{2,2} & \cdots & y_{2,m} \\ \vdots & \vdots & \vdots & \vdots \\ y_{n,1} & y_{n,2} & \cdots & y_{n,m} \end{array}$$

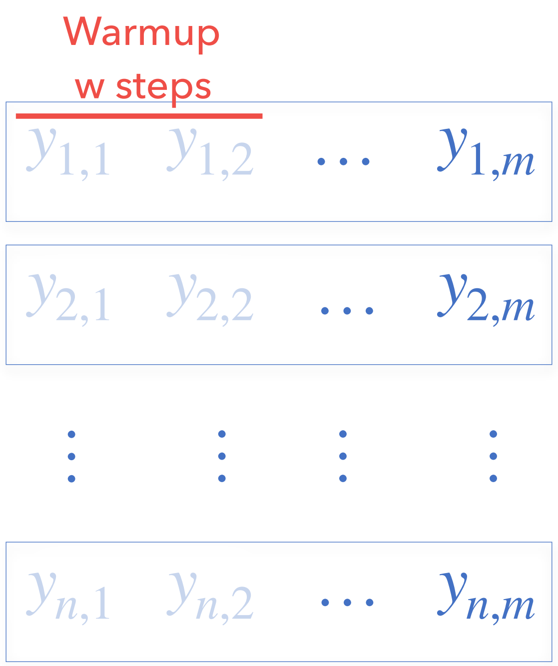
Agents wealth at steady state



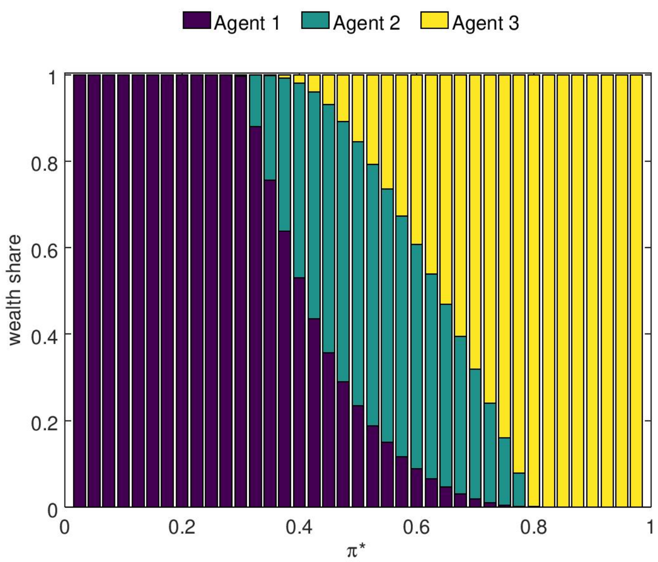
Steady-State Analysis by autoRD: Market Selection

Simple repeated betting market from Kets et al, AAAI 2014

- 1 event realises at every step with a fixed probability π^*
- 3 Fractional Kelly bettors with a belief on π^* and place bets accordingly



Agents wealth at steady state



Steady-State Analysis by autoRD: Market Selection

Simple repeated betting market from Kets et al, AAAI 2014

- 1 event realises at every step with a fixed probability π^*
- 3 Fractional Kelly bettors with a belief on π^* and place bets accordingly

Warmup w steps

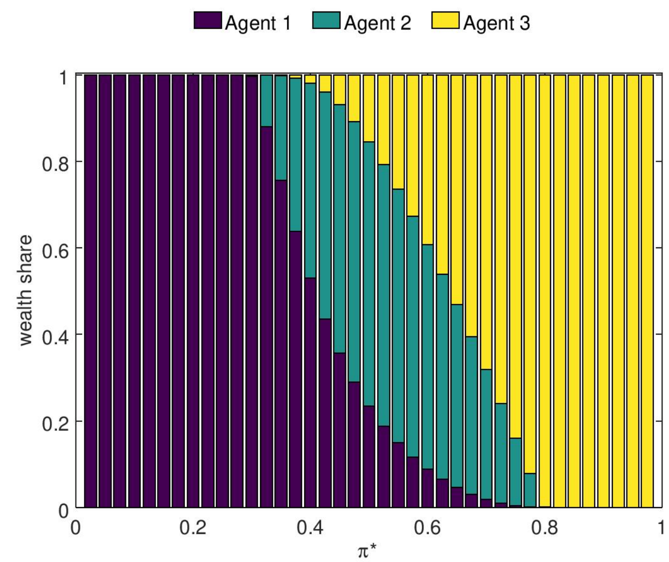
$$y_{1,1} \quad y_{1,2} \quad \cdots \quad y_{1,m} \qquad \sum_{t=w+1}^{m} \frac{y_{1,t}}{m-w} = \overline{Y}_1(w)$$

$$y_{2,1} \quad y_{2,2} \quad \cdots \quad y_{2,m} \qquad \vdots \qquad = \overline{Y}_2(w)$$

$$\vdots \quad \vdots \quad \vdots \quad \vdots \qquad \vdots \qquad \vdots$$

$$y_{n,1} \quad y_{n,2} \quad \cdots \quad y_{n,m} \qquad \vdots \qquad = \overline{Y}_n(w)$$

Agents wealth at steady state



Steady-State Analysis by autoRD: Market Selection

Simple repeated betting market from Kets et al, AAAI 2014

- 1 event realises at every step with a fixed probability π^*
- 3 Fractional Kelly bettors with a belief on π^* and place bets accordingly

Warmup w steps

$$y_{1,1} \quad y_{1,2} \quad \cdots \quad y_{1,m} \qquad \sum_{t=w+1}^{m} \frac{y_{1,t}}{m-w} = \overline{Y}_1(w)$$

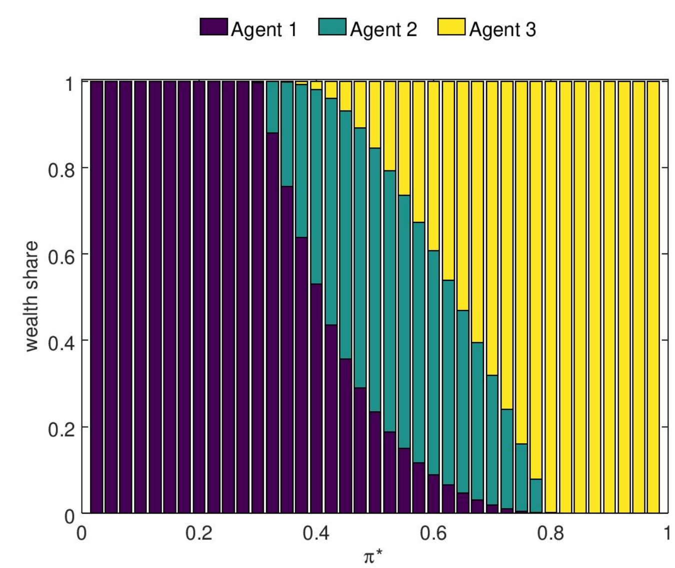
$$y_{2,1} \quad y_{2,2} \quad \cdots \quad y_{2,m} \qquad \vdots \qquad = \overline{Y}_2(w)$$

$$\vdots \quad \vdots \quad \vdots \quad \vdots \qquad \vdots \qquad \vdots$$

$$y_{n,1} \quad y_{n,2} \quad \cdots \quad y_{n,m} \qquad \vdots \qquad = \overline{Y}_n(w)$$

$$\sum_{i=1}^{n} \frac{\overline{Y}_i(w)}{n} = \overline{Y}(w) \approx E[Y] = \lim_{t \to \infty} E[Y_t]$$

Agents wealth at steady state



Steady-State Analysis by autoRD: Market Selection

Simple repeated betting market from Kets et al, AAAI 2014

- 1 event realises at every step with a fixed probability π^*
- 3 Fractional Kelly bettors with a belief on π^* and place bets accordingly

Warmup
w steps

$$y_{1,1} \quad y_{1,2} \quad \cdots \quad y_{1,m} \qquad \sum_{t=w+1}^{m} \frac{y_{1,t}}{m-w} = \overline{Y}_1(w)$$

$$y_{2,1} \quad y_{2,2} \quad \cdots \quad y_{2,m} \qquad \vdots \qquad = \overline{Y}_2(w)$$

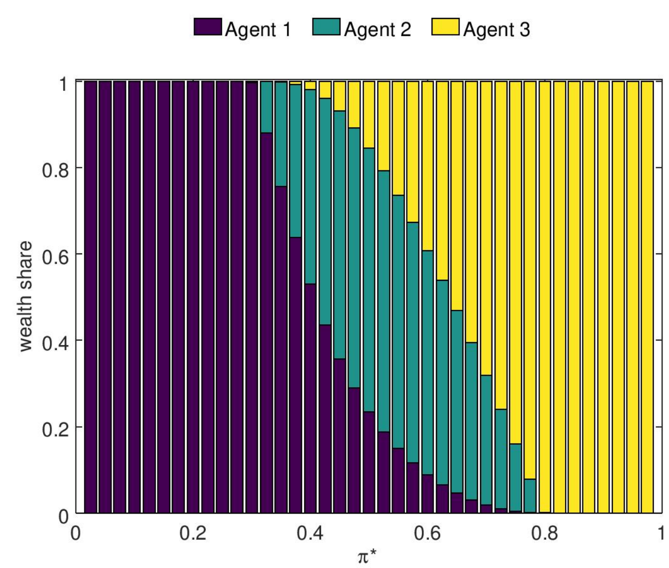
$$\vdots \quad \vdots \quad \vdots \quad \vdots \qquad \vdots \qquad \vdots$$

$$y_{n,1} \quad y_{n,2} \quad \cdots \quad y_{n,m} \qquad \vdots \qquad = \overline{Y}_n(w)$$

$$\sum_{i=1}^{n} \frac{\overline{Y}_i(w)}{n} = \overline{Y}(w) \approx E[Y] = \lim_{t \to \infty} E[Y_t]$$

Replication and deletion (RD), [Law,Kelton2015]

Agents wealth at steady state



Steady-State Analysis by autoRD: Market Selection

Simple repeated betting market from Kets et al, AAAI 2014

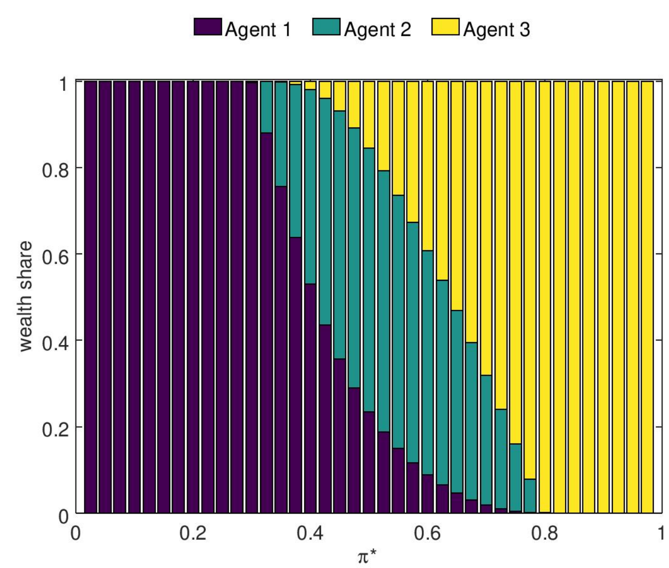
- 1 event realises at every step with a fixed probability π^*
- 3 Fractional Kelly bettors with a belief on π^* and place bets accordingly

Warmup
w steps

$$\begin{matrix} y_{1,1} & y_{1,2} & \cdots & y_{1,m} & \sum_{t=w+1}^{m} \frac{y_{1,t}}{m-w} & = \overline{Y}_1(w) \\ y_{2,1} & y_{2,2} & \cdots & y_{2,m} & \vdots & = \overline{Y}_2(w) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{n,1} & y_{n,2} & \cdots & y_{n,m} & \vdots & = \overline{Y}_n(w) \end{matrix}$$

$$\sum_{i=1}^{n} \frac{\overline{Y}_i(w)}{n} = \overline{Y}(w) \approx E[Y] = \lim_{t \to \infty} E[Y_t]$$

Agents wealth at steady state



Replication and deletion (RD), [Law,Kelton2015]

What are the correct values for w, m, n?

