PREDICTING MARKET TRENDS BY MEANS OF AGENT-BASED MODELS*

Przemysław Repetowicz

Probability Dynamics, Custom House Quay, Dublin 1, Ireland

(Received February 18, 2020)

Financial markets are examples of complex systems being driven by both endogenous processes as well as exogenous shocks. One can obtain insight into the dynamics of those systems by assuming market participants (agents) belonging to particular groups and being characterized by a particular demand and supply functions. The agents themselves are likely to switch to a different group under persuasion of their peers (herding behavior). The mathematical description of this phenomenon in terms of a non-trivial Markov model, combined with the clearing of the market by a market maker provides analytical results for the probability distribution of the asset prices and returns. That in turn allows for trading strategies to be designed and implemented in real life. In the course of the article, I will review several agent-based models as well as mention the practicalities of quantitative trading.

DOI:10.5506/APhysPolBSupp.13.597

1. Introduction

— We use a *bottom-up* approach to describe the financial market.

— We start from the rules of behavior and interaction between traders (agents) to reproduce market features on the macro-scale (stylized facts, trends and patterns).

— Our models are inspired by *statistical physics* (kinetic spin models, cellular automata, path integrals) however, in addition, we use modern tools from *probability and statistics* (Markov chains, convergence of stochastic processes) along with *financial economics* (Walras equilibrium theory, behavioral economics).

— We describe *convergence* of economic systems towards equilibrium rather than only their behavior at equilibrium.

* Plenary talk presented at the 45$^{\text{th}}$ Congress of Polish Physicists, Kraków, September 13–18, 2019.

2. Objectives

— Describe and *predict bubbles* in the financial or housing markets.
— Develop new reliable trading strategies (statistical arbitrage) to be used in turbulent times (subprime mortgage crisis).
— Develop methods for *hedging large fluctuations* (volatility trading) to avoid financial disasters (Long Term Capital Management, Northern Rock).

3. Model assumptions

— The market is composed of *agents* who buy, sell or hold the asset.
— Agents are divided into several groups depending on their behavior, their access to information and their *demand* or *supply* for the asset.
— We consider three groups; *the fundamentalists* (price asset according to its "fundamental" value), *the trend followers* (watch the trend and invest accordingly) and *noise traders* (bet randomly).
— Agents can switch between groups *endogenously* with likelihoods that depend on the numbers of agents.
— The total number of agents N is finite and the distribution of the numbers of agents, *i.e.* the numbers of fundamentalists n_{F}, of trend followers n_{TF} and of noise traders n_{NT}, evolves in discrete time as a zero-memory process (*Markov chain*).
— The asset price is settled by a *Walrasian auctioneer* who matches the supply and demand of traders.

4. The results

A *Non-linear, non-stationary Markov chain*: The full time-dependent solution for the evolution of a generic class of Markov chains with occupation dependent transition probabilities[1].

B *Non-stationary distribution of price returns*: The distribution of returns as a function of time.

C *Parameter estimation*: Maximum likelihood and the method of moments.

A: Let $(n_t^{(\mathrm{T})}, n_t^{(\mathrm{F})}) = (n, N-n)$ subject to $n_{\mathrm{T}} + n_{\mathrm{F}} = N = \mathrm{const}$ be the numbers of agents in the different groups at time t and let $f_t(n) := \mathbb{P}\left(n_t^{(\mathrm{T})} = n\right)$. Then:

[1] The above result can be applied in such fields as non-linear diffusion processes or modeling distributions of wealth in the society, modelling spreading of epidemics, *etc.*

Master equation (conditional probabilities)

$$\begin{aligned} f_{t+1}(n) &= \pi\,(n-1\to n)\;f_t(n-1) \\ &\quad +\pi\,(n+1\to n)\;f_t(n+1) \\ &\quad +(1-\pi\,(n\to n+1)-\pi\,(n\to n-1))\;f_t(n)\,. \end{aligned} \tag{4.1}$$

