



TECHNISCHE
UNIVERSITÄT
WIEN

S E M I N A R A R B E I T

Stock price formation: useful insights from a multi-agent reinforcement learning model

ausgeführt am

Institut für
Finanz- und Versicherungsmathematik
TU Wien

unter der Anleitung von

**Associate Prof. Dipl.-Ing. Dr.techn.
Gerhold Stefan**

durch

Holzbauer Paul

Matrikelnummer: 11815131

Wien, Februar 2021

Inhaltsverzeichnis

1	Einleitung	1
2	Aufbau des Modells	2
2.1	Grundwerte der Agenten	3
2.2	Das order Buch	3
2.3	Handelsalgorithmen	4
2.3.1	Algorithmus \mathcal{F}^i	4
2.3.2	Algorithmus \mathcal{T}^i	6
3	Handelsstrategie der Agenten	8
4	Einflüsse auf den Markt	10
4.1	Best agent herding	12
4.2	Worst agent herding	15
4.3	Noise traders	16
5	Fazit	19
	Literaturverzeichnis	20
	Abbildungsverzeichnis	21

1 Einleitung

Multi-agent Systeme sind seit vielen Jahren Teil der quantitativen Finanzforschung. Sie wurden speziell verwendet um Auswirkungen auf die Marktstruktur, also die Einzelheiten des Austauschs auf den Märkten, studieren zu können. Hierbei haben financial multi-agent Systeme gezeigt, dass es spezielle Muster gibt, welche für alle Gruppen von Finanzprodukten angemessen sind.

Diese Forschung wurde lange Zeit infrage gestellt, da die Herausforderung bestand Agenten realistisch und menschenähnlich zu erstellen. Früher hat man sich auf so genannte zero-intelligence agents verlassen, also Agenten die nach speziellen und vorher festgelegten Regeln handeln. Dabei wurden natürlich dynamische Entwicklungen wie zum Beispiel Erkenntnisse aus der Entscheidungstheorie oder Spieltheorie vernachlässigt, welche natürlich für einen Markt essentiell sind. Jedoch verglichen mit anderen Modellen haben multi-agent Systeme, zwei bedeutende Vorteile:

- es werden deutlich weniger Annahmen, wie eine Gaußverteilung oder Markteffizienzhypothese, benötigt
- sie stellen auf natürliche Art und Weise spezielle wichtige Phänomene angemessen an das komplexe System dar

In den letzten Jahren gab es auch enorme Fortschritte im Bereich des machine learnings. Insbesondere wurden diese auch im Fachgebiet des reinforcement learnings, also jenem Bereich, wo Agenten selbstständig Strategien entwickeln um ihren Gewinn zu maximieren, erzielt.

Diese Seminararbeit basiert auf dem Artikel *Stock price formation: useful insights from a multi-agent reinforcement learning model* von J. Lussange, S. Bourgeois-Gironde, S. Palminteri, and B. Gutkin. [1]

2 Aufbau des Modells

Das Modell basiert auf einer bestimmten Anzahl I an Agenten die jeweils unabhängig voneinander mit J Aktien handeln. Zu jedem Zeitpunkt $t \in \{1, \dots, T\}$ kann gehandelt werden, wobei T im Vorhinein zufällig ausgewählt werden kann oder man versucht ein bestimmtes Zeitintervall wie eine Woche T_w , ein Monat T_m oder ein Jahr T_y zu simulieren. Zu erwähnen ist hierbei, dass $T_w = 5$ ist, da nahezu alle Börsen der Welt am Wochenende geschlossen haben und somit es auch keine Möglichkeit für einen Handel gibt. Dementsprechend reduziert sich auch die jeweilige Anzahl für ein Monat auf 21 Tage und für ein Jahr auf 252 Tage.

Zu Beginn werden alle Agenten zum Zeitpunkt $t = 0$ mit ihren jeweils eigenen Parametern (siehe 2.1) initialisiert und sie erhalten ein Portfolio der Form $A_{equity}^i(t) = \sum_{j=0}^J Q^{i,j}(t) \cdot P^j(t)$. $Q^{i,j}(t)$ beschreibt die Anzahl der Aktien j von Agent i zum Zeitpunkt t und $P^j(t)$ den Markt Preis der Aktie j zum Zeitpunkt t . Danach werden nun alle Marktpreise mit dem Wert $P^j(t = 0) = \text{€}100$ initialisiert und eine Simulation erstellt J Zeitreihen $\mathcal{T}^j(t)$, die auf den gegebenen Grundwerte der jeweiligen Aktie basiert. Die Agenten kennen diese natürlich nicht und daher approximieren sie den Wert $\mathcal{T}^j(t)$ der Aktie j und wenden das Konzept der Kointegration an. Dadurch erhalten sie eine neue Zeitreihe $\mathcal{B}^{i,j}(t)$. Diese Liste repräsentiert also nun die selbst geschätzten und eigens approximierten Grundwerte der Aktie j für den i -ten Agent. Somit muss sich nun jeder Agent auf zwei Informationsquellen für die jeweilige Aktienstrategie verlassen:

- Die erste Informationsquelle ist die Chartanalyse. Hierbei wird versucht, dass man mithilfe des bisherigen Graphen mögliche Trends erkennt und damit die Zukunft möglichst genau vorhersagen kann.
- Die zweite Informationsquelle ist die Fundamentalanalyse. Bei dieser Methode ist das Ziel wiederum die Zukunft genau vorherzusagen, jedoch werden diesmal die Grundwerte der Aktien und bekannte wirtschaftliche Kennzahlen als Basis verwendet.

Nun bekommen alle Agenten die Möglichkeit Informationen über 1000 Zeitschritte hinweg zu sammeln und mit ihrem Lernprozess zu beginnen. Danach werden alle

Werte auf ihren Ursprungswert zurückgesetzt und die Agenten erhalten anschließend die Ergebnisse übermittelt, die sie in ihrer Lernphase erzielt haben.

2.1 Grundwerte der Agenten

Jeder Agent benötigt nun ein paar Parameter mit denen er, wie oben erwähnt, initialisiert wird damit der Handel beginnen kann. Sei dafür $\sim \mathcal{U}()$ definiert als die stetige Gleichverteilung und $\sim \mathcal{U}\{\}$ als die diskrete Gleichverteilung. Diese Parameter der Agenten sind wie folgt definiert:

- maximaler Drawdown $l^i \sim \mathcal{U}(50\%, 60\%)$: dieser Wert gibt den maximalen Verlust des i-ten Agenten an. Ist der Verlust höher so gilt der Agent als bankrott
- Reflexivitätsparameter $\rho^i \sim \mathcal{U}(0, 100\%)$: hierbei wird bestimmt wie sehr sich der Agent auf seine jeweilige Fundamental- oder Chartanalyse verlässt, abhängig von seiner jeweiligen Kursvorhersage
- Investitionszeitspanne $\tau^i \sim \mathcal{U}\{T_w, 6T_w\}$: hiermit wird festgelegt, nach wievielen Zeitschritten der Agent seine Aktie wieder verkauft
- trading Fenster $\mathcal{W}^i \sim \mathcal{U}\{T_w, \tau^i\}$: mit diesem Wert wird die optimale Zeit bestimmt um ein Angebot an das order Buch zu senden
- Gedächtnisintervall $h^i \sim \mathcal{U}\{T_w, T - \tau^i - 2T_w\}$: dieser gibt die Zeitspanne an, die ein Agent für seinen Lernprozess benötigt
- Transaktionswert $g^i \sim \mathcal{U}(0.2, 0.8)$: durch diesen Parameter wird die maximale Differenz vom eigens berechneten Kurswert bestimmt, bei dem der Agent bereit ist einer Transaktion zuzustimmen
- reinforcement Lernrate $\beta^i \sim \mathcal{U}(0.05, 0.20)$: dieser Wert wird auf zwei Algorithmen, die später erwähnt werden, angewendet und beschreibt die Lernkurve des jeweiligen Agenten