Steady-State Analysis by autoRD: Market Selection

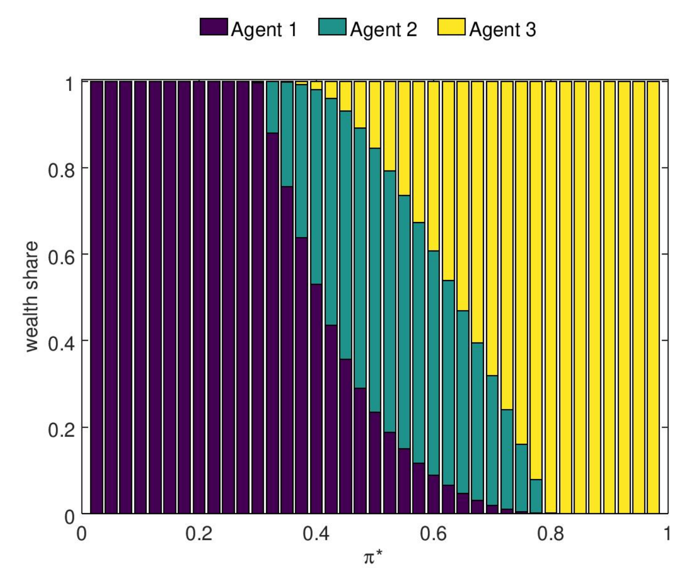
Simple repeated betting market from Kets et al, AAAI 2014
- 1 event realises at every step with a fixed probability π^*
- 3 Fractional Kelly bettors with a belief on π^* and place bets accordingly

Warmup w steps

$$
\begin{matrix}
y_{1,1} & y_{1,2} & \cdots & y_{1,m} & \sum_{t=w+1}^{m} \frac{y_{1,t}}{m-w} = \bar{Y}_1(w) \\
y_{2,1} & y_{2,2} & \cdots & y_{2,m} & \vdots \quad = \bar{Y}_2(w) \\
\vdots & \vdots & \vdots & \vdots & \vdots \quad \vdots \\
y_{n,1} & y_{n,2} & \cdots & y_{n,m} & \vdots \quad = \bar{Y}_n(w)
\end{matrix}
$$

$$\sum_{i=1}^{n} \frac{\bar{Y}_i(w)}{n} = \bar{Y}(w) \approx E[Y] = \lim_{t\to\infty} E[Y_t]$$

Agents wealth at steady state



Replication and deletion (RD), [Law,Kelton2015]

What are the correct values for w, m, n?

THESE ARE DIFFICULT QUESTIONS

ARE THEY CRUCIAL?

Steady-State Analysis by autoRD: Market Selection

Simple repeated betting market from Kets et al, AAAI 2014

- 1 event realises at every step with a fixed probability π^*
- 3 Fractional Kelly bettors with a belief on π^* and place bets accordingly

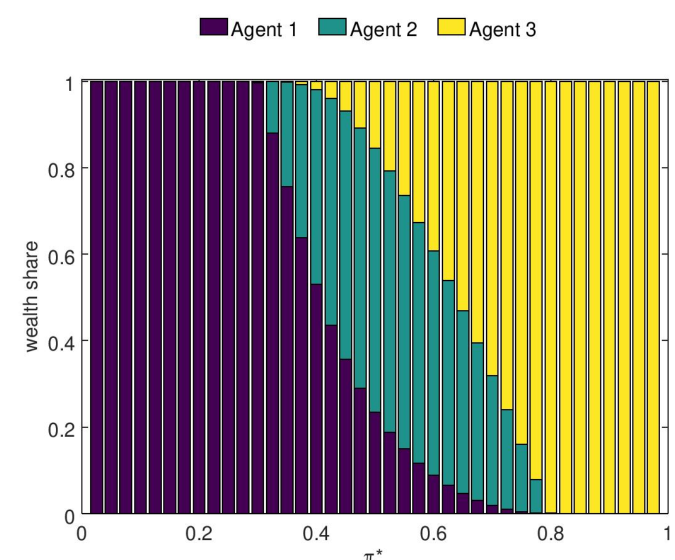
Warmup
w steps

$$
\begin{matrix}
y_{1,1} & y_{1,2} & \cdots & y_{1,m} & \sum_{t=w+1}^{m} \frac{y_{1,t}}{m-w} = \overline{Y}_1(w) \\
y_{2,1} & y_{2,2} & \cdots & y_{2,m} & \vdots \quad = \overline{Y}_2(w) \\
\vdots & \vdots & \vdots & \vdots & \vdots \quad \vdots \\
y_{n,1} & y_{n,2} & \cdots & y_{n,m} & \vdots \quad = \overline{Y}_n(w)
\end{matrix}
$$

$$\sum_{i=1}^{n} \frac{\overline{Y}_i(w)}{n} = \overline{Y}(w) \approx E[Y] = \lim_{t\to\infty} E[Y_t]$$

Replication and deletion (RD), [Law,Kelton2015]

Agents wealth at steady state



Wrong by [Kets et al AAAI2014]

What are the correct values for w, m, n?

THESE ARE DIFFICULT QUESTIONS

ARE THEY CRUCIAL?

YES

Steady-State Analysis by autoRD: Market Selection

Simple repeated betting market from Kets et al, AAAI 2014

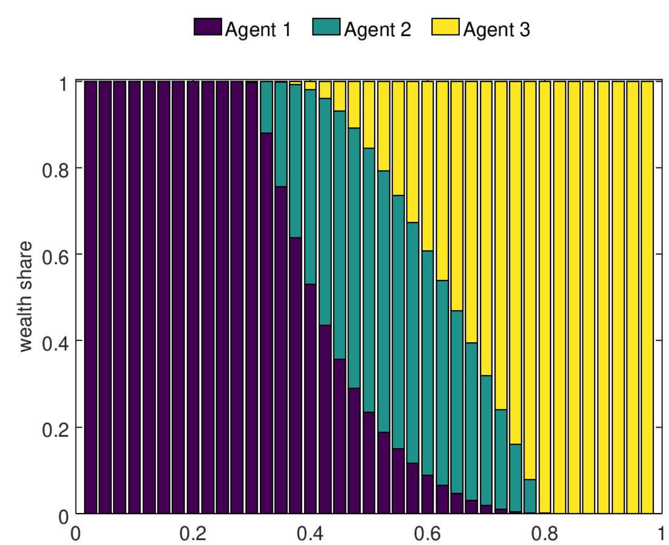
- 1 event realises at every step with a fixed probability π^*
- 3 Fractional Kelly bettors with a belief on π^* and place bets accordingly

Warmup w steps

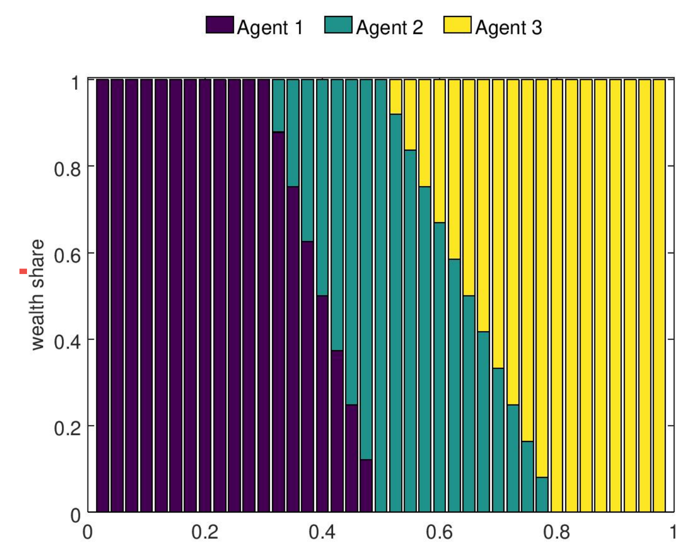
$$\begin{matrix} y_{1,1} & y_{1,2} & \cdots & y_{1,m} & \sum_{t=w+1}^{m} \frac{y_{1,t}}{m-w} = \bar{Y}_1(w) \\ y_{2,1} & y_{2,2} & \cdots & y_{2,m} & = \bar{Y}_2(w) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{n,1} & y_{n,2} & \cdots & y_{n,m} & = \bar{Y}_n(w) \end{matrix}$$

$$\sum_{i=1}^{n} \frac{\bar{Y}_i(w)}{n} = \bar{Y}(w) \approx E[Y] = \lim_{t\to\infty} E[Y_t]$$

Agents wealth at steady state



Wrong by [Kets et al AAAI2014]



Correct by MultiVeStA
Same as analytical solution from [Bottazzi, Giachini, Quantitative Finance 2019]

Replication and deletion (RD), [Law,Kelton2015]

What are the correct values for w, m, n?

THESE ARE DIFFICULT QUESTIONS

ARE THEY CRUCIAL?

YES

Steady-State Analysis by autoRD: Market Selection

Simple repeated betting market from Kets et al, AAAI 2014

- 1 event realises at every step with a fixed probability π^*
- 3 Fractional Kelly bettors with a belief on π^* and place bets accordingly

Warmup
w steps

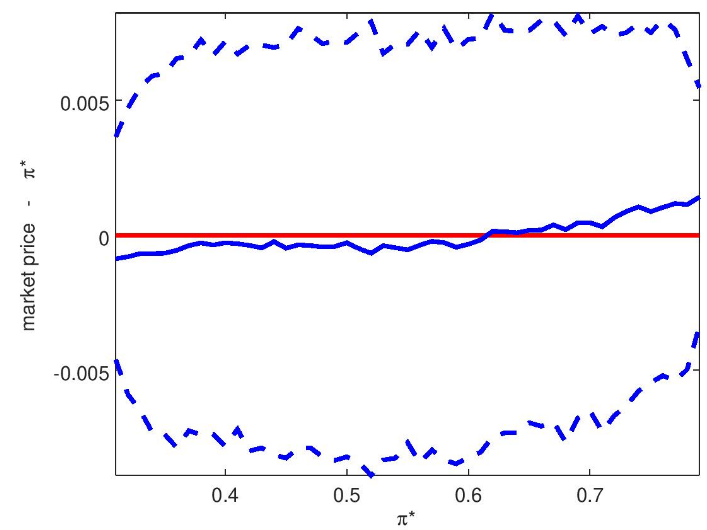
$$\begin{array}{ccccc} y_{1,1} & y_{1,2} & \cdots & y_{1,m} & \sum_{t=w+1}^{m} \frac{y_{1,t}}{m-w} = \overline{Y}_1(w) \\ y_{2,1} & y_{2,2} & \cdots & y_{2,m} & = \overline{Y}_2(w) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{n,1} & y_{n,2} & \cdots & y_{n,m} & = \overline{Y}_n(w) \end{array}$$

$$\sum_{i=1}^{n} \frac{\overline{Y}_i(w)}{n} = \overline{Y}(w) \approx E[Y] = \lim_{t \to \infty} E[Y_t]$$

Replication and deletion

What are the correct values for w, m, n?

Does the market price match π^*?