Following Kirman (ensure $n=0,N$ absorbing states), we take

$$\begin{aligned} \pi\,(n\to n+1) &= (N-n)\,(a_+ + bn)\,, \\ \pi\,(n\to n-1) &= n\,(a_- + b(N-n))\,, \end{aligned} \tag{4.2}$$

where $a_\pm = A_\pm/N$ and $b=B/N^2$ (ensure probabilities positive).
Solution (small-t expansion)

$$f_t(n) = \sum_{p=0}^{t} \delta_{n,p} \sum_{j=0}^{2t-2} \mathfrak{A}_p^{(j)}(t)\frac{1}{N^j}\,, \tag{4.3}$$

where

$$\mathfrak{A}_p^{(0)}(t) := \binom{t}{p}(1-A_+)^{t-p}A_+^p\,, \tag{4.4}$$

$$\mathfrak{A}_p^{(j)}(t) := \sum_{q_1=\lceil \frac{j}{2}\rceil}^{2j-1} \sum_{q=0}^{1+q_1} \binom{t}{1+q_1}\binom{t-1-q_1}{p-q}(1-A_+)^{t-1-q_1-p+q}\,A_+^{p-q+1}\mathfrak{W}^{(j)}_{q_1-\lceil\frac{j}{2}\rceil+1,1+q}\,, \tag{4.5}$$

where

$$\mathfrak{W}^{(j)}_{q,q_1} := \frac{1}{(q_1-1)!}\frac{\mathrm{d}^{q_1-1}}{\mathrm{d}z^{q_1-1}}\mathcal{W}^{(j)}_{q+\lceil j/2\rceil}(z)\bigg|_{z=0} \quad \text{for} \quad \begin{matrix} q=1,2j-\lceil j/2\rceil \\ q_1=1,1+q+\lceil j/2\rceil \end{matrix}\,, \tag{4.6}$$

$$\begin{aligned} &\mathcal{W}_q^{(j)}(z) = \\ &\left[A_+\,(q-1)\mathcal{W}^{(j-1)}_{q-2}(z)+\left(\mathcal{W}^{(j-1)}_{q-1}(z)\right)'\right](1-z)\,(A_-+B+z(A_+-B)) \\ &+\left[A_+(q-1)\mathcal{W}^{(j-2)}_{q-2}(z)+\left(\mathcal{W}^{(j-2)}_{q-1}(z)\right)'+A_+^2(q-1)_{(2)}z\mathcal{W}^{(j-2)}_{q-3}(z)\right. \\ &\left.+\,2A_+(q-1)z\left(\mathcal{W}^{(j-2)}_{q-2}(z)\right)'+z\left(\mathcal{W}^{(j-2)}_{q-1}(z)\right)''\right](-B)\,(1-z)^2 \end{aligned} \tag{4.7}$$

for $j\geq 3$ and $\lceil (j+2)/2\rceil \leq q \leq 2j$. Subject to initial values on quantities $\mathcal{W}_{\cdot}^{(1)}$ and $\mathcal{W}_{\cdot}^{(2)}$.

1. Normalization

$$\sum_{p=0}^{t} \mathfrak{A}_p^{(j)}(t) = \delta_{j,0}\,. \tag{4.8}$$

2. Stationary state

$$\lim_{t\to\infty} f_t(n) = \binom{N}{n} \frac{\left(\frac{a_+}{b}\right)^{(n)}\left(\frac{a_-}{b}\right)^{(N-n)}}{\left(\frac{a_+}{b}+\frac{a_-}{b}\right)^{(N)}}\,. \tag{4.9}$$

We plot the solution (4.3) in Fig. 1 and the time-evolution of the distribution in Fig. 2.