2.2 Das order Buch

Der nächste essentielle Teil des Modells ist das order Buch. Hier können alle Agenten festlegen welche Aktien sie kaufen oder verkaufen möchten. Zusätzlich besteht auch noch die Möglichkeit eine Höhe für den maximalen Kaufpreis oder minimaler Verkaufspreis für die jeweilige Aktie zu bestimmen. Klarerweise existiert für jede

handelbare Aktie auch ein eigenes order Buch, also empfängt das order Buch der Aktie j auch nur die Angebote, welche die jeweilige Aktie j betreffen. Nachdem alle Agenten ihre Kauf- und Verkaufswünsche, gemeinsam mit der gewünschten Anzahl an das order Buch übermittelt haben, beginnt dieses mit einem Sortierprozess und ordnet alle Kaufwünsche nach fallenden Preisen und die Verkaufswünsche nach steigenden Angeboten. Nun beginnt der Handel. Hierbei werden die obersten beiden Angebote miteinander verglichen und kontrolliert ob der angegebene Kaufpreis über dem Verkaufspreis liegt. Gilt dies nicht findet kein Handel statt. Andernfalls wird im nächsten Schritt der endgültige Verkaufspreis ermittelt, welcher durch das arithmetische Mittel der beiden Werten definiert ist. Dieser Prozess findet solange statt solange der höchste noch ausstehende Kaufpreis über dem niedrigsten noch vorhandenen Verkaufspreis liegt. Alle anderen Angebote werden zurückgewiesen, da kein Handel mehr stattfinden kann. Zum Schluss bestimmt das order Buch noch den neuen Marktwert $P^j(t+1)$ der Aktie j zum Zeitpunkt $t+1$, welcher definiert ist als das letzte und niedrigste arithmetische Mittel. Außerdem wird zusätzlich noch das trading Volumen $V^j(t+1)$ bestimmt, welches durch die Anzahl der Aktien die während dieses Zeitschritts gehandelt wurden festgelegt ist.[1]

2.3 Handelsalgorithmen

Nun wissen wir wie sich der Preis einer Aktie entwickelt, kommen wir nun zu den selbstständig lernenden Algorithmen, die das Handeln der Agenten beeinflussen. Hierbei gibt es 2 unterschiedliche Algorithmen: \mathcal{F}^i welcher die optimale Preisvorhersage lernt und \mathcal{T}^i welcher auf Basis des vorherigen Algorithmus effizientes Handeln lernt.

2.3.1 Algorithmus \mathcal{F}^i

Mithilfe dieses ersten Algorithmus beobachtet der Agent die langfristige ($s_0^{\mathcal{F}}$) und die kurzfristige Volatilität ($s_1^{\mathcal{F}}$) des Aktienkurses. Der Unterschied zwischen der eigenen Schätzung und dem aktuellen Marktwert wird mit ($s_2^{\mathcal{F}}$) bezeichnet. Ausgehend davon lernt der Agent seine Preisvorhersage in der vorgegebenen Investitionszeitspanne τ^i zu optimieren, indem er drei verschiedene Möglichkeiten ausprobiert um die Zukunft exakt vorherzusagen: 1. mithilfe eines einfachen ökonomischen Instruments, welches entweder auf dem Mean-Reversions-Effekt, dem Mittelwert oder dem aktuellen Markttrend basiert ($a_0^{\mathcal{F}}$) 2. wie weit man in die Vergangenheit schauen soll ($a_1^{\mathcal{F}}$) und 3. wie stark die Grundwerte der Aktien Einfluss auf die Kursentwicklung nehmen ($a_2^{\mathcal{F}}$).

Zuerst berechnet jeder Agent die Varianzen σ_L und σ_S des Preises über das Intervall

$[t - 3\tau^i, t]$ bzw. $[t - \tau^i, t]$. Der Wert $\sigma_L(t)$ wird nun mit den früher ermittelten und in aufsteigender Reihenfolge sortierten Werten verglichen, welche so gekürzt sind, dass sie der Größe des Gedächtnisintervalls h^i entsprechen. Liegt der neu berechnete Wert im untersten Viertel, dann gilt $s_0^{\mathcal{F}} = 0$. Ist der Wert höher als das 3. Quantil, so gilt $s_0^{\mathcal{F}} = 2$. Für alle anderen Werte setzt man $s_0^{\mathcal{F}} = 1$. Dadurch erkennt der Agent wie die langfristige Volatilität des Preises $P^j(t)$ der Aktie j aussieht. Für σ_S sieht der Vorgang ähnlich aus. Der neu ermittelte Wert wird mit den vorher ermittelten und aufsteigend sortierten Werten verglichen und es gilt wieder $s_1^{\mathcal{F}} = 0$ wenn der Wert in das unterste Viertel fällt, $s_1^{\mathcal{F}} = 2$ falls der Wert über der 75% Grenze liegt und $s_1^{\mathcal{F}} = 1$ andernfalls. Zusammen betrachtet helfen diese Werte dem Agenten nun, dass er eine bessere Vorstellung der Marktstruktur bekommt und somit auch eine genauere Vorhersage treffen kann.

Nun definiert der Algorithmus eine Menge $\mathcal{A}^{\mathcal{F}} = \{a_0^{\mathcal{F}}, a_1^{\mathcal{F}}, a_2^{\mathcal{F}}\}$ wobei die Elemente jeweils die Werte 0, 1 oder 2 annehmen können. Danach werden die Mittelwerte $P_{[t-2T, t-T]}^j(t)$ und $P_{[t-T, t]}^j(t)$ berechnet, wobei $T = (1 + a_1^{\mathcal{F}})\tau^i/2$. Mithilfe des ökonomischen Instruments von oben kann nun folgendes für die Fälle $a_0^{\mathcal{F}} = 0, 1, 2$ berechnet werden:

- $P^{i,j}(t) = P^j(t) + P_{[t-2T, t-T]}^j(t) - P_{[t-T, t]}^j(t)$
- $P^{i,j}(t) = \frac{1}{2}P_{[t-2T, t-T]}^j(t) + \frac{1}{2}P_{[t-T, t]}^j(t)$
- $P^{i,j}(t) = P^j(t) - P_{[t-2T, t-T]}^j(t) + P_{[t-T, t]}^j(t)$