Wrong CI on wrong w, m, n
from [Kets et al AAAI2014]

THESE ARE DIFFICULT QUESTIONS
ARE THEY CRUCIAL?
YES

Steady-State Analysis by autoRD: Market Selection

Simple repeated betting market from Kets et al, AAAI 2014

- 1 event realises at every step with a fixed probability π^*
- 3 Fractional Kelly bettors with a belief on π^* and place bets accordingly

Warmup
w steps

$$\begin{matrix} y_{1,1} & y_{1,2} & \cdots & y_{1,m} \\ y_{2,1} & y_{2,2} & \cdots & y_{2,m} \\ \vdots & \vdots & \vdots & \vdots \\ y_{n,1} & y_{n,2} & \cdots & y_{n,m} \end{matrix} \qquad \sum_{t=w+1}^{m} \frac{y_{1,t}}{m-w} = \overline{Y}_1(w), \quad \vdots = \overline{Y}_2(w), \quad \vdots, \quad \vdots = \overline{Y}_n(w)$$

$$\sum_{i=1}^{n} \frac{\overline{Y}_i(w)}{n} = \overline{Y}(w) \approx E[Y] = \lim_{t\to\infty} E[Y_t]$$

Replication and deletion

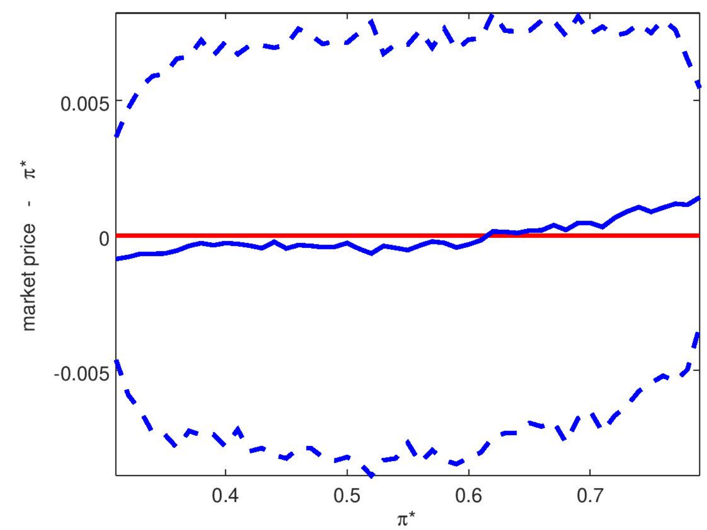
What are the correct values for w, m, n?

THESE ARE DIFFICULT QUESTIONS

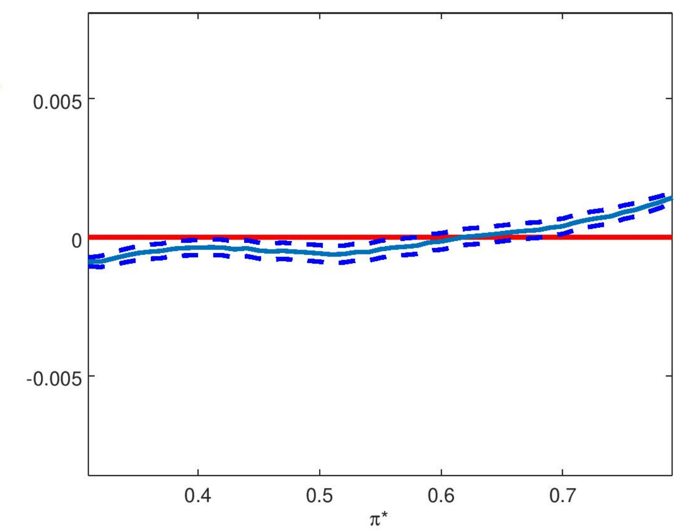
ARE THEY CRUCIAL?

YES

Does the market price match π^*?



Wrong CI on wrong w, m, n
from [Kets et al AAAI2014]



Correct CI on wrong w, m, n
from [Kets et al AAAI2014]

Steady-State Analysis by autoRD: Market Selection

Simple repeated betting market from Kets et al, AAAI 2014
- 1 event realises at every step with a fixed probability π^*
- 3 Fractional Kelly bettors with a belief on π^* and place bets accordingly

Warmup
w steps

$$\begin{matrix} y_{1,1} & y_{1,2} & \cdots & y_{1,m} & \sum_{t=w+1}^{m} \frac{y_{1,t}}{m-w} = \overline{Y}_1(w) \\ y_{2,1} & y_{2,2} & \cdots & y_{2,m} & = \overline{Y}_2(w) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ y_{n,1} & y_{n,2} & \cdots & y_{n,m} & = \overline{Y}_n(w) \end{matrix}$$

$$\sum_{i=1}^{n} \frac{\overline{Y}_i(w)}{n} = \overline{Y}(w) \approx E[Y] = \lim_{t \to \infty} E[Y_t]$$

Replication and deletion

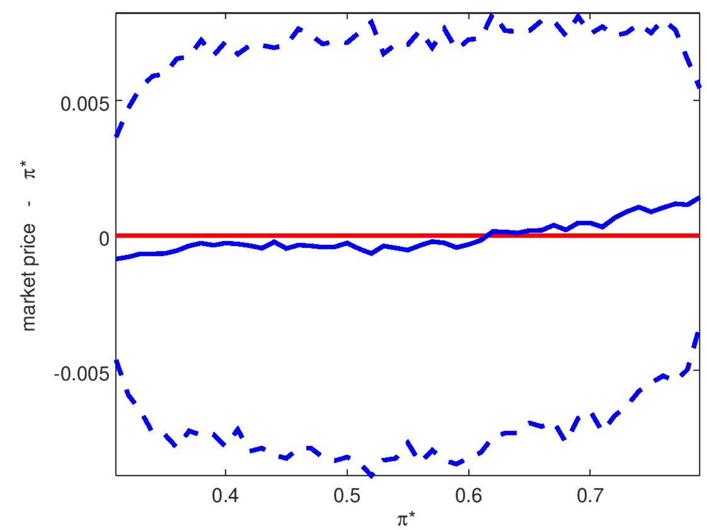
What are the correct values for w, m, n?

THESE ARE DIFFICULT QUESTIONS

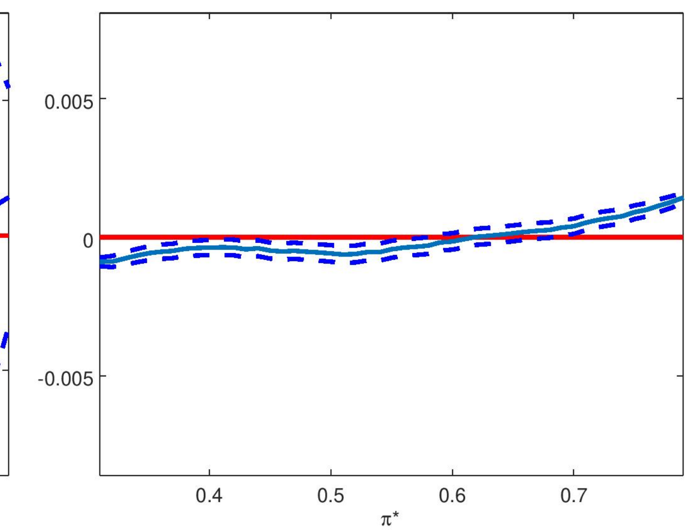
ARE THEY CRUCIAL?

YES

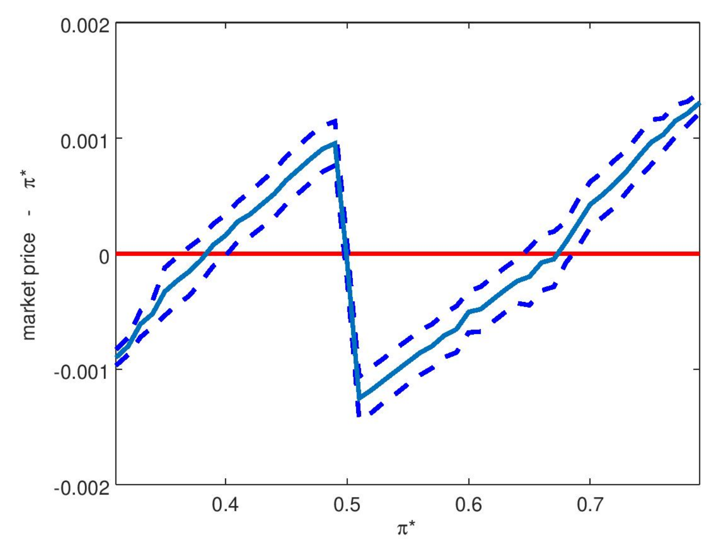
Does the market price match π^*?



Wrong CI on wrong w, m, n
from [Kets et al AAAI2014]



Correct CI on wrong w, m, n
from [Kets et al AAAI2014]



Correct by MultiVeStA
Same as analytical solution
from [Bottazzi, Giachini, Quantitative Finance 2019]

OUTLINE

1. Motivation, vision, and proposal
    1. Automated analysis with statistical guarantees for ABMs
    2. The MultiVeStA Statistical Model Checker
2. Transient Analysis of a large-scale financial macro ABM
    1. Estimation of expected outcome and Confidence Interval
    2. Counterfactual analysis for different model configurations
3. **Steady-state analysis of a prediction market model**
    1. Steady-state analysis by Replication and Deletion (RD)
    2. **Warmup estimation**
    3. Steady-state analysis by Batch Means (BM)
    4. A methodology for ergodicity analysis based on RD and BM
4. Conclusions & Future works

Warmup Estimation in the ABM Community

Typical approach used in the ABM community
Based on Welch's graphical method [Welch1983]

$$\begin{matrix} y_{1,1} & y_{1,2} & \cdots & y_{1,m} \\ y_{2,1} & y_{2,2} & \cdots & y_{2,m} \\ \vdots & \vdots & \vdots & \vdots \\ y_{n,1} & y_{n,2} & \cdots & y_{n,m} \end{matrix}$$

1. Do n simulations of a *given large length* m
2.
3.
4.1.
4.2.

Warmup Estimation in the ABM Community

Typical approach used in the ABM community
Based on Welch's graphical method [Welch1983]

$$\begin{matrix} y_{1,1} & y_{1,2} & \cdots & y_{1,m} \\ y_{2,1} & y_{2,2} & \cdots & y_{2,m} \\ \vdots & \vdots & \vdots & \vdots \\ y_{n,1} & y_{n,2} & \cdots & y_{n,m} \end{matrix}$$

1. Do n simulations of a *given large length* m
2. Compute $\overline{Y}_t$ for each $t \in [1,m]$
3.

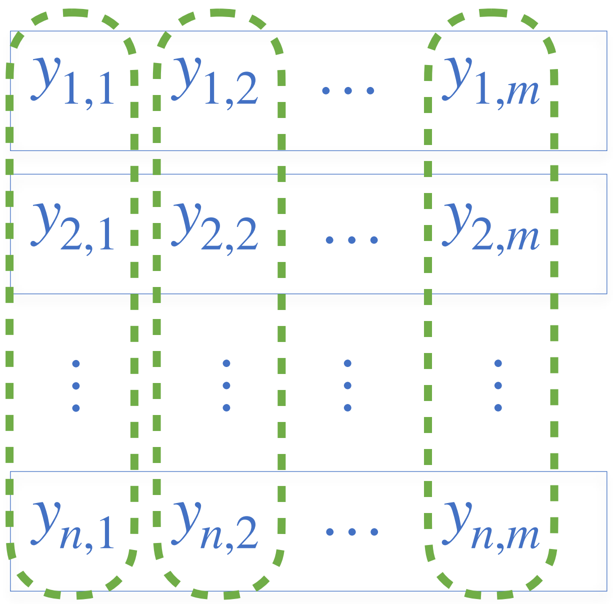
4.1.

