Do short-selling constraints matter?*

Francesca Cornelli
London Business School and CEPR

Bilge Yilmaz
Wharton School

ABSTRACT

We analyze the impact of short selling constraints in a setting in which there is uncertainty about the amount of informed traders. We show that price converges to the fundamental value if informed traders do not face significant short selling restrictions. In contrast, when short sellers face large costs, the prices may remain bounded away from fundamentals even when there are arbitrarily many informed traders with noisy private signals. In this case, long run equilibrium prices are determined by (i) the ratio between buy and sell orders; and (ii) prior beliefs. We then analyze a richer framework, where we show that not only rational expectation equilibrium prices do not converge to the fundamental, but also they are path dependent: the equilibrium price will depend on the specific sequence of buy and sell orders (in other words, on a random component).

*We would like to thank Harjoat Bharma, Pedro Saffi, Ingrid Werner and seminar participants at Aalto University, Boston College, Copenhagen Business School, Frankfurt, INSEAD, MIT, Oslo BI, Stanford, the Stockholm School of Economics, UC Davis, UCLA, University of Amsterdam and the UBC Summer Symposium.

I. Introduction

During the recent financial crisis, short sales came under attack as a potential source of destabilization. In 2008 both the SEC in US and the FSA in UK prohibited short-selling on specific companies, in order to stabilize their market price. In December 2008, however, the SEC Chairman Christopher Cox released an interview where he admitted he regretted using the short-selling ban, saying that "preliminary findings point to several unintended market consequences and side effects caused by the ban".[1] In Europe there has been a recent history of trying to ban or regulate short-selling (or sometimes only naked short-selling). Although these are examples of temporary bans introduced for extreme circumstances, which is not what we will discuss in this paper, they are symptomatic of the current debate: while governments and regulators are seeing short-sales with suspicion, markets claim that short-sales help liquidity and price discovery.

The academic literature (reviewed in the next section) has studied whether short sales restrictions hinder or not price discovery. The argument in favor of short-sales is that it allows investors who have a pessimistic view on a stock but do not own that stock, to trade it anyway. Therefore, constraining short-sales will bias the price. Most papers that have modeled this phenomenon to show that prices are biased upwards, such as Hong and Stein (2009), have relied on some form of myopia on the investors side. On the opposite hand of the spectrum, Diamond and Verrecchia ((1987) have shown that in a fully rational model prices will converge to the fundamental value even in presence of a short-sale ban. Prices may converge more slowly, but there is no long term effect on the market price. At the same time, empirical studies have had mixed evidence: while some studies show that costs of short-selling may lead to mispricing, other studies have shown that a reduction in the costs of short-selling has not led in a move downwards of prices.

This paper shows that short-sales constraints may affect long run prices even when investors are fully rational. We show two effects. First, we start from the Diamond and Verrecchia (1987) model and introduce one new feature: there is uncertainty on the number of informed traders. This seems a very reasonable assumption: even if one knew all the existing institutional investors, these investors are not following all the stocks, and therefore one does not know how many investors informed there are for a specific stock. This simple and realistic

[1] "SEC chief has regrets over short-selling ban", Reuters, December 31st, 2008.

feature turns out to be very powerful: we can show that in such case, as long as the signal observed by the informed investors is not too precise, if short selling is costly the rational expectation equilibrium prices do not converge any more to the fundamental values, but they instead converge to a value that is determined by (i) ratio between buy and sell orders; and (ii) the initial priors of the model.

The intuition for this result is simple: in models like Diamond and Verrecchia (1987) market makers know how many informed investors there are: by observing how many investors are buying and how many are selling, they can also figure out how many would be short selling if there were no constraints. Therefore, the information that these informed investors have will ultimately be incorporated in the price. But if market makers do not know how many informed investors could potentially be trading, then such inference becomes more difficult and not all the information may end up being incorporated in the prices.

We also show that short-sales constraints may lead to prices biased both upwards or downwards. Thus the intuition that in presence of short-sales constraints prices will be biased upwards is not true when everybody is rational. This is very intuitive: if everybody knows prices will be biased upwards, rational agents will adjust prices downwards. When there is no additional uncertainty, as in the Diamond and Verrecchia (1987) model, they will adjust correctly in the long run and the price will not be biased, while in presence of additional uncertainty as in our model they may both over or under-adjust. The fact that prices may be biased in both directions can help explain the different results the empirical literature has had in testing for the effects of short-sales constraints. If prices are not biased systematically upwards, there is no reason to expect a statistically significant movement in the prices when short sales constraints are removed or added. Although we show that prices can be biased both upwards and downwards, part of the original intuition still holds, since the model is not symmetric: the price is more likely to be biased when the fundamental value is low. In fact, when the fundamental value is low, short sales constraints are most likely to be binding and thus agents are most likely to underestimate the number of people who are short sales constrained.

We then introduce a more complex model and we show that not only the rational expectation equilibrium prices are different from the fundamental value, but they will also depend on the specific realization of the sequence of buy and sell orders. In other words, the price cannot be predicted even if we know the priors but it will be path dependent. These results show that the effect of short-sales constraints is more difficult to assess than one may have expected, and may also depends on random outcomes. They also suggested that initial trading conditions matters and there may be incentives to affect them in one direction rather than another.

Another way to state the contribution of this paper is that we show that if the uncertainty the market faces is multi-dimensional, then the prices may converge to levels that are determined by prior beliefs. Therefore, fundamental values may have no impact even though there are arbitrarily many informed traders who have private information about fundamentals. Note however that multi-dimensional uncertainty in our setting is a necessary but not sufficient condition. For each of the various cases we consider, we always start by showing that if short sales are not costly, prices converge to the fundamental values. Thus, even with multi-dimensional uncertainty it is the short-sales constraints that lead to non-convergence.

We only focus on equilibrium long run prices and therefore have nothing to say about short-run phenomena. In particular, this paper does not study whether it may be optimal to impose bans on short-sales to avoid crisis in the short run. Similarly, we do not study the role of short sales constraints in bubbles formation. Other papers derive conditions for bubbles to arise when there are short sales constraints (see, for example, Harrison and Kreps (1978)). Allen, Morris and Postlewaite (1993) show that necessary conditions for a bubble to arise are that (1) each agent must be short sale constrained in the future with some positive probability and (2) agents' trade is not common knowledge. We do not focus on whether temporary bubbles may emerge, but on whether in the long run the price will converge to the fundamental.

Finally, our model can be useful to study IPOs, where it has been argued that it is difficult to short a stock for a month following an IPO. Many people, starting with Miller (1977), have argued that this is one reason why we observe prices rising so much after an IPO (see for example, Aggarwal, Krigman and Womack (2003)).

In the next section we briefly discuss the related literature and the existing empirical evidence on short-sales and the effects of constraints on short-sales. In Section 3 we then introduce our basic model, where the value of the underlying asset can take values 0 or 1. We start analyzing what happens when there are no short-sales constraints at all and show that in such case prices convergence to the fundamental value. We then introduce a small constraint (in the form of a cost ϵ to short sell) and finally a large cost effectively totally banning short sales. We show that when large costs are introduced the price does not converge to the fundamental value. In Section 4 we then extend the model in the sense that the underlying asset can take 3 values: 0, 1 and 2. We derive a similar result to the central one in Section 3 (the no-convergence result) and then show some examples in which the price is path dependent. Section 5 concludes.

II. Related Literature and Empirical Evidence

There is a very large literature, mostly empirical, on short selling. As already mentioned in the Introduction, most of the theoretical models who show that short sale constraints affect prices in equilibrium have some myopic component where the optimistic investors do not fully update for the absence of the pessimistic ones, due to the short sale constraints (see for example Miller (1977), Hong and Stein (2003), and Duffie, Garleanu and Pedersen (2002)). Rational expectations models as in Diamond and Verrecchia (1987) predict instead that the long-run equilibrium prices will not be affected by short sale constraints.

Bai, Chang and Wang (2006) is the only other paper that shows that short sales constraints may actually lead to prices which differ from the fundamental value in a rational model. They introduce risk aversion and show that if prices do not aggregate all the information then trading these assets is riskier, and thus their price will decrease. This is thus a different effect. Moreover, their prediction is that prices will be downward biased (while in our model they can be both downwards or upwards biased) and do not obtain that the price is path-dependent.

Several empirical papers show that short sellers are well informed and that they earn excess returns (see, among others, Cohen, Diether and Malloy (2005), Desai, Krishnamurthy and Venkataraman (2006), Boehmer, Jones and Zhang, (2008a), Diether, Lee and Werner (2009) and Boehmer and Wu (2012)). This is consistent with prices not incorporating all the available information. Moreover, Karpoff and Lou (2011) study short-selling in companies that are later found guilty of financial misrepresentation. They find that short-sales helped reduce the price inflation due to the earnings misstatements: Such information would not have been incorporated in the prices if it were not for short-sellers. All these papers show the important role played by short sales and that short-sellers are informed, but do not analyze the impact of short sale constraints (in other words, they do not show what would happen in a world with or without constraints).

