

May 2022

## Modeling Empirical Stock Market Behavior Using a Hybrid Agent-Based Dynamical Systems Model

Daniel A. Cline

*Binghamton University*, [dcline1@binghamton.edu](mailto:dcline1@binghamton.edu)

Grant T. Aguinaldo

*Binghamton University*, [gaguina1@binghamton.edu](mailto:gaguina1@binghamton.edu)

Christian Lemp

*Binghamton University*, [wlemp1@binghamton.edu](mailto:wlemp1@binghamton.edu)

Follow this and additional works at: <https://orb.binghamton.edu/nejcs>



Part of the [Finance and Financial Management Commons](#), [Non-linear Dynamics Commons](#), and the [Numerical Analysis and Computation Commons](#)

---

### Recommended Citation

Cline, Daniel A.; Aguinaldo, Grant T.; and Lemp, Christian (2022) "Modeling Empirical Stock Market Behavior Using a Hybrid Agent-Based Dynamical Systems Model," *Northeast Journal of Complex Systems (NEJCS)*: Vol. 4 : No. 2 , Article 1.

DOI: [10.22191/nejcs/vol4/iss2/1](https://doi.org/10.22191/nejcs/vol4/iss2/1)

Available at: <https://orb.binghamton.edu/nejcs/vol4/iss2/1>

This Article is brought to you for free and open access by The Open Repository @ Binghamton (The ORB). It has been accepted for inclusion in Northeast Journal of Complex Systems (NEJCS) by an authorized editor of The Open Repository @ Binghamton (The ORB). For more information, please contact [ORB@binghamton.edu](mailto:ORB@binghamton.edu).

# Modeling Empirical Stock Market Behavior Using a Hybrid Agent-Based Dynamical Systems Model

Daniel A. Cline<sup>1\*</sup>, Grant T. Aguinaldo<sup>1</sup>, and Christian Lemp<sup>1</sup>

<sup>1</sup> Department of Systems Science and Industrial Engineering

Binghamton University, Binghamton, NY, USA

\* dcline1@binghamton.edu

## Abstract

We describe the development and calibration of a hybrid agent-based dynamical systems model of the stock market that is capable of reproducing empirical market behavior. The model consists of two types of trader agents, fundamentalists and noise traders, as well as an opinion dynamic for the latter (optimistic vs. pessimistic). The trader agents switch types stochastically over time based on simple behavioral rules. A system of ordinary differential equations is used to model the stock price as a function of the states of the trader agents. We show that the model can reproduce key stylized facts (e.g., volatility clustering and fat tails) while providing a behavioral interpretation of how the stock market itself can cause periods of high volatility and large price movements, even when the economic value of the stock grows at a constant rate.

## 1 Introduction

Modeling the dynamics of financial asset prices often involves tradeoffs between tractability and empirical consistency. As we seek to capture the empirical behavior of asset price movements, our models often become more complex and, therefore, less tractable. Furthermore, while some models are mostly concerned with reproducing empirical market behaviors, others are primarily concerned with providing plausible explanations for the causes of these behaviors. The model proposed in this paper seeks to accomplish both, while maintaining simplicity as much as possible.

Empirical behaviors of asset prices show remarkably similar patterns across financial markets and asset types [1, 2]. These statistical properties have come to be known as stylized facts, some of the most notable being:

- **Uncorrelated returns:** The time series of returns shows virtually no autocorrelation for all lags greater than zero.<sup>1</sup> However, returns are not independent.
- **Volatility clustering:** The variance of returns shows temporal dependence. In particular, large changes in prices cluster together in certain time periods, while small changes in prices cluster together in others. This manifests itself in significant autocorrelations in the absolute values of returns. Furthermore, the autocorrelation function is a decreasing function of the number of lags.
- **Fat tails:** The unconditional distribution of returns is leptokurtic with positive excess kurtosis, meaning the tails of the distribution have much more weight assigned to them than the tails of a normal distribution. Put simply, large movements in price are much more common than what Gaussian noise can produce.
- **Asymmetric returns:** In the case of stock markets, the unconditional distribution of returns shows negative skewness. In particular, large downward movements in price are more common than large upward movements in price.

Due to its simplicity, geometric Brownian motion (GBM) is often used to model the price dynamics of financial assets [3]. GBM is an equation-based model that assumes price movements are independent and lognormally distributed (i.e., the log returns of the time series of prices follow a normal distribution). The benefit of this assumption is that closed-form solutions are readily available for many financial derivatives [4]. However, from the definitions above, we see that GBM precludes volatility clustering and fat tail behavior, so its assumptions are consistently violated in real financial markets.

To better account for empirical market behavior, researchers have put considerable effort into developing more sophisticated equation-based models that are capable of reproducing stylized facts. Two common approaches are autoregressive conditional heteroskedasticity (ARCH) models [5] and stochastic volatility models [6]. However, while these models are able to reproduce certain stylized facts, they do not provide a behavioral explanation for their underlying cause.

More recently, applying agent-based models (ABM) to financial markets has allowed for behavioral explanations of stylized facts [7–9]. In this approach, a micro model of the market is constructed with many trader agents, each following certain rules or trading strategies, such that their simulated interactions collectively lead to emergent macro market behaviors that are consistent with stylized facts. This allows for plausible behavioral explanations of the underlying causes of stylized facts, since traders' actions are directly incorporated into the model. However,

---

<sup>1</sup>The  $n$ -lag autocorrelation coefficient for a time series  $r_t$  is given by  $\text{corr}(r_t, r_{t-n})$ .

models proposed in the literature are often quite complex, utilizing concepts such as variable memory, percolation theory, networks, adaptive learning, classifiers, and genetic programs.<sup>2</sup> Due to their complexity, these models are often difficult to implement and calibrate.

Weighing the tradeoffs between equation-based models and agent-based models, we sought to build a hybrid agent-based dynamical systems model where the behaviors of discrete trader agents, following a herding mechanism originally introduced to explain the feeding behavior of ants [11], are distilled into an evolving market price using a system of ordinary differential equations to model the role of a market maker. We show that interactions and feedbacks among these simple trader agents lead to the emergence of macro behaviors that are statistically consistent with stylized facts found in real market data.