4.2.

$$\sum_{i=1}^{n} \frac{y_{i,t}}{n} = \overline{Y}_t$$

Warmup Estimation in the ABM Community

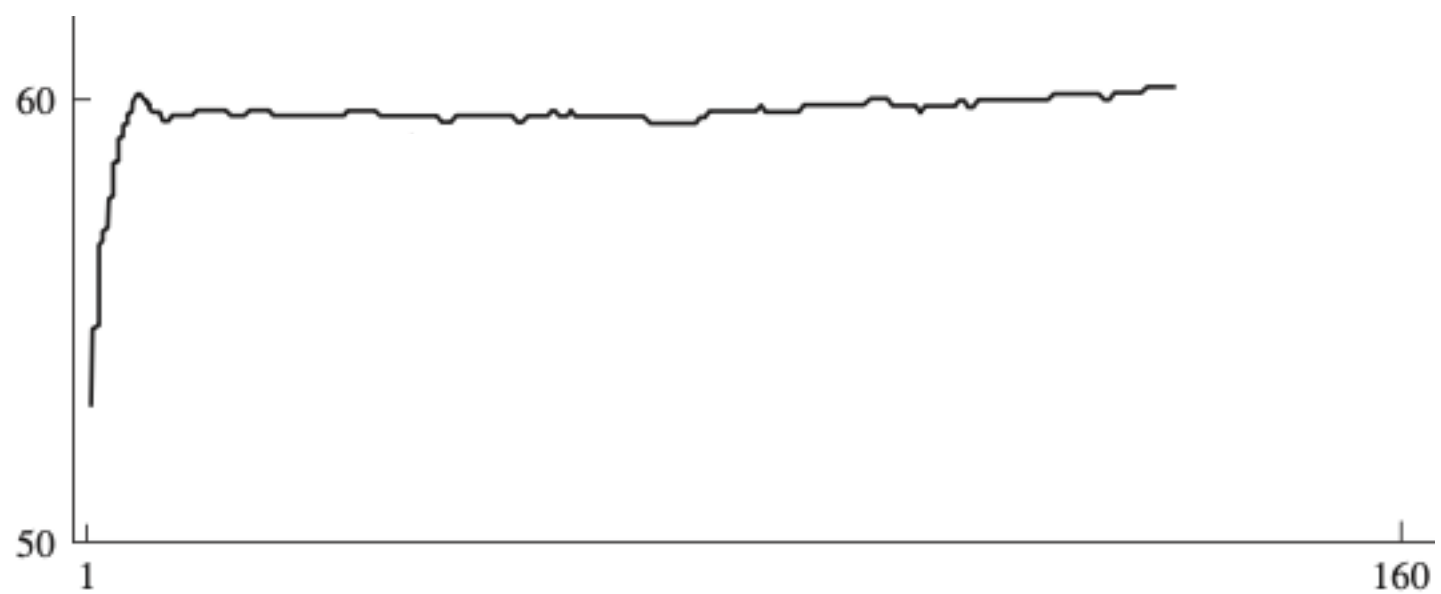
Typical approach used in the ABM community
Based on Welch's graphical method [Welch1983]



1. Do n simulations of a *given large length* m
2. Compute $\overline{Y}_t$ for each $t \in [1,m]$
3. Plot all $\overline{Y}_t$ Can smooth by averaging with a window of neighbour time points

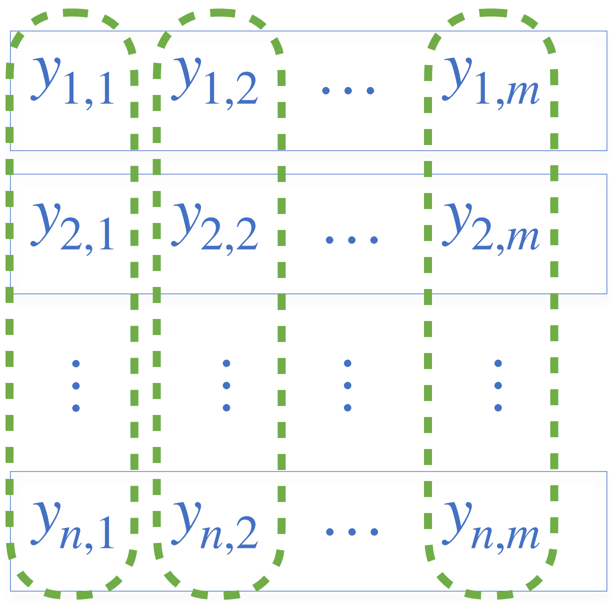
4.1.

4.2.



Example from a reference book on simulation [Law,Kelton2015]

Warmup Estimation in the ABM Community

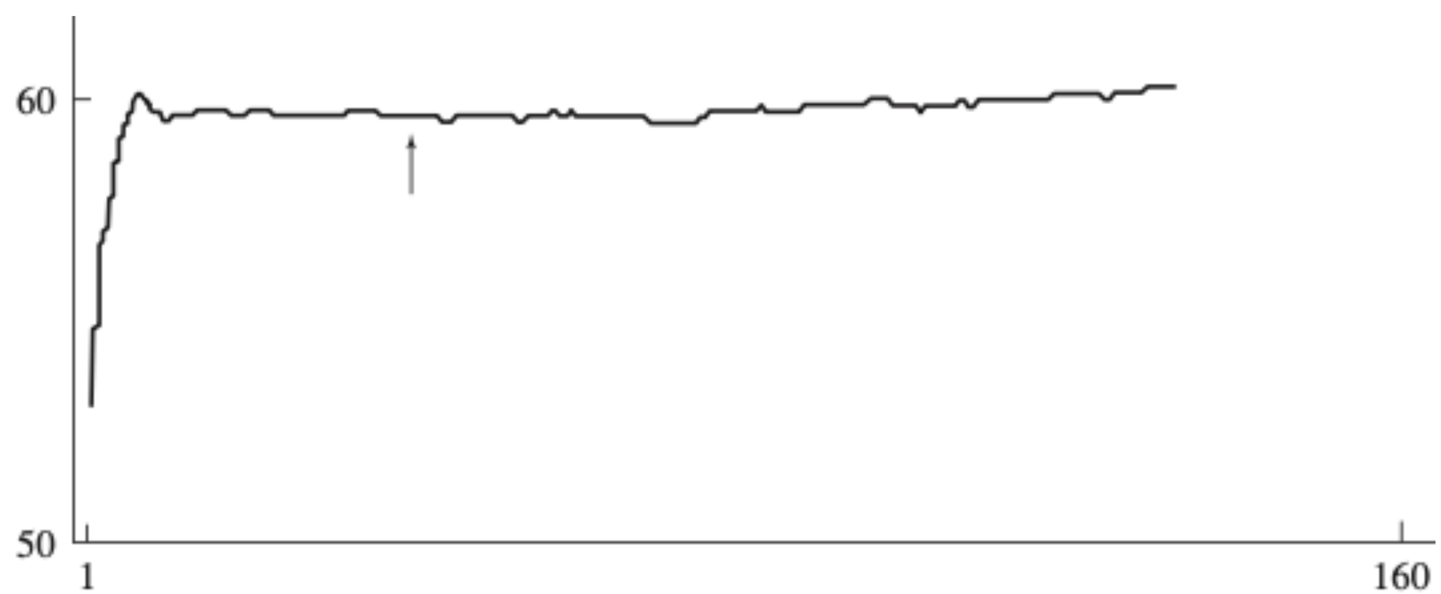


Typical approach used in the ABM community
Based on Welch's graphical method [Welch1983]

1. Do n simulations of a *given large length* m
2. Compute $\overline{Y}_t$ for each $t \in [1,m]$
3. Plot all $\overline{Y}_t$ Can smooth by averaging with a window of neighbour time points

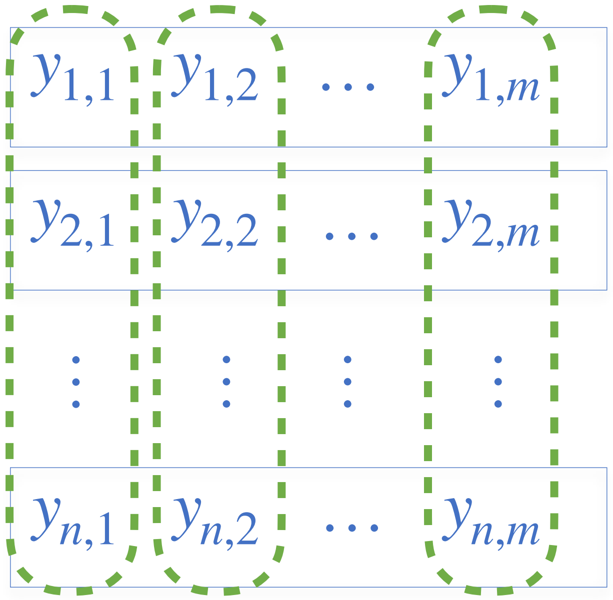
4.1. Choose the time point w after which the plot *seems to converge*

4.2.



Example from a reference book on simulation [Law,Kelton2015]

Warmup Estimation in the ABM Community

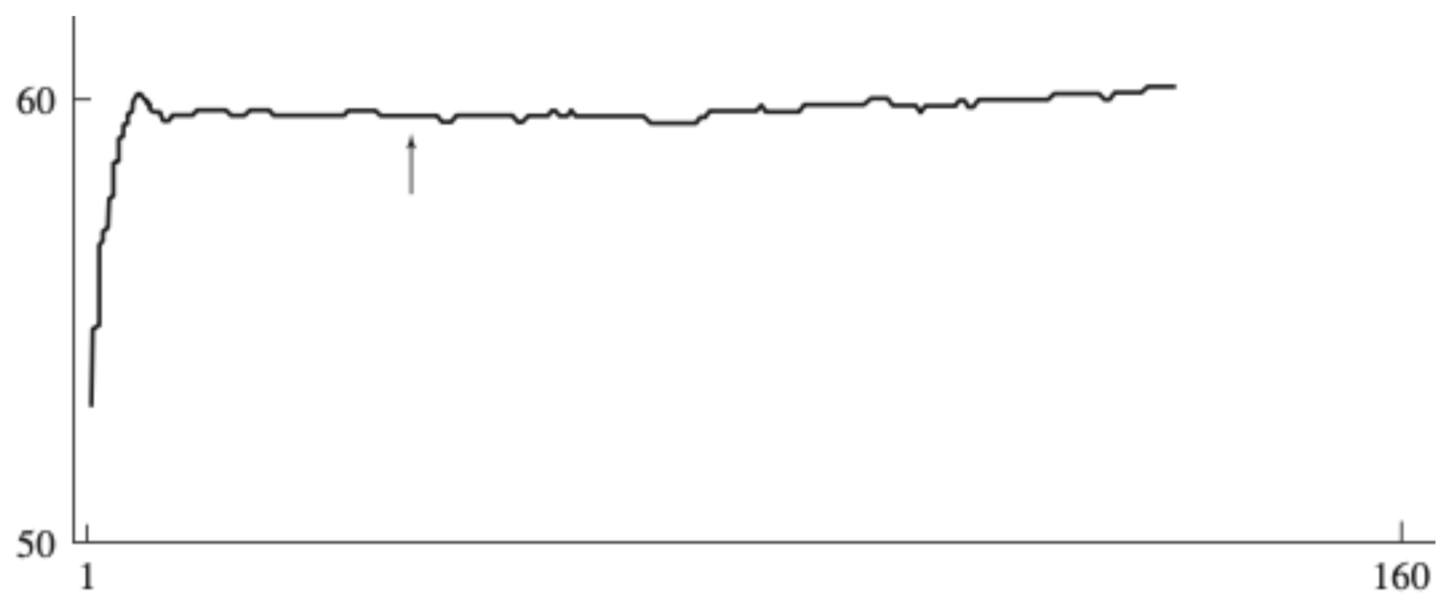


Typical approach used in the ABM community
Based on Welch's graphical method [Welch1983]

1. Do n simulations of a *given large length* m
2. Compute $\overline{Y}_t$ for each $t \in [1,m]$
3. Plot all $\overline{Y}_t$ Can smooth by averaging with a window of neighbour time points

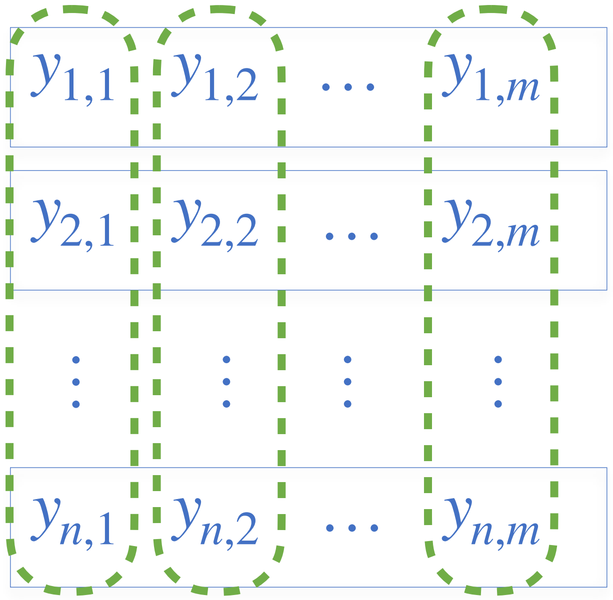
4.1. Choose the time point w after which the plot *seems to converge*