Some papers have been identifying effects of short sale constraints. Nagel (2005) argues that stock with low institutional holdings are the ones where the short-sale constraints are most likely to be binding and finds that indeed these are the stocks showing return anomalies. Bris, Goetzmann and Zhu (2007) study 46 equity markets around the world with different degrees of short sales restrictions and find that short sales constraints reduce market efficiency. Saffi and Sigurdsson (2011) use the lending supply as a proxy for short-sales constraints. They find not only that short sale constraints reduce price efficiency, but also that lower constraints are associated with a greater degrees of negative skewness and fewer occurrences of extreme price increases, but no link with price decreases. Finally, Campello, Matta and Saffi (2018) show how the equity lending market affects corporate behaviour.

A series of recent papers have used recent regulation changes and bans during the crisis to assess the effect of short sales constraints. Boehmer, Jones and Zhang (2012) and Pagano and Beber (2013) study the effects of the 2008 and 2009 bans on short sales. Both papers find that the bans brought a deterioration of market quality. Boehmer, Jones and Zhang (2008b) however study a change in the SEC regulation that reduced the cost to short sell (the uptick rule). Since for some stocks the uptick rule had already been removed on a trial phase, they can look at the effect on the prices of the shares that were affected by the regulation change and compare it to the effect on the prices of the shares that were not affected, but do not find significant difference between the two sets of shares. Similarly, Kaplan, Moskowitz and Sensoy (2012) use a randomized experiment to study the effect on prices of a change in the cost of short-selling and find no evidence consistent with overpricing.

Our results can help to reconciliate these different findings, since we find that prices can be both upwards and downwards biased.

Our paper is also related to a literature introducing multidimensional uncertainty in the Glosten and Milgrom (1985) model. Jacklin, Kleidon and Pfleiderer (1992) modify the Glosten and Milgrom (1985) model introducing uncertainty about the extent of portfolio insurance. Thus the market maker can mistakenly interpret a trade for portfolio insurance as an informed trade. They show that prices are biased and the price falls when the amount of portfolio insurance is revealed. Thus, while they introduce uncertainty about the type of demand, we introduce uncertainty about the unobserved demand (the constrained short-seller). Avery and Zemsky (1998) start from the informational cascades model of Bikchandani, Hirshleifer and Welch (1992). In this last paper, the sequential nature of the Glosten and Milgrom model creates informational cascades. Avery and Zemsky (1998) first show that if prices are allowed to change to incorporate all public information then informational cascades are impossible and prices converge to the fundamental value. Then they introduce an additional source of uncertainty (event uncertainty) and show that herd behaviour is once again possible and prices may not converge to the fundamental. In our case, however, the introduction of an additional form of uncertainty is not enough to distort rational expectation prices, unless short sale constraints are also present.

III. The Model

Our model follows Glosten and Milgrom (1985) and Diamond and Verrecchia (1987). There are two assets that traders may own. These two assets are ex-ante identical in terms of distribution of fundamental values, fraction of informed traders that may trade, etc. Fraction $\rho \in (0,1)$ of traders owns asset 1 and $1-\rho$ owns asset 2. In our analysis we will focus and analyze what happens to asset 1 only. This is without loss of generality, since asset 2 analysis would be exactly symmetric.[2]

Each asset has a fundamental value $v \in \{0,1\}$, where $\Pr(v=1)=\lambda$. The market makers are risk-neutral and face no inventory costs or constraints. There is also a continuum of risk neutral traders that can be of two types: liquidity (or noise) and informed. The fraction $\mu \in [0,1]$ of all traders is informed, i.e., fraction $1-\mu$ of traders are liquidity traders.[3] Each informed trader, i, receives an independent (conditional on v) signal $s_i \in \{0,1\}$; $\Pr(s_i=1|v=1)=\Pr(s_i=0|v=0)=\phi \in (\frac{1}{2},1)$. The posterior belief of a trader with a high signal is $\beta(v=1|s_i=1)=\frac{\lambda\phi}{\lambda\phi+(1-\lambda)(1-\phi)}>\lambda$.

In line with Glosten and Milgrom (1985), each trader can trade only one unit of asset 1. Note that our results do not depend on the single trade size assumption and we can handle multiple trade sizes.[4]

In this standard model we introduce an additional element of uncertainty: the fraction of informed traders μ is not known. We assume instead that $f(\mu)>0$ for all $\mu \in [0,1]$ is the

[2]To be precise, Glosten and Milgrom (1985) and Diamond and Verrecchia (1987) have one asset only. Our results do not depend on having two assets, we could obtain the same results with one asset (and in fact we are focusing on one asset only). The reason we introduce two assets is that it is more natural to think that every trader owns some asset. Moreover, it will make more natural to discuss the implications of short-selling constraints, as it will be clearer later.

[3]Note that the values of the two assets are not correlated and the fraction of informed investors for asset 1 is independent from the action of informed investors for asset 2.

[4]If there are multiple trade sizes, then the informed trader maximizes his profit. In equilibrium, the informed trader is indifferent between different trade sizes and plays a mixed strategy.

probability density function of μ. This is the key difference between our setting and Glosten and Milgrom (1985) and Diamond and Verrecchia (1987).

A half of the liquidity traders buys and the other half of them sells an asset, i.e., liquidity shock is symmetric (this assumption is only for simplification). If a liquidity trader needs to buy, he will buy one of the two assets, we assume that a liquidity trader buys asset 1 with probability ρ.[5] If instead the liquidity trader needs to sell, in absence of short selling costs he is absolutely indifferent between selling the asset he owns and short selling. We assume that he sells asset 1 with probability ρ.[6]

The period is denoted by $t \in \{1, 2, ..., T, ..., \infty\}$. Market makers are competitive and uninformed: following Glosten and Milgrom (1985) they set prices such that in each period they make zero profit. They post bid and ask prices for one share, then the traders decide whether they would like to trade. And nature picks one trader (out of all the traders who are willing to trade).[7] One unit trade takes place at the posted price. Let H^t stand for all of the publicly available information at the beginning of period t and $h^t \in \{-1, 1\}$ is the order in period t. Therefore, the ask price is $A_t = \Pr(v = 1|H^t, h^t = 1)$ and bid price is $B_t = \Pr(v = 1|H^t, h^t = -1)$.

Our aim is to determine the value to which the price will converge as the number of trades observed becomes infinitely large, i.e. as $t \to \infty$. As a first step, we want to show that in the absence of short selling constraints prices converge to the fundamental value. In order to determine that, we first determine the optimal strategy of an informed trader in Lemma 1.

LEMMA 1: *An informed trader always wants to trade and he buys if and only if $s_i = 1$.*

The proof is immediate and is omitted.

Given the informed and noise traders' trading strategies, the fraction of buy orders converges to a level determined by the fundamentals. Let q_v stand for this level, where $v \in \{0, 1\}$.

---

[5] An alternative assumption would be to assume that when the noise traders need to buy they will buy the asset they do not own already, for example for hedging reasons. This makes the analysis more cumbersome but ultimately delivers the same results.

[6] For our main result, it is not necessary for the liquidity trader to buy or sell asset 1 with the same probability. However, using probability ρ for both buying and selling avoids introduction of additional notation.

[7] Note that an outcome of no trade is possible only when there are no traders who would like to trade. Alternatively, we can allow the market makers to observe the fraction of traders who do not trade by allowing nature to pick these traders. Our results hold as long as there is residual uncertainty about the type of the trader who prefers not to trade.

Finally, let P_t denote the probability that the true state of nature is $v = 1$, in other words P_t is the conditional expectation of the asset's value at time t, given all the public information. As in Diamond and Verrecchia (1987), P_t can also be defined as the price of the asset at time t. We are now able to determine what P_t converges to as $t \to \infty$.