This paper is structured as follows: In Section 2, we construct a model that draws on findings in the behavioral and empirical finance literature to understand how the behaviors of different types of traders can affect asset prices in a given market. In Section 3, we describe our approach for running numerical simulations and calibrating the model to historical data. A statistical analysis of the results is also provided. In Section 4, we conclude with a brief discussion of our work and directions for future work. The working code for all results in this paper can be found on GitHub.<sup>3</sup>

## 2 Model

We consider a speculative market for a single stock, where buy and sell orders are placed by multiple traders and prices are set by a market maker based on the demand generated by the traders' orders. The model we propose is a hybrid agent-based dynamical systems model of the stock market, where traders are represented by agents whose aggregate states are processed by a market maker, which takes the form of a dynamical system with stochastic parameters. Our approach builds on the simple model proposed in [12], which draws heavily on [11, 13, 14]. Our goal was to construct a model that better reproduces the stylized facts found in real market data while maintaining simplicity as much as possible.

The model consists of two types of trader agents (traders), fundamentalists and noise traders (i.e., chartists). As in [12], each noise trader is designated either an optimist or pessimist at any given time, and noise traders randomly switch between the two based on a majority opinion dynamic originally proposed in [11]. Traders also switch between fundamentalists and noise traders, and do so as a function of

---

<sup>2</sup>A well-known example of one such model is the Santa Fe artificial stock market [10].

<sup>3</sup><https://github.com/dcline1/MarketABM/blob/master/MarketABM.ipynb>

the divergence between the stock's market price and its fundamental value.

At the center of the model is a market maker, who sets the market price of the stock as a function of the excess demand generated by the traders. Similar to the approach originally proposed in [14], the market maker is modeled as a system of ordinary differential equations (ODE) with stochastic parameters. However, instead of assuming equilibrium dynamics at each time step, we use the system of differential equations to directly model the dynamical system through time.

## 2.1 Fundamental Value

We start with the fundamental value of the stock, which we take to be its true economic value. In practice, financial analysts apply fundamental analysis (e.g., dividend discount models and pricing metrics, such as price-earnings ratios) to determine the fundamental value of a stock. Since analysts often disagree on the true value of a stock, the fundamental value here should be thought of as a measure of the average of many individual estimates of the fundamental value (i.e., the market's overall view of the fundamental value of the stock).

While it is common in the ABM literature to take the fundamental value of the stock to be constant through time [12, 15], this is not a realistic assumption since stock prices tend to grow exponentially, and over the long run, market prices are driven by the fundamental value. Therefore, in order to directly compare results from the model with real market data, we model the fundamental value of the stock using the following exponential function

$$\frac{dV_t}{dt} = \mu V_t, \quad (1)$$

where  $V_t$  is the fundamental value at time  $t$  and  $\mu$  is the growth rate. Note that the growth rate of the fundamental value is deterministic and constant in Eq. (1). We chose this simplifying assumption because our goal was to study the endogenous behavior of the market, where price movements are entirely dictated by the interactions among individual traders. Allowing for stochastic growth rates would make the model less interpretable, since attributing price movements to trader behaviors vs. exogenous shocks to the fundamental value would be difficult. By removing uncertainty in the fundamental value, the market itself is entirely responsible for the stock price behavior produced by the model.

## 2.2 Trader Types

We consider two types of trader agents that are found in virtually every financial market, fundamentalists and noise traders.<sup>4</sup> Fundamentalist traders buy and sell

<sup>4</sup>These are referred to as  $\alpha$ -traders and  $\beta$ -traders, respectively, in [13].

based on perceived mispricings in the market (i.e., whether the asset is over or undervalued). They do so using information about the fundamental value of the stock and do not consider market sentiment in their trading decisions. In contrast, noise traders are driven by herd instincts and tend to follow the majority opinion of the market (i.e., overall market sentiment) when making trading decisions. Herd behavior, which is driven by diffusion of opinion (i.e., contagion) rather than knowledge of market fundamentals, is well documented in the literature [16]. As in [12], we designate each noise trader either an optimist or pessimist.

The model consists of  $N_t^f$  fundamentalist traders,  $N_t^c$  noise traders, and  $N = N_t^f + N_t^c$  traders total. Noise traders are further split into  $N_t^o$  optimist traders and  $N_t^p$  pessimist traders, where  $N_t^c = N_t^o + N_t^p$ . To avoid absorbing states, boundary conditions are also imposed to ensure that there is at least one of each type of trader at all times (i.e.,  $N_t^f > 0$ ,  $N_t^o > 0$ , and  $N_t^p > 0$ ). Note that while the total number of traders is fixed throughout the simulation, the number of fundamentalists, optimists, and pessimists are dynamic and change stochastically through time based on simple behavioral rules (described in Section 2.4).

Since we ultimately wish to calibrate the model to market data, we would like the parameters of our model to be (relatively) invariant to the total number of traders. Therefore, rather than directly use the number of traders in our model specification, we use the proportion of traders that are of a given type by defining the following variables

$$Y_t^f = \frac{N_t^f}{N}, \quad Y_t^c = \frac{N_t^c}{N}, \quad Y_t^o = \frac{N_t^o}{N}, \quad Y_t^p = \frac{N_t^p}{N}. \quad (2)$$

This gives the following relationships

$$\begin{aligned} Y_t^c &= Y_t^o + Y_t^p, \\ 1 &= Y_t^f + Y_t^c. \end{aligned}$$

Note that, due to the boundary conditions,  $Y_t^f$ ,  $Y_t^c$ ,  $Y_t^o$ , and  $Y_t^p$  are all strictly between 0 and 1. We also define the following stochastic term based on market sentiment

$$X_t = \frac{N_t^o - N_t^p}{N_t^c}, \quad (3)$$

which is positive when there are more optimists than pessimists and negative when there are more pessimists than optimists. Note that  $X_t$  provides a measure of the overall mood of the market and is always strictly between  $-1$  (pessimistic) and  $1$  (optimistic). Furthermore, as will be discussed in Section 2.4, the herding behavior of noise traders imposes (stochastic) attractors on  $X_t$  at  $1$  and  $-1$ , with intermittent regime switches between the two.

### 2.3 Market Maker

In practice, traders submit buy and sell orders to a market maker, who executes trades by matching orders from all traders and settles on a market price. Market makers also maintain their own inventory of the stock, providing additional liquidity to the market by depleting their inventory when buy orders outnumber sell orders and by accumulating inventory when sell orders outnumber buy orders. From a modeling standpoint, this means that the market price does not necessarily reflect a perfect equilibrium of supply and demand among traders at all times, but rather it tends towards equilibrium in what is known as sluggish price adjustment [13, 14]. As in [12], we use an ordinary differential equation to model the behavior of the market maker.