4.2. If no such point exists, go back to point 1 and compute n new simulations



Example from a reference book on simulation [Law,Kelton2015]

Warmup Estimation in the ABM Community



Typical approach used in the ABM community
Based on Welch's graphical method [Welch1983]

1. Do n simulations of a *given large length* m
2. Compute $\overline{Y}_t$ for each $t \in [1,m]$
3. Plot all $\overline{Y}_t$ Can smooth by averaging with a window of neighbour time points

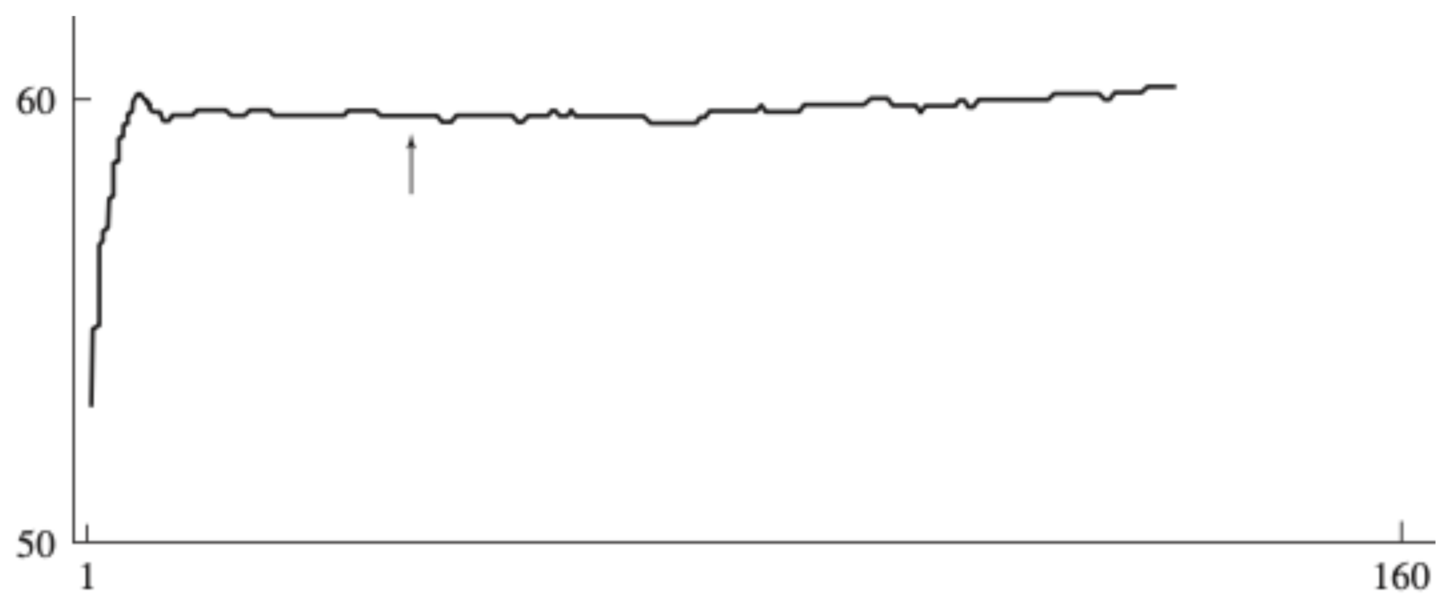
4.1. Choose the time point w after which the plot *seems to converge*

4.2. If no such point exists, go back to point 1 and compute n new simulations

No statistical guarantees

Not automatic:

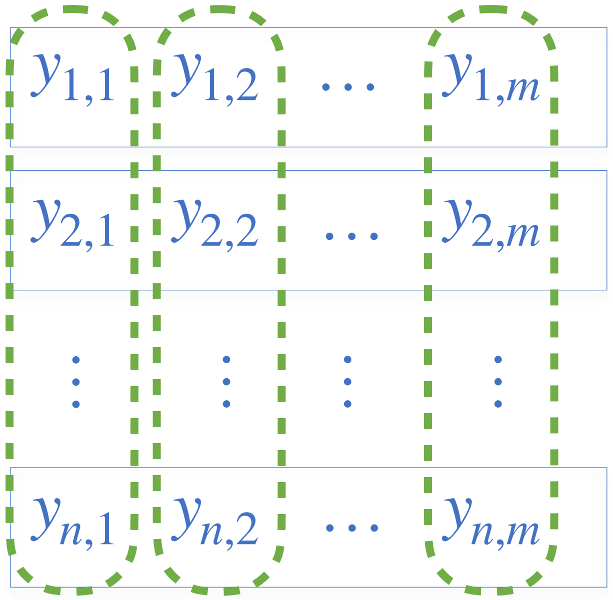
- **Depends on chosen parameters**
- **Decision is human-made**



Example from a reference book on simulation [Law,Kelton2015]

Warmup Estimation in the ABM Community

Typical approach used in the ABM community
Based on Welch's graphical method [Welch1983]



1. Do n simulations of a *given large length* m
2. Compute $\overline{Y}_t$ for each $t \in [1,m]$
3. Plot all $\overline{Y}_t$ Can smooth by averaging with a window of neighbour time points

4.1. Choose the time point w after which the plot *seems to converge*

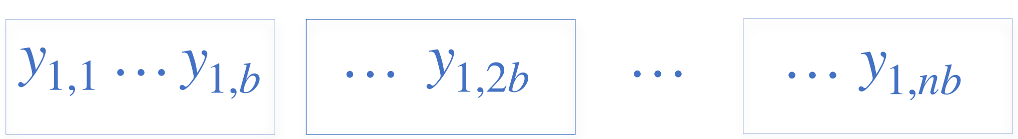
4.2. If no such point exists, go back to point 1 and compute n new simulations

No statistical guarantees

Not automatic:

- **Depends on chosen parameters**
- **Decision is human-made**

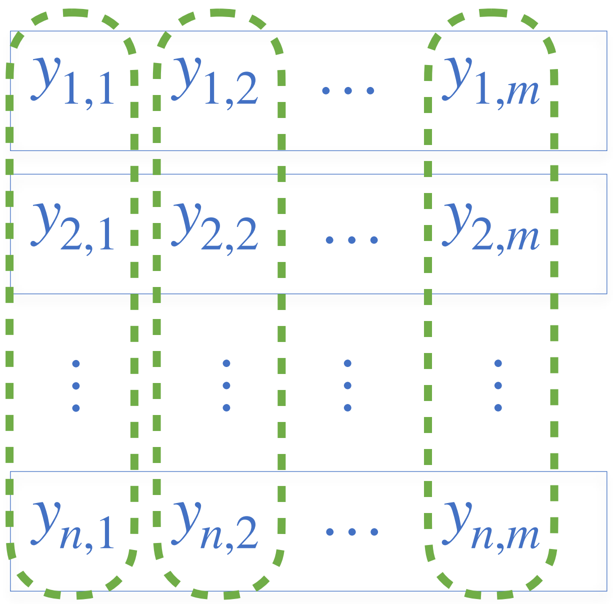
A more recent approach by the ABM community [Grazzini2012]



1. Do *1 long simulation* of a *given large length* m
2. Divide it in a *given number wi of* windows of consecutive steps
3.
4.
5.

Warmup Estimation in the ABM Community

Typical approach used in the ABM community
Based on Welch's graphical method [Welch1983]



1. Do n simulations of a *given large length* m
2. Compute $\overline{Y}_t$ for each $t \in [1,m]$
3. Plot all $\overline{Y}_t$ Can smooth by averaging with a window of neighbour time points

4.1. Choose the time point w after which the plot *seems to converge*

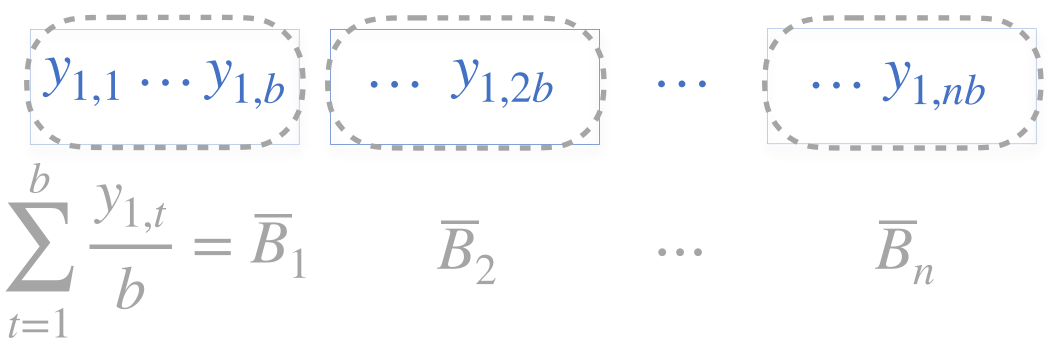
4.2. If no such point exists, go back to point 1 and compute n new simulations

No statistical guarantees

Not automatic:

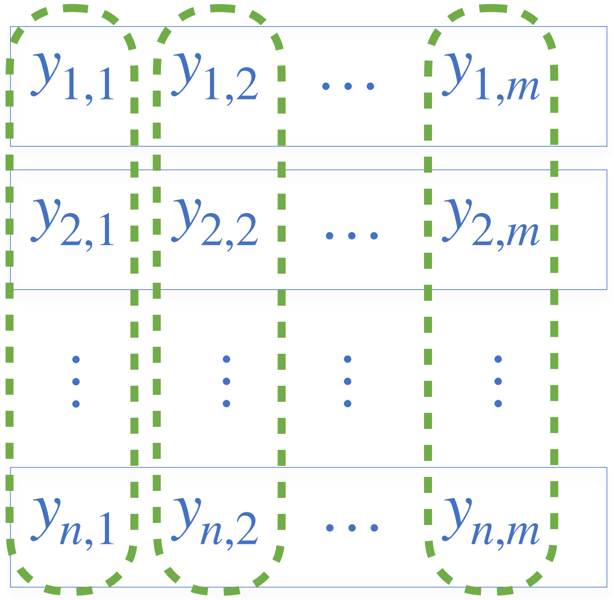
- **Depends on chosen parameters**
- **Decision is human-made**

A more recent approach by the ABM community [Grazzini2012]



1. Do *1 long simulation* of a *given large length* m
2. Divide it in a *given number wi of* windows of consecutive steps
3. Compute the mean $\overline{B}_i$ within each window
4.
5.

Warmup Estimation in the ABM Community



Typical approach used in the ABM community
Based on Welch's graphical method [Welch1983]

1. Do n simulations of a *given large length* m
2. Compute $\overline{Y}_t$ for each $t \in [1,m]$
3. Plot all $\overline{Y}_t$ Can smooth by averaging with a window of neighbour time points

4.1. Choose the time point w after which the plot *seems to converge*

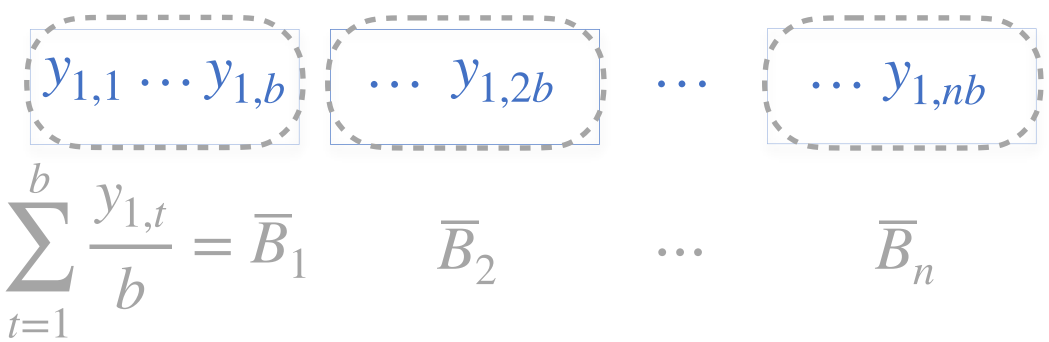
4.2. If no such point exists, go back to point 1 and compute n new simulations

No statistical guarantees

Not automatic:

- **Depends on chosen parameters**
- **Decision is human-made**

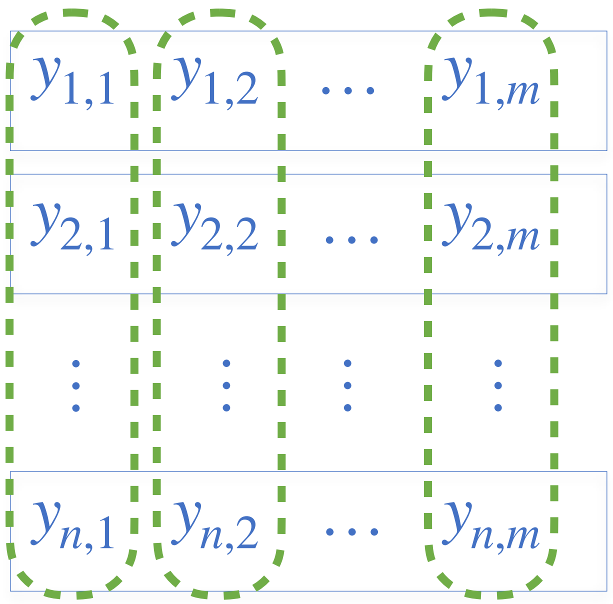
A more recent approach by the ABM community [Grazzini2012]



1. Do *1 long simulation* of a *given large length* m
2. Divide it in a *given number wi of* windows of consecutive steps
3. Compute the mean $\overline{B}_i$ within each window
4. Perform a randomness test on the computed means
5.