PROPOSITION 1: *In the absence of short selling constraint the price P_t converges to v as $t \to \infty$.*

PROOF: First note that $1 - (1-\mu)(1-\rho)$ fraction of traders would like to trade asset 1. Therefore, as t gets arbitrarily large the fraction of buy orders converges to $q_1 = \frac{\mu\phi + \frac{(1-\mu)\rho}{2}}{1-(1-\mu)(1-\rho)}$ if $v = 1$. Similarly, the fraction of buy orders converges to $q_0 = \frac{\mu(1-\phi)+\frac{(1-\mu)\rho}{2}}{1-(1-\mu)(1-\rho)}$ if $v = 0$. Note that

$$q_1(\mu = 0) = q_0(\mu = 0) = \frac{1}{2}$$

Next, we differentiate both q_1 and q_0 with respect to μ and show that $\frac{\partial q_1}{\partial \mu} > 0 > \frac{\partial q_0}{\partial \mu}$ for all $\mu \in [0, 1]$.

$$\frac{\partial q_1}{\partial \mu} = \frac{(\phi - \frac{\rho}{2})(1-(1-\mu)(1-\rho)) - (1-\rho)(\mu\phi + \frac{(1-\mu)\rho}{2})}{(1-(1-\mu)(1-\rho))^2}$$

Simplifying leads to $\frac{\partial q_1}{\partial \mu} = \frac{(2\phi-1)\rho}{2(\mu+\rho-\mu\rho)^2} > 0$. Similarly,

$$\frac{\partial q_0}{\partial \mu} = \frac{(1-\phi-\frac{\rho}{2})(1-(1-\mu)(1-\rho)) - (1-\rho)(\mu(1-\phi) + \frac{(1-\mu)\rho}{2})}{(1-(1-\mu)(1-\rho))^2}$$

Simplifying leads to $\frac{\partial q_0}{\partial \mu} = -\frac{(2\phi-1)\rho}{2(\mu+\rho-\mu\rho)^2} < 0$. Therefore, we have $q_1 > \frac{1}{2} > q_0$ for all $\mu \in (0, 1)$. Consequently, whenever the market maker observes a convergence to $q > \frac{1}{2}$ he concludes that $v = 1$ and similarly he concludes that $v = 0$ whenever he observes a convergence to $q < \frac{1}{2}$. Therefore, although the market makers do not know the realization of μ, in the long run they observe different fraction of buy orders depending on whether the true value v is 0 or 1 and thus learn the true value of v. ■

Although we introduced multi-dimensional uncertainty, in absence of short sales constraints this is not enough and the price is still converging to the fundamental value as in Diamond and Verrecchia (1987). That is because without short sales constraints all investors are trading and therefore it is easier to learn over time how many are the informed investors.

III.1. Introducing minimal constraints on short selling. In the previous section we assumed that there is no cost in short-selling. However, in real life traders have to borrow the underlying share and therefore bear a cost. This is equivalent to assume that short-selling involves a small ϵ cost. In such a situation, liquidity traders will never short sell asset 1, because they can always sell the asset they own (asset 2) at 0 cost. In Diamond and Verrecchia (1987) there is only one asset, and they assume that liquidity traders short sell if there are no constraints and do not short sell at all if there are constraints. In our context, with two assets, it is easier to justify why noise traders do not short-sell, even with an ϵ cost.

The arbitrarily small cost, ϵ, however, will not deter an informed trader from short selling and we show in the next proposition that even in this situation the price will still converge to the fundamental value.

PROPOSITION 2: *In the absence of short selling constraints on the informed traders (or if short selling costs are arbitrarily small) the price P_t converges to v.*

PROOF: We first study the strategy of the liquidity traders: A fraction $(1-\mu)(1-\rho)$ of traders are liquidity traders who do not own asset 1. Among these half of them, i.e., fraction $\frac{(1-\mu)(1-\rho)}{2}$, will sell an asset. However, given that they do not own the asset 1, they will not trade asset 1. The other half, i.e., another fraction $\frac{(1-\mu)(1-\rho)}{2}$, will buy an asset. But $(1-\rho)$ of these will buy asset 2. Therefore, due to noise traders who do not own asset 1, fraction $\frac{(1-\mu)(1-\rho)}{2}(1+(1-\rho))$ of traders will not trade asset 1. Similarly, the fraction $(1-\mu)\rho$ of traders are liquidity traders who own asset 1. Half of these, i.e., $\frac{(1-\mu)\rho}{2}$, will buy an asset. But $(1-\rho)$ of these will buy asset 2. Therefore, the fraction $\frac{(1-\mu)\rho(1-\rho)}{2}$ of traders will buy asset 2. In sum, fraction $\frac{(1-\mu)(1-\rho)}{2}(1+(1-\rho)+\rho)$ of traders are noise traders who will not trade asset 1. Note that $\frac{(1-\mu)(1-\rho)}{2}(1+(1-\rho)+\rho)$ simplifies to $(1-\mu)(1-\rho)$.

Next, we consider the $v=1$ case. For sufficiently small ϵ arbitrarily large fraction of informed traders buy. Therefore, for sufficiently small ϵ, as t gets arbitrarily large the fraction of buy orders converges to $q_1 = \frac{\mu\phi+\frac{(1-\mu)\rho}{2}}{1-(1-\mu)(1-\rho)}$ if $v=1$. Similarly, the fraction of buy orders converges to $q_0 = \frac{\mu(1-\phi)+\frac{(1-\mu)\rho}{2}}{1-(1-\mu)(1-\rho)}$ if $v=0$. The rest of the proof follows immediately from proof of the previous proposition. ■

III.2. High Short Sale Constraints. To keep the analysis simple we will impose the most extreme constraint and rule out short selling by imposing a cost equal to 1. (For our results to hold, the constraint must prohibit a sufficiently large fraction of the informed short selling.) We will show in the next proposition that if the signal observed by the informed traders is not too high then the price will not converge to the fundamental value.

Let us first consider the $v = 1$ case. A fraction $\mu(1-\rho)(1-\phi)$ of traders are informed traders who received a signal that the value of the shares is 0 and would like to sell asset 1 but cannot sell it since they do not own asset 1. Moreover, from the previous subsection we know that fraction $(1-\mu)(1-\rho)$ of traders are liquidity traders who will not trade asset 1. Therefore, $1-\mu(1-\rho)(1-\phi)-(1-\mu)(1-\rho) = 1-(1-\rho)[\mu(1-\phi)+(1-\mu)]$ is the fraction of traders who would like to trade asset 1. Similarly, the fraction of traders that would like to trade when $v = 0$ is $1-(1-\rho)[\mu\phi+(1-\mu)]$.

If $v = 1$, then the fraction of buys (out of all trades) q_1 converges to $\frac{\mu\phi+\frac{(1-\mu)\rho}{2}}{1-(1-\rho)[\mu(1-\phi)+(1-\mu)]}$. Similarly, if $v = 0$ then the fraction of buys (out of all trades) q_0 converges to $\frac{\mu(1-\phi)+\frac{(1-\mu)\rho}{2}}{1-(1-\rho)[\mu\phi+(1-\mu)]}$.

Now let us look at the problem from the point of view of the market makers. They observe that the fraction of buys is converging to $q \in (0,1)$. Let μ_1^q stand for the solution of $\frac{\mu\phi+\frac{(1-\mu)\rho}{2}}{1-(1-\rho)[\mu(1-\phi)+(1-\mu)]} = q$. Therefore, μ_1^q is the fraction of traders who are informed that would lead to a fraction q of buy orders when $v = 1$. Similarly, we define μ_0^q. Then the price converges to $\Pr(v = 1|q) = \frac{\lambda \Pr(\delta_1=q)}{\lambda \Pr(\delta_1=q)+(1-\lambda)\Pr(\delta_0=q)} = \frac{\lambda f(\mu_1^q)}{\lambda f(\mu_1^q)+(1-\lambda) f(\mu_0^q)}$. As we show below, as long as the informed traders are not too "informed" then prices may never converge to v.

Proposition 3: *Assume short selling is not allowed. Then, if the precision of the informed traders' signal is sufficiently high, $\phi > \frac{1}{(1+\rho)}$, then the price always converge to v. Otherwise, if (i) $v = 0$; or (ii) $v = 1$ and $\mu \leq \frac{\phi+\phi\rho-1}{1-3\phi+\phi\rho}$; then prices never converge to v. In those cases, the price converges to $\frac{\lambda f(\mu_1^q)}{\lambda f(\mu_1^q)+(1-\lambda) f(\mu_0^q)}$.*

Proof: We start our argument with by analyzing the functions q_1 and q_0. First note that

$$q_1(\mu = 0) = q_0(\mu = 0) = \frac{\frac{\rho}{2}}{1-(1-\rho)} = \frac{1}{2}$$

and

$$q_1(\mu = 1) = \frac{\phi}{1-(1-\rho)(1-\phi)} > q_0(\mu = 1) = \frac{1-\phi}{1-(1-\rho)\phi} > 0$$

Next, we differentiate both q_1 and q_0 with respect to μ.

$$\frac{\partial q_1}{\partial \mu} = \frac{(\phi - \frac{\rho}{2})(1-(1-\rho)[\mu(1-\phi)+(1-\mu)]) - (1-\rho)\phi(\mu\phi + \frac{(1-\mu)\rho}{2})}{(1-(1-\rho)[\mu(1-\phi)+(1-\mu)])^2}$$

Simplifying leads to $\frac{\partial q_1}{\partial \mu} = \frac{\rho(\phi-\rho+\phi\rho)}{2(1-(1-\rho)[\mu(1-\phi)-(1-\mu)])^2} > 0$ since $\phi > \frac{1}{2}$.

