Short Selling Constraints and their Effects on Market Efficiency:
Insights from Agent-based Modeling

Björn-Christopher Witte, Christopher Kah
Department of Economics, University of Bamberg, Germany

Abstract:
In this paper we examine the influence of short selling constraints on the efficiency of financial markets using a dynamic multi-agent model which follows the chartist-fundamentalist approach. Key property of the model is that each agent manages an own inventory of assets. Apart from the general short selling constraint, we investigate different tick rules including the uptick rule. We find that short selling constraints can reduce volatility and decrease the frequency of extreme price falls and rises. On the other hand, the rules deteriorate market distortion. In particular, the market becomes overvalued. The effects are more pronounced, if the constraint is more restrictive.

Keywords:
Agent-based modeling, chartist-fundamentalist approach, market efficiency, short selling constraints, uptick rule.

JEL Classification:
G14; G18, G28.

________________________
Björn-Christopher Witte
Department of Economics
University of Bamberg
Feldkirchenstr. 21
96045 Bamberg
Germany
Phone: [+49] 178 140 3640
Fax: [+49] (951) 863 2635
1. **Introduction**

The Securities and Exchange Commission (SEC) and many other financial market regulation authorities around the world (temporarily) prohibited short selling for certain stocks after the Lehman bankruptcy in fall 2008. According to the official statement by the SEC, the aim of their ban was “to prevent substantial disruption in the securities markets” and to attenuate “a crisis of confidence, without a fundamental underlying basis” (SEC 2008). The measure was motivated by the assumption that short selling was responsible for a part of the financial market instability observed:

“Recent market conditions have made us concerned that short selling in the securities of a wider range of financial institutions may be causing sudden and excessive fluctuations of the prices of such securities in such a manner so as to threaten fair and orderly markets” (SEC 2008)

Academic research provides nonuniform evidence for the role of short selling and short selling constraints on market efficiency. Important findings can be related to two aspects of market efficiency: (1) the distortion of the market, and (2) market risk, whereby the latter refers to the overall volatility of the market and to statistic properties of the distribution of returns, particularly the occurrence of extreme declines.

The majority of studies concerning (1) report short selling constraints to produce overvaluation and speculative bubbles (e.g., Jones & Lamont, 2002, Ofek & Richardson 2003, Bris et al., 2007, Chang et al. 2007, Cohen et al., 2007). With reference to (2), results are ambiguous. Boehmer et al. (2009) document that the SEC’s short selling ban in 2008 increased intraday volatility of the respective stocks listed on NYSE and Nasdaq. In contrast, Chang et al. (2007) report evidence from the Hong-Kong Stock market which indicates that short selling constraints might reduce volatility. Other findings are that short selling constraints increase the skewness of the return distribution meaning that the weight of extreme negative returns declines (Bris et al., 2007, Chang et al., 2007).

The research method of the studies quoted above consists in analyses of empirical data. These studies face a common difficulty: In order to attribute some difference of market efficiency to short selling constraints directly, one had to conduct controlled experiments, i.e., one had to compare two markets which solely differ in terms of the variable to be investigated: the existence the constraints.

In our study, we investigate the effects of short selling constraints in an artificial laboratory. To this purpose, we build a multi-agent model of a generic security market. As a research method we use simulation experiments. Advantages of this technique are the possibility to generate large data sets, to control all variables precisely, and to highlight cause-and-effects chains relatively easily. Financial market models have been used successfully to gain insights about the effects of regulatory policies, such as transaction taxes (Westerhoff & Dieci, 2006) or central bank interventions (Szpiro, 1994). For a review of these studies see Westerhoff (2008).

Agent-based financial market models usually reproduce financial markets as systems of behaviorist agents who rely on simple heuristics for being incapable to derive perfectly rational actions. Regularly such heuristics consist in simplified versions of technical and fundamental trading strategies. Respective models have shown that the interaction of the heterogeneous traders can create complex price dynamics which replicates some of the most important statistical properties of financial markets (see e.g., Chen et al., 2009).

Acknowledgement: The authors thank Frank Westerhoff for his helpful comments and two anonymous proof-readers. Special thanks go to Leanne Ussher for her interesting suggestions.
The outstanding feature of our model is that each agent possesses an individual security inventory that stores the number of shares owned by the respective agent. If the inventory is negative, the agent has built up short positions in the past. The model design enables us to test different variants of short selling constraints. Each constraint is implemented by imposing a characteristic restriction under which sells can be executed. In case of a general constraint, shorting is banned categorically, i.e., sells are only carried out if the inventory of the individual agent would not become negative. In case of a conditional constraint, shorting is prohibited only under certain market conditions. This applies to the so-called “tick rules”, under which shorting some asset is disallowed only if the most recent price change is below some defined limit. The prominent uptick rule (Rule 10a-1, Securities Exchange Act of 1934, removed by Rule 201 Regulation SHO in 2007) is a specific case of this principle. Here, the limit is set to zero, i.e., shorting is prohibited if the most recent tick has been negative.

In our study, we investigate the general short selling constraint as well as tick rules with different positive and negative limits including the uptick rule. The evaluation criteria for the respective policy are its effects for market stability and efficiency. The direct comparison of different types of short selling constraints may be regarded as one of the novel contributions of our analysis. Methodologically, the explicit representation of individual agents and their asset inventories is relatively new in the field of simple financial market models.

Our results indicate that the consequences of short selling constraints for market efficiency and stability are ambiguous. On the one hand, the constraints, no matter if general or limit-based, increase market distortion as they produce overvaluation by preventing part of the pessimistic expectations to be reflected in prices. On the other hand, the constraints reduce volatility and lower the frequency of extreme price movements, in the negative and, more surprisingly, in the positive domain. The effects are most pronounced for the general constraint. However, the general constraint tends to block less transaction than most conditional constraints as it protects investors from being affected by constraints durably. Furthermore, we observe that short selling constraints lead to an assimilation of inventory sizes.

The remainder of this article is structured as follows: In section 2, we introduce a nonlinear, stochastical financial market model in which shorting is not restricted and outline its dynamic properties. Section 3 focuses on the implementation of general and conditional short selling constraints and describes measures to capture market efficiency and stability. In section 4 we present the numerical results of our Monte Carlo analysis concerning the effects of the different constraints for market efficiency and stability. In section 5, we explain the numerical results and discuss their sensitivity with respect to the assumptions made. Section 6 summarizes the insights gained and highlights needs for future research.

2. The Basic Model
In this section, we present the design of our model market without short selling constraints being implemented. Therefore we firstly outline the theoretical foundation of financial market models like ours.

2.1 Theoretical Approach
Behavioral finance (survey by Shleifer, 2000) stresses that economic agents act boundedly rational. As one of the pioneer contributors in the field, Simon (1955) argued that agents lack the cognitive abilities to derive optimal actions, even though they strive to find them. Psychological evidence reports that in such a situation agents rely on simple heuristics, sometimes called “rules of thumb” (Kahneman et al., 1986, Shiller, 1999, and Hirshleifer, 2001). This notion is essential for the modeling of investor behavior: If the simple heuristics can be identified, the
behavior of investors can be formalized to some degree of accuracy. Models following the so-called chartist-fundamentalist approach (surveys by Hommes, 2006 and Westerhoff, 2008) rest on this principle. The approach is based on the well-documented observation that financial investors use either technical or fundamental strategies (Menkhoff, 1997, Lui & Mole, 1998). Traders applying fundamental rules are called fundamentalists. Fundamentalists seek to calculate the intrinsic value of some asset by methods of corporate or economic assessment. Their aim is to generate profits by exploiting mispricing (for representative literature see Stein, 1988, and Greenwald et al., 2001). Traders using technical rules are called chartists. Chartists try to identify typical patterns in price movements, such as trends, in the belief that such patterns contain information about the future price development (for representative literature see Murphy, 1999, and Pring, 2002). Models following the chartist-fundamentalist approach have shown that the pure interaction between the different traders groups generates complex dynamics of prices which mimics real financial markets quite accurately.

The design of our model is rooted in the chartist-fundamentalist approach, in principle. However, we do not treat each of the two groups as a whole but distinguish between individual agents. This makes possible to implement the constraints where they apply: on the level of individual behavior.

2.2 Model Framework
In the following a model of a generic security market with N independent agents and one traded asset is proposed. Short selling constraints are absent. The model rests on the following logic: Each agent decides in every period if she uses the technical or the fundamental strategy. The probability with which some trader chooses a specific strategy depends on the market distortion. The transactions made by each agent are accumulated in her inventory. Finally, the price in the next period is computed from the total excess demand in the market. The model can be formalized as follows.

Assuming new fundamental information to appear randomly, the evolution of the fundamental value of the asset, $F_t$, can be modeled as a random walk. It follows that:

$$ F_t = F_{t-1} + \alpha_t, $$

where $\alpha_t$ is a normally distributed random variable with mean 0 and constant standard deviation $\sigma^2$.

The mechanism of price adaption follows the market maker approach, as, for instance, proposed by Farmer & Joshi (2002). With $P_t$ referring to the log of the asset price, the principle can be formalized as:

$$ P_{t+1} = P_t + a \times D_t. $$

The market maker is a large market participant who absorbs imbalances between supply and demand. If there is a positive (negative) excess demand, $D_t$, in the market, she reacts by a price rise (cut) in the next period. The positive parameter $a$ determines the intensity of price adjustment. $D_t$ is the total excess demand generated in period $t$, i.e. the aggregate of the individual orders $D_{i,t}$ summed over all N agents:

$$ D_t = \sum_{i=1}^{N} D_{i,t}. $$

The fundamental strategy prescribes to exploit discrepancies between the intrinsic value of an asset and its price. Similar to Westerhoff & Dieci (2006) and others, the demand $D_{i,t}$ of a trader who uses this strategy can be written as:

$$ D_{i,t}^F = b(F_t - P_t) + \beta_{i,t} $$

The equation implies that fundamentalists buy (sell) if they perceive the asset to be
undervalued (overvalued). $\beta_{i,t}$ is a normally distributed random variable with mean 0 and constant standard deviation $\sigma_\beta$. The stochastic component accounts for the fact that each agent might use a slightly different trading rule or commit mistakes when determining the asset’s true value, $F_t$. The parameter $b$ regulates the aggressiveness of the rule.

The technical strategy is based on the identification of price trends and their extrapolation. Hence, the individual demand of some agent $i$ who uses the technical strategy, $D^C_{i,t}$, can be expressed as:

$$D^C_{i,t} = c(P_t - P_{t-1}) + y_{i,t}$$

The equation is derived from a similar formulation used in Westerhoff (2005). It states that chartists buy (sell) more, the higher (lower) the preceding price change has been. The parameter $c$ determines the aggressiveness of the rule, $y_{i,t}$ is a normally distributed random variable with mean 0 and constant standard deviation $\sigma_\gamma$. The stochastic component allows for the large variety of technical rules existent.

Switching between strategies follows a stochastic rule. Each trader chooses the technical strategy with some probability $\rho^C_t$, which depends on the market distortion. Otherwise she adopts the fundamental strategy. The probability $\rho$ can be written as:

$$\rho^C_t = \frac{1}{1 + d(F_t - P_t)^2}.$$  

The equation states that agents rely on the technical strategy less likely if the discrepancy between price and value of the asset traded is greater. Very similar formulations of the switching mechanism can be found in Westerhoff (2001, 2003) or He & Westerhoff (2005) among others. The underlying argument can be summarized as follows: Under the assumption that prices return to value sooner or later, the fundamental strategy promises large profits to exploit, if mispricing is great. Simultaneously, trend extrapolation, i.e. technical strategy, is risky, as a fundamental correction must be expected. The parameter $d$ is a measure of agents’ sensitivity to a given market distortion. The higher $d$, the quicker agents switch to the fundamental strategy. An agent is either a chartist or a fundamentalist. Accordingly, the probability for some agent $i$ to choose the fundamental strategy, $\rho^F_t$, results as:

$$\rho^F_t = 1 - \rho^C_t.$$  

After each agent has submitted her orders, his inventory is updated with his demand:

$$I_{i,t} = I_{i,t-1} + D_{i,t}.$$  

The equation above implies that each order is actually executed. When introducing short selling constraints we will refute this assumption.

### 2.3 Dynamic Properties

Figure 1 illustrates a typical run of the model dynamics over 2000 simulation steps. One simulation step represents one trading day. The top panel shows the evolution of the asset price (black) and its intrinsic value (gray), the second panel depicts the respective price returns $(P_t - P_{t-1})$, and the third panel shows the number of fundamentalists (black) and chartists (gray) in every period. We can see that prices fluctuate around the fundamental value quite erratically, with intervals of high and low volatility. Turbulent intervals correspond with a high presence of chartists, whereas in tranquil intervals

---

1 The difference is that in our model the respective equation gives the demand of an individual chartist and not of the group of all chartist as in Westerhoff (2005).

2 The distinctive feature of our formulation is that we derive probabilities for the strategy choice of individual agents instead of weights of agent groups.
fundamentalists prevail. In general, the dynamics is formed by the interplay of the two trading strategies. Chartists extrapolate trends and drive prices away from value at times. However, the more a bubble inflates, the more traders switch to the fundamental strategy to exploit the mispricing. The fundamental trading finally causes the bubble to burst and prices realign.

The lower panel in figure 1 displays the evolution of the inventory sizes of five different agents. At the start of the simulation run, the inventory of each agent contains 10,000 shares. Since agents trade independently of each other, the inventory sizes drift apart as the simulation is running. Furthermore, some of the inventories plunge into the negative domain indicating that the respective agent has built up a short position.

-- Figure 1 about here --

The setting of parameters and initial values is summarized by table 1. The model is configured such that the dynamics represents a relatively nervous market. Returns range from −17.9% to +16.1%, with an average absolute return of 2.8% for the simulation run shown.

-- Table 1 about here --

As mentioned before, price dynamics in real financial markets exhibit common statistical features, often termed “stylized facts” (surveys by Guillaume et al., 1997, Mantegna & Stanley, 2000, Cont, 2001, Lux & Ausloos, 2002, Johnson, et al., 2003, Sornette, 2004, and Lux, 2009). The accuracy of behavior of a financial market model is usually evaluated by testing for the replication of such features. Econometric tests have shown that our model replicates some of the most important stylized facts: (i) excess volatility (variance of prices higher than variance of intrinsic value), (ii) bubbles and crashes, (iii) volatility clustering (alteration of tranquil and turbulent intervals), (iv) uncorrelated returns, and (v) heavy tails in the return distribution. Except for (iv) and (v), the relevant panels in figure 1 illustrate these facts. Important statistic indicators are furthermore depicted by figure 8, which is based on a run with 10,000 simulation steps. The top panel illustrates the autocorrelation of periodical returns (P_t−P_{t−1}) for different lags. The gray horizontal lines represent the one-percent quantile of significance. We observe that none of the bars exceeds one of these lines meaning that model returns are not correlated significantly. The middle panel depicts the same but for absolute returns (|P_t−P_{t−1}|). The absolute returns are correlated significantly up to a lag of about 40 periods, which reflects volatility clustering. The bottom panel illustrates the distributions of returns. The black curve indicates how the returns would be distributed if following a normal distribution with the same mean and the same variance. The comparison turns out considerable heavy tails in the return distribution.

3. Using the Model to Evaluate Short Selling Constraints

We now describe how the model can be applied to evaluate the consequences of general and limit-based short selling constraints for market efficiency. In a first step, we explain how the model has been modified to implement the different constraint rules. Secondly, we define the measures used to evaluate market efficiency.

3.1 Implementation of Short Selling Constraints

Under a general short selling constraint, agents are not allowed to build up short positions at any time. Therefore, all individual inventories are required to be non-negative, i.e.:

\[
l_{i,t} \geq 0 \ \forall \ i, t.
\]  

To guarantee this restriction not to be violated, the individual demand functions (eqs. 4 and 5) must be adapted. The
individual demand function of an agent $i$ who chooses the fundamental strategy in some period $t$ is rewritten as:

$$D^F_{i,t} = \min\{-S^F_{i,t}, I_{i,t-1}\}, \quad (4')$$

where $S^F_{i,t}$ denotes the fundamental trading signal in period $t$. $S^F_{i,t}$ is analogue to eq. (4):

$$S^F_{i,t} = b(F_t - P_t) + \beta_{i,t}. \quad (9)$$

The new demand function of an agent $i$ who selects the technical strategy follows as:

$$D^C_{i,t} = \min\{-S^C_{i,t}, I_{i,t-1}\},$$

where $S^C_{i,t}$ denotes the technical trading signal in period $t$. $S^C_{i,t}$ is analogue to eq. (5):

$$D^F_{i,t} = \begin{cases} 
\min\{-S^F_{i,t}, I_{i,t-1}\}, & \text{if } S^F_{i,t} < 0 \wedge \Delta p_t < \text{limit} \wedge I_{i,t-1} > 0 \\
0, & \text{if } S^F_{i,t} < 0 \wedge \Delta p_t < \text{limit} \wedge I_{i,t-1} \leq 0 \\
S^F_{i,t}, & \text{otherwise.}
\end{cases} \quad (4'')$$

and

$$D^C_{i,t} = \begin{cases} 
\min\{-S^C_{i,t}, I_{i,t-1}\}, & \text{if } S^C_{i,t} < 0 \wedge \Delta p_t < \text{limit} \wedge I_{i,t-1} > 0 \\
0, & \text{if } S^C_{i,t} < 0 \wedge \Delta p_t < \text{limit} \wedge I_{i,t-1} \leq 0 \\
S^C_{i,t}, & \text{otherwise.}
\end{cases} \quad (5'')$$

where $\Delta p_t$ is the price change from $t-1$ to $t$:

$$\Delta p_t = P_t - P_{t-1}. \quad (11)$$

Eqs. (4'') and (5'') stipulate that, if her trading signal in $t$ is negative, agent $i$ can sell at maximum the number of stocks in her inventory, provided that shorting is prohibited in the respective period. Otherwise, agent’s actual buys/sells correspond to the trading signal she receives.\(^3\)

$$S^C_{i,t} = c(P_t - P_{t-1}) + \gamma_{i,t}. \quad (10)$$

Eqs. (4’) and (5’) simply state that if her trading signal in $t$ is negative, agent $i$ can sell at maximum the number of stocks in her inventory.

Under a limit-based short selling constraint, shorting the asset is not prohibited categorically. Rather, the restriction is in effect in some period $t$ only if the preceding tick has fallen below a specific critical limit. Assuming that a tick corresponds to the price change from one simulation step to the next, $\Delta p_t$, eqs. 12 and 14 can be modified into:

3.2 Measures of Market Efficiency and Stability

Following related studies (e.g., Hermsen et al. 2009), market efficiency is reflected in several indicators. Firstly, the price of the asset should be near to its intrinsic value; the asset should be neither over- nor undervalued. Secondly, market risk should be small. In particular, price variance should be low, and extreme price changes should occur rarely. These considerations disemboque in the following measures.

The absolute distortion of the market is defined as the average absolute difference of price and intrinsic value:

$$D^{abs} = \frac{1}{T} \sum_{t=1}^{T} |P_t - F_t|, \quad (12)$$

\(^3\) We assume that the market maker is never affected by short selling constraints, particularly because her inventory is too large to be exhausted under normal conditions. Therefore, her inventory is not considered here.
The absolute distortion indicates the accuracy of the asset’s valuation. The higher $D$, the higher the average mispricing in the market.

In contrast to the absolute distortion, the raw distortion differentiates between over- and undervaluation. It is defined as the average raw difference between price and intrinsic value:

$$D_{raw} = \frac{1}{T} \sum_{t=1}^{T} (P_t - F_t).$$  (13)

The more positive (negative) the raw distortion, the greater the tendency of the market towards overvaluation (undervaluation).

Price volatility $V$ is defined as the average of absolute returns:

$$V = \frac{1}{T} \sum_{t=1}^{T} |\Delta p_t|.$$  (14)

$V$ indicates financial market risk. The higher $V$, the higher the variability of prices.

The tail index of the return distribution is another indicator of financial market risk. The Hill estimator $\alpha^H$ (Lux & Ausloos 2006), gives an approximation of this index. It is calculated as follows:

$$\alpha^H = \left( \frac{1}{k} \sum_{i=1}^{l} (\ln|\Delta p_{L-i+1}| - \ln|\Delta p_{L-k}|) \right)^{-1}$$  (15)

where $k$ denotes the number of observations in the tail and the returns $\Delta p_t$ are sorted in a descending order: $\Delta p_{L} > \Delta p_{L-1} > \Delta p_{L-2} > \ldots > \Delta p_{1}$. We use the common tail fraction of 5%.

The tail index focuses on extraordinary events. A lower (higher) $\alpha^H$ indicates more (less) probability mass in the tails of the return distribution and, thus, a higher (lower) frequency of extreme price movements.

In addition to the tail index, the skewness of the return distribution focuses on the relation between extreme positive and extreme negative returns. The skewness of some distribution $X$ is the normalized third central moment of $X$:

$$S(X) = E \left[ \frac{(X - \mu_X)^3}{\sigma_X^3} \right],$$  (16)

where $E$ stands for expectation, $\mu_X$ is the mean of $X$, and $\sigma_X$ is the standard deviation of $X$. The more positive (negative) the skewness of the return distribution, the larger the proportion of probability mass in the right (left) tail of the distribution.

The tail index as well as the skewness indicator interprets extreme returns relative to the variance of the distribution. However, policy makers might also be interested in the occurrence of returns greater than some absolute percentage. Therefore we count the number of price rises (falls) greater (lower) than some limit $\delta$ ($-\delta$) per 1000 simulation steps. For the number of extreme price rises, $R^{+\delta}_{1000}$, this leads to:

$$R^{+\delta}_{1000} = \frac{|X|}{1000} \text{ with } X = \{\Delta p_1, \Delta p_2, \ldots, \Delta p_T\} \land \Delta p_t > \delta .$$  (17)

And for the number of extreme negative returns, $R^{-\delta}_{1000}$:

$$R^{-\delta}_{1000} = \frac{|X|}{1000} \text{ with } X = \{\Delta p_1, \Delta p_2, \ldots, \Delta p_T\} \land \Delta p_t < -\delta .$$  (18)

For the present analysis we set $\delta$ to 5%. Other values, e.g. 3% or 7%, yield similar results.

4. **Monte Carlo Analysis**

The consequences of the different short selling constraints are tested by a Monte Carlo analysis. In a first step, we describe the setting of the simulation experiments performed. The results are presented afterwards. The dynamic mechanisms driving the results will be explained and discussed in section 5.
4.1 Simulation Design
To evaluate the consequences of short selling constraints, one could simply turn the constraint on or off and measure the differences observable. However, this procedure would ignore that the constraint can affect investors more or less. Hence, it could be possible that certain effects appear only or even reverse if the impact of the constraints is relatively strong (or relatively low, respectively). Generally, the strength of the impact of a short selling constraint is determined by the proportion of traders who would like to sell more shares than they possess. This proportion tends to be greater, the higher the number of shares traded in the market. In our model this number results as the initial inventory size of each agent \( I_{v0} \) times the number of agents \( N \). Hence, by reducing (raising) \( I_{v0} \), we can increase (decrease) the impact of the short selling constraint.

Based on the consideration above, we stress test each constraint rule for a set of different initial inventories \( I_{v0} \). The respective inventories are: 80,000, 70,000, 60,000, 50,000, 40,000, 30,000, 20,000, 10,000, 5,000, and 2,000 shares per agent. Beside the general short selling constraint, we test five limit-based rules. The respective limits are: \(+1\%\), \(0\%\) (uptick rule), \(-1\%\), \(-3\%\), and \(-5\%\). For each combination of specific rule and specific inventory size, we performed twenty simulation runs with 7000 observation steps each. The measures were started at step 2001 to guarantee an appropriate distribution of inventory sizes.

Beside the large data experiments, we test the short selling constraints on singular simulation runs. The exemplary runs are meant to demonstrate the impact of the different constraints and to support our understanding about why certain effects appear.

4.2 Numerical Results
To begin with, we explore the impact of short selling constraints on a single simulation run. Figure 2 shows the same run as depicted in figure 1, except that this time the general short selling constraint is set into effect.

-- Figure 2 about here --

Focus the bottom panel first. The panel displays the inventories sizes of the same agents as in figure 1. The general constraint prevents each agent from building up short positions at any time. Hence, in contrast to the basic model without constraints (figure 1), none of the inventories ever falls below the zero level. Besides, the panel indicates that the inventories of the “richest” agents have reduced. Thus, the range of inventories has tightened form the bottom and from the top. The upper two panels give an impression of some effects of the constraint concerning market efficiency and stability. Compared to the basic model, the dynamics of prices is shifted upwards significantly (top panel). Most of the time the price moves above the intrinsic value, i.e., the asset tends to be overvalued. On the other hand, the dynamics seems to have tranquilized (middle panel): First, returns seem to have decreased on average. Second, extreme returns seem to have reduced in number and in size.

-- Figure 3 about here --

Figure 3 illustrates the same simulation run as figure 2, except that this time the uptick rule, representative for the limit-based rules, binds. The bottom panel shows that the uptick rule, like the other limit rules tested, does not ban shorting categorically. At some steps of time, singular investors have build up negative inventories, i.e. short position. However, the short positions are significantly lower than in the case without constraints (figure 1). The other effects appear to be very similar to the ones observed with the general constraint (figure 2): The market tends to be overvalued (upper panel), whereas the variance of returns seems to reduce (middle panel).
The large data experiments confirm and extend the observations made so far. Figure 4 and 5 illustrate the indicators of market efficiency as introduced in section 3.2. In each panel, the initial inventories \( \text{Inv}_{i0} \) are ordered from the highest setting \((\text{Inv}_{i0}=80,000)\) on the left to the lowest \((\text{Inv}_{i0}=2,000)\) on the right. The left-most data point corresponds to the reference case where no constraint (“no c.”) binds. The constraint rules are marked as follows: categorical, solid black; limit +0.01, large dashes; limit 0 (uptick rule), solid gray; limit -0.01, medium dashes; limit -0.03, small dashes; limit -0.05, dotted.

Look at figure 4 first. The figure shows the measures of market distortion. The left panel displays the absolute distortion and the right panel the raw distortion.

--- Figure 4 about here ---

The left panel indicates that short selling constraints tend to increase the absolute distortion of the market. The right panel shows that this mispricing is largely due to overvaluation: Without constraints the raw distortion is about zero on average, i.e., the asset tends to be over- and underpriced to equal degrees. In contrast, if short selling constraints govern, the raw distortion turns positive, i.e., overvaluations dominate. For the smallest starting inventory the average overvaluation is even equal to the absolute distortion, i.e., the mispricing completely consists in overvaluation. Comparing the different constraints, we find that the effect is the greater (lower), the more (less) restrictive the rule is designed. Accordingly, the categorical constraint takes the greatest effect, while the conditional constraint with limit −0.05 takes the lowest. Besides, the effects are more pronounced, the lower the size of agents’ initial inventories, which confirms our assumption made in section 4.1. We conclude that short selling constraints exacerbate market distortion as they produce overvaluation. The impairment is the worse, the more often the rule binds.

Figure 5 displays the indicators of market risk: The top panels highlight the price volatility (left) and the Hill estimator of the tail index of the return distribution (right). The panels in the middle illustrate the number of price falls lower than −5% (left) and the number of price rises greater than +5% (right). The skewness of the return distribution is depicted at the bottom.

--- Figure 5 about here ---

The results show that all short selling constraints tested reduce price volatility, i.e., the average price change becomes lower. At the same time, the Hill tail estimator rises, except for the less restrictive rule with limit −0.05, i.e., the tails in the return distribution become less heavy. Focusing extreme returns larger than 5%, we observe that the constraints reduce the number of respective price falls and the number of respective price rises. The effect on the skewness indicator is relatively sensitive to the design of the rule and to the size of the initial inventories: Normally, short selling constraints tend to produce positive skewness, i.e., extreme price rises occur more often than extreme price falls. Yet, the effect tends to disappear, if the rule affects agents very strongly, i.e., if the rule is rather restrictive and/or agents’ initial inventory is low. In general, it holds that the effects are more distinct, the more restrictive the constraint and the lower the initial inventory of agents. As a conclusion, the results show that short selling constraints reduce market risk. The improvement is the greater, the more often the prohibition applies.

Having discovered the impact of short selling constraints on the efficiency measures, we want to highlight some other interesting effects that become apparent in the artificial laboratory. Some of these observations will be valuable, furthermore, in order to retrieve explanations for the results concerning market efficiency. Figure 6 and 7 reveal the respective results.
Watch figure 6 first. The left panel shows the average number of chartists in some period. Apparently short selling constraints tend to reduce the weight of chartists. Again, the effect is the greater, the more restrictive the rule and the lower the initial inventory of agents. The right panel illustrates the proportion of transactions blocked by the different short selling constraints. Blocked transactions are short sells which would be executed without constraints but actually cannot because the prohibition applies. We expect that more restrictive rules block more transactions. In particular, the general constraint is supposed to have the highest impact because it always prohibits shorting whereas the conditional rules only sometimes bind. Surprisingly, the panel reveals that the general constraint prevents fewer deals than most conditional constraints. Only for small inventories the two less strict rules (limit −0.03 and −0.05) fall short. The reason for the counterintuitive finding is that conditional constraints enable agents to build up negative inventories, i.e. short positions. If an agent has a short position of the asset traded, she can only sell by further shorting. Thus, whenever the prohibition applies, the respective transaction is rejected. First when the agent has built up a sufficient long position, she is not affected anymore. Yet, this can take some time, such that many trades have been blocked in the meantime. In contrast, the general constraint prevents any short positions. As a result, agents are not “trapped” durably, and fewer transactions fail.

Figure 7 concentrates on agents’ inventories. We conducted two simulation series, one without short selling constraints and one with the categorical constraint. Each series consists of 20 runs with 5000 simulations steps each. The two boxplots capture the average distribution of agents’ inventory sizes at the end of the run for the respective series (for example, the upper end of the whisker represents the mean maximum inventory averaged over all runs in the series). The comparison of the box plots confirms that the constraint leads to the assimilation of inventory sizes, as already observed above. The constraint has reduced the distance between the 25th and the 75th percentile of the distribution. Furthermore, the maximum inventory has decreased while the minimum inventory has increased.

5. Interpretation

Having reported the numerical results of the Monte Carlo analysis, this section aims to provide an explanation of the different consequences of short selling constraints. Moreover, we discuss the robustness of the results obtained. In particular, we explore if and how far the effects observed depend on the design of our model.

5.1 Explication of Numerical Results

The numerical results of short selling constraints on market efficiency and stability can be summarized as follows: On the one hand, short selling constraints increase mispricing because they produce overvaluation. On the other hand, the constraints reduce market risk, which implies a reduction of price volatility and extreme price movements. In general, the effects are the stronger, the more agents are affected by the constraint. Agents are more affected, firstly, if the constraint is more restrictive and, secondly, if they are equipped with less shares.

Asking about the causal chains between short selling constraints and the changes concerning market efficiency, the model reveals the following: Short selling constraints prevent that part of the investors who would like to sell (for being

---

4 Of course, the average inventory is constant, as neither the number of shares traded nor the number of agents changes.
pessimistic about the future development of prices) can execute their transactions as they depend on shorting. Hence, the balance between supply and demand shifts towards the latter, and prices rise, leading to overvaluation. As a consequence, the market distortion deteriorates.

The simultaneous improvement of market risk is driven by three channels. Channel (1) can be attributed to the short selling constraint directly. The constraint cuts part of the sells, which causes price drops to lose strength. As a result, there is lower probability mass in the left tail of the return distribution, the Hill tail estimator rises, and the skewness indicator turns positive. Besides, the number of price falls greater than some specific percentage declines. Channel (2) is due to fundamental trading activity: The overvaluation of the asset gives fundamentalists incentives to sell. The additional sells by fundamentalists weaken every price rise. Hence, extreme price rises occur less often. Channel (3) rests on a shift of trader groups: Due to the rise of market distortion, fundamental (technical) strategy becomes more (less) attractive (see eq. 9). As a consequence, the proportion of chartists declines (see figure 6). Chartists represent a major cause for market turbulence (see also Day and Huang 1990). Thus, their numerical decline stabilizes the dynamics of prices leading to an improvement of all risk indicators.

In addition, channel (3) explains why short selling constraints lead to a reduction of positive inventories: Due to the overvaluation, traders with long positions tend to sell more than to buy. (Whenever the asset is overvalued, the fundamental strategy prescribes to sell). Hence, long positions tend to reduce.

5.2 Discussion
A common criticism against model studies is that their results are highly sensitive to the assumption made. In particular, many findings seem to depend on the rationality of agent behavior. This point has lead to rivaling conclusions about the relationship between short selling constraints and market distortion. As mentioned before, empirical studies confirm that constraints produce overvaluation. However, those studies face a methodological difficulty. The argument dates back to Fama (1970) who termed it “the joint hypothesis problem”: In order to measure market distortion, one needs a computational method that determines the intrinsic value of the asset. Therefore, the finding of market distortion can originate either from true market inefficiency or from a method that does not correctly account for all information relevant for value. (One could suspect, for example, that short selling constraints increase the value of the respective assets because they reduce market risk. Hence, a higher price level after the introduction of constraints would not point to overvaluation).

The joint hypothesis problem stresses the importance to propose a strong argument about why short selling constraints should produce overvaluation. Empirical studies often refer to an idea originally purported in the model study by Miller (1977) (e.g., Chang et al. 2007). The reasoning is analogue to the one valid in our model: Short selling constraints prevent part of the pessimistic investors from trading on their belief of a future price drop whereas optimistic investors are not affected. Due to this asymmetry the asset price will reflect positive prospects to a higher degree than negative ones, i.e., the market tends towards overvaluation.. Diamond & Verrechia (1987) challenge this argument. The authors argue that the existence of short selling constraints is common knowledge. In particular, investors would regard that some pessimistic agents cannot trade on their private, negative information. Based on this assumption a rational expectation model is constructed. The model shows that despite of short selling constraints, prices remain unbiased.
Our study mediates between the models of Miller (1977) and Diamond & Verrechia (1987). On the one hand, the agents in our model do react to the constraint. Agents recognize the overvaluation and respond, firstly, by choosing fundamental analysis more often to exploit arbitrage opportunities, and, secondly, by reducing their average net demand. On the other hand, the reaction of agents does not remove the overvaluation. The reason is that the power of fundamentalists is too low. In other words, the potential of arbitrage is limited. The observation of limited arbitrage is one of the building blocks of behavioral finance (for a survey see Shleifer, 2000). For example, Shleifer (1997) reports that arbitrage is conducted by a relatively small number of traders managing the capital of others. Those traders fail to eliminate mispricing completely, because the amount of capital is too low, or because they are forced to deviate from their strategy if being under pressure, among other reasons.

We conclude that our finding receives some empirical support: Short selling constraints, no matter if general or limit-based, tend to increase the level of prices. Investors react to this upward tendency by arbitrage trading. However, the forces of arbitrage are too weak to let prices realign. Therefore, the market remains overvalued on average.

It remains to ask about the robustness of our results concerning market risk. We found that short selling constraints lead to a reduction of market risk because of three channels: (1) the obstruction of sells due to constraints, (2) the reaction of fundamentalists to overvaluation, and (3) shifts of strategy weights due to overvaluation, and. To figure out if the risk finding is robust, we have to determine if these channels are likely to occur in reality.

Channel (1) is a direct effect of the prohibition. Therefore we assume the channel to be quite unproblematic. Indeed, empirical studies show that trading volume reduces significantly when shorting is prohibited because most transactions involving short sells are banned (e.g., Boehme et al. 2009). Channel (2) is sensitive to the assumption that traders are able to recognize mispricing at least to some degree. Regarding the myriad of methods to calculate fundamental value (for examples see e.g., Greenwald et al., 2001), we have strong evidence that traders at least get some ideas about the true value of an asset. Channel (3) depends on the assumption that investors do not learn (or at least learn slowly) if the general conditions in the market have changed. The argument can be expressed as follows: The heuristic of strategy choice (eq. 6) is based on the price level to which traders expect prices to return in the long run. Typically this is the intrinsic value of the asset. Yet, if the market tends to be overvalued, the level is above. Rational investors would react by adapting their heuristic of strategy choice. In the end, the proportions of technical and fundamental traders would be hardly affected. If real investors adjust their heuristic of strategy choice or do not, depends on their rationality. We suppose that at least some investors do react to the new level of prices, meaning that channel (3) is weaker in reality than in our model. Nevertheless, we conclude that at least two of the three channels producing a decrease of market risk are quite robust (channel 1 and 2).  

6. Conclusion
More than one year after the 2008 short selling ban by the SEC it is still an open question if the measure stabilized financial

---

5 An exception might be if investors assume the constraint only to hold temporally. In this case, expecting prices to return to value is still rational. Hence, the heuristics of strategy choice needs not to be adapted.
6 All three channels are likely to be active if traders assume the constraint to hold only temporally (as it probably applies to the constraint imposed by the SEC in 2008.
markets. Empirical studies report that short selling constraints create overvaluation. Less agreement has been achieved with respect to the consequences for market risk. Agent-based models have been proven to be a successful tool to explore regulative policies in an artificial laboratory. Such models view financial markets as systems consisting of heterogeneous agents that rely on simple heuristics, and can replicate real market dynamics quite accurately. Agent-based models enable to isolate specific cause-and-effect chains and to test regulative measures in variable configurations.

In our study we develop a multi-agent model of a financial market to stress test different variants of short selling constraints. In the basic model every investor manages an own inventory of shares. A negative inventory represents a short position. The traders choose between fundamental and technical trading rules. The basic model is then extended by short selling constraints. We investigate the general short selling constraint, which prohibits shorting categorically, and different limit-based constraints (like the uptick rule), which disallow shorting only if the preceding price change falls below a defined limit. The results confirm that short selling produces overvaluation, since part of the negative expectations is not reflected in prices anymore. On the other hand, market risk declines. The latter finding contains several aspects: Firstly, market volatility reduces. Secondly, the heavy tails in the return distribution decline, i.e., extreme price movements occur less often. The latter is true for extreme price falls and, more surprisingly, for extreme rises. Besides, the skewness of the return distribution turns positive indicating that the decline of extreme price falls is over proportional. The effects mentioned arise no matter if the constraint is general or limit-based, but their intensity grows if the rule is more restrictive.

Although the results appear to be quite robust, future agent-based models should focus several aspects to refine our understanding of the relationship between short selling constraints and market efficiency. Such aspects are: Differences between temporal and durable constraints, the influence of borrowing fees, pressure to cover the short position (particularly so-called “short squeezes”), the influence of budget constraints, or the explicit consideration of the role of lenders.
References


Figure 1: Typical dynamics of the basic model (no short selling constraints). From top to bottom, the panels show the evolution of the log of the price (black) and the log intrinsic value (gray); the corresponding return time series; the number of fundamentalists (gray) and chartists (black); and the inventory sizes of ten exemplary agents. For parameter setting and initial values see table 1.
Figure 2: Typical model dynamics with general short selling constraint implemented. From top to bottom, the panels show the evolution of the log of the price (black) and the log intrinsic value (gray); the corresponding return time series; and the inventory sizes of ten exemplary agents. Parameter setting and initial values as in table 1.
Figure 3: Typical model dynamics with uptick rule (limit=0) activated. Organization of panels as in figure 2. Parameter setting and initial values as in table 1.
Figure 4: Measures of market distortion for different variants of short selling constraints. The left panel shows the absolute distortion and the right panel the raw distortion, each for different values of agents’ initial inventory $Inv_0$. The different constraints are: categorical (solid black), limit +0.01 (large dashes), limit 0 (uptick rule) (solid gray), limit -0.01 (medium dashes), limit -0.03 (small dashes) and limit -0.05 (dotted).
Figure 5: Measures of market risk for different variants of short selling constraints. From top to bottom and left to right: the volatility of prices, the Hill estimator of the tail index of the return distribution, the number of price falls smaller than −5%, the number of price rises greater than +5%, and the skewness of the return distribution. The organization of the singular panels is analogue to figure 4.
Figure 6: The number of chartists (left) and the proportion of transactions blocked by the different constraints (right). The organization of the singular panels is analogue to figure 4.
Figure 7: Distribution of inventory sizes after 5000 simulation steps, either without short selling constraint (left) or with general constraint (right). The bottom (top) of the box is the 25th (75th) percentile. The mid vertical line represents the median and the black dot indicates the respective mean. The upper (lower) end of the whisker stands for the minimum (maximum) inventory. All indicators are averaged over a series of 20 runs for each case. The dashed lines highlight the level of the 25th and 75th percentile of the distribution without constraints as well as the level of the mean and the median.
Figure 8: Tests for some important stylized facts. The upper two panels show the autocorrelation of returns (top) and absolute returns (middle) for different lags. The bottom panel depicts the return distribution. The measures are based on a run with 10,000 simulation steps.
<table>
<thead>
<tr>
<th>Param.</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>0.0001</td>
<td>Strength of price adaption</td>
</tr>
<tr>
<td>b</td>
<td>9</td>
<td>Aggressiveness of fundamentalists</td>
</tr>
<tr>
<td>c</td>
<td>0.1</td>
<td>Aggressiveness of chartists</td>
</tr>
<tr>
<td>d</td>
<td>100</td>
<td>Speed of strategy switching</td>
</tr>
<tr>
<td>$\sigma^a$</td>
<td>0.01</td>
<td>SD of periodical change of fundamental value</td>
</tr>
<tr>
<td>$\sigma^\beta$</td>
<td>50</td>
<td>SD of random component of fundamental demand</td>
</tr>
<tr>
<td>$\sigma^\gamma$</td>
<td>800</td>
<td>SD of random component of chartist demand</td>
</tr>
<tr>
<td>$D_{i,0}$</td>
<td>0</td>
<td>Demand of agent $i$ in period 0</td>
</tr>
<tr>
<td>$F_0$</td>
<td>0</td>
<td>Log fundamental value in period 0</td>
</tr>
<tr>
<td>$I_{i,0}$</td>
<td>10,000</td>
<td>Size of starting inventory of agent $i$</td>
</tr>
<tr>
<td>$P_0$</td>
<td>0</td>
<td>Log price in period 0</td>
</tr>
</tbody>
</table>

Table 1: Setting of parameters and initial values, with brief descriptions of their functions.