Warmup Estimation in the ABM Community

Typical approach used in the ABM community
Based on Welch's graphical method [Welch1983]



1. Do n simulations of a *given large length* m
2. Compute $\overline{Y}_t$ for each $t \in [1,m]$
3. Plot all $\overline{Y}_t$ Can smooth by averaging with a window of neighbour time points

4.1. Choose the time point w after which the plot *seems to converge*

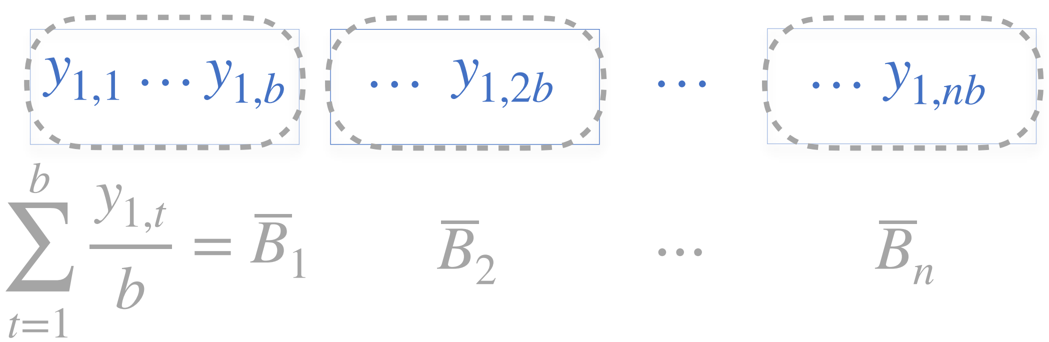
4.2. If no such point exists, go back to point 1 and compute n new simulations

No statistical guarantees

Not automatic:

- **Depends on chosen parameters**
- **Decision is human-made**

A more recent approach by the ABM community [Grazzini2012]



1. Do *1 long simulation* of a *given large length* m
2. Divide it in a *given number wi of* windows of consecutive steps
3. Compute the mean $\overline{B}_i$ within each window
4. Perform a randomness test on the computed means
5. If the test passes, then the transient is completed

Warmup Estimation in the ABM Community

Typical approach used in the ABM community
Based on Welch's graphical method [Welch1983]

$$\begin{array}{cccc} y_{1,1} & y_{1,2} & \cdots & y_{1,m} \\ y_{2,1} & y_{2,2} & \cdots & y_{2,m} \\ \vdots & \vdots & \vdots & \vdots \\ y_{n,1} & y_{n,2} & \cdots & y_{n,m} \end{array}$$

1. Do n simulations of a *given large length m*
2. Compute $\overline{Y}_t$ for each $t \in [1,m]$
3. Plot all $\overline{Y}_t$ Can smooth by averaging with a window of neighbour time points

4.1. Choose the time point w after which the plot *seems to converge*

4.2. If no such point exists, go back to point 1 and compute n new simulations

No statistical guarantees

Not automatic:

- **Depends on chosen parameters**
- **Decision is human-made**

A more recent approach by the ABM community [Grazzini2012]

$$y_{1,1} \cdots y_{1,b} \qquad \cdots\; y_{1,2b} \qquad \cdots \qquad \cdots\; y_{1,nb}$$

$$\sum_{t=1}^{b} \frac{y_{1,t}}{b} = \overline{B}_1 \qquad \overline{B}_2 \qquad \cdots \qquad \overline{B}_n$$

1. Do *1 long simulation* of a *given large length m*
2. Divide it in a *given number wi of* windows of consecutive steps
3. Compute the mean $\overline{B}_i$ within each window
4. Perform a randomness test on the computed means
5. If the test passes, then the transient is completed

Has statistical guarantees

Not automatic:

- ***"with appropriate settings the tests can detect non-stationarity"***
- **Decision depends on chosen parameters**

Warmup Estimation in the ABM Community

This is a **rediscovery** of part of the Batch Means (BM) method

- First proposal in [Conway1963]
- First automatic version in [Law,Carson1979]
- Approach for steady state analysis
  - Alternative to Replication and Deletion based on 1 long simulation

A more recent approach by the ABM community [Grazzini2012]

$$y_{1,1} \cdots y_{1,b} \quad \cdots\ y_{1,2b} \quad \cdots \quad \cdots\ y_{1,nb}$$

$$\sum_{t=1}^{b} \frac{y_{1,t}}{b} = \overline{B}_1 \qquad \overline{B}_2 \qquad \cdots \qquad \overline{B}_n$$

1. Do *1 long simulation* of a ***given large length*** m
2. Divide it in a ***given number wi*** *of* windows of consecutive steps
3. Compute the mean $\overline{B}_i$ within each window
4. Perform a randomness test on the computed means
5. If the test passes, then the transient is completed

Has statistical guarantees

Not automatic:

- ***"with appropriate settings the tests can detect non-stationarity"***
- **Decision depends on chosen parameters**

Warmup Estimation in the ABM Community

This is a rediscovery of part of the Batch Means (BM) method

- First proposal in [Conway1963]
- First automatic version in [Law,Carson1979]
- Approach for steady state analysis
  - Alternative to Replication and Deletion based on 1 long simulation

Our automated warmup estimation procedure builds on BM-related results

- We also propose a simple novel version of BM for steady-state analysis
- Based on [Law,Carson1979] [Steiger et al 2005] [Tafazzoli et al 2011] [Gilmore et al 2017]

A more recent approach by the ABM community [Grazzini2012]

$y_{1,1} \cdots y_{1,b}$ | $\cdots\ y_{1,2b}$ | $\cdots$ | $\cdots\ y_{1,nb}$

$\sum_{t=1}^{b} \frac{y_{1,t}}{b} = \overline{B}_1 \quad \overline{B}_2 \quad \cdots \quad \overline{B}_n$

1. Do *1 long simulation* of a ***given large length m***
2. Divide it in a ***given number wi*** *of* windows of consecutive steps
3. Compute the mean $\overline{B}_i$ within each window
4. Perform a randomness test on the computed means
5. If the test passes, then the transient is completed

Has statistical guarantees

Not automatic:

- ***"with appropriate settings the tests can detect non-stationarity"***
- **Decision depends on chosen parameters**

Warmup Estimation by autoWarmup: our Automated Proposal

1. Do *1 long simulation* of length *m*
2. Divide it in *B batches* of consecutive steps
3. Compute the mean $\overline{B}_i$ within each batch

$$y_{1,1} \cdots y_{1,b} \quad \cdots\ y_{1,2b} \quad \cdots \quad \cdots\ y_{1,Bb}$$

$$\sum_{t=1}^{b} \frac{y_{1,t}}{b} = \overline{B}_1 \qquad \overline{B}_2 \qquad \cdots \qquad \overline{B}_B$$

Warmup Estimation by autoWarmup: our Automated Proposal

0. Set *m*= *B***b*,
   B=128 is the number of simulation batches
   b=16 is the number of steps in each batch

$y_{1,1} \cdots y_{1,b}$ | $\cdots\ y_{1,2b}$ | $\cdots$ | $\cdots\ y_{1,Bb}$

$$\sum_{t=1}^{b} \frac{y_{1,t}}{b} = \overline{B}_1 \qquad \overline{B}_2 \qquad \cdots \qquad \overline{B}_B$$

1. Do *1 long simulation* of length *m*
2. Divide it in *B batches* of consecutive steps
3. Compute the mean $\overline{B}_i$ within each batch

Warmup Estimation by autoWarmup: our Automated Proposal

0. Set *m*= *B***b*,
   B=128 is the number of simulation batches
   b=16 is the number of steps in each batch

16 $y_{1,1} \cdots y_{1,b}$ 16 $\cdots\ y_{1,2b}$ $\cdots$ 16 $\cdots\ y_{1,Bb}$

$$\sum_{t=1}^{b} \frac{y_{1,t}}{b} = \overline{B}_1 \qquad \overline{B}_2 \qquad \cdots \qquad \overline{B}_B$$

1. Do *1 long simulation* of length *m*
2. Divide it in *B batches* of consecutive steps
3. Compute the mean $\overline{B}_i$ within each batch

Warmup Estimation by autoWarmup: our Automated Proposal

0. Set *m*= *B***b*,
   - *B*=128 is the number of simulation batches
   - *b*=16 is the number of steps in each batch

16	16		16
$y_{1,1} \cdots y_{1,b}$	$\cdots\ y_{1,2b}$	$\cdots$	$\cdots\ y_{1,Bb}$
$\sum_{t=1}^{b} \frac{y_{1,t}}{b} = \overline{B}_1$	$\overline{B}_2$	$\cdots$	$\overline{B}_B$

1. Do *1 long simulation* of length *m*
2. Divide it in *B batches* of consecutive steps
3. Compute the mean $\overline{B}_i$ within each batch
4. Perform statistical tests to check if *m* is large enough
   - Discard the first 4 batches
   - Perform a normality test on the computed means
   - Check for low lag-1 autocorrelation on the means

Warmup Estimation by autoWarmup: our Automated Proposal

0. Set *m*= *B***b*,
   - *B*=128 is the number of simulation batches
   - *b*=16 is the number of steps in each batch

16 $y_{1,1} \cdots y_{1,b}$ 16 $\cdots\ y_{1,2b}$ $\cdots$ 16 $\cdots\ y_{1,Bb}$

$\sum_{t=1}^{b} \frac{y_{1,t}}{b} = \overline{B}_1 \quad \overline{B}_2 \quad \cdots \quad \overline{B}_B$

1. Do *1 long simulation* of length *m*
2. Divide it in *B batches* of consecutive steps
3. Compute the mean $\overline{B}_i$ within each batch
4. Perform statistical tests to check if *m* is large enough
   - Discard the first 4 batches
   - Perform a normality test on the computed means
   - Check for low lag-1 autocorrelation on the means

5.1 If all tests pass, we conclude that the warmup has ended

Warmup Estimation by autoWarmup: our Automated Proposal

0. Set *m*= *B**b*,
   B=128 is the number of simulation batches
   b=16 is the number of steps in each batch

16 $y_{1,1} \cdots y_{1,b}$ 16 $\cdots y_{1,2b}$ $\cdots$ 16 $\cdots y_{1,Bb}$

$\sum_{t=1}^{b} \frac{y_{1,t}}{b} = \overline{B}_1$ $\overline{B}_2$ $\cdots$ $\overline{B}_B$

1. Do *1 long simulation* of length *m*
2. Divide it in *B batches* of consecutive steps
3. Compute the mean $\overline{B}_i$ within each batch
4. Perform statistical tests to check if *m* is large enough
   - Discard the first 4 batches
   - Perform a normality test on the computed means
   - Check for low lag-1 autocorrelation on the means