Somit definiert nun $a_0^{\mathcal{F}}$ welche Art des ökonomischen Tools gewählt wird und $a_1^{\mathcal{F}}$ bestimmt über welchen Zeithorizont die Werte ermittelt werden. Der dritte Wert $a_2^{\mathcal{F}}$ bestimmt in welche Richtung der Mittelwert zwischen den $P^{i,j}(t)$ gewichtet wird. Dadurch ergibt sich die Preisvorhersage $H^{i,j}(t)$ des Agenten:

$$H^{i,j}(t) = \alpha P^{i,j}(t) + (1 - \alpha) \mathcal{B}^{i,j}(t) \quad \alpha \in \mathbb{R} \quad (2.1)$$

Danach berechnet der Algorithmus \mathcal{F}^i die relative Änderung zwischen der früheren Schätzung $H^{i,j}(t-\tau^i)$ welche vor τ^i Zeitschritten berechnet wurde und den tatsächlich eingetretenen Wert $P^j(t)$. Alle berechneten Werte werden wiederum aufsteigend sortiert und so zusammengekürzt, dass die Größe dem Gedächtnisintervall h^i entspricht. Nun wird der zuvor berechnete Wert mit der Zahlenfolge verglichen und der Rückgabewert wird durch die Liste $\mathcal{R}^{\mathcal{F}} = \{4, 2, 1, -1, -2, -4\}$ festgelegt. Welcher Wert ausgewählt wird, wird wiederum durch einen Vergleich mit den vergangen

Werten bestimmt. Falls die neue Berechnung einen Wert im Intervall $[0\%, 5\%)$ ergibt, erhält der Agent den höchsten Wert. Für die Intervalle $[5\%, 25\%)$ und $[25\%, 50\%)$ stehen die Ziffern 2 bzw. 1. Für den Rest gilt ähnliches also wird für das Intervall $[50\%, 75\%)$ der Wert -1, für $[75\%, 95\%)$ -2 und für $[95\%, 100\%]$ die Zahl -4 festgelegt.

Zum Schluss wird nun die Strategie verbessert. Dafür wird nach $\tau^i/T_m + 2$ Schritten die optimale Strategie berechnet, wie der Agent vor τ^i Zeitschritten handeln hätte sollen. Zusätzlich dazu wird die optimale Strategie mithilfe der Lernkurve β^i und den neu erhaltenen Informationen verbessert.

2.3.2 Algorithmus \mathcal{T}^i

Mit dem zweiten Algorithmus \mathcal{T}^i überwacht der Agent die Veränderungen des Aktienkurses, gemäß den Werten die der erste Algorithmus berechnet hat ($s_0^{\mathcal{T}}$), die Volatilität ($s_1^{\mathcal{T}}$) und das trading Volumen ($s_2^{\mathcal{T}}$). Ausgehend davon lernt der Algorithmus \mathcal{T}^i seine Investitionen zu optimieren indem er zwei Möglichkeiten ausprobiert: 1. soll er eine bestimmte Menge seiner Aktien halten, verkaufen oder kaufen ($a_0^{\mathcal{T}}$) und 2. wie flexibel im Preis sollte der Agent sein wenn er gewillt ist einen Handel abzuschließen ($a_1^{\mathcal{T}}$).

Zu Beginn wird der Wert $\mu = (H^{i,j}(t) - P^j(t))/P^j(t)$ berechnet und je nachdem ob der Wert positiv oder negativ ist mit den aufsteigend geordneten Mengen μ_- bzw. μ_+ verglichen. Falls der ermittelte Wert negativ ist und zu den unteren 95% von μ_- gehört gilt $s_0^{\mathcal{T}} = 0$, ansonsten $s_0^{\mathcal{T}} = 1$. Für ein positives μ gilt $s_0^{\mathcal{T}} = 1$ wenn der Wert unter dem 5% Quartil liegt und $s_0^{\mathcal{T}} = 2$ sonst. Daher beschreibt μ nun ob es einen sinkenden, stabilen oder steigenden Preis der Aktie j gibt. Nun wird der vorher ermittelte Wert σ_L aufgerufen. Falls der Wert im unteren Drittel von der Auflistung aller σ_L zu finden ist setzt man $s_1^{\mathcal{T}} = 0$. Für einen Wert im obersten Drittel gilt $s_1^{\mathcal{T}} = 2$ und für alle anderen $s_1^{\mathcal{T}} = 1$. Somit beschreibt $s_1^{\mathcal{T}}$ im zweiten Algorithmus nun die langfristige Volatilität. Weiters beobachtet der Agent wie sich das trading Volumen $V^j(t)$ entwickelt und ordnet auch dieses in aufsteigender Reihenfolge. Nun gilt $s_2^{\mathcal{T}} = 0$ falls $V^j(t) = 0$, $s_2^{\mathcal{T}} = 1$ falls es im untersten Drittel liegt und $s_2^{\mathcal{T}} = 2$ andernfalls. Durch diesen Wert erkennt ein Agent nun wie aktiv der Markt im Moment ist, was wiederum Einfluss auf den Kauf- und Verkaufspreis, der an das order Buch gesendet wird, hat.

Nun wird wieder eine Menge $\mathcal{A}^{\mathcal{T}} = \{a_0^{\mathcal{T}}, a_1^{\mathcal{T}}\}$ definiert, wobei diese beiden Werte abhängig vom Ergebnis des ersten Algorithmus \mathcal{F}^i sind. Falls $a_0^{\mathcal{T}} = 1$ ist schick der Agent kein Angebot an das order Buch. Wenn $a_0^{\mathcal{T}} = 0$ ist gibt es ein Verkaufsangebot des Agenten, welches gemäß der Formel für den ask Preis (siehe unten) $P_{ask}^{i,j}(t)$

berechnet wird. Beim dritten Fall $a_0^{\mathcal{F}} = 2$ wird ein Kaufangebot an das order Buch gesendet, welches wiederum nach der Formel für den bid Preis (siehe unten) $P_{bid}^{i,j}(t)$ berechnet wird. Der bid Preis ist für $a_1^{\mathcal{F}} = 0, 1, 2$ definiert als:

- $P_{bid}^{i,j}(t) = \min[H^{i,j}(t), P_j(t)] + g^i S^i(t - 1)$
- $P_{bid}^{i,j}(t) = \min[H^{i,j}(t), P_j(t)]$
- $P_{bid}^{i,j}(t) = \min[H^{i,j}(t), P_j(t)] - g^i S^i(t - 1)$

wobei g^i für den Transaktionswert und $S^j(t - 1)$ für den Markt spread stehen. Der Teil $\pm g^i S^j(t - 1)$ bestimmt also wie sehr ein Agent an seinem vorher festgelegten Kauf- und Verkaufspreis festhält.

Danach wird der ask Preis für $a_1^{\mathcal{F}} = 0, 1, 2$ definiert:

- $P_{ask}^{i,j}(t) = \max[H^{i,j}(t), P_j(t)] - g^i S^i(t - 1)$
- $P_{ask}^{i,j}(t) = \max[H^{i,j}(t), P_j(t)]$
- $P_{ask}^{i,j}(t) = \max[H^{i,j}(t), P_j(t)] + g^i S^i(t - 1)$

Abhängig vom aktuellen Aktienkurs $P^j(t)$ wird nun die Differenz zwischen dem aktuellen Gesamtvermögen und dem möglichen Gesamtvermögen berechnet, welches der Agent gehabt hätte, falls er vor τ^i Zeitschritten kein Angebot gesendet hätte: $Q_{OB}^{i,j}(t - \tau^i)[P^j(t) - P_{OB}^{i,j}(t - \tau^i)]$. Hier beschreibt $Q_{OB}^{i,j}(t - \tau^i)$ die Anzahl und $P_{OB}^{i,j}(t - \tau^i)$ den Handelspreis welcher vom order Buch zum Zeitpunkt $t - \tau^i$ ermittelt wurde. Zu bemerken ist, dass die Handelsangebote, die von den Agenten an das order Buch übermittelt werden, von den hier erzielten Werten abweichen können, da manchmal der gewünschte Kaufpreis unter dem geringsten Verkaufspreis liegt und der tatsächliche Preis bei den Angeboten das arithmetische Mittel der beiden Preise ist. Nun wird der Wert mit der Liste von den früheren Werten verglichen, welche aufsteigend sortiert ist und der Rückgabewert wird wiederum durch die Menge $\mathcal{R}^{\mathcal{F}} = \{4, 2, 1, -1, -2, -4\}$ beschrieben. Der Auswahlprozess für die Rückgabewert funktioniert analog zu jenem im ersten Algorithmus. Zum Schluss wird die Strategie mit der Lernkurve β^i auf die gleiche Weise wie beim Algorithmus \mathcal{F} verbessert. [2]

3 Handelsstrategie der Agenten

Nachdem wir nun wissen wie das Modell aufgebaut ist, wollen wir nun beobachten ob es einen Zusammenhang zwischen einem Erfolg und der gewählten Handelsstrategie gibt. Hierfür muss zuerst festgestellt werden ob sich Agenten eher für die Fundamental- oder Chartanalyse entscheiden. Zur Erinnerung, jeder Agent wurde mit dem Reflexionsparameter $\rho^i \sim \mathcal{U}(0, 100\%)$ initialisiert, der die anfängliche Tendenz bestimmt. Mithilfe des Wertes a_2^F wird der Parameter idealisiert und der Agent lernt ob er sich eher auf die Fundamental- oder Chartanalyse verlassen sollte, basierend auf den Werten für die Volatilitäten und dem Unterschied zwischen seiner Vorhersage und dem tatsächlich eingetretenen Marktpreis. Alle folgenden Simulationen wurden mit den Parametern $I = 500$, $J = 1$ und $T = 2875$ durchgeführt. In Abbildung 3.1. kann man erkennen, dass zum Zeitpunkt $t = T$ die besten 10% aller Agenten eher eine Fundamentalanalyse und die schlechtesten 10% eher Chartanalyse als Vorhersage verwenden. Somit kann gefolgert werden, dass Agenten die eine Fundamentalanalyse verwenden tendenziell besser abschneiden als jene mit einer Chartanalyse.

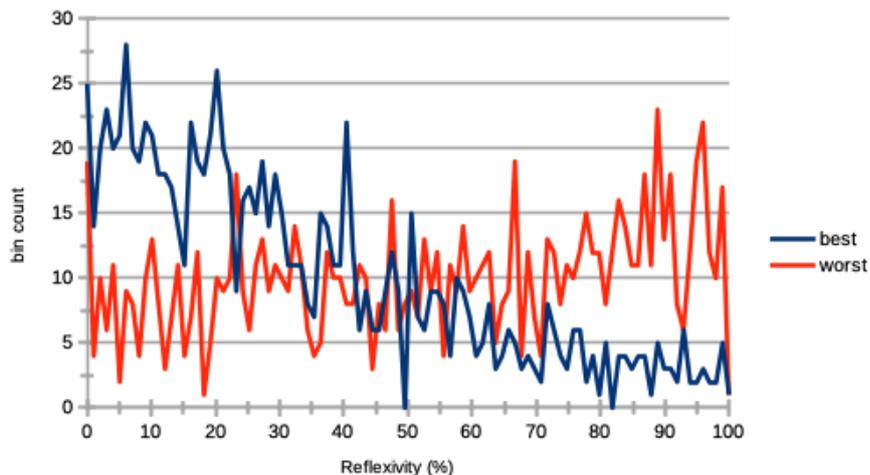


Abbildung 3.1: Reflexionsparameter für die besten bzw. schlechtesten Agenten

Der Reflexionsparameter nimmt weiters auch Einfluss auf die Preisvorhersage und dadurch hat er auch Auswirkungen auf den bid-ask spread. Der bid-ask spread ist die Differenz zwischen dem höchsten Preis den ein Anleger bereit ist zu zahlen und dem tiefsten Preis, bei dem die Teilnehmer bereit sind die Aktie zu verkaufen. Betrachtet man nun den Transaktionswert g^i , also jene Abweichung vom Angebot bei dem der Agent bereit ist den Handel einzugehen, kann man erkennen, dass die besten Agenten eher einen kleineren Wert haben und die schlechtesten einen höheren Wert, wie man in Abbildung 3.2. sehen kann. [1]

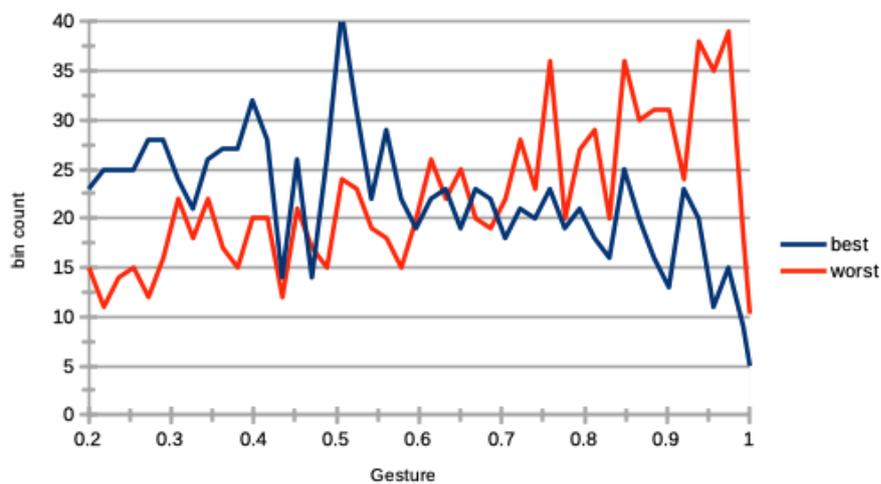


Abbildung 3.2: Transaktionswert der besten bzw. schlechtesten Agenten

4 Einflüsse auf den Markt

Betrachten wir nun welche Auswirkungen es auf den Markt geben kann, wenn ein gewisser Prozentsatz entweder eine bessere Lernkurve hat oder versucht wird den besten beziehungsweise schlechtesten Agenten nachzuahmen. Zum Schluss werden noch Auswirkungen auf den Markt festgestellt, wenn ein immer größer werdender Prozentsatz zufällig handelt, wie es in der Realität üblich ist.

Zur Erinnerung, jeder Agent wurde anfangs mit einer Lernrate $\beta^i \sim \mathcal{U}(0.05, 0.20)$ initialisiert. Für die erste hier durchgeführte Studie erhält nun ein bestimmter Prozentsatz $p = 0\%, 20\%, 40\%, 60\%, 80\%$ eine verdoppelte Lernrate. In Abbildung 4.1. können wir nun sehen wie sich er Mittelwert der Wochen-, Monats- und 6 Monatsvolatilität, für genau diese Prozentsätze, verändert. Alle übrigen $100 - p$ Agenten haben weiterhin nur den ursprünglichen Faktor β^i als Lernkurve. Hier kann beobachtet werden, dass die Volatilität des Preises für alle 3 Zeitintervall relativ stabil bleibt.

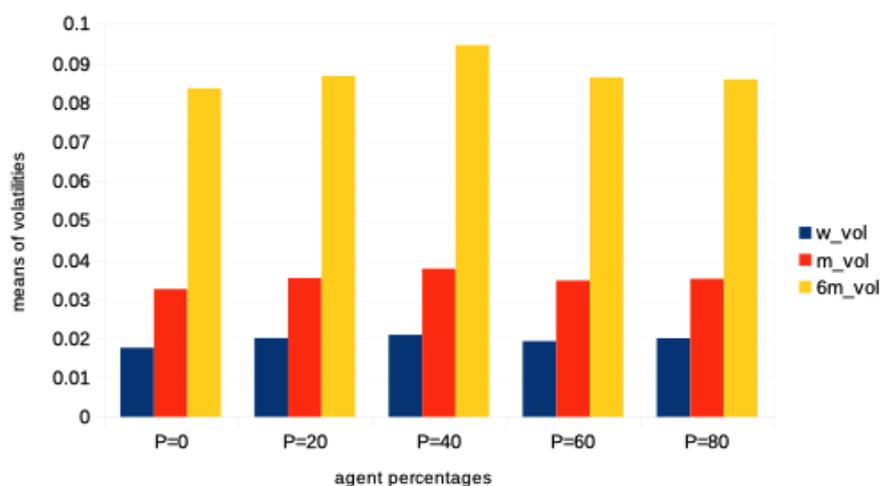


Abbildung 4.1: Volatilität des Preises, wenn ein Prozentsatz p eine doppelte Lernrate β^i erhält.

Betrachten wir nun den Fall, dass die Lernkurve aller Agenten mit dem Faktor $\zeta = 0.5, 1, 1.5, 2, 2.5$ multipliziert wird. Auch hier sehen wir in Abbildung 4.2., ähnlich wie im vorherigen Experiment, dass die mittlere Volatilität des Preises für alle drei Beobachtungszeiträume stabil ist. Ebenfalls gilt, dass die Volatilität von einer Woche in blau, von einem Monat in rot und von 6 Monaten in gelb dargestellt wird.

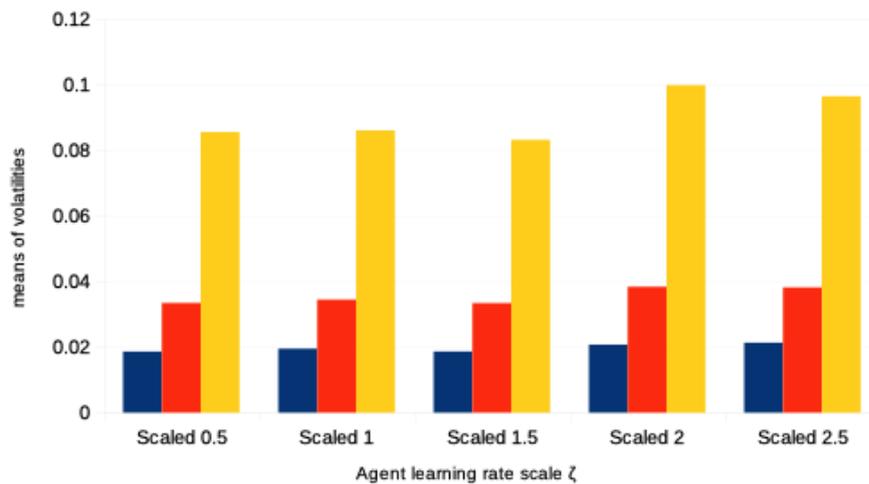


Abbildung 4.2: Volatilität des Preises, wenn alle Agenten eine mit ζ skalierte Lernrate haben.

Die nächste wichtige Frage ist, ob und wenn ja, wie sich die Anzahl der Markt crashes verändert. Von einem Markt crash wird gesprochen, wenn der Marktpreis innerhalb eines Zeitschritts um mehr als 20% fällt. Sowohl in Abbildung 4.3. als auch Abbildung 4.4. kann man sehen, dass die Anzahl der crashes für beide Szenarien stark steigt. Dies ist insofern verwunderlich, da man vermuten hätte können, dass der Markt stabiler wird, wenn entweder immer mehr Agenten ein größeres Wissen ansammeln oder das Wissen aller Teilnehmer steigt.

Da die Anzahl der Markt crashes deutlich gestiegen ist, liegt die Vermutung nahe, dass auch mehr Agenten bankrott werden, also ihren maximalen Drawdown l^i überschreiten. In Abbildung 4.5. können wir ein Ergebnis sehen, welches auf einen leichten Anstieg der bankrotten Teilnehmer hindeutet. Einen ähnlichen Verlauf zeigt auch Abbildung 4.6., indem der Mittelwert aller bankrotten Agenten, die eine mit ζ multiplizierte Lernrate besitzen, dargestellt ist.

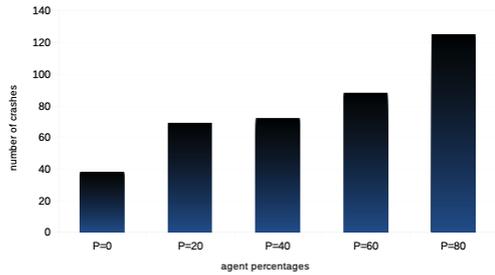


Abbildung 4.3: Anzahl der Marktcrashes, wenn ein Prozentsatz p eine doppelte Lernrate β^i erhält.

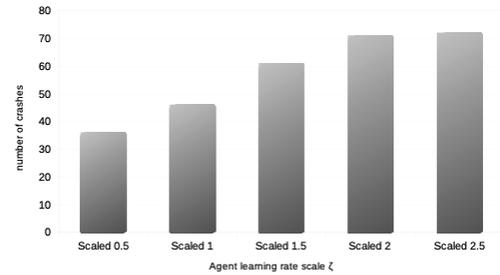


Abbildung 4.4: Anzahl der Marktcrashes, wenn alle Agenten eine mit ζ multiplizierte Lernrate besitzen.

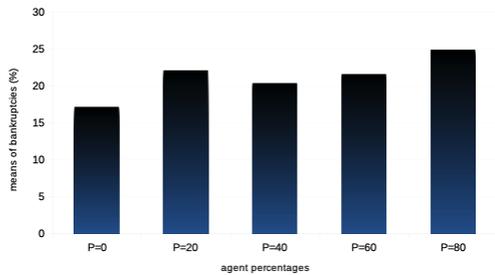


Abbildung 4.5: Prozent der bankrotten Agenten, wenn ein Prozentsatz p eine doppelte Lernrate β^i erhält.

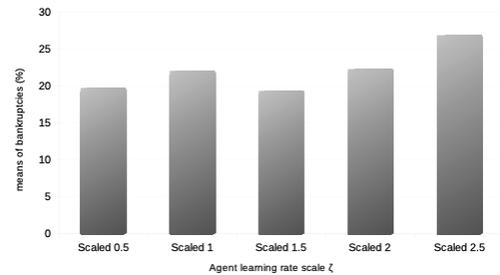


Abbildung 4.6: Prozent der bankrotten Agenten, wenn alle eine mit ζ multiplizierte Lernrate besitzen.

4.1 Best agent herding

Im nächsten Experiment wird untersucht, welche Auswirkungen es geben kann, wenn ein bestimmter Prozentsatz $p = 0\%, 20\%, 40\%, 60\%, 80\%$ die Handlungen des im letzten Zeitschritt besten Agenten imitiert. In diesem Modell wird es so verwirklicht, dass dieser Prozentsatz dann zum Zeitpunkt $t + 1$ dieselben Angebote sendet wie der beste Agent zum Zeitpunkt t . Dies ist für die Realität natürlich nicht untypisch, weil jeder versucht seinen Gewinn zu maximieren und man sich daher gerne am besten orientiert.

Wie auch in den vorherigen Beispielen analysieren wir zunächst die Auswirkungen auf die Volatilität des Preises, die in Abbildung 4.7 dargestellt ist, wobei die y-Achse

für eine bessere Lesbarkeit logarithmisch skaliert ist. Die schwarze strichlierte Linie repräsentiert die realen Daten und man kann sehen, dass sich der rote ($p = 0\%$), gelbe ($p = 20\%$) und grüne ($p = 40\%$) Graph diesem sehr gut annähert. Für die höheren Prozent jedoch, wird die Volatilität des Preises stark erhöht.

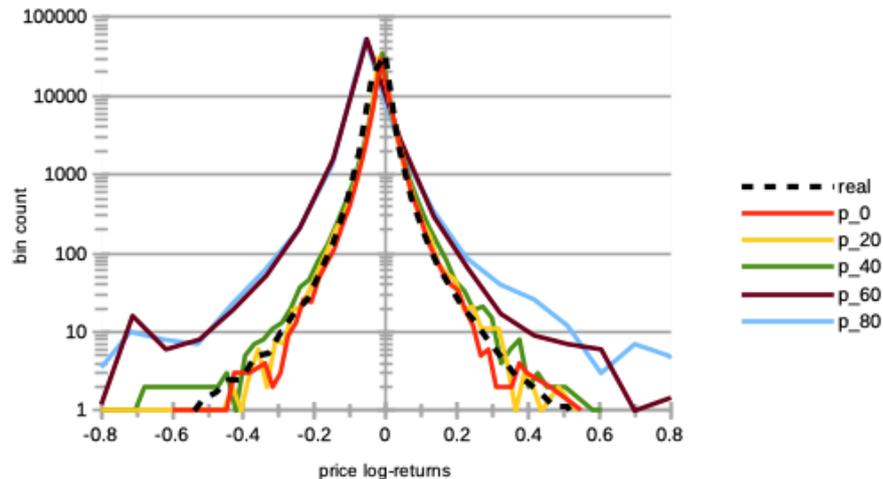


Abbildung 4.7: Volatilität des Preises bei best agent herding

Betrachten wir nun die Veränderung des trading Volumens. Von vorn herein könnte hier die Annahme getroffen werden, dass es stark erhöht wird, weil der Teilnehmer auf einen ähnlich guten Ausgang wie im vorherigen Schritt hofft. Um daher seinen Gewinn zu maximieren, wird auch die Risikobereitschaft erhöht und damit auch mit einem höheren Vermögen gehandelt. In Abbildung 4.8. sehen wir anhand der blauen Balken, welche das best agent herding repräsentieren, dass genau der gegenteilige Fall eintritt und das Geldvolumen drastisch reduziert wird.

Betrachten wir nun die Anzahl der Marktcrashes unter der Annahme, dass ein Prozentsatz p dem besten Agenten aus dem vorherigen Zeitintervall nachahmt. Auch hier erhöht sich die Anzahl drastisch, was uns die blauen Balken in Abbildung 4.9. zeigen.

Als letzte Auswirkung unter der best agent herding Annahme beobachten wir die Entwicklung des bid-ask spread, welcher die Differenz zwischen dem höchsten Preis den ein Anleger bereit zu zahlen ist und dem tiefsten Preis, bei dem ein Anleger bereit ist die Aktie zu verkaufen, beschreibt. Hierbei besteht die Annahme dass der

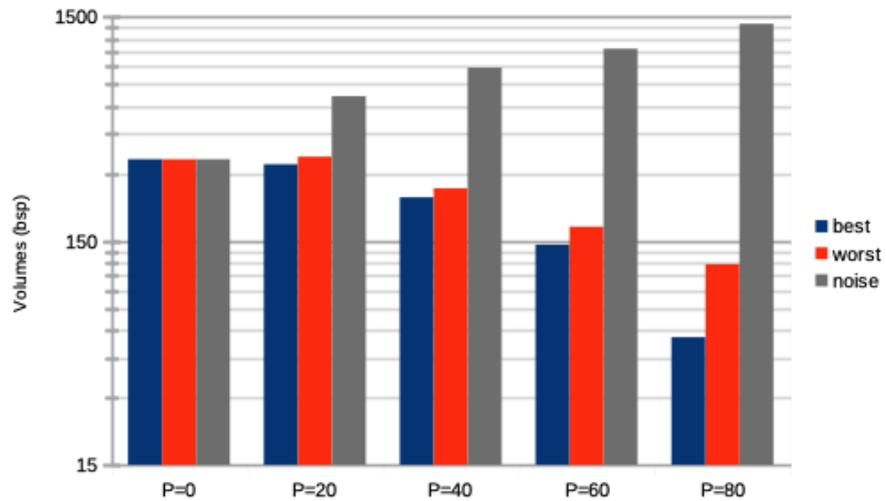


Abbildung 4.8: Trading Volumen für steigendes p aufgeschlüsselt nach best agent herding (blau), worst agent herding (rot) und noise traders (grau) mit logarithmisch skaliertem y-Achse

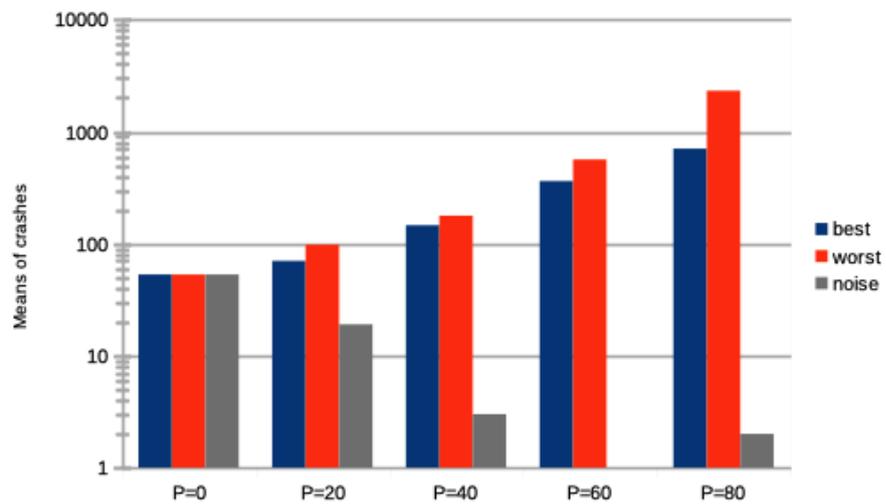


Abbildung 4.9: Anzahl der Marktcrashes für steigendes p aufgeschlüsselt nach best agent herding (blau), worst agent herding (rot) und noise traders (grau) mit logarithmisch skaliertem y-Achse

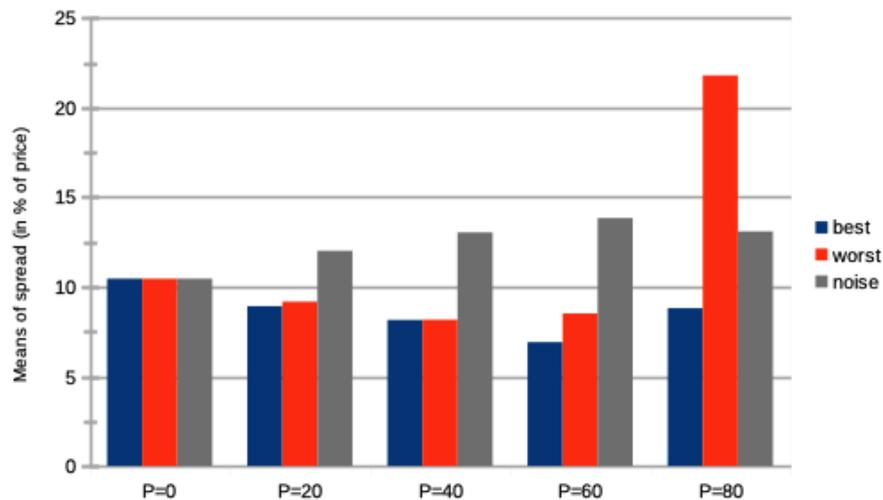


Abbildung 4.10: Durchschnittlicher bid-ask spread in % für steigendes p aufgeschlüsselt nach best agent herding (blau), worst agent herding (rot) und noise traders (grau) mit logarithmisch skaliertem y-Achse

durchschnittliche Wert stetig sinkt, da die besten Agenten meistens auch die genaueren Vorhersagen treffen und somit die Diskrepanz zwischen den einzelnen Angeboten sinken sollte. In den blauen Graphen der Abbildung 4.10. sieht man, dass sich diese Annahme zunächst bewahrheitet es jedoch zwischen $p = 60\%$ und $p = 80\%$ einen leichten Anstieg gibt.

Zu erwähnen ist, dass trotz der unterschiedlichen Parameter die durchschnittliche Anzahl der bankrotten Agenten stabil bleibt und sich für alle Prozentsätze im Intervall $[19.51, 26.01]$ befindet.

4.2 Worst agent herding

Nachdem wir den Einfluss des besten Agenten behandelt haben, kommen wir nun zum Gegenteil: Wie würde sich der Markt entwickeln, wenn ein Prozentsatz $p = 0\%$, 20% , 40% , 60% , 80% dem schlechtesten Agenten folgt. Hier ist anzumerken, dass als schlechtesten Agent nicht jener gilt der bereits bankrott ist, sondern derjenige mit dem geringsten Nettowert. Dieses Ereignis ist für die Realität sehr unrealistisch, da kein Marktteilnehmer sein Vermögen einem sehr schlechten Agenten anvertraut, bei dem die Wahrscheinlichkeit eines Verlusts hoch ist. Jedoch besteht das Interesse bei dieser Modellannahme nicht an den einzeln erzielten Ergebnissen, sondern eher an den Gemeinsamkeiten und Unterschieden zum best agent herding.

Wie auch bei den vorherigen Experimenten, betrachten wir auch hier zuerst die

Volatilität des Preises. In Abbildung 4.11. sieht man für $p \leq 60\%$ einen ähnlichen Verlauf wie beim best agent herding. Für den Fall, dass 80% aller Agenten dem schlechtesten folgen springt dieser Wert aber nochmals signifikant in die Höhe.

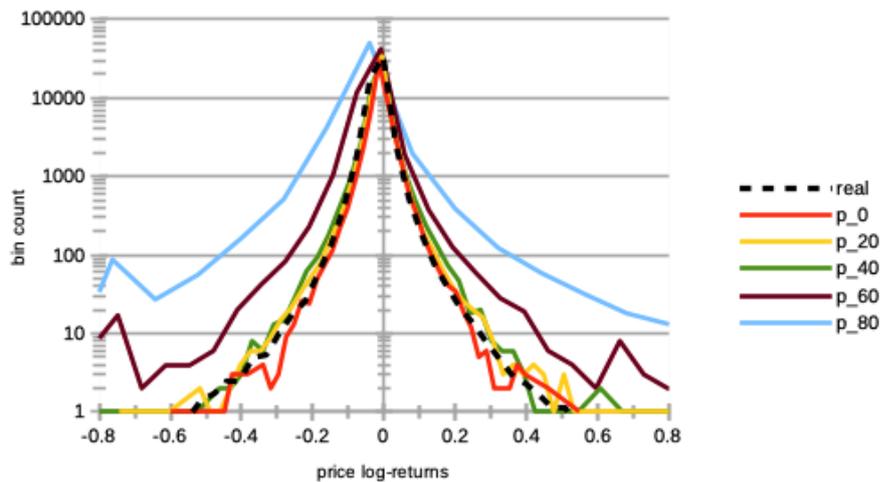


Abbildung 4.11: Volatilität des Preises bei worst agent herding

Für das trading Volumen, welches in Abbildung 4.8. dargestellt ist, können wir anhand der roten Balken erkennen, dass auch hier der Wert drastisch sinkt, jedoch nicht so stark wie beim best agent herding. Für die Anzahl der Markt crashes, siehe Abbildung 4.9., gilt ähnliches, wobei hier, wie wenig verwunderlich, der Durchschnitt stark von unter 100 bei $p = 0\%$ auf deutlich über 1000 bei $p = 80\%$ ansteigt.

Wie verhält sich nun der bid-ask spread im Beispiel des worst agent herding. Dieser sinkt bis $p = 40\%$, steigt dann zuerst leicht an und wächst ab $p = 80\%$ stark. Für die Anzahl der bankrotten Agenten, ergibt sich erwartungsgemäß ein sehr hoher Wert: für $p > 20\%$ werden stets über 70% bankrott.