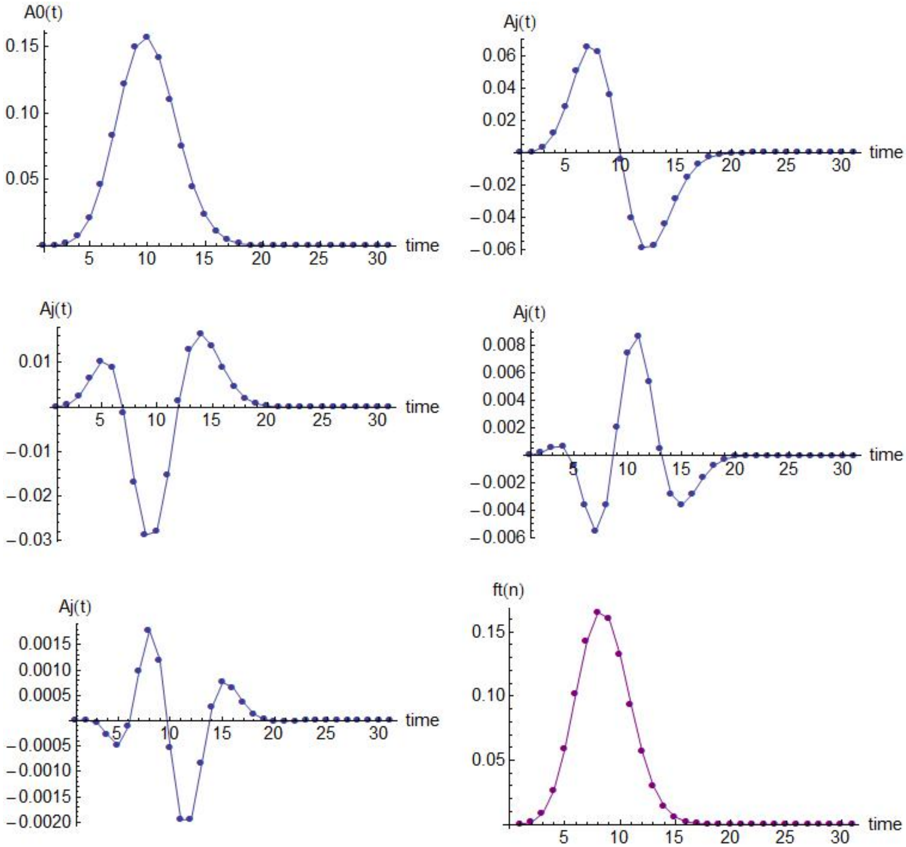

Fig. 1. The leading order along with the higher order terms in the distribution of the number of agents $f_t(n)$. Here, we took $t = 30$, $N = 40$ and $(A_+, A_-, B) = (0.3, 0.2, 0.2)$.

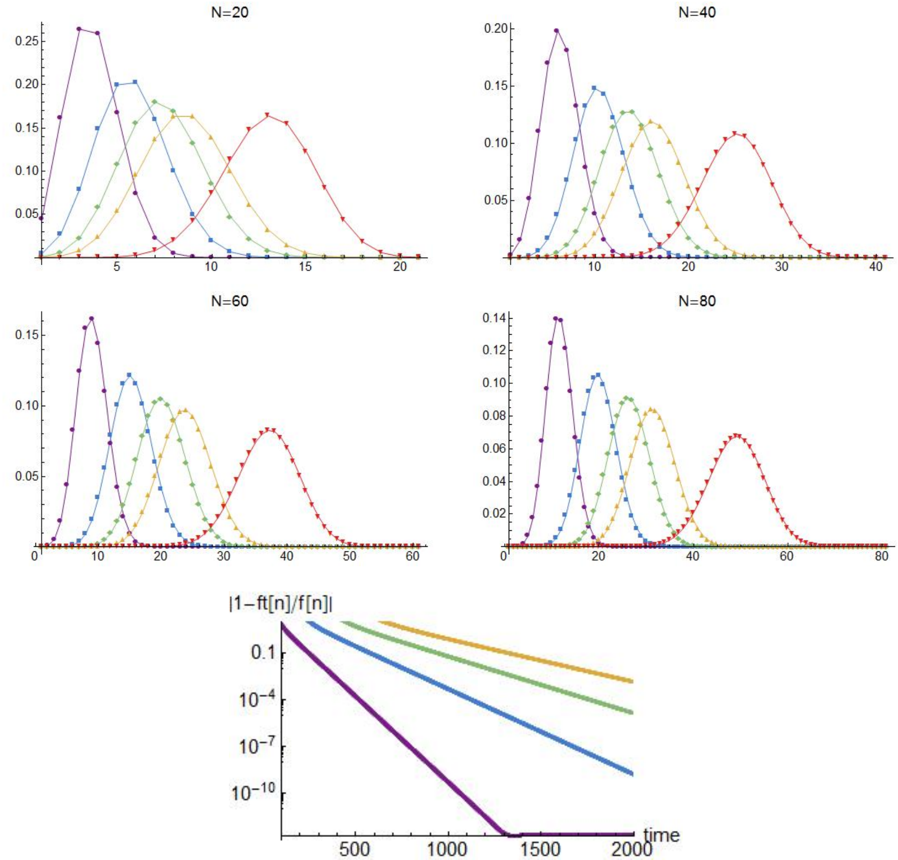

Fig. 2. (Color online) Top: The agent distribution $f_t(n)$ from equation (4.3) as a function of time t. Here, for each value of N, we take $t = N/2, N, 3/2N, 2N$ (from left to right: violet, blue, green, yellow and red). The model parameters are $(A_+, A_-, B) = (0.3, 0.2, 0.2)$ as before. Bottom: The distance between the current distribution and the steady state (the coloring as before except for the red which is missing).

Proof outline

1. Insert Ansatz (4.3) into master equation (4.1) and work out recurrence relations for modes $\mathfrak{A}_p^{(j)}(t)$.
2. Take a Z-transform with respect to variable p obtaining recurrence relation with differentiation with respect to z.
3. Solve the resulting recurrence equations for $j = 0, 1, \ldots$ and observe the pattern in the solutions (as a function of j) and prove it by induction.
4. Invert the Z-transform.

5. Mathematica file with proofs available upon request.
6. Stationary state obtained from detailed balance

$$\pi(n+1 \to n)\, f(n+1) = \pi(n \to n+1)\, f(n)\,.$$

Time evolution of the distribution

Solution (large-t expansion)

$$f_t(n) = \sum_{M=0}^{N} C_M\ (1-E_M)^t\ X_n^{(M,N)}\,, \tag{4.10}$$

where $E_M := b(\epsilon_- + \epsilon_+ + (M-1))\,M$ and

$$\begin{aligned}(\epsilon_- + \epsilon_+)^{(N)}\, X_n^{(M,N)} &:= \sum_{m=0}^{n} (-1)^{n-m} \\ &(-\epsilon_- - M+1)^{(m)} (\epsilon_-)^{(N-m)} \binom{-\epsilon_- - \epsilon_+ + 1 - M}{n-m} \binom{\epsilon_- + \epsilon_+ + M + N - 1}{m} \\ &= \binom{N}{n} \frac{\epsilon_-^{(N-n)} (\epsilon_+ + M)^{(n-M)}}{(N-M+1)^{(M)}}\, \mathcal{P}^{(M)}(\epsilon_-, \epsilon_+)\,,\end{aligned} \tag{4.11}$$

where $\left(\mathcal{P}^{(M)}(\epsilon_-, \epsilon_+)\right)_{M=0}^{1} = (1, -n\,\epsilon_- + (N-n)\,\epsilon_+)\ .$

We plot the solution (4.10) along with its large-N asymptotics in Fig. 3.

Proof outline

1. Rewrite master equation into ODE form

$$f_{t+1}(n) - f_t(n) = \Delta_n \left[\pi(n+1 \to n)\, f_t(n+1) - \pi(n \to n+1)\, f_t(n)\right]\ . \tag{4.12}$$

2. Assume factorization $f_t(n) := T^{(E)}(t)\, X^{(E)}(n)$, then

$$T^{(E)}(t) = T(0)\ (1+E)^t\ , \tag{4.13}$$

$$\begin{aligned}E\, \tilde{X}^{(E)}(z) &= (1-z) \\ &\times \underbrace{\left[-Na_+ + ((a_- + b(N-1)) + (a_+ - b(N-1))z)\frac{\mathrm{d}}{\mathrm{d}z} - bz(1-z)\frac{\mathrm{d}^2}{\mathrm{d}z^2}\right]}_{\hat{H}_z} \\ &\times \tilde{X}^{(E)}(z)\,,\end{aligned} \tag{4.14}$$

where $\tilde{X}^{(E)}(z) := \sum_{n=0}^{N} X^{(E)}(n) z^n$ is the Z-transform of $X^{(E)}(n)$.

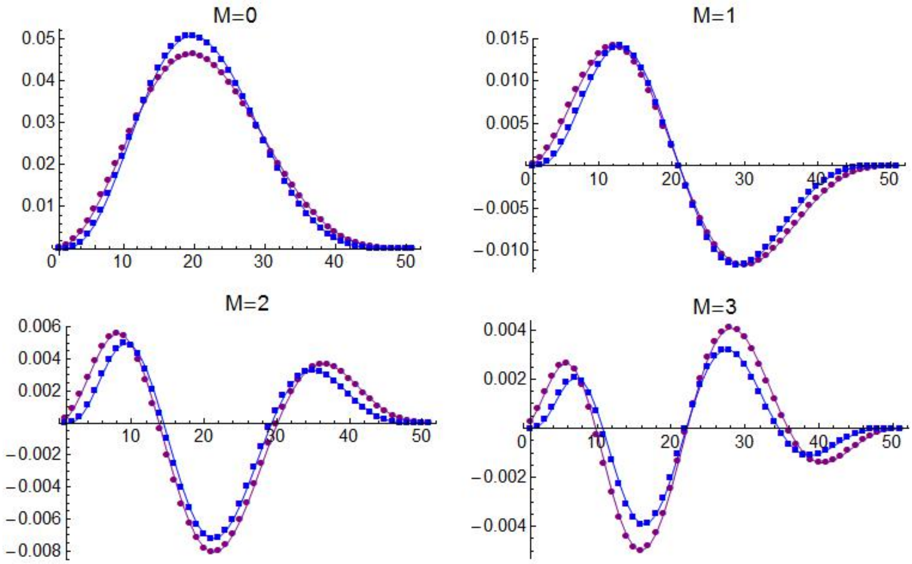

Fig. 3. (Color online) The finite-N modes (dots/purple) along with the large-N asymptotics (squares/blue) for $M = 0, \ldots, 5$ from the top to the bottom, respectively. I took $N = 50$ and $(a_+, a_-, b) = (0.4, 0.6, 0.1)$. Note that $\epsilon_\pm := a_\pm / b$.

3. Define $\mathit{\Delta} := (b - a_- - a_+)^2 - 4bE$ and solve ODE (4.14)

$$\tilde{X}^{(E)}(z) \simeq (1-z)^{\frac{-a_- - a_+ + b - \sqrt{\Delta}}{2b}} \times F_{2,1}\left[\begin{matrix} \frac{-a_- + a_+ + b - \sqrt{\Delta}}{2b}, & -N + \frac{a_- - a_+ + b - \sqrt{\Delta}}{2b} \\ 1 - \frac{a_-}{b} - N & \end{matrix} ; z\right]. \tag{4.15}$$

4. Boundary conditions (probability flux conservation)

$$\sum_{n=0}^{N} X^{(E)}(n) = \delta_{E,0} \Longleftrightarrow E = -(a_- + a_+ + b(M-1))M\,, \quad M \in \mathbb{N} \wedge M \leq N \tag{4.16}$$

$$\Longrightarrow$$

$$\tilde{X}^{(E)}(z) = (1-z)^M \frac{(\frac{a_-}{b})^{(N)}}{(\frac{a_-}{b} + \frac{a_+}{b})^{(N)}} \times F_{2,1}\left[\begin{matrix} \frac{a_+}{b} + M, & -N + M \\ 1 - \frac{a_-}{b} - N & \end{matrix} ; z\right]. \tag{4.17}$$

5. Invert Z-transform

$$(\epsilon_- + \epsilon_+)^{(N)} X^{(E)}(n) = \binom{N}{n} \frac{\epsilon_-^{(N-n)}(\epsilon_+ + M)^{(n-M)}}{(N-M+1)^{(M)}} \mathcal{P}_n^{(M)}(\epsilon_-, \epsilon_+), \tag{4.18}$$

where

$$\left(\mathcal{P}_n^{(p)}(\epsilon_-, \epsilon_+)\right)_{p=0}^1 = (1, -n\,\epsilon_- + (N-n)\,\epsilon_+) \tag{4.19}$$

and

$$\mathcal{P}_n^{(M)}(\epsilon_-, \epsilon_+) := \sum_{p=0}^{M\wedge n} (-1)^p \binom{M}{p} (N-n)_{(M-p)}\, n_{(p)}\, (\epsilon_- - n + N)^{(p)} (\epsilon_+ + n)^{(M-p)}. \tag{4.20}$$

B: The time evolution of the distribution of returns.
We generalize existing approaches [1–4] and model the *excess demand* of fundamentalists, trend followers and speculators as $\log(\frac{p_\mathrm{F}}{p_t})$, $\frac{\dot{p}_t}{p_t}$ and ξ_t. Here, $p_\mathrm{F} = \mathrm{const}$ is the "fundamental value" of the price and ξ_t is a random variable with independent increments.
The Walrasian equilibrium yields

$$n_\mathrm{TF} \frac{p_{t+\Delta t} - p_t}{p_t} = n_\mathrm{F} \log\left(\frac{p_\mathrm{F}}{p_t}\right) + n_\mathrm{NT} \xi_t\,. \tag{4.21}$$

Define $r_t := \frac{p_{t+\Delta t} - p_t}{p_t}$. In the following, we analyze limit cases.

Trend followers and fundamentalists only

Denote $\epsilon^\pm := a^\pm / b$. Evolution of returns (take (4.21) and differentiate)

$$r_{t+\Delta t} = r_t \left(1 - \frac{z}{1-z}\right)\Big|_t\,, \tag{4.22}$$

where $z := n_\mathrm{F}/N$ is the fraction of fundamentalists.

The pdf of returns[2]

$$\rho_{r_t}(x) = \int_0^1 \delta\left(x - r_0\left(1 - \frac{z}{1-z}\right)^{\frac{t}{\Delta t}}\right) \underbrace{\frac{z^{-1+\epsilon_+}(1-z)^{-1+\epsilon_-}}{B(\epsilon_-,\epsilon_+)}}_{\text{stationary agents' pdf}} \mathrm{d}z$$

$$= \frac{\left(1 - \left(\frac{x}{r_0}\right)^{\frac{1}{n}}\right)^{\epsilon_+ -1} \left(2 - \left(\frac{x}{r_0}\right)^{\frac{1}{n}}\right)^{-\epsilon_- -\epsilon_+}}{B(\epsilon_-,\epsilon_+)\left| n r_0 \left(\left(\frac{x}{r_0}\right)^{\frac{1}{n}}\right)^{n-1}\right|}, \tag{4.23}$$

where $n := t/\Delta t$ is odd.

The moments[3]

$$E\left[r_t\right] = r_0 E\left[\left(1 - \frac{z}{1-z}\right)^n\right] = r_0 \sum_{p=0}^{n} \binom{n}{p} (-1)^p \frac{B(\epsilon_+ + p, \epsilon_- - p)}{B(\epsilon_+,\epsilon_-)}, \tag{4.24}$$

$$E\left[r_t\right] = r_0^2 E\left[\left(1 - \frac{z}{1-z}\right)^{2n}\right] = r_0^2 \sum_{p=0}^{2n} \binom{2n}{p} (-1)^p \frac{B(\epsilon_+ + p, \epsilon_- - p)}{B(\epsilon_+,\epsilon_-)}. \tag{4.25}$$

Speculators and fundamentalists only

Then the evolution of returns

$$r_t \simeq \log \frac{p_{t+\Delta t}}{p_t} = \frac{1-z}{z} \Delta \xi_t\,, \tag{4.26}$$

where $z := n_{\mathrm{F}}/N$ is the fraction of fundamentalists.

The pdf

$$\rho_{r_t}(x) = \int_{-1}^{1}\int_0^1 \delta\left(x - \frac{1-z}{z}\eta\right) \underbrace{\frac{z^{-1+\epsilon_+}(1-z)^{-1+\epsilon_-}}{B(\epsilon_-,\epsilon_+)}}_{\text{agents' pdf}} \underbrace{\upsilon_{\Delta\xi_t}(\eta)}_{\text{noise increment' pdf}} \mathrm{d}z\mathrm{d}\eta$$

---

[2] For x negative, replace $(x/r_0)^{1/n}$ by $-(|x|/r_0)^{1/n}$.

[3] The 1st and the 2nd moments exist iff $\epsilon_- > n$ and $\epsilon_- > 2n$, respectively.

$$= \int_{-1}^{1} \frac{\operatorname{sign}(\eta)}{x\, B(\epsilon_-, \epsilon_+)} \left(\frac{\eta}{\eta + x}\right)^{\epsilon_+} \left(\frac{x}{\eta + x}\right)^{\epsilon_-} 1_{\eta\, x \geq 0}\, \upsilon_{\Delta\xi_t}(\eta) \mathrm{d}\eta$$

$$\underbrace{=}_{\upsilon_{\Delta\xi_t}(\eta) = p 1_{\eta > 0} + q 1_{\eta < 0}} \frac{|x|^{-\epsilon_+ - 1}\, {}_2F_1\left(\epsilon_+ + 1, \epsilon_- + \epsilon_+; \epsilon_+ + 2; -\frac{1}{|x|}\right)}{(\epsilon_+ + 1) B(\epsilon_-, \epsilon_+)}$$

$$\times \left(p\, 1_{x \geq 0} + q\, 1_{x < 0}\right) . \tag{4.27}$$

Let $(\tilde{\mu}_j)_{j=1}^{\infty}$ denote the moments of $\Delta\xi_t$.

The first moment (trend)[4]

$$\mu_1 := E\left[r_t\right] = E\left[\frac{1 - z}{z}\right] E\left[\Delta\xi_t\right] = \frac{\epsilon^-}{\epsilon^+ - 1} \tilde{\mu}_1 . \tag{4.28}$$

The variance (volatility)[5]

$$E\left[(r_t - \mu_1)^2\right] = \left(\frac{\epsilon^-}{\epsilon^+ - 1}\right)^2 \left[\frac{\epsilon^+ - 1}{\epsilon^+ - 2} \frac{\epsilon^- + 1}{\epsilon^-} \tilde{\mu}_2 - (\tilde{\mu}_1)^2\right] . \tag{4.29}$$

C: Parameter estimation.

— Method of moments:

1. Equate moments in (4.28), (4.29) to empirical moments.
2. Solve for parameters $\epsilon_\pm$.
3. From the moments compute get *trend* and *volatility* estimates.
4. Employ the Kelly criterion [5] and use the parameters in your trading strategy.

— Method of maximum likelihood:

1. Find estimators of parameters $\epsilon_\pm$ by maximizing the log-likelihood–use expression (4.27) for the pdf of returns.
2. Find errors for parameters $\epsilon_\pm$ by Taylor expanding log-likelihood around maximum above.
3. Compare results to those from above.
4. If results match, proceed as in the last point above.

[4] Exists for $\epsilon_+ > 1$ only.
[5] Exists for $\epsilon_+ > 2$ only.

5. Conclusions

— Basing firstly on types of trader's behavior (competition between *greed* and *fear*), then secondly on tendencies to change behavior (*imitation, herding*) and, finally, on the auctioneer balancing *supply and demand*, we described quantitatively the whole non-stationary distribution of price returns.

— The above description is encapsulated in three interrelated models and in a recipe to test the models with data and to estimate (calibrate) the relevant model parameters.

— The result can (will) be used in a statistical arbitrage trading strategy.

— The analytical derivations and other relevant documentation (Mathematica code) is available upon request.

REFERENCES

[1] A. Kirman, *Q. J. Econ.* **108**, 137 (1993).

[2] S. Alferano, T. Lux, F. Wagner, *Comput. Econ.* **26**, 19 (2005).

[3] N.T. Chan, B. LeBaron, A.W. Lo, T. Poggio, «Agent-based Models of Financial Markets: A Comparison with Experimantal Markets», unpublished working paper, MIT, 1999.

[4] F. Wagner, *Physica A* **322**, 607 (2003), `arXiv:cond-mat/0207280`.

[5] J.L. Kelly, *Bell Syst. Tech. J.* **35**, 917 (1956).