5.1 If all tests pass, we conclude that the warmup has ended

5.2 If one test fails
   - Double *b* squeezing the batches in the first *B/2* ones
   - Double *m* by performing *m* new simulation steps
   - Go back to step 3

Warmup Estimation by autoWarmup: our Automated Proposal

0. Set *m*= *B***b*,
   - *B*=128 is the number of simulation batches
   - *b*=16 is the number of steps in each batch

16 16 16

$$\underbrace{y_{1,1} \cdots y_{1,b}}_{16} \quad \underbrace{\cdots\ y_{1,2b}}_{16} \quad \cdots \quad \underbrace{\cdots\ y_{1,Bb}}_{16}$$

$$\sum_{t=1}^{b} \frac{y_{1,t}}{b} = \overline{B}_1 \qquad \overline{B}_2 \qquad \cdots \qquad \overline{B}_B$$

1. Do *1 long simulation* of length *m*
2. Divide it in *B batches* of consecutive steps
3. Compute the mean $\overline{B}_i$ within each batch
4. Perform statistical tests to check if *m* is large enough
   - Discard the first 4 batches
   - Perform a normality test on the computed means
   - Check for low lag-1 autocorrelation on the means

5.1 If all tests pass, we conclude that the warmup has ended

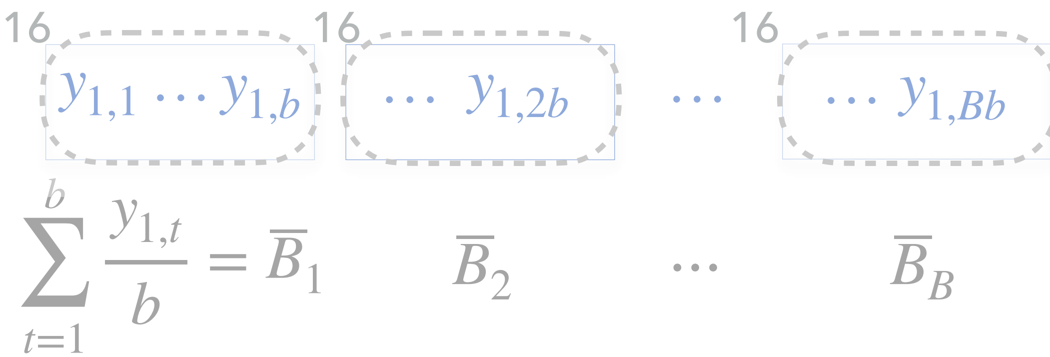
32 32

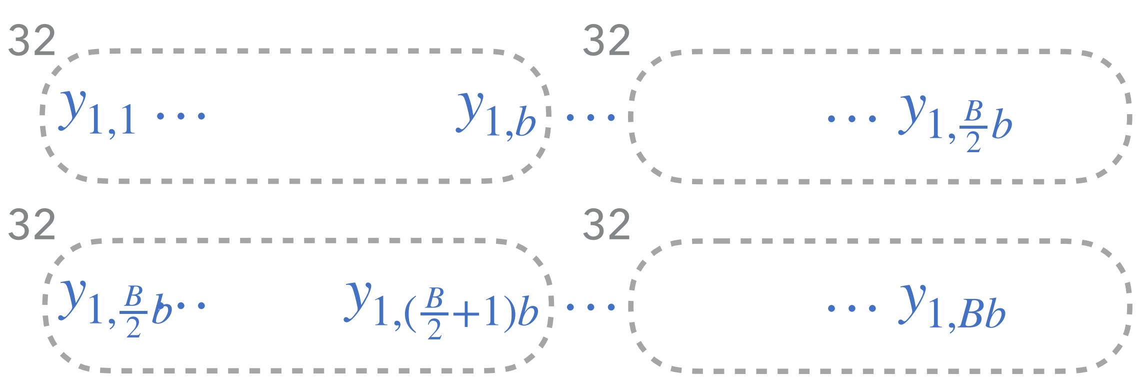
$$\underbrace{y_{1,1} \cdots \qquad y_{1,b}}_{32} \cdots \underbrace{\cdots\ y_{1,\frac{B}{2}b}}_{32}$$

5.2 If one test fails
   - Double *b* squeezing the batches in the first *B/2* ones
   - Double *m* by performing *m* new simulation steps
   - Go back to step 3

Warmup Estimation by autoWarmup: our Automated Proposal

0. Set *m*= *B**b*,
   B=128 is the number of simulation batches
   b=16 is the number of steps in each batch





1. Do *1 long simulation* of length *m*
2. Divide it in *B batches* of consecutive steps
3. Compute the mean $\overline{B}_i$ within each batch
4. Perform statistical tests to check if *m* is large enough
   Discard the first 4 batches
   Perform a normality test on the computed means
   Check for low lag-1 autocorrelation on the means

5.1 If all tests pass, we conclude that the warmup has ended

5.2 If one test fails
   Double *b* squeezing the batches in the first *B/2* ones
   Double *m* by performing *m* new simulation steps
   Go back to step 3

Steady State Analysis by autoBM: our BM-Based Proposal

0. Set *m*= *B***b*,
   B=128 is the number of simulation batches
   b=16 is the number of steps in each batch

$y_{1,1} \cdots y_{1,b}$ $\quad \cdots y_{1,2b}$ $\quad \cdots \quad \cdots y_{1,Bb}$

$$\sum_{t=1}^{b} \frac{y_{1,t}}{b} = \overline{B}_1 \qquad \overline{B}_2 \qquad \cdots \qquad \overline{B}_B$$

$$\sum_{j=l+1}^{n} \frac{\overline{B}_j}{n-l} = \overline{B}(l) \approx E[Y] = \lim_{t\to\infty} E[Y_t]$$

1. Do *1 long simulation* of length *m*
2. Divide it in *B batches* of consecutive steps
3. Compute the mean $\overline{B}_i$ within each batch
4. Perform statistical tests to check if *m* is large enough
   - Discard the first 4 batches
   - Perform a normality test on the computed means
   - Check for low lag-1 autocorrelation on the means

5.1 If all tests pass, we conclude that the warmup has ended
   - Compute the grand mean (mean of the means)
   - Compute the width *d* of the CI of grand mean
   - Adjust *d* according to the residual correlation in the means

5.2 If one test fails
   - Double *b* squeezing the batches in the first *B/2* ones
   - Double *m* by performing *m* new simulation steps
   - Go back to step 3

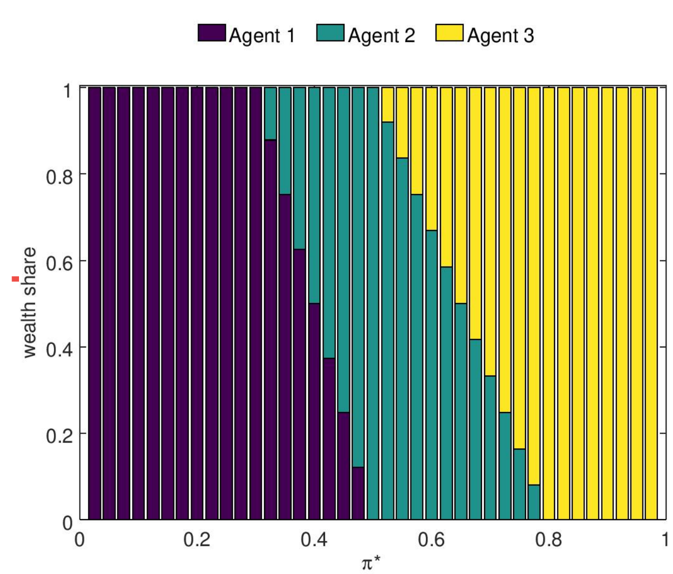
Steady-State Analysis: How To Do It in MultiVeStA?

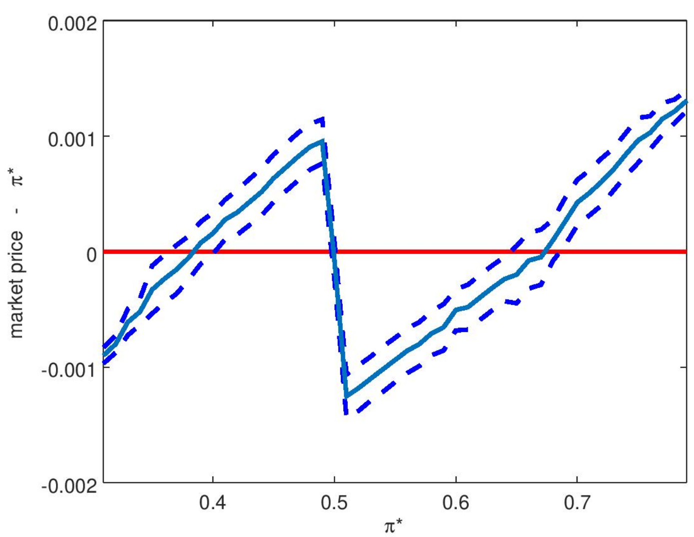
Simple repetitive betting market from Kets et al, AAAI 2014

- 1 event realises at every step with a fixed probability π^*
- 3 Fractional Kelly bettors. Have a belief on π^* and place bets accordingly

A query to study the wealth of each agent and the market price at steady-state

```
obs(o) = s.eval(o) ;
//Only one of the three commands below should be used
eval warmup(E[ obs(0) ],E[ obs(1) ],E[ obs(2) ],E[ obs("price") ]) ;
eval autoBM(E[ obs(0) ],E[ obs(1) ],E[ obs(2) ],E[ obs("price") ]) ;
eval autoRD(E[ obs(0) ],E[ obs(1) ],E[ obs(2) ],E[ obs("price") ]) ;
```

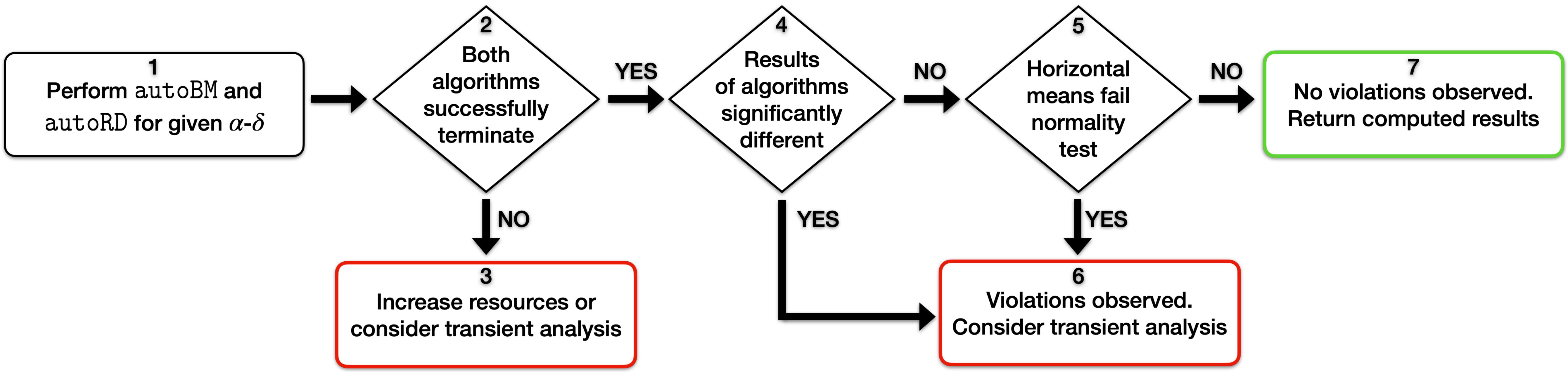




OUTLINE

1. Motivation, vision, and proposal
    1. Automated analysis with statistical guarantees for ABMs
    2. The MultiVeStA Statistical Model Checker
2. Transient Analysis of a large-scale financial macro ABM
    1. Estimation of expected outcome and Confidence Interval
    2. Counterfactual analysis for different model configurations
3. **Steady-state analysis of a prediction market model**
    1. Steady-state analysis by Replication and Deletion (RD)
    2. Warmup estimation
    3. Steady-state analysis by Batch Means (BM)
    4. **A methodology for ergodicity analysis based on RD and BM**
4. Conclusions & Future works

A Methodology for Ergodicity Diagnostics



OUTLINE

1. Motivation, vision, and proposal
   1. Automated analysis with statistical guarantees for ABMs
   2. The MultiVeStA Statistical Model Checker
2. Transient Analysis of a large-scale financial macro ABM
   1. Estimation of expected outcome and Confidence Interval
   2. Statistical comparison of different model configurations
3. Steady-state analysis of a prediction market model
   1. Steady-state analysis by Replication and Deletion (RD)
   2. Warmup estimation
   3. Steady-state analysis by Batch Means (BM)
   4. A methodology for ergodicity analysis based on RD and BM
4. **Conclusions & Future works**

CONCLUSIONS

- Fully automated framework for statistical analysis of ABMs
  - Transient analysis with statistical tests to compare model configurations
  - Warmup estimation
  - Steady-state analysis by Replication and Deletion and by Batch Means
  - Ergodicity diagnostics
- Tool-supported one-click analysis:
  - Less manual error-prone tasks => more reproducibility & reliability
  - Automatically parallelise simulations: 15 days => 15 hours
  - Implemented in the statistical analyser MultiVeStA
    - Supports simulators written in Java, Python, R, C++, JMAB, NetLogo
- Validated on two models from the literature:
  - Large-scale macro financial ABM, Small-scale prediction market model
  - We obtained new insights on the considered models
  - We avoid analysis errors from previous publications
- Our approach is rooted in results from:
  - Communities of Simulation, Computer Science, Operations Research
  - We aim at strengthening the cross-fertilisation of these communities with the ABM one

FUTURE WORK

- Add more techniques
  - Detection of multiple stationary points
  - Advanced SMC techniques to
    - Handle rare events, Reduce number of simulations required
  - More!? Model calibration, Sensitivity analysis, …
- Apply the approach to further models and domains
  - Any JMAB model is now natively supported
  - We have integrated a 'classic' ABM model, Islands model [FagioloDosi2003]
  - We wish to natively support further frameworks for ABM modelling
    - LSD, JASMINE, Mesa, …
- Would you like to use MultiVeStA to analyse your models?
  - Just contact us andrea.vandin@santannapisa.it
- Interested in projects related to MultiVeStA?
  - Just contact us! andrea.vandin@santannapisa.it

REFERENCES

[Secchi,Seri2017] Secchi, D. and R. 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), 94–121.

[Caiani et al2016] Caiani, A., A. Godin, E. Caverzasi, M. Gallegati, S. Kinsella, and J. E. Stiglitz (2016). Agent based-stock flow consistent macroeconomics: Towards a benchmark model. Journal of Economic Dynamics and Control 69, 375–408

[Welch1947] Welch, B. L. (1947). The generalization ofstudent's' problem when several different population variances are involved. Biometrika 34(1/2), 28–35.

[Chow2002] Chow, S.-C., J. Shao, and H. Wang (2002). A note on sample size calculation for mean comparisons based on noncentral t-statistics. Journal of biopharmaceutical statistics 12(4), 441–456.

[Law,Kelton2015] Law, A. M. and D. M. Kelton (2015). Simulation Modeling and Analysis (5th ed.). McGraw-Hill Higher Education, http://www.averill-law.com/simulation-book/

[Kets et al2014] Kets, W., D. M. Pennock, R. Sethi, and N. Shah (2014). Betting strategies, market selection, and the wisdom of crowds. In Twenty-Eighth AAAI Conference on Artificial Intelligence.

[Bottazzi,Giachini2019] Bottazzi, G. and D. Giachini (2019b). Far from the madding crowd: Collective wisdom in prediction markets. Quantitative Finance 19(9), 1461–1471

[Conway1963] Conway, R. W. (1963). Some tactical problems in digital simulation. Management Science 10(1), 47–61.

[Law,Carson1979] Law, A. M. and J. S. Carson (1979). A sequential procedure for determining the length of a steady-state simulation. Operations Research 27(5), 1011–1025.

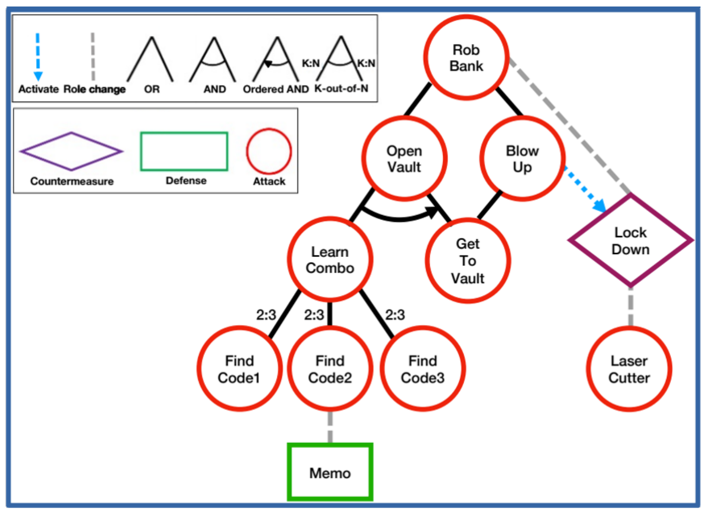
[Steiger et al 2005] Steiger, N. M., E. K. Lada, J. R. Wilson, J. A. Joines, C. Alexopoulos, and D. Goldsman (2005). Asap3: A batch means procedure for steady-state simulation analysis. ACM Transactions on Modeling and Computer Simulation (TOMACS) 15(1), 39–73.

[Tafazzoli et al 2011] Tafazzoli, A., J. R. Wilson, E. K. Lada, and N. M. Steiger (2011). Performance of skart: A skewness-and autoregression-adjusted batch means procedure for simulation analysis. INFORMS Journal on Computing 23(2), 297–314.

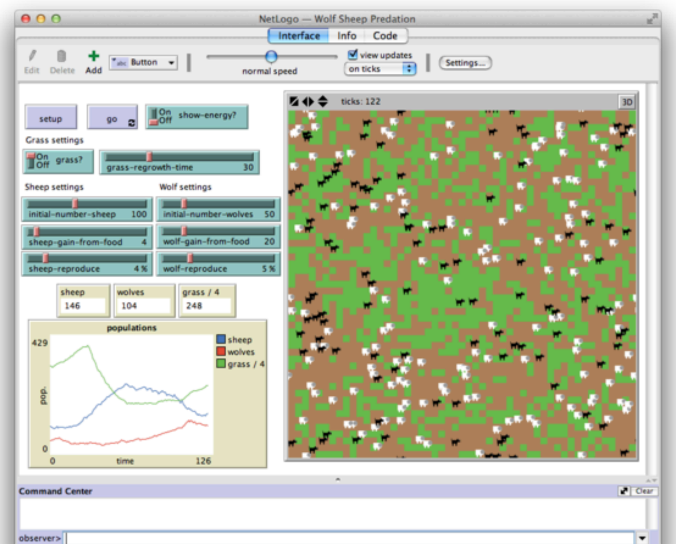
[Gilmore et al 2017] Gilmore, S., D. Reijsbergen, and A. Vandin (2017). Transient and steady-state statistical analysis for discrete event simulators. In Integrated Formal Methods - 13th International Conference, IFM 2017, Proceedings, pp. 145–160.

Would you like to join the MultiVeStA family?

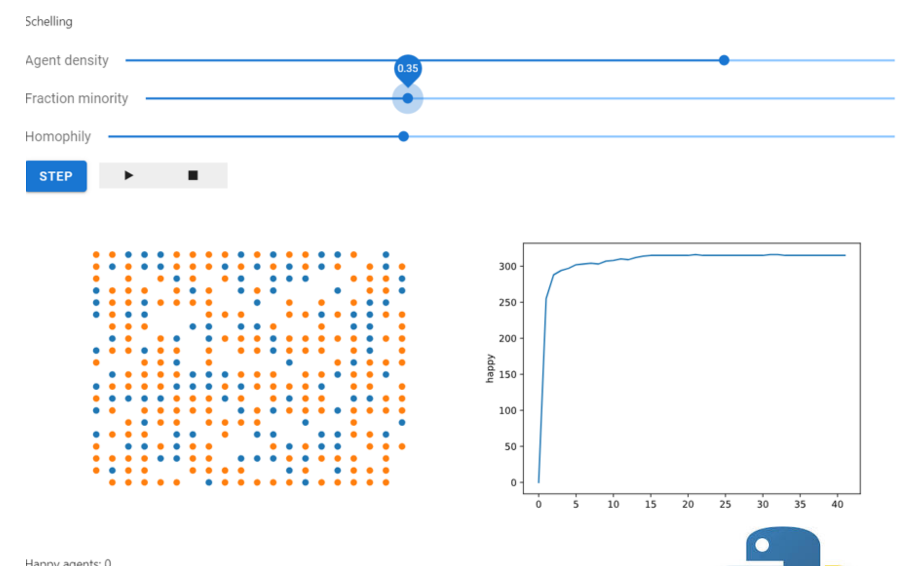
- **Projects available**
  - As an exam for this course
  - As starting points for Master projects?
  - As starting points for longer collaborations!?



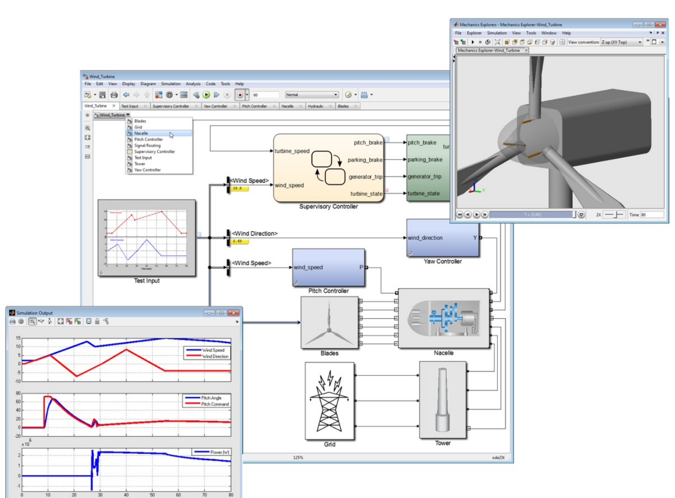
RisQFLan - Security



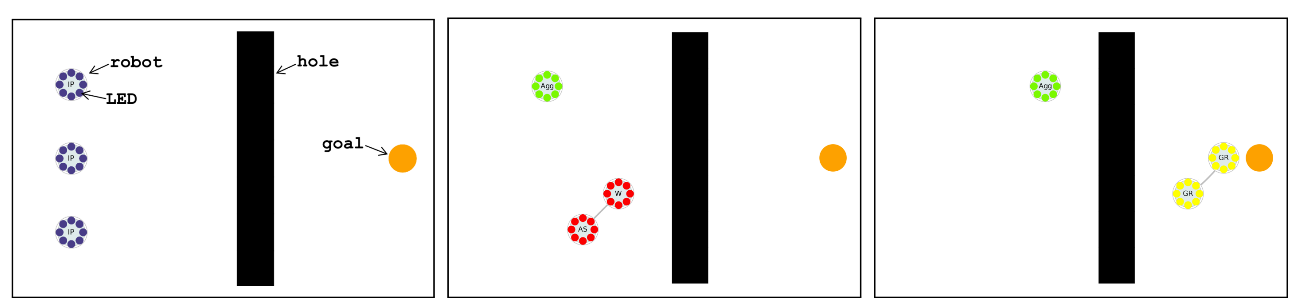
NetLogo multi-agent modeling
millions of students/teachers/researchers



Mesa: ABM in Python



Matlab Simulink



Maude - rewriting logic

More...

THANK YOU FOR YOUR ATTENTION!

QUESTIONS?

FEEDBACK?

JEDC Paper

https://www.sciencedirect.com/science/article/abs/pii/S0165188922001634

Tool and models available at:

https://bit.ly/MultiVeStATool