Since we are most interested in returns (e.g., percentage changes in price) as opposed to dollar value changes in price, we model the dynamics of the market price of the stock,  $S_t$ , using an ODE of the following form

$$\frac{1}{S_t} \frac{dS_t}{dt} = g(D_f + D_c), \quad (4)$$

where  $D_f$  measures the excess demand from fundamentalists,  $D_c$  measures the excess demand from noise traders, and  $g(\cdot)$  is an arbitrary function that controls the speed of market clearing as a function of the total excess demand from all traders. Here excess demand can be thought of as the number of buy orders relative to the number of sell orders at a given time. Note that since Eq. (4) models the relative price change ( $dS_t/S_t$ ), returns are independent of the absolute price level of the stock.

We assume that fundamentalists buy and sell shares of stock based on perceived mispricings in the market, measured by the relative (i.e., percentage) difference between the market price and the fundamental value

$$D_f = \alpha Y_t^f \left( \frac{V_t - S_t}{S_t} \right), \quad (5)$$

where  $\alpha$  is a measure of the volume of trades made (e.g., the number of shares traded) by fundamentalist traders, and is fixed throughout the simulation. This is a simplified version of the more general excess demand function given in [13] and is similar to the excess demand function for fundamentalists given in [12]. However, rather than use the simple difference between market price and fundamental value, we use the relative difference so that the demand for stocks is a function of potential rates of return and not the dollar value of the stock.<sup>5</sup> When the fundamental value

<sup>5</sup>Eq. (5) is consistent with the mean-reversion drift term in the Ornstein–Uhlenbeck process.

is not constant, using an absolute deviation instead of a percentage deviation will distort demand through time, since the magnitude of the price grows as a function of time.

Noise traders, on the other hand, trade based on their opinion about the future. The excess demand function for noise traders proposed in [12] is of the form  $\beta Y_t^c X_t$ , where  $\beta$  represents the volume (i.e., size) of trades made by noise traders and  $X_t$  is the market sentiment index given by Eq. (3). However, this demand function assumes that demand is symmetric for optimists and pessimists, and as a result, we found that it could not generate the large negative skewness and excess kurtosis found in real financial markets.

Since market selloffs are often accompanied by increases in trading volume [17], we chose to model trading volume as an increasing function of pessimistic market sentiment. To keep the function as simple as possible, we model this increase in volume linearly as a function of  $(1 - X_t)$  using the following excess demand function for noise traders

$$D_c = \beta Y_t^c \left( \frac{1 - X_t}{2} \right) X_t = \beta Y_t^p X_t. \quad (6)$$

Note that as the number of pessimists increases, so does  $(1 - X_t)$ , which works to push prices lower during market selloffs due to larger trade orders. Therefore, this excess demand function induces negative skewness and high excess kurtosis in the distribution of returns.

The price adjustment function  $g(D_f + D_c)$  in Eq. (4) is often taken to be  $\kappa(D_f + D_c)$  for some constant  $\kappa > 0$ , where  $\kappa$  is a parameter that controls the speed of price adjustment per unit time. For simplicity, [12] takes the limit as this price adjustment constant goes to infinity (i.e., instantaneous market clearing), which allows them to solve for the equilibrium price when  $dS_t/dt = 0$ . In this paper, we assume that  $\kappa$  is finite (i.e., sluggish price adjustment) so that the system is generally not at equilibrium and the ODE defines a dynamical system through time.

Putting it all together, we get the following ODE for the market maker

$$\frac{dS_t}{dt} = \alpha Y_t^f (V_t - S_t) + \beta Y_t^p S_t X_t, \quad (7)$$

where we have chosen to absorb  $\kappa$  into the  $\alpha$  and  $\beta$  parameters to reduce the number of parameters in the model. Since Eq. (7) involves parameters that are themselves stochastic processes ( $Y_t^f$ ,  $Y_t^p$ , and  $X_t$ ), we simulate it using the forward Euler method with a simulation time step  $\Delta t$  corresponding to one trading day.

Note that since the first expression on the right-hand side is a mean-reverting term, the model pulls the market price back to its fundamental value when it diverges too far. This is consistent with empirical findings of long-run market mean



reversion [18]. Furthermore, the second expression on the right-hand side includes the stochastic component  $X_t$ , which has (stochastic) attractors at 1 and  $-1$ , so it tends to pull the market price of the stock above (below) the fundamental value for long periods of time based on market sentiment until a regime switch reverses sentiment and pulls the market price below (above) the fundamental value. The recurrent regime switches between optimists and pessimists is the primary driver of volatility clustering in the model.

#### 2.4 Switching Behavior

We assume all traders of a given type follow the same simple rules. We also assume that market information is disseminated to all traders equally (e.g., through media reports) without consideration of trader proximity.<sup>6</sup> Similar to [12], for each time step, each optimist has a probability  $p_{op}$  of switching to a pessimist and each pessimist has a probability  $p_{po}$  of switching to an optimist, where the switches are Bernoulli and are assigned the following probabilities

$$p_{op} = \nu_1 \Delta t \frac{N_t^p}{N_t^c}, \quad p_{po} = \nu_1 \Delta t \frac{N_t^o}{N_t^c}, \quad (8)$$

where  $\nu_1$  controls the velocity of opinion diffusion among noise traders and we multiply by  $\Delta t$  to make  $\nu_1$  invariant to the size of the time step. We see from these definitions that when the majority of noise traders are optimists (pessimists), the few remaining pessimists (optimists) will have a relatively high probability of switching to optimists (pessimists), while the optimists (pessimists) will have a relatively low probability of switching to pessimists (optimists). This contagion of opinion leads to herd behavior, as noise traders are either increasingly drawn into an optimistic market or sell out of a pessimistic market. In other words, the system is attracted to the states where the majority of traders are either optimists ( $X_t \rightarrow 1$ ) or pessimists ( $X_t \rightarrow -1$ ) and there is a low probability of switching out of these states. However, given the stochastic nature of the model, the overall opinion of the market does occasionally switch between the two, resulting in intermittent regime switches in market sentiment.