4.3 Noise traders

Für die letzte Studie wird der Einfluss von noise traders auf den Markt betrachtet. Noise traders sind Agenten, die weder dem besten noch dem schlechtesten folgen, sondern zufällig handeln. Hierbei kann man erwarten, dass bei steigender Anzahl der Markt stabiler wird, da mehr liquide Mittel zur Verfügung gestellt werden.

Zuerst analysieren wird wieder die Volatilität des Preis für eine Prozentsatz $p = 0\%, 20\%, 40\%, 60\%, 80\%$ an noise traders. Die ersten Unterschiede sind bereits bei der gelben Linie für $p = 20\%$ sichtbar, da der Graph hier deutlich unter den realen Werten liegt und diese, im Gegensatz zum best/worst agent herding, nicht angenähert werden. Dieser Trend setzt sich auch bei steigendem p fort, wie in Abbildung 4.12. ersichtlich ist.

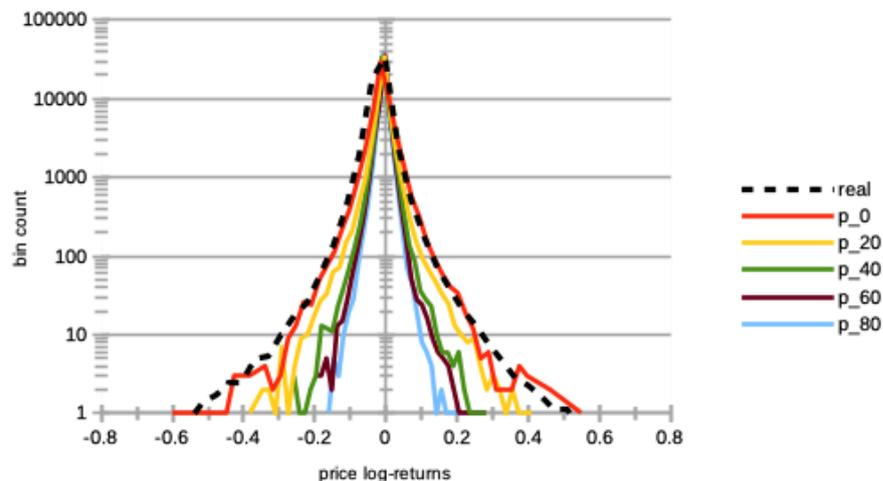


Abbildung 4.12: Volatilität des Preises bei noise traders

Auch beim trading Volumen in Abbildung 4.8. sieht man anhand der grauen Balken, dass sich auch dieser Wert im Kontrast zu den vorherigen beiden Gruppen entwickelt und somit stark ansteigt.

Wir wissen bereits, dass die Anzahl der Markt crashes beim best und worst agent herding stark zunehmen. Abbildung 4.9. zeigt, dass für eine immer größer werdende Gruppe an noise traders das Gegenteil gilt, womit unsere Annahme am Beginn bestätigt wird. Für den bid-ask spread, in Abbildung 4.10. durch die grauen Balken dargestellt, kann eine langsame und nahezu gleichbleibende Steigerung festgestellt werden.

Wie auch schon bei den vorherigen Werten, gibt es auch bei der Anzahl der bankrotten Agenten ein gutes Ergebnis. Ähnlich zu den beiden vorherigen Experimenten werden am Anfang $\approx 23\%$ der Agenten bankrott. Dieser Wert sinkt stetig bei zunehmender Anzahl an noise traders und erreicht für $p = 80\%$ einen Wert von unter 19% .

Dies ist wiederum bemerkenswert, da somit Agenten die zufällig handeln unwahrscheinlicher bankrott gehen, als jene die dem besten aus dem vorherigen Zeitschritt als Vorbild haben.

Abschließend kann also festgestellt werden, dass es sowohl für den Markt als auch für die Stabilität des Preises nicht vorteilhaft ist, wenn mehr Marktteilnehmer den besten nachahmen. Hingegen wird der Markt gestärkt, wenn genug Personen zufällig handeln.[1]

5 Fazit

Es wurde ein Modell entwickelt, indem eine beliebige Anzahl an unabhängigen Agenten selbstständig Aktien handeln kann. Zuerst wurden die Algorithmen, welche für den Handel zuständig sind beschrieben und danach untersucht ob es einen Zusammenhang zwischen der Wahl der Handelsstrategie und dem Erfolg gab. Hierbei stellte sich heraus, dass Teilnehmer welche eine Chartanalyse betreiben tendenziell schlechtere Ergebnisse erzielten als jene mit einer Fundamentalanalyse. Danach gab es verschiedenste Szenarien und man betrachtete die Auswirkungen auf den Markt. Dabei konnte festgestellt werden, dass sich die Volatilität des Preises, bei Erhöhung der Lernrate, kaum ändert, jedoch sowohl die Anzahl der Marktcrashes als auch jener Prozentsatz der Agenten, die ihren maximalen Drawdown überschritten, angestiegen sind. Bei der Studie in der verschiedenste Agenten nachgeahmt wurden, konnte man erkennen, dass es für die Stabilität des Preises, das Trading Volumen, die Anzahl der Marktcrashes und die Anzahl der bankrotten Teilnehmer von Vorteil ist, wenn ein größerer Prozentsatz zufällig handelt.

Literaturverzeichnis

- [1] J. Lussange, S. Bourgeois-Gironde, S. Palminteri, and B. Gutkin, arXiv:1910.05137v1 (2020)
- [2] J. Lussange, S. Bourgeois-Gironde, S. Palminteri, and B. Gutkin, arXiv:1909.07748v5 (2020).

Abbildungsverzeichnis

3.1	Reflexionsparameter für die besten bzw. schlechtesten Agenten	8
3.2	Transaktionswert der besten bzw. schlechtesten Agenten	9
4.1	Volatilität des Preises, wenn ein Prozentsatz p eine doppelte Lernrate β^i erhält.	10
4.2	Volatilität des Preises, wenn alle Agenten eine mit ζ skalierte Lernrate haben.	11
4.3	Anzahl der Markt crashes, wenn ein Prozentsatz p eine doppelte Lernrate β^i erhält.	12
4.4	Anzahl der Markt crashes, wenn alle Agenten eine mit ζ multiplizierte Lernrate besitzen.	12
4.5	Prozent der bankrotten Agenten, wenn ein Prozentsatz p eine doppelte Lernrate β^i erhält.	12
4.6	Prozent der bankrotten Agenten, wenn alle eine mit ζ multiplizierte Lernrate besitzen.	12
4.7	Volatilität des Preises bei best agent herding	13
4.8	Trading Volumen für steigendes p aufgeschlüsselt nach best agent herding (blau), worst agent herding (rot) und noise traders (grau) mit logarithmisch skaliertem y-Achse	14
4.9	Anzahl der Markt crashes für steigendes p aufgeschlüsselt nach best agent herding (blau), worst agent herding (rot) und noise traders (grau) mit logarithmisch skaliertem y-Achse	14
4.10	Durchschnittlicher bid-ask spread in % für steigendes p aufgeschlüsselt nach best agent herding (blau), worst agent herding (rot) und noise traders (grau) mit logarithmisch skaliertem y-Achse	15
4.11	Volatilität des Preises bei worst agent herding	16
4.12	Volatilität des Preises bei noise traders	17