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

LEM Working Paper available at:
www.lem.sssup.it/WPLem/2020-31.html

Tool and models available at:
https://bit.ly/MultiVeStATool

Research Webinar, IMT School for Advanced Studies Lucca 17/03/2021

ABOUT ME

PISA
END 2019 Tenure-track Assistant Professor @ Sant'Anna

COPENHAGEN
2019-NOW 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

http://people.compute.dtu.dk/anvan/

ABOUT ME

Departments of Computer Science

Institute of Economics
Excellence Institute EMBedDS

PISA
END 2019 Tenure-track Assistant Professor @ Sant'Anna

COPENHAGEN
2019-NOW 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

http://people.compute.dtu.dk/anvan/

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

‘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.

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

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

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 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

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

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

Evolution of bankruptcies

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Evolution of bankruptcies

Confidence Intervals width MultiVeStA VS by 100 sims

Intermediate CIs width by MultiVeStA

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

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

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

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

Evolution of bankruptcies

Confidence Intervals width
MultiVeStA VS by 100 sims

OK, BUT...

IS ALL THIS STUFF IMPORTANT ???

by MultiVeStA

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

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

97.5% CI
MultiVeStA
$\delta = 0.5$

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

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

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

97.5% CI
MultiVeStA
$\delta = 0.5$

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

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

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]

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) ;
```

δ=0.005

δ=0.5

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 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

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

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

Warmup
w steps

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

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

Warmup
w steps

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

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

Warmup
w steps

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

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

Agents wealth at steady state

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

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

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

Replication and deletion

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

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

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

Replication and deletion

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

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

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

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

Replication and deletion

Agents wealth at steady state

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

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

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

Replication and deletion

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

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

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

Replication and deletion

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

THESE ARE DIFFICULT QUESTIONS
ARE THEY CRUCIAL?
YES

Agents wealth at steady state

Wrong by [Kets et al AAAI2014]

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

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

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

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

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]

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

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

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

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

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

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

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]

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

1. Do n simulations of a *given large length* m
2.
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]

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

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

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

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]

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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 16 16

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

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

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

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

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

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

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

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

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
- 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, …
- Would you like to use MultiVeStA to analyse your models?
  - Just contact us

THANK YOU FOR YOUR ATTENTION!

QUESTIONS?

FEEDBACK?

LEM Working Paper available at:
www.lem.sssup.it/WPLem/2020-31.html

Tool and models available at:
https://bit.ly/MultiVeStATool