The number of fundamentalists and noise traders is fixed in [12]. Because of this, the behavior of the stock price very closely mirrors the value of  $X_t$  in their model, remaining above  $V_t$  by a roughly constant amount when  $X_t$  is close to 1 and remaining below  $V_t$  by a roughly constant amount when  $X_t$  is close to  $-1$ . We chose to incorporate the ability of trader agents to switch between fundamentalists

<sup>6</sup>See [19] for an example of an ABM with opinion diffusion over a network.

and noise traders based on the following transition probabilities<sup>7</sup>

$$p_{fc} = \nu_2 \Delta t e^{-\lambda \rho}, \quad p_{cf} = \nu_2 \Delta t (1 - e^{-\lambda \rho}), \quad (9)$$

where  $p_{fc}$  is the probability that a fundamentalist trader switches to a noise trader,  $p_{cf}$  is the probability that a noise trader switches to a fundamentalist trader,  $\lambda > 0$  is a free parameter, and

$$\rho = \frac{|V_t - S_t|}{V_t} \quad (10)$$

is the absolute percentage deviation of the stock price from the fundamental value. Again, we multiply by  $\Delta t$  to make  $\nu_2$  invariant to the size of the time step. By allowing for switching between fundamentalists and noise traders, we vary the weights (i.e.,  $Y_t^f$  and  $Y_t^p$ ) applied to the first and second terms on the right-hand side of Eq. (7), which allows for richer price behavior.

Intuitively,  $p_{fc}$  is greatest when there is no deviation from the fundamental value, in which case there is no perceived advantage to being a fundamentalist trader, so traders increasingly switch to noise traders. On the other hand,  $p_{cf}$  grows as the price diverges from the fundamental value, so noise traders increasingly switch to fundamentalist traders to take advantage of the perceived mispricing.

Note that  $p_{fc} = p_{cf}$  when  $e^{-\lambda \rho} = 1 - e^{-\lambda \rho}$ , which gives

$$\lambda = -\frac{1}{\rho} \ln \frac{1}{2}. \quad (11)$$

Therefore, the choice of  $\lambda$  determines the level of absolute percentage price deviation above which traders are more inclined to switch to fundamentalists from noise traders. In other words, the choice of  $\lambda$  controls how far the market price will diverge from the fundamental value before fundamentalists begin to dominate and pull it back toward its fundamental value.

## 2.5 Visualization of Model

To illustrate the model's overall dynamics, each model component is plotted over time in Figure 1. The top plot shows the percentage of noise traders through time who are optimists and pessimists ( $Y_t^o/Y_t^c$  and  $Y_t^p/Y_t^c$ , respectively). The middle plot shows the percentage of traders through time who are fundamentalists and noise traders ( $Y_t^f$  and  $Y_t^c$ , respectively). The bottom plot shows both the fundamental value and market price of the stock generated by the model ( $V_t$  and  $S_t$ , respectively).

<sup>7</sup>See [15] for an alternative switching model based on expected payoff differentials.

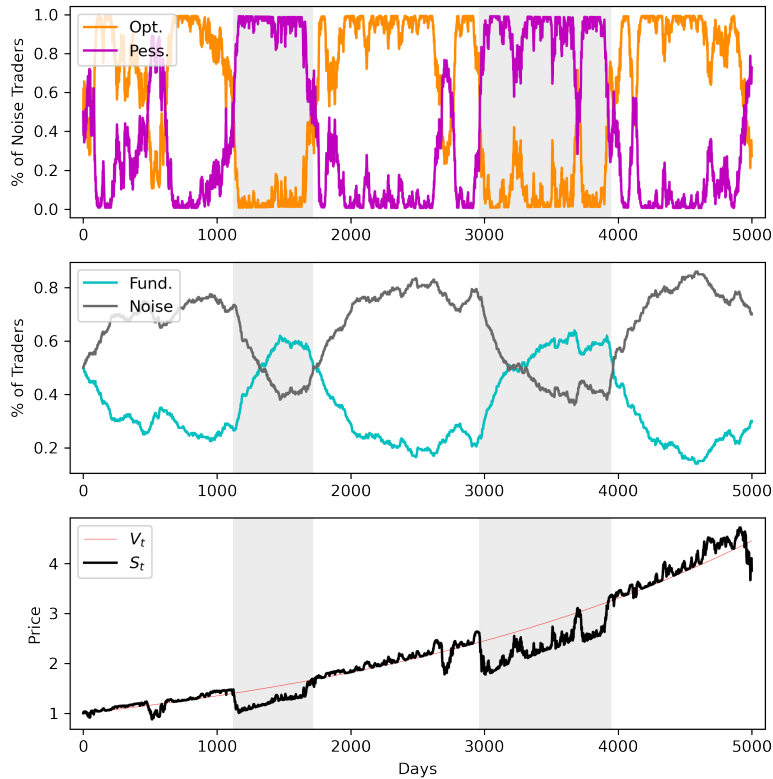


Figure 1: Daily time series dynamics of the hybrid agent-based dynamical systems model. Top: optimists vs. pessimists, Middle: fundamentalists vs. noise traders, Bottom: simulated stock price vs. fundamental value.

The stochastic attractor behavior of  $X_t$  in Eq. (3) is evident in the top plot, as at any given time, the majority of noise traders are either optimists ( $X_t \rightarrow 1$ ) or pessimists ( $X_t \rightarrow -1$ ). Note that while the dynamics of  $X_t$  are independent of the stock price, the stock price relative to the fundamental value is highly dependent on them. This can be seen by comparing the top and bottom plots and noting that the stock price is generally above the fundamental value when the market is optimistic and it is generally below the fundamental value when the market is pessimistic.

It is also evident that the number of fundamentalist traders tends to grow during prolonged periods of market pessimism, as seen in the shaded regions, since the heavier trading volume for pessimistic noise traders in Eq. (6) tends to depress the market price, leading to more conversions from noise traders to fundamentalists due to Eqs. (9) and (10). This makes intuitive sense, as there's more incentive for fundamentalists to take advantage of mispricing in the market when the herding behavior of noise traders depresses the market price.

Conversely, the number of noise traders tends to grow during prolonged periods of market optimism, as seen in the unshaded regions, since the lighter trading volume for optimistic noise traders in Eq. (6) gives outside weight to fundamentalist traders in driving the price, initially keeping the market price closely aligned with the fundamental value. But as the number of noise traders grows, their larger relative influence causes the market price to diverge further from the fundamental value until a regime shift leads to a large drop in price. The cycle then repeats itself.

### 3 Numerical Results

In this section, we compare stock price behavior from (1) a historical data set, (2) geometric Brownian motion, and (3) the hybrid agent-based dynamical systems model proposed in this paper. By calibrating to historical data, we can directly compare the statistical properties of all three time series and draw conclusions about the proposed model.

#### 3.1 Historical Returns

We used the closing prices of the SPDR S&P 500 ETF Trust (symbol, SPY) from March 2, 2001 to January 15, 2021 for our historical data set.<sup>8</sup> Closing prices were converted to daily log returns as follows

$$r_t = \log\left(\frac{S_t}{S_{t-1}}\right), \quad (12)$$

where  $S_t$  is the stock price at time (date)  $t$ . This gave us a time series of returns for 5,000 trading days. Summary statistics for the historical returns are given in the SPY column of Table 1. The time series and histogram for  $r_t$  are also plotted in Figures 2(a) and 2(b), along with autocorrelation plots for  $r_t$  and  $|r_t|$  out to 40 lags (days) in Figures 3(a) and 3(b).

#### 3.2 Geometric Brownian Motion

The dynamics of the stock price,  $S_t$ , assuming geometric Brownian motion are given by the following stochastic differential equation (SDE)

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \quad (13)$$

where  $W_t$  is simple Brownian motion with distribution  $W_t \sim N(0, t)$ . Comparing this functional form to Eq. (1), we see that GBM is just an exponential with Gaussian noise added. This SDE was discretized using the forward Euler method with

<sup>8</sup><https://finance.yahoo.com/quote/SPY/history?p=SPY>

Statistic	SPY	GBM	ABM
Mean (annual)	0.0753	0.0723	0.0697
St. Dev. (annual)	0.1966	0.1975	0.1940
Skewness	-0.3030	-0.0216	-0.6024
Excess Kurtosis	12.8985	-0.0637	8.2427
Jarque-Bera			
Test statistic	34,737	1.2	14,457
p-value	0.0000	0.5437	0.0000
Ljung-Box( $ r $ , 15)			
Test statistic	8,425	16.9	8,964
p-value	0.0000	0.3250	0.0000

Table 1: Time series summary statistics

a time step of  $\Delta t = 1/252$ , corresponding to 252 trading days in a year (i.e., daily time steps). The two parameters  $\mu$  and  $\sigma$  correspond to the annualized mean and standard deviation of  $S_t$ , so these parameters were set to the annualized average return ( $\mu = 0.0753$ ) and standard deviation ( $\sigma = 0.1966$ ) from our historical SPY data set.

Summary statistics for the simulated time series of daily log returns for a 5,000 day simulation period are given in the GBM column of Table 1. The time series and histogram of returns are also shown in Figures 2(c) and 2(d), along with the autocorrelation plots for  $r_t$  and  $|r_t|$  in Figures 3(c) and 3(d).

### 3.3 Agent-Based Dynamical Systems Model

The full model consists of the system of ordinary differential equations given by Eqs. (1) and (7), which we discretize using the forward Euler method with a time step of  $\Delta t = 1/252$  (i.e., daily). An array keeps track of the type of each trader agent, where agents can switch types every time step according to realizations of Bernoulli random variables with probabilities given by Eqs. (8) and (9), along with the boundary conditions given in Section 2.2.

We ran extensive numerical experiments across a wide range of parameter values to determine valid parameter ranges that produced model behavior reasonably close to historical market data. When running these experiments, instead of specifying values for  $\nu_1$  and  $\nu_2$  directly, we chose to specify values for  $\nu_1 \Delta t$  and  $\nu_2 \Delta t$ , since these values must be between 0 and 1 to ensure Eqs. (8) and (9) give valid probabilities. The narrowed parameter space is given in Table 2.

Given the parameter space, we sought to calibrate the model to the historical

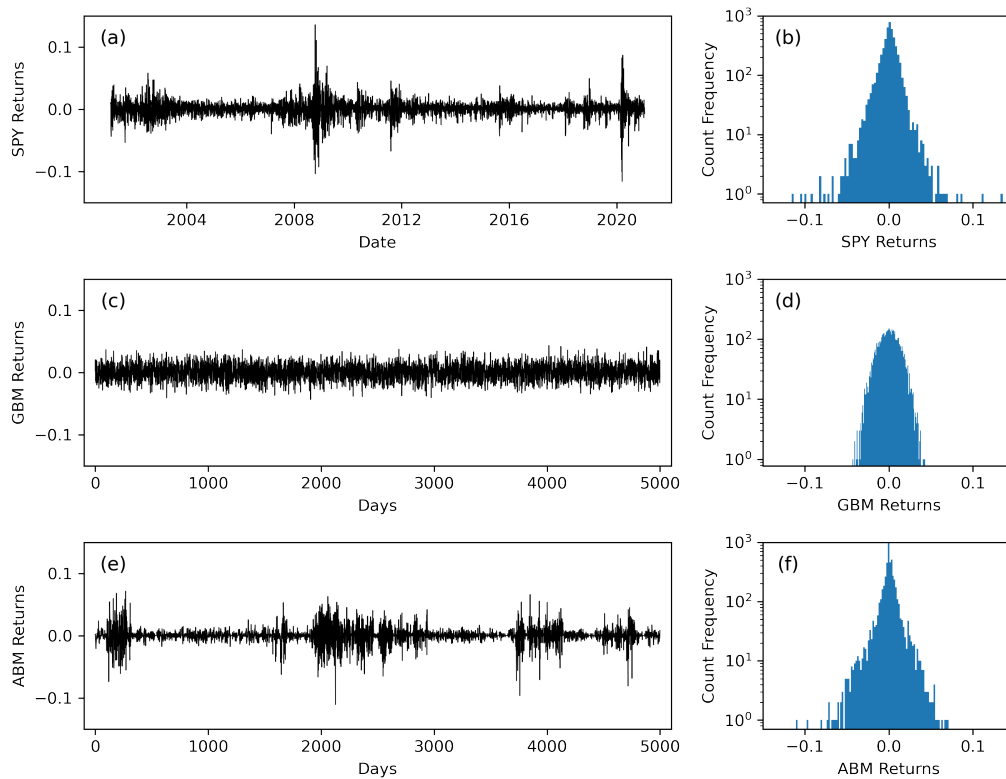


Figure 2: Time series and histograms of daily log returns,  $r_t$ . Top: historical data set (SPY), Middle: geometric Brownian motion (GBM), Bottom: hybrid agent-based dynamical systems model (ABM).

SPY data set in order to test its ability to replicate the stylized facts found in a real financial market. Since our goal was to demonstrate the capabilities of the model, rather than perfect calibration, we used a relatively simple calibration process, which nonetheless produced promising results. More elaborate calibration approaches for similar models can be found in [20, 21]. Bayesian optimization methods could also be considered [22].

Our numerical experiments revealed material interdependencies among parameters in the model, with various different combinations of parameter values resulting in similar statistical behavior. Therefore, we elected to fix certain parameters in the model and calibrate others. Specifically, we set  $N = 200$ ,  $N_0^f = 100$ , and  $N_0^c = 100$ , with an initial probability of 0.5 for designating each noise trader an optimist or pessimist (i.e.,  $E[N_0^o] = E[N_0^p] = N_0^c/2$ ).<sup>9</sup> While the total number of

<sup>9</sup>These values are consistent with the parameters used for numerical simulations in [12, 21].

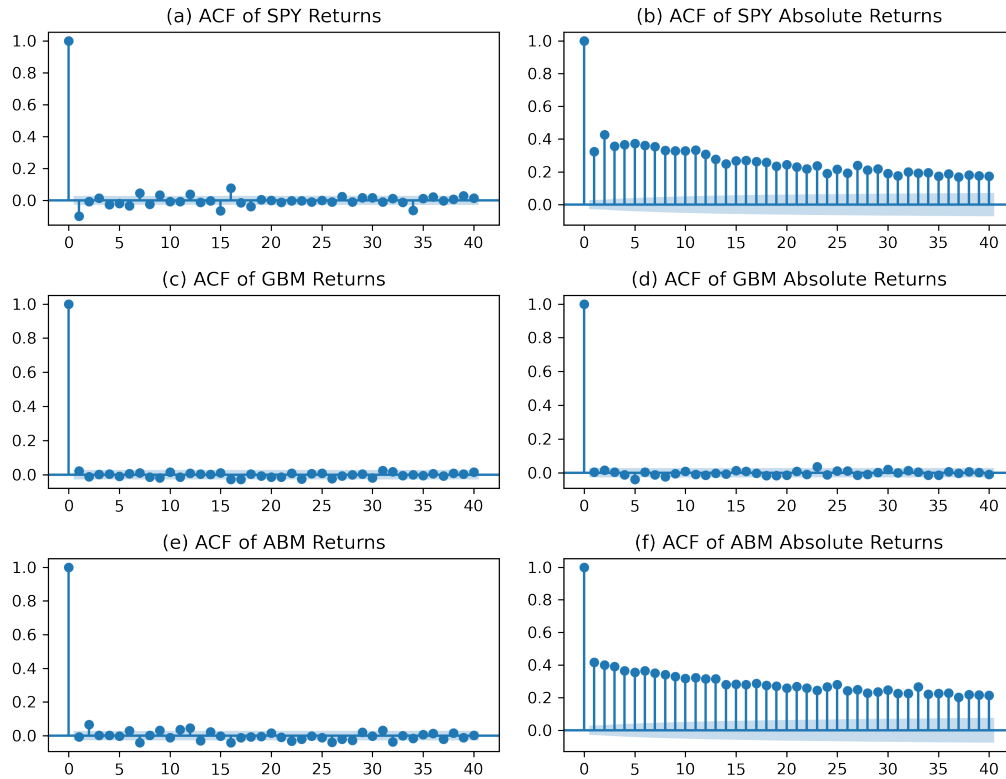


Figure 3: Autocorrelation functions of time series of daily log returns,  $r_t$ , and absolute values of daily log returns,  $|r_t|$ . Top: historical data set (SPY), Middle: geometric Brownian motion (GBM), Bottom: hybrid agent-based dynamical systems model (ABM).

traders has some effect on model behavior, it does not materially affect the values of other parameters in the model, since Eqs. (2) and (3) are functions of proportions of trader types as opposed to the actual number of traders.

Since  $V_t$  drives the long-run return of the stock through the mean-reversion term in Eq. (7), which pulls  $S_t$  towards  $V_t$ , we set  $\mu$  in Eq. (1) equal to the annualized average return from our historical data set ( $\mu = 0.0753$ ). We also fixed  $\alpha = 800$  and  $\nu_2 \Delta t = 0.003$ .

Since smaller values of  $\lambda$  correspond to larger price deviations in Eq. (11), noise traders tend to dominate the simulation as  $\lambda$  is decreased, which leads to greater variability in the simulated time series and larger calibration errors. Conversely, larger values of  $\lambda$  correspond to smaller price deviations in Eq. (11), so fundamentalists tend to dominate as  $\lambda$  is increased, which leads to lower kurtosis. We fixed  $\lambda = 6.9$ , corresponding to an absolute percentage price deviation in Eq. (11) of  $\rho \approx 0.1$ , or about 10%, above which fundamentalists tend to dominate, which gave

Parameter	Low	High
$\alpha$	700	1,000
$\beta$	100	$\alpha/2$
$\nu_1\Delta t$	0.1	0.9
$\nu_2\Delta t$	0.001	0.01
$\lambda$	4	14

Table 2: Suggested ranges for model parameters

good calibration results by balancing the two.

The fixed parameter values above leave ample flexibility in the model so that calibration only needs to be performed on two remaining parameters,  $\beta$  and  $\nu_1$ . To adequately cover the remaining parameter space, we created the following parameter sets used in our calibration

$$\begin{aligned}\beta &= \{100 + 25i \mid i = 0, 1, \dots, 12\}, \\ \nu_1\Delta t &= \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\},\end{aligned}\quad (14)$$

corresponding to the ranges given in Table 2. Finally, since we are primarily interested in returns (i.e., percentage changes in price), we set the initial values of  $S_0$  and  $V_0$  to 1. The vector of simulated prices can easily be converted to any starting price by simply multiplying it by the desired starting price.

We define our objective function for the calibration routine as follows

$$\begin{aligned}f &= 10e^2(\sigma) + 0.1e^2(skew) + e^2(kurt) + 0.1e^2(rng) \\ &\quad + 10(ACF_3(r))^2 + e^2(ACF_3(|r|)),\end{aligned}\quad (15)$$

where  $\sigma$  is the annualized standard deviation of the daily returns,  $skew$  is the skewness of the daily returns,  $kurt$  is the excess kurtosis of the daily returns,  $rng$  is the range of the daily returns (i.e., the maximum daily return over the time series minus the minimum daily return),  $ACF_3(r)$  is the average of the first 3 autocorrelations of the daily returns time series,  $ACF_3(|r|)$  is the average of the first 3 autocorrelations of the absolute values of the daily returns time series, and  $e^2(\cdot)$  is the squared relative (i.e., percentage) error between the time series produced by the model and the historical time series, with the historical value in the denominator. Since estimated higher order moments like kurtosis have high variance, including the range in the objective function resulted in a better fit in the tails. We also took the autocorrelations of market returns to be zero, which is why we did not apply the error function to  $ACF_3(r)$ . Note that the mean return,  $\mu$ , is specified directly in the model so we did not include it in the objective function. Finally, since the statistical measures



are reported in different units, we chose the weights heuristically to achieve a good fit without any one component dominating the calibration.

To avoid local minima, we performed a grid search on the cartesian product of the sets of parameters given in Eq. (14) to find the set of parameter values that achieved the lowest average objective function value. In order to mitigate simulation bias and obtain more robust results, the model was calibrated by running 10 replications for each set of parameter values and taking the average of the objective function value for the 10 runs, where the same set of 10 starting seeds was used for each parameter set. For each replication, we ran the model for 6,000 time steps with a warmup period of 1,000, where the first 1,000 days of the simulation were discarded to avoid transient phases. This left 5,000 samples from which to calculate the objective function for each replication, which is consistent with the size of our historical SPY data set.

We found that the minimum average objective function value was achieved for  $\beta = 150$  and  $\nu_1 \Delta t = 0.7$ . Summary statistics of the daily log returns for a single replication of the calibrated model for a 5,000 day simulation period are given in the ABM column of Table 1. The time series and histogram of returns are plotted in Figures 2(e) and 2(f). The autocorrelation plots for  $r_t$  and  $|r_t|$  are shown in Figures 3(e) and 3(f).

### 3.4 Analysis of Results

We see very clear evidence of volatility clustering for the SPY and ABM time series in Figures 2(a) and 2(e), where returns have long periods of small price movements punctuated by bursts of much larger movements. This is contrasted with the returns for GBM in Figure 2(c), which show uniform volatility throughout. We also see clear evidence of fat tail (leptokurtic) behavior in that occasionally there are much larger daily price movements for the SPY and ABM time series compared to GBM. This is further confirmed in the histograms in Figures 2(b) and 2(f), which have very similar shape. Note the much larger weight given to the tails compared to the histogram for GBM returns in Figure 2(d), which are normally distributed by construction. We can also see evidence of negative skew for SPY and ABM in Figures 2(b) and 2(f), with more weight given to negative returns relative to positive returns.

These observations are in line with the summary statistics in Table 1. While all time series produced similar standard deviations for returns, only SPY and ABM show non-trivial negative skewness and large positive kurtosis. Furthermore, given the negligible p-values from the Jarque-Bera tests, which test for normality using a test statistic that incorporates both skewness and kurtosis, we strongly reject the null hypothesis that the SPY and ABM daily log returns are normally distributed.

As expected, we fail to reject the null hypothesis for the normally distributed GBM returns.

Moving on to Figure 3, from the left side, we see that all three time series show negligible autocorrelations for daily returns. However, we again see very clear evidence of volatility clustering for SPY and ABM in Figures 3(b) and 3(f) with substantial autocorrelations among the absolute values of returns. So while the returns for SPY and ABM are uncorrelated, they are certainly not independent. We also note that the values of the autocorrelations of the absolute values of returns align closely between SPY and ABM, where both are decreasing functions of the number of lags. As expected, there is negligible autocorrelation among the absolute values of returns for GBM in Figure 3(d), since GBM returns are independent by construction.

These observations are consistent with the Ljung-Box results in Table 1, which test for significant autocorrelations over  $n$  lags. As is standard in the empirical finance literature, we tested for autocorrelations in the absolute values of returns,  $|r_t|$ , over  $n = 15$  daily lags. The null hypothesis of zero autocorrelation for all  $n$  lags is strongly rejected for SPY and ABM, but not for GBM.

Note that the above results are for a single replication of the model. To give a more complete picture of the general behavior of the model, we also ran 100 replications of the fully calibrated model, calculated summary statistics for each replication (i.e., each sample path), and looked at their distribution. Histograms for the mean, standard deviation, skewness, and excess kurtosis for each of the 100 replications are plotted in Figure 4. The corresponding monte carlo statistics are given in Table 3.

We see that the mean, standard deviation, and skewness of the historical SPY data set all fall within the 95% confidence intervals produced by the model, while on average the excess kurtosis generated by the model is lower than the excess kurtosis of the historical data set.<sup>10</sup> Matching higher-order moments like kurtosis is generally more difficult than lower-order moments due to high sensitivity to outliers. However, we do see that the model can produce excess kurtosis that matches and even significantly exceeds the excess kurtosis of the historical SPY data set. As these results show, the model overall is quite capable of reproducing the many stylized facts found in real financial markets.

<sup>10</sup>For comparison, we also ran the same analysis on the functional form given in [12], specifically  $\frac{dS_t}{dt} = \alpha Y^f (V_t - S_t) S_t + \beta Y^c S_t X_t$ , where  $Y^f$  and  $Y^c$  are constants. While 95% CIs for mean and standard deviation were similar to those produced by Eq. (7), the calibrated model produced less pronounced average skewness (-0.1251 with a range of -0.3649 to 0.0928) and much smaller average excess kurtosis (3.2708 with a range of 1.4156 to 6.2518).

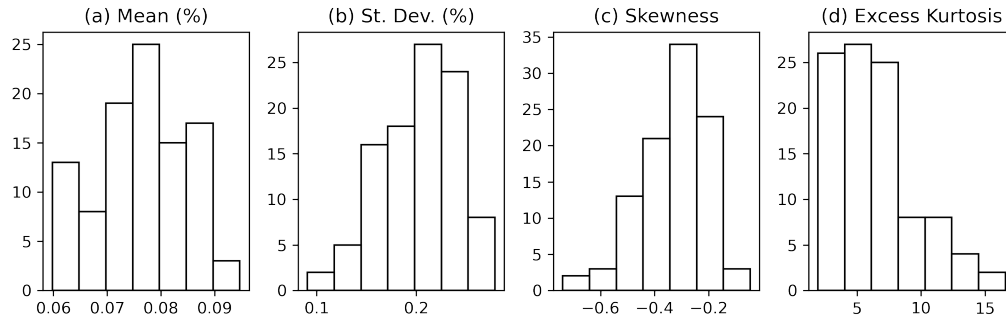


Figure 4: Histograms of time series statistics for 100 replications of ABM

Statistic	Average	95% CI	St. Dev.	Min	Max
Mean (annual)	0.0764	(0.0747, 0.0781)	0.0084	0.0598	0.0947
St. Dev. (annual)	0.2021	(0.1944, 0.2099)	0.0389	0.0908	0.2792
Skewness	-0.3256	(-0.3509, -0.3003)	0.1269	-0.7399	-0.0470
Excess Kurtosis	6.3694	(5.7458, 6.9929)	3.1270	1.9537	16.5794

Table 3: Monte carlo statistics for 100 replications of ABM

#### 4 Conclusion

This paper proposes an agent-based model that can accurately replicate stylized facts in real financial market data, while allowing for a behavioral explanation of the drivers of universal characteristics of financial markets. The model we propose uses simple behavioral heuristics expressed through transition probabilities and trading volumes. Central to the model is the herding behavior of noise traders, who increasingly enter optimistic markets and sell out of pessimistic markets, where an element of panic selling is captured via larger trading volumes when the market is pessimistic. On the other hand, fundamentalist traders tend to pull the market price back to its fundamental economic value, and they do so with greater trading volume for larger divergences in price (i.e., higher expected returns). Furthermore, as the market diverges from the fundamental value, noise traders become aware of the significant mispricing and increasingly switch to fundamentalist traders, betting that the market will revert towards its fundamental value. Finally, when the market is close to its fundamental value, little will be gained from a reversion to market fundamentals, so fundamentalists increasingly switch to noise traders to follow the short-term trend.

The model reproduces empirical stock price behavior while being built on as-

sumptions where the underlying economic value of the stock is growing at a deterministic and constant rate. In other words, the model presented in this paper shows that the market itself can cause volatility clustering, negative skew, and fat tail behavior, even when the fundamental value of the stock shows no such behavior.

While agent-based models are often employed in modeling complex systems, the absence of robust methods to calibrate and validate these models is frequently reported in the literature [23]. We took a simple approach to calibration in this paper, but work can be undertaken to better explore the parameter space and develop superior calibration methodologies. Furthermore, similar to [12], given the relative simplicity of the model, work can be undertaken to develop analytical expressions where possible, which could lead to further insights.

#### Acknowledgements

We wish to thank Hiroki Sayama and two anonymous reviewers whose critical comments and suggestions helped improve this manuscript. All errors are our own.

#### References

- [1] Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1:223–236.
- [2] Hutson, E., Kearney, C., & Lynch, M. (2007). Volume and skewness in international equity markets. *Journal of Banking & Finance*, 32:1255–1268.
- [3] Hull, J. C. (2009). *Options, Futures, and Other Derivatives*, 7th Ed. Pearson Prentice Hall.
- [4] Black, F. & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654.
- [5] Engle, R. F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4):987–1007.
- [6] Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–343.
- [7] Lux, T. & Alfarano, S. (2016). Financial power laws: Empirical evidence, models, and mechanisms. *Chaos, Solitons & Fractals*, 88(C):3–18.

- [8] Wang, L., Ahn, K., Kim, C., & Ha, C. (2018). Agent-based models in financial market studies. *Journal of Physics: Conf. Ser.*, 1039:012022.
- [9] Bookstaber, R. (2017). Agent-based models for financial crises. *Annual Review of Financial Economics*, 9(1):85-100.
- [10] LeBaron, B. (2001). A builder's guide to agent-based financial markets. *Quantitative Finance*, 1:254–261.
- [11] Kirman, A. (1993). Ants, Rationality, and Recruitment. *The Quarterly Journal of Economics*, 108:137-156.
- [12] Alfarano, S. & Lux, T. (2007). A noise trader model as a generator of apparent financial power laws and long memory. *Macroeconomic Dynamics*, 11(S1):80–101.
- [13] Day, R. H. & Huang, W. (1990). Bulls, Bears, and Market Sheep. *Journal of Economic Behavior and Organization*, 14:299-329.
- [14] Lux, T. (1995). Herd Behaviour, Bubbles and Crashes. *The Economic Journal*, 105:881-896.
- [15] Lux, T. & Marchesi, M. (2000). Volatility clustering in financial markets: a microsimulation of interacting agents. *International Journal of Theoretical and Applied Finance*, 3(4):675-702.
- [16] Shiller, R. J. (2000). *Irrational Exuberance*. Princeton University Press.
- [17] Li, C., Huang, W., Wang, W., & Chia., W. (2021). Price Change and Trading Volume: Behavioral Heterogeneity in Stock Market. *Computational Economics*, 2021.
- [18] Campbell, J. Y. & Shiller, R. J. (2001). Valuation Ratios and the Long-Run Stock Market Outlook: An Update. *NBER Working Paper Series*, No. 8221.
- [19] Zubillaga, B. J., Vilela, A. L. M., Wang, C., Nelson, K. P., & Stanley, H. E. (2022). A three-state opinion formation model for financial markets. *Physica A: Statistical Mechanics and its Applications*, 588:126527.
- [20] Winker, P., Gilli, M., & Jeleskovic, V. (2007). An objective function for simulation based inference on exchange rate data. *Journal of Economic Interaction and Coordination*, 2:125-145.

- [21] Chen, Z. & Lux. T. (2018). Estimation of Sentiment Effects in Financial Markets: A Simulated Method of Moments Approach. *Computational Economics*, 52:711-744.
- [22] Shahriari, B., Swersky, K., Wang, Z., Adams, R. P., & De Freitas, N. (2016). Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proceedings of the IEEE*, 104(1):148-175.
- [23] Lamperti, F., Roventini, A., & Sani, A. (2018). Agent-based model calibration using machine learning surrogates. *Journal of Economic Dynamics and Control*, 90:366–389.