This means that $\frac{\partial q_1}{\partial \mu} > 0$ for all $\mu \in [0,1]$ and therefore q_1 is always increasing in μ. Similarly, we can now show that either $\frac{\partial q_0}{\partial \mu} > 0$ for all $\mu \in [0,1]$ or $\frac{\partial q_0}{\partial \mu} \leq 0$ for all $\mu \in [0,1]$. In other words, q_0 is either always increasing or always decreasing in μ.

$$\frac{\partial q_0}{\partial \mu} = \frac{(1-\phi-\frac{\rho}{2})(1-(1-\rho)[\mu\phi+(1-\mu)]) - (1-\rho)(1-\phi)(\mu(1-\phi) + \frac{(1-\mu)\rho}{2})}{(1-(1-\rho)[\mu\phi+(1-\mu)])^2}$$

Simplifying leads to $\frac{\partial q_0}{\partial \mu} = \frac{\rho(1-\phi-\phi\rho)}{2(1-(1-\rho)[\mu\phi-(1-\mu)])^2}$. Note that the sign of the numerator determines the sign of $\frac{\partial q_0}{\partial \mu}$. Furthermore, the numerator does not depend on μ. Therefore, either $\frac{\partial q_0}{\partial \mu} > 0$ for all $\mu \in [0,1]$ or $\frac{\partial q_0}{\partial \mu} \leq 0$ for all $\mu \in [0,1]$. Solving $1-\phi-\phi\rho > 0$, leads to $\phi < \frac{1}{(1+\rho)}$. The condition $\phi < \frac{1}{(1+\rho)}$ implies that $\frac{1-\phi}{1-(1-\rho)\phi} > \frac{1}{2}$. Consequently, if $\phi < \frac{1}{(1+\rho)}$ $\frac{\partial q_0}{\partial \mu}$: whenever market maker observes the fraction of buy orders, q, converging to an element of ($\frac{1}{2}$, $\frac{1-\phi}{1-(1-\rho)\phi}$), he cannot determine the state of the world. Then price converges to $\Pr(v=1|q) = \frac{\lambda f(\mu_1^q)}{\lambda f(\mu_1^q)+(1-\lambda)f(\mu_0^q)}$.

It is immediate that, if $\phi < \frac{1}{(1+\rho)}$ and $v=0$ then prices never converge to v. And similarly if $\phi < \frac{1}{(1+\rho)}$ and $v=1$ then there exists an upper bound on μ such that for lower value of μ the prices will not converge to v. This upper bound is the solution of $\frac{\mu\phi + \frac{(1-\mu)\rho}{2}}{1-(1-\rho)[\mu(1-\phi)+(1-\mu)]} = \frac{1-\phi}{1-(1-\rho)\phi}$. Therefore, the upper bound is $\frac{1-\phi-\phi\rho}{3\phi-\phi\rho-1}$. Thus, if we have $v=1$, $\phi < \frac{1}{(1+\rho)}$ and $\mu \leq \frac{1-\phi-\phi\rho}{3\phi-\phi\rho-1}$ then prices never converge to v. ■

The intuition of this result is the following. The price will not converge when the orders observed could arise both when $v=0$ or $v=1$. In such case, the market makers who observe these orders cannot determine which is the fundamental value. Consider the case where ρ is very small: this is a case in which the condition for convergence to the fundamental value, $\phi > \frac{1}{(1+\rho)}$, is not satisfied. Let us look at the fraction of buy orders when $v=0$ and μ is high: although the fundamental value is low, there are a lot of informed investors (and remember we are in the case in which the informativeness of the signal ϕ is not too high, so many of them will have observed a high signal) and thus the number of buy orders can be quite high.

Consider now the fraction of buy orders when $v = 1$ and μ is low: although the informed investors are buying, there is very few of them (and the noise traders are only buying with probability ρ, which is low) and thus in the end the fraction of buy orders for the case of $v = 1$ and μ low is not necessarily higher than for the case of $v = 0$ and μ high. In other words, it is possible to find values of μ for which the fraction of buy orders is the same both for $v = 0$ or $v = 1$ and the market makers cannot determine whether $v = 0$ or $v = 1$. Let us now look at the fraction of sell orders. A small ρ means that only a small fraction of informed traders own asset 1. Since short selling is banned, only those who own asset 1 can sell it, and there are not many of them. Thus there will be not a lot of difference in the sell orders between the case of $v = 0$ and $v = 1$ and it will be even easier to find values of μ such that the fraction of sell orders will be the same whether $v = 0$ or $v = 1$.

One thing that should come clear from the intuition is the importance of the uncertainty about μ. It is because the market makers do not know μ that they cannot distinguish between orders in the cases of $v = 0$ and $v = 1$. This is exactly the intuition we gave in the introduction: once we introduce uncertainty about how many informed investors are around, it is much more difficult to update just by observing the orders in the market.

Note also that every period the market makers are updating their priors not only about v but also about μ. We show in the following corollary that when the prices do not converge to the fundamentals is also the case in which market makers never learn about μ.

COROLLARY 1: *If the precision of informed traders' signal is sufficiently low, $\phi \leq \frac{1}{(1+\rho)}$, then for any fraction of buy orders $q \leq \frac{1-\phi}{1-(1-\rho)\phi}$ the market makers can never learn μ completely. In particular, the distribution on μ converges to a binary distribution where $\lim_{t\to\infty} \Pr(\mu_1^q) = \frac{\lambda f(\mu_1^q)}{\lambda f(\mu_1^q)+(1-\lambda)f(\mu_0^q)}$ and $\lim_{t\to\infty} \Pr(\mu_0^q) = \frac{(1-\lambda)f(\mu_0^q)}{\lambda f(\mu_1^q)+(1-\lambda)f(\mu_0^q)}$.*

PROOF: For any given fraction of buy orders there could be at most two different fractions of informed traders. In particular, if we observe q fraction of buy orders and $v = 1$ then μ must be μ_1^q. Similarly, if we observe q fraction of buy orders and $v = 0$ then μ must be μ_0^q. Therefore, as the fraction of buy orders converge to q then the posterior beliefs about the fraction of informed traders converge to a binary distribution. ■

One important implication of the Proposition 4 is that in presence of short-sales bans prices can be biased both upwards or downwards, depending on what are the parameters' values. This is easy to see: given that P has to be in the interval $[0, 1]$, when the price does not converge to the fundamental it will be necessarily biased upward if the true value is 0 and downward if the true value is 1. Thus the general intuition that short sales constraints bias prices upwards is not true. The intuition is the following. A short sale ban will not allow people with negative news to trade. However, market makers know that and will adjust the price according to their priors. If they believe there may be more short-sale constrained people than there actually are, they may over-adjust, and the price will be biased downwards. Thus there may be an over- or under-adjustment depending on what their priors are relative to the true values of μ and v (in other words, given different parameters value we will find a different bias). Therefore, the conflicting results described in Section 2, where some papers found an effect of short-sales constraints while other papers do not, are not anymore surprising. Those papers were looking for evidence that short-sales constraints led to overpricing. But in the context of this model, different stocks will have different priors and parameters value, thus some stocks will be overvalued and other stocks will be undervalued. When we look for an effect on the prices for the aggregate of all stocks, there is no reasons why we should necessarily find a statistically significant effect.

A related point is that the likelihood of mispricing due to short sale constraints is higher if the fundamental value is low. This is due to the fact that when $v = 0$ the prices may fail to converge to v for any realization of the amount of informed traders whereas prices converge to the v when $v = 1$ and the fraction of informed traders is sufficiently large.

Finally, note that as the ownership of asset 1 (ρ) increases, it is easier to achieve convergence to the fundamental. This is why as ρ increases it is more likely that people who want to sell asset 1 actually own it, so short sales constraints are less binding.

IV. What is the minimum necessary cost?

So far we have showed that if costs are very low, short sale constraints do not have an effect on the long run prices. If instead costs are very high, this is equivalent to a ban on short sales and it affects the long run rational expectation equilibrium prices. The question thus arises of whether the costs that would lead to the equivalent of a ban are prohibitively high or not.

In this section we assume the cost of short selling C is a generic value which is not infinitesimally small. In such case an informed trader will short sell if and only if the profit from such trade is higher than the cost of short selling. The first immediate conclusion is that the analysis becomes more complex. In fact, the profit from the trade will depend from the bid ask prices the market maker will set, which in turn depend on the trade history up to that point. Therefore the decision whether to trade or not will depend on the history, i.e. it will be path dependent. We thus proceed in the following way. First we show that, for given parameters of the model, there exists a minimum cost such that, independently from the history, an informed trader will never want to short sell (in other words, we are back to the situation of the previous section, i.e. to a situation which is equivalent to a ban on short sales). Second, we show through a calibration that such minimum cost is not particularly high.

We start analyzing the case where costs are small enough so that informed traders can shortsell. As we showed at the beginning, the ask price is $A_t = \Pr(v = 1|H^t, h^t = 1)$ and bid price is $B_t = \Pr(v = 1|H^t, h^t = -1)$. The bid and ask price will depend on the previous history up to period t H^t and the order in period t h^t:

$$\begin{aligned} A_t &= \frac{\Pr(v = 1|H^t)\Pr(v = 1 \wedge h^t = 1)}{\Pr(v = 1|H^t)\Pr(v = 1 \wedge h^t = 1) + (1 - \Pr(v = 1|H^t))\Pr(v = 0 \wedge h^t = 1)} \\ &= \frac{P_{t-1}[\mu\phi + (1-\mu)\frac{1}{2}]}{P_{t-1}[\mu\phi + (1-\mu)\frac{1}{2}] + (1 - P_{t-1})[\mu(1-\phi) + (1-\mu)\frac{1}{2}]} \\ &= \frac{P_{t-1}[\mu\phi + (1-\mu)\frac{1}{2}]}{P_{t-1}\mu(2\phi - 1) + \mu(1-\phi) + (1-\mu)\frac{1}{2}} \end{aligned}$$

$$\begin{aligned} B_t &= \frac{\Pr(v = 1|H^t)\Pr(v = 1 \wedge h^t = -1)}{\Pr(v = 1|H^t)\Pr(v = 1 \wedge h^1 = -1) + (1 - \Pr(v = 1|H^t))\Pr(v = 0 \wedge h^t = -1)} \\ &= \frac{P_{t-1}[\mu(1-\phi) + (1-\mu)\frac{1}{2}]}{P_{t-1}[\mu(1-\phi) + (1-\mu)\frac{1}{2}] + (1 - P_{t-1})[\mu\phi + (1-\mu)\frac{1}{2}]} \\ &= \frac{P_{t-1}[\mu(1-\phi) + (1-\mu)\frac{1}{2}]}{P_{t-1}\mu(1 - 2\phi) + \mu\phi + (1-\mu)\frac{1}{2}} \end{aligned}$$

Note that the price P_{t-1} is a sufficient statistics for the history H^t. Also, note that $A_t > B_t$. Next we calculate the expected value of the company, given history up to period t H^t and signal s:

$$E[v|H^t, s=1] = \Pr[v=1|H^t, s=1] = \frac{P_{t-1}\phi}{P_{t-1}\phi + (1-P_{t-1})(1-\phi)} = \frac{P_{t-1}\phi}{P_{t-1}(2\phi-1)+(1-\phi)}$$

and

$$E[v|H^t, s=-1] = \frac{P_{t-1}(1-\phi)}{P_{t-1}(1-\phi)+(1-P_{t-1})\phi} = \frac{P_{t-1}(1-\phi)}{P_{t-1}(1-2\phi)+\phi}$$

Note that $E[v|H^t, s=1] > A_t > B_t > E[v|H^t, s=-1]$. In other words, if trading costs are sufficiently low, the informed trader will want to trade. The expected profit of short selling following a bad signal is

$$B_t - E[v|H^t, s=-1] = \frac{P_{t-1}(\mu(1-\phi)+(1-\mu)\frac{1}{2})}{P_{t-1}\mu(1-2\phi)+\mu\phi+(1-\mu)\frac{1}{2}} - \frac{P_{t-1}(1-\phi)}{P_{t-1}(1-2\phi)+\phi}$$

One can verify that $B_t - E[v|H^t, s=-1]$ converges to zero as long as P_{t-1} goes to either 0 or 1. Furthermore, note that $B_t - E[v|H^t, s=-1]$ is a concave function. Therefore, the expected profit due to short selling converges to zero if price converges to v and has a maximum possible value given by Π^*. This has two important implications. First, no matter what the cost is (as long as it is not infinitesimal) it will constrain short sales as the price approaches the true value. Second, if the cost of short selling C is higher than the value Π^* then the informed trader will never trade and we are back in the same situation as in the previous section. In what follows, we show in a calibration that the value Π^* is actually not too high.

IV.1. Numerical solutions for $\mu = \frac{1}{2}$ and $\phi = \frac{2}{3}$. Recall that expected profit of short selling following a bad signal, $B_t - E[v|H^t, s=-1]$, is given by

$$\frac{P_{t-1}(\mu(1-\phi)+(1-\mu)\frac{1}{2})}{P_{t-1}\mu(1-2\phi)+\mu\phi+(1-\mu)\frac{1}{2}} - \frac{P_{t-1}(1-\phi)}{P_{t-1}(1-2\phi)+\phi}.$$

Using $\mu = \frac{1}{2}$ and $\phi = \frac{2}{3}$, and simplifying terms leads to $\frac{3P_{t-1}-3P_{t-1}^2}{2P_{t-1}^2-11P_{t-1}+14}$. Figure 1 shows the expected profit of short selling as a function of previous peroid's price, P_{t-1}. Note that Π^* is less than 0.09 for $\mu = \frac{1}{2}$ and $\phi = \frac{2}{3}$.

V. Multiple Values

Our main result in the previous section is that short selling constraints may prevent prices from converging to the fundamental value in spite of having arbitrarily many informed traders. However, even in such cases, the price converges to a level that is characterized by the parameters of the model with no room for a stochastic outcome. To see this, first note that the price is calculated by market makers who use prior beliefs and figure out likelihood of $v = 1$. In the learning process, market makers update simultaneously about v and μ having observed a sequence of buy-and-sell orders. However, in the long run what matters is not the specific sequence of buy-and-sell orders. Instead, in the long run what matter is the ratio between buy and sell orders. And this ratio is determined by the fundamentals and the distribution on information technology, type of traders and liquidity trading. In this sense, if an economist knows both realization of fundamentals, v and μ, and the prior beliefs, she can perfectly forecast the long run equilibrium price. In other words, the long run equilibrium price is not path dependent.

A natural question is what happens if the fundamental value can take more than two values, say $v \in \{0, 1, 2\}$. This adds an additional complication to our analysis. In particular, the likelihood of having a buy or sell order depends on the price. To see this consider a fully symmetric information structure, i.e. the signal precision is constant across v and $E[v|s = 2] > 1 = E[v|s = 1] > E[v|s = 0]$. When the price is between 0 and 1, an informed trader will buy as long as his signal is $s = 1$ or $s = 2$ whereas an informed trader with signal $s = 0$ will sell. On the other hand, when the price is between 1 and 2, an informed trader will buy only if he has $s = 2$ signal. So a buy order is more likely to come from an informed trader when the price is less than 1 with respect to the case when the price is greater than 1. Consequently, the learning process is not identical for the market makers across these two intervals.

With this richer fundamental value space, our main result that short selling constraints may prevent prices from converging to the fundamental value still holds. More strikingly, the long run equilibrium price depends on the sequence of buy-and-sell orders. For example, when there are many sell orders at the beginning of trading, the price may move into the $(0, 1)$ interval and remain in that interval so that the long run equilibrium may converge to a level in this interval independently of v. However, when there are several buy orders at the beginning of trading, the price may move to the $(1, 2)$ interval leading to a different learning process for the market makers and leading to a different convergence. Therefore, long run equilibrium prices can be stochastic and in particular depend on the realization of orders in the short run.

In the remainder of this section we will formally prove these observations. Given that our main objective is to show path dependence, we will focus on a simple information structure and identify two different sequences of orders that lead to different long run equilibrium prices. In particular, we will identify two extreme paths that will simplify our arguments. However, before we proceed we formally extend our model to the multiple values setting.

Each of the two assets has a fundamental value $v \in \{0, 1, 2\}$, where $\Pr(v = k) = \lambda_k$. For simplicity, we impose uniform priors, i.e., $\lambda_i = \frac{1}{3}$.[8] Each informed trader, i, receives an independent signal (conditional on v) $s_i \in \{0, 1, 2\}$; $\Pr(s_i = k|v = k) = \phi \in (\frac{1}{2}, 1)$, and $\Pr(s_i = k|v \neq k) = \frac{1-\phi}{2}$.[9] Consequently, the posterior belief of a trader with signal k is $\beta(v = k|s_i = k) = \phi$ and $\beta(v \neq k|s_i = k) = \frac{1-\phi}{2}$ for all k. All other features of the model are identical to those in Subsection 3.2. We will now derive a series of results which are immediate and thus proofs will be omitted.

Lemma 2: *The first period bid price will be below* 1 *and the ask price will be above* 1*. In the first period an informed trader with a signal* $s_i = 2$ *will buy and an informed trader with a signal* $s_i = 0$ *will sell while an informed trader with signal* $s_i = 1$ *will not trade.*

The above observation is immediate from the facts that (i) the expected value given prior beliefs is 1; and (ii) the expected value conditional on signal $s_i = 0$ ($s_i = 2$) is smaller (greater) than 1.

The next result follows from the fact that the expected value can never increase after observing a sell order.

Lemma 3: *There exists* t' *such that for all* $\hat{t} > t'$*, if* $h^t = -1$ *for all* $t \leq \hat{t}$ *then both* $A_{\hat{t}+1}$ *and* $B_{\hat{t}+1}$ *will be in the* $(0, 1)$ *interval.*

Similarly, we have the symmetric result as well:

[8] The assumption of uniform priors will keep the notation as lean as possible.

[9] This specific information structure simplifies the terms and yet is sufficient to exhibit path dependence.

Lemma 4: *There exists t' such that for all $\hat{t} > t'$, if $h^t = 1$ for all $t \leq \hat{t}$ then both $A_{\hat{t}+1}$ and $B_{\hat{t}+1}$ will be in the $(1,2)$ interval.*

In the Appendix, we show that in the absence of any short selling constraint, also in this extended model the price converges to the fundamental value.

In the next proposition, we extend our main result to this setting, i.e., prices may never converge to the fundamental value when there are significant short sale constraints. We use the two paths highlighted in Lemmas 7 and 8 to show two cases where prices may not converge to the fundamental values.

Proposition 4: *If f is an increasing function and the precision of the informed traders' signal is sufficiently low, $\phi < \frac{1}{(1+\rho)}$, then there exists a sequence of buy and sell orders such that price converges to $\frac{3f(\mu_2^q)}{f(\mu_0^q)+2f(\mu_2^q)} < 1$ independent of v in one sequence. Similarly, if f is a decreasing function and the precision of the informed traders' signal is sufficiently low $\phi < \frac{1-\rho}{(1+\rho)}$, then there exists a sequence of buy and sell orders such that price converges to $\frac{f(\mu_0^q)+2f(\mu_2^q)}{2f(\mu_0^q)+f(\mu_2^q)} > 1$ independent of v.*

Proof: To prove the proposition we just need to find two sequences such that the two statements in the proposition are true. Thus, we start by considering a sequence beginning with sufficiently many sell orders. Using Lemma 7, this implies that transaction prices will initially be between 0 and 1. We assume that the transaction prices will always remain between 0 and 1 and then characterize the sufficient conditions for it to hold.

If the price remains between 0 and 1, all the informed investors who observed a signal that the value of the shares is 2 or 1 will want to buy, and those who observed a signal that the value of the shares is 0 will want to sell. Since the probability of observing these signals depends on the value of the fundamental, we need to consider each value of the fundamental as a separate case. Let us first consider the $v = 2$ case. A fraction $\mu(1-\rho)(\frac{1-\phi}{2})$ of traders are informed traders who received a signal that the value of the shares is 0 and would like to sell asset 1 but cannot sell it since they do not own a share. Moreover, from the proof of Proposition 2, we know that fraction $(1-\mu)(1-\rho)$ of traders are liquidity traders who will not trade asset 1. Therefore, $1-\mu(1-\rho)(\frac{1-\phi}{2})-(1-\mu)(1-\rho) = 1-(1-\rho)[\mu(\frac{1-\phi}{2})+(1-\mu)]$ is the fraction of traders who would like to trade asset 1. Similarly, the fraction of traders that would like to trade when $v = 1$ is $1-(1-\rho)[\mu(\frac{1-\phi}{2})+(1-\mu)]$. Finally, the fraction of traders that would like to trade when $v = 0$ is $1-(1-\rho)[\mu\phi+(1-\mu)]$.

If $v = k$ for $k \in \{1,2\}$, then the fraction of buys (out of all trades) q_k converges to $\frac{\mu(\frac{1+\phi}{2})+\frac{(1-\mu)\rho}{2}}{1-(1-\rho)[\mu(\frac{1-\phi}{2})+(1-\mu)]}$. Similarly, if $v = 0$ then the fraction of buys (out of all trades) q_0 converges to $\frac{\mu(1-\phi)+\frac{(1-\mu)\rho}{2}}{1-(1-\rho)[\mu\phi+(1-\mu)]}$. Note that q_0 is exactly the same function q_0 as we used in the proof of Proposition 4 (while the function q_k is different).

First note that for $k \in \{1,2\}$

$$q_k(\mu=0) = q_0(\mu=0) = \frac{\frac{\rho}{2}}{1-(1-\rho)} = \frac{1}{2}$$

and

$$q_k(\mu=1) = \frac{\frac{1+\phi}{2}}{1-(1-\rho)(\frac{1+\phi}{2})} > q_0(\mu=1) = \frac{1-\phi}{1-(1-\rho)\phi} > 0$$

(since $\phi > \frac{1}{2}$).

Next, we differentiate q_k with respect to μ and first show that $\frac{\partial q_k}{\partial \mu} > 0$ for all $\mu \in [0,1]$.

$$\frac{\partial q_2}{\partial \mu} = \frac{(\frac{1+\phi-\rho}{2})(1-(1-\rho)(\mu(\frac{1-\phi}{2})+(1-\mu))) - (1-\rho)(\frac{1+\phi}{2})(\mu(\frac{1+\phi}{2})+\frac{(1-\mu)\rho}{2})}{(1-(1-\rho)(\mu(\frac{1-\phi}{2})+(1-\mu)))^2}$$

Note that the denominator is always positive and simplifying the numerator leads to $\frac{1}{4}\rho\,(\phi-\rho+\phi\rho+1) > 0$ implying that $\frac{\partial q_k}{\partial \mu} > 0$. Recall that we already know from the proof of Proposition 4 that if $\phi < \frac{1}{(1+\rho)}$, then $\frac{\partial q_0}{\partial \mu} > 0$ for all $\mu \in [0,1]$. Consequently, if $\phi < \frac{1}{(1+\rho)}$ whenever market maker observes the fraction of buy orders, q, converging to an element of $(\frac{1}{2}, \frac{1-\phi}{1-(1-\rho)\phi})$, he cannot determine the state of the world.

Now that we showed that in this case the price does not converge to the fundamental value, we want to show to what price it will converge. Let μ_k^q stand for the solution of $\frac{\mu(\frac{1+\phi}{2})+\frac{(1-\mu)\rho}{2}}{1-(1-\rho)[\mu(\frac{1-\phi}{2})+(1-\mu)]} = q$, where $k \in \{1,2\}$. Therefore, μ_k^q is the fraction of traders who are informed that would lead to a fraction q of buy orders when $v \in \{1,2\}$. Similarly, we determine μ_0^q. Then the posterior beliefs converge to $\Pr(v = k|q) = \frac{f(\mu_k^q)}{f(\mu_0^q)+f(\mu_1^q)+f(\mu_2^q)}$. Therefore, the price converges to the expected value of v, conditional on observing q, which is equal to $\frac{f(\mu_1^q)+2f(\mu_2^q)}{f(\mu_0^q)+f(\mu_1^q)+f(\mu_2^q)}$. However, recall that for our arguments to hold the price has to be in the $(0,1)$ interval. Therefore, we must have $\frac{f(\mu_1^q)+2f(\mu_2^q)}{f(\mu_0^q)+f(\mu_1^q)+f(\mu_2^q)} < 1$ for not having price converge to

v. Note that $f(\mu_1^q) = f(\mu_2^q)$ since $\mu_1^q = \mu_2^q$. Therefore, last inequality becomes $\frac{3f(\mu_2^q)}{f(\mu_0^q)+2f(\mu_2^q)} < 1$, which can be rewritten as $f(\mu_2^q) < f(\mu_0^q)$. Given that $\mu_2^q \leq \mu_0^q$, $f(\mu_2^q) < f(\mu_0^q)$ is satisfied by f that is an increasing function. Thus, as long as f is an increasing function we must have $\frac{3f(\mu_2^q)}{f(\mu_0^q)+2f(\mu_2^q)} < 1$.

It is immediate that, if $\phi < \frac{1}{(1+\rho)}$ and $v = 0$ then prices never converge to v. And similarly if $\phi < \frac{1}{(1+\rho)}$ and $v = 1$ then there exists an upper bound on μ such that for lower value of μ the prices will not converge to v. This upper bound is the solution of $\frac{\mu(\frac{1+\phi}{2})+\frac{(1-\mu)\rho}{2}}{1-(1-\rho)[\mu(\frac{1-\phi}{2})+(1-\mu)]} = \frac{1-\phi}{1-(1-\rho)\phi}$. Therefore, the upper bound is $\frac{1-\phi-\phi\rho}{\phi(2-\rho)}$. Thus, if we have $v \in \{1, 2\}$, $\phi < \frac{1}{(1+\rho)}$ and $\mu \leq \frac{1-\phi-\phi\rho}{\phi(2-\rho)}$ then prices never converge to v.

We now want to prove the second part of the proposition, so we choose a different sequence, where the price will again not converge to the fundamental, but will converge to a price which is different from the case we just analyzed. Therefore, we consider a sequence in which there are sufficiently many buy orders initially. So we first suppose that prices remain always between 1 and 2 and then characterize the sufficient conditions for it to hold.

Let us first consider the $v = 2$ case. A fraction $\mu(1-\rho)(1-\phi)$ of traders are informed traders who received a signal $s_i \in \{0, 1\}$ and would like to sell asset 1 but cannot sell it since they do not own a share. Moreover, from the previous subsection we know that fraction $(1-\mu)(1-\rho)$ of traders are liquidity traders who will not trade asset 1. Therefore, $1-\mu(1-\rho)(1-\phi)-(1-\mu)(1-\rho) = 1-(1-\rho)[\mu(1-\phi)+(1-\mu)]$ is the fraction of traders who would like to trade asset 1. On the other hand, when $v = 1$, the fraction of traders that receive a signal $s_i \in \{0, 1\}$ and would like to sell asset 1 but cannot sell it since they do not own a share is $\mu(1-\rho)(\frac{1+\phi}{2})$. Therefore, $1-(1-\rho)[\mu(\frac{1+\phi}{2})+(1-\mu)]$ is the fraction of traders who would like to trade asset 1. Similarly, the fraction of traders that would like to trade when $v = 0$ is $1-(1-\rho)[\mu(\frac{1+\phi}{2})+(1-\mu)]$. If $v = k$ for $k \in \{0, 1\}$, then the fraction of buys (out of all trades) q_k converges to $\frac{\mu(\frac{1-\phi}{2})+\frac{(1-\mu)\rho}{2}}{1-(1-\rho)[\mu(\frac{1+\phi}{2})+(1-\mu)]}$. Similarly, if $v = 2$ then the fraction of buys (out of all trades) q_2 converges to $\frac{\mu\phi+\frac{(1-\mu)\rho}{2}}{1-(1-\rho)[\mu(1-\phi)+(1-\mu)]}$. Note that q_2 in this case is exactly the same function as q_1 in Proposition 4 and therefore we know from the proof of Proposition 4 that $\frac{\partial q_2}{\partial \mu} > 0$ for all $\mu \in [0, 1]$.

Note that for $k \in \{0, 1\}$

$$q_2(\mu = 0) = q_k(\mu = 0) = \frac{\frac{\rho}{2}}{1-(1-\rho)} = \frac{1}{2}$$

and

$$q_2(\mu=1)=\frac{\phi}{1-(1-\rho)(1-\phi)}>q_k(\mu=1)=\frac{\frac{1-\phi}{2}}{1-(1-\rho)(\frac{1+\phi}{2})}>0$$

Next, we differentiate q_k with respect to μ.

$$\frac{\partial q_k}{\partial \mu}=\frac{(\frac{1-\phi-\rho}{2})(1-(1-\rho)(\mu(\frac{1+\phi}{2})+(1-\mu)))-(1-\rho)(\frac{1-\phi}{2})(\mu(\frac{1-\phi}{2})+\frac{(1-\mu)\rho}{2})}{(1-(1-\rho)(\mu(\frac{1+\phi}{2})+(1-\mu)))^2}$$

Note that the denominator is always positive and simplifying the numerator leads to $-\frac{1}{4}\rho\left(\phi+\rho+\phi\rho-1\right)$ Therefore, if $\phi<\frac{1-\rho}{1+\rho}$, then $\frac{\partial q_k}{\partial \mu}>0$ for all $\mu\in[0,1]$. Consequently, if $\phi<\frac{1-\rho}{1+\rho}$ whenever the market maker observes the fraction of buy orders, q, converging to an element of ($\frac{1}{2}$, $\frac{\frac{1-\phi}{2}}{1-(1-\rho)(\frac{1+\phi}{2})}$), he cannot determine the state of the world.

We now determine the price to which this sequence will converge. Let μ_k^q stand for the solution of $\frac{\mu(\frac{1-\phi}{2})+\frac{(1-\mu)\rho}{2}}{1-(1-\rho)[\mu(\frac{1+\phi}{2})+(1-\mu)]}=q$, where $k\in\{0,1\}$. Therefore, μ_k^q is the fraction of traders who are informed that would lead to a fraction q of buy orders when $v\in\{0,1\}$. Similarly, we define μ_2^q. Therefore, price converges to $\frac{f(\mu_1^q)+2f(\mu_2^q)}{f(\mu_0^q)+f(\mu_1^q)+f(\mu_2^q)}$. As long as $\frac{f(\mu_1^q)+2f(\mu_2^q)}{f(\mu_0^q)+f(\mu_1^q)+f(\mu_2^q)}>1$ we may have a possibility of not having price converge to v. Note that $f(\mu_0^q)=f(\mu_1^q)$ since $\mu_0^q=\mu_1^q$. Therefore, last inequality becomes $\frac{f(\mu_0^q)+2f(\mu_2^q)}{2f(\mu_0^q)+f(\mu_2^q)}>1$, which can be rewritten as $f(\mu_2^q)>f(\mu_0^q)$. Given that $\mu_2^q\leq\mu_0^q$, $f(\mu_2^q)>f(\mu_0^q)$ is satisfied by f that is a decreasing function. Then price converges to $\frac{f(\mu_0^q)+2f(\mu_2^q)}{2f(\mu_0^q)+f(\mu_2^q)}$. And as long as f is a decreasing function we must have $\frac{3f(\mu_2^q)}{f(\mu_0^q)+2f(\mu_2^q)}>1$.

It is immediate that, if $\phi<\frac{1}{(1+\rho)}$ and $v\in\{0,1\}$ then prices never converge to v. And similarly if $\phi<\frac{1-\rho}{1+\rho}$ and $v=2$ then there exists an upper bound on μ such that for lower value of μ the prices will not converge to v. This upper bound is the solution of $\frac{\mu\phi+\frac{(1-\mu)\rho}{2}}{1-(1-\rho)(\mu(1-\phi)+(1-\mu))}=\frac{\frac{1-\phi}{2}}{1-(1-\rho)(\frac{1+\phi}{2})}$. Therefore, the upper bound is $\frac{1-\phi-\rho-\phi\rho}{5\phi-\rho(1+\phi)-1}$. Thus, if we have $v=2$, $\phi<\frac{1-\rho}{1+\rho}$ and $\mu\leq\frac{1-\phi-\rho-\phi\rho}{5\phi-\rho(1+\phi)-1}$ then prices never converge to v.

■

Proposition 9 shows that along the paths highlighted in Lemmas 7 and 8 prices may not converge to the fundamental. However, it could be that in both cases the prices will converge to the same price. This would mean that the final convergence price would depend on the initial conditions (such as prior beliefs, liquidity, etcetera), but not on the specific path undertaken. Instead, the following proposition shows that the two paths will achieve convergence to two different prices. In particular, we give an example where the value of the parameters and the prior beliefs are identical, but there exists a sequence such that the price converges to the fundamental values and another sequence where the price converges to a value different from v. In other words, the rational expectations long run equilibrium price depends on the specific path taken, besides all the other parameters.

PROPOSITION 5: *If $\phi \in (\frac{1-\rho}{1+\rho}, \frac{1}{1+\rho})$ and f is an increasing function, then there exists two different sequences of orders such that in one the price will converge to v and in the other the price will never converge to v.*

PROOF: Note that if there are many buy orders initially and yet $\phi > \frac{1-\rho}{1+\rho}$, then market maker can determine whenever $v = 2$. Therefore, when $\phi \in (\frac{1-\rho}{1+\rho}, \frac{1}{1+\rho})$, there exists a sequence of buy and sell orders that starts with many buy orders and lead to long run price converging to v whenever $v = 2$. In contrast, in the proof of Proposition 9 we established that the price never converged to v for $\phi \in (\frac{1-\rho}{1+\rho}, \frac{1}{1+\rho})$ as long as there is an initial sequence of sell orders. ■

VI. Implications for IPOs

As mentioned in the introduction, in the period after an IPO it is difficult to take a short position. We should mention that Edwards and Hanley (2010) find active short selling in the first days following an IPO. However, as Gibbs and Hao (2018) point out, this shows that short sales are not completely impossible, but it is consistent with short sales still be constrained. Most researchers seem to agree that following an IPO and before a lockup expiration short sales are possible but more difficult, and Patatoukas, Sloan and Wang (2018) present evidence that short sales are constrained after an IPO. In such case, the model in this paper has important implications for IPOs. In this Section we discuss such implications. A model that develops the specific issues of IPOs is beyond the scope of this paper.

The first implication is that the price set by the underwriter at the IPO has an effect on the long run value of the shares. There is a large literature on the IPO pricing and why the price may be set strategically. But most of the literature focuses on how the underwriter may set the price in order to influence the share price in the first days of trading. Especially in a rational investors setting, the initial IPO price could not influence the long run price, when all information is revealed and the price is equal to the fundamental value. This paper instead suggests that the IPO price could influence the long run price in two ways. First, if the initial

IPO price influences the prior of investors on what the price is, this paper shows how the price will converge to a value determined by the initial priors of the investors. Second, even if the underwriter cannot influence the investors expectations, the version of our model where the fundamental can take more than 2 values shows that the rational expectation price will be path dependent and thus will depend (among other things) on the initial conditions.[10]

Another important implication is about the effect of price stabilization. It is well known that shortly after the IPO the underwriter may engage in price stabilization. However, price stabilization can only take place for a short period after the IPO. This paper shows that stabilizing the price even for a short period may have permanent effects on the long run price. In fact, we showed in the previous section that learning will take place differently depending on the area where the price is. By doing price stabilization, the underwriter can make sure that the price stays in the area which facilitate the most positive learning.

The asymmetry highlighted above can also explains some of the abrupt changes in the path we sometimes observe in IPOs (see, for example, Aggarwal, Krigman, and Womack 2002). If the price keeps changing through the learning and suddenly moves from one area to another, then the path can change in an abrupt way, because learning is taking place differently.

VII. Conclusion

We have shown that if we introduce uncertainty about the number of informed investors in the Diamond and Verrecchia model, convergence to the fundamental is not achieved any more. Moreover, as we increase the number of values the underlying asset can have, we show that the rational expectation equilibrium price is path-dependent.

These results can help in the present debate regarding the optimality of imposing constraints on short-sales constraints. Note also that the empirical evidence shows that short-sales constraints can be relaxed by the use of options or CDS. Given that options and derivatives are costly, this would be equivalent to moving, for example, from a situation in which short-sales are banned, to one in which they are possible but costly (which is obviously a milder constraint).

---

[10] One should keep however in mind that such implications will be valid only until there are some constraints on short-sales.

VIII. Appendix

In this section, we focus on the model extension considered in Section 4 and show that an analogue of Proposition 2 holds for the multiple value setting. More specifically, if there are no short selling constraints, then prices converge to v, the fundamental value, where $v \in \{0, 1, 2\}$.

PROPOSITION 6: *In the absence of short selling constraint the price converges to v as $t \to \infty$.*

PROOF: Let us suppose this is not true and assume the long-run price is in the $(1, 2)$ interval. As before, the fraction of traders who would like to trade asset 1 is $1 - (1 - \mu)(1 - \rho)$ when there are no short selling constraints. Consequently, for a given value of v the fraction of buy orders will converge to

$$
\begin{aligned}
q_2 &= \frac{\mu\phi + \frac{(1-\mu)\rho}{2}}{1 - (1-\mu)(1-\rho)} \\
q_k &= \frac{\mu\frac{1-\phi}{2} + \frac{(1-\mu)\rho}{2}}{1 - (1-\mu)(1-\rho)}, \text{ where } k \in \{0, 1\}
\end{aligned}
$$

As in Proposition 1 it is easy to show that $q_2 > \frac{1}{2} > q_k$ for all $\mu > 0$ where $k \in \{0, 1\}$. Therefore, if the fraction of buy orders is greater than $\frac{1}{2}$, then the price converges to 2. If the fraction of buy orders is less than $\frac{1}{2}$, then the price cannot remain in the $(1, 2)$ interval as the expected value is less than 1.

Next, consider the case where long-run price level in the $(0, 1)$ interval. In this case, the fraction of buy orders will converge to

$$
\begin{aligned}
q_k &= \frac{\mu\frac{1+\phi}{2} + \frac{(1-\mu)\rho}{2}}{1 - (1-\mu)(1-\rho)}, \text{ where } k \in \{1, 2\} \\
q_0 &= \frac{\mu\frac{1-\phi}{2} + \frac{(1-\mu)\rho}{2}}{1 - (1-\mu)(1-\rho)}
\end{aligned}
$$

Similar to the previous case, we have $q_k > \frac{1}{2} > q_0$ for all $\mu > 0$ where $k \in \{1, 2\}$. Therefore, if the fraction of buy orders is less than $\frac{1}{2}$, then the price converges to 0. If the fraction of buy orders is greater than $\frac{1}{2}$, then the price cannot remain in the $(0, 1)$ interval as the expected value is greater than 1.

The only possibility left for prices not to converge to the fundamental value is the case where the price converges to 1 although $v \neq 1$. From the arguments above, we know that price cannot converge to 1 from above when $v = 2$, since $v = 2$ leads to convergence to 2 in the $(1, 2)$ interval . Convergence to 1 from below is not possible either since price overshoots 1. Similarly, price cannot converge to 1 when $v = 0$. The price oscillating around 1 is not possible either given that $q_2 > \frac{1}{2} > q_0$ always holds. ■

References

AGGARWAL, R. K., L. KRIGMAN, AND K. L. WOMACK , 2002, "Strategic IPO underpricing, information momentum, and lockup expiration selling," *Journal of Financial Economics*, 66, 105-137.

ALLEN, F., S. MORRIS AND A. POSTLEWAITE , 1993, "Finite Bubbles with Short Sale Constraints and Asymmetric Information", *Journal of Economic Theory*, 61, pages 206-229

AVERY, C. AND P. ZEMSKY , 1998, Multi-Dimensional Uncertainty and Herd Behavior in Financial Markets, *American Economic Review*, 88, pages 724748.

BAI, Y., E. C. CHANG AND J. WANG , 2006, "Asset Prices Under Short-Sale Constraints", mimeo, MIT.

BEBER, A., M. PAGANO , 2013, "Short-Selling Bans Around the World: Evidence from the 2007-09 Crisis," *Journal of Finance*, 68, pages 343-381.

BIKHCHANDANI, S., D. HIRSHLEIFER AND I. WELCH , 1992, "A Theory of Fads, Fashion, Custom and Cultural Change as Informational Cascades," *Journal of Political Economy*, 100, pages 992-1027.

BOEHMER, E., C.M. JONES AND X. ZHANG , 2008a, "Which Shorts are Informed?," *Journal of Finance*, 63, pages 491-527.

BOEHMER, E., C.M. JONES AND X. ZHANG , 2008b, "Unshackling Short Sellers: The Repeal of the Uptick Rule," *mimeo*, Columbia University.

BOEHMER, E., C.M. JONES AND X. ZHANG , 2013, "Shackling Short Sellers: The 2008 Shorting Ban," *Review of Financial Studies*, 26, pages 1363-1400.

BOEHMER, E., AND J. WU , 2012, "Short Selling and the Price Discovery Process," *Review of Financial Studies*, 26, pages 287-322.

BRIS, A., W.N. GOETZMANN AND N. ZHU , 2007, "Efficiency and the Bear: Shorts Sales and Markets Around the World," *Journal of Finance*, 62, pages 1029-1079.

COHEN, L., K.B. DIETHER AND C.J. MALLOY , 2007, "Supply and Demand Shifts in the Shorting Market," *Journal of Finance*, 62, pages 2061-2096.

item[CAMPELLO, M., R. MATTA AND P.A.C. SAFFI], 2018, "THE RISE OF THE EQUITY LENDING MARKET: IMPLICATIONS FOR CORPORATE POLICIES," *mimeo*, CORNELL UNIVERSITY.

DIAMOND, W. D. AND R. E. VERRECCHIA , 1987, "CONSTRAINTS ON SHORT-SELLING AND ASSET PRICE ADJUSTMENT TO PRIVATE INFORMATION", *Journal of Financial Economics* 18(2), PAGES 273-311.

Diether, K.B., K. Lee and I.M. Werner , 2009, "Short-Sale Strategies and Return Predictability," *Review of Financial Studies* 22(2), pages 575-607.

Duffie, D., N. Garleanu and L.H. Pedersen , 2002, "Securities Lending, shorting, and pricing," *Journal of Financial Economics* 66, pages 307-339.

Edwards, A.K. and K.W. Hanley , 2010, "Short Selling in Initial Public Offerings," *Journal of Financial Economics* 98, pages 21-39.

Gibbs, M. and Q. Hao , 2018, "Short Selling around the Expiration of IPO Share Lockup," *Journal of Banking and Finance* 88, pages 30-43.

Glosten, L. and P. Milgrom , 1985, "Bid, ask and transaction prices in a specialist market with heterogenously informed traders", *Journal of Financial Economics* 14, pages 71-100.

Harrison, M.J. and D.M. Kreps , 1978, Speculative Investors Behavior in a Stock Market with Heterogeneous Expectations," *Quarterly Journal of Economics* 92, 323-336.

Hong, H. and J. Stein , 2003, "Differences of Opinion, Short-Sales Constraints and Market Crashes," *Review of Financial Studies*, 16, pages 487-525.

Jacklin, C.J., A.W. Kleidon and P. Pfleiderer , 1992, "Underestimation of Portfolio Insurance and the Crash of October 1987," *Review of Financial Studies*, 5, pages 35-63.

Kaplan, S.N., T.J. Moskowitz and B.A. Sensoy , 2012, "The Effects of Stock Lending on Security Prices: An Experiment," *Journal of Finance*, forthcoming.

Karpoff, J.M. and X. Lou , 2010, "Short Sellers and Financial Misconduct," *Journal of Finance*, 65, pages 1151-1168.

Miller, E. , 1977, "Risk, Uncertainty and Divergence of Opinion", *Journal of Finance*, 32, pages 1879-1913.

Nagel, S. , 2005, "Short sales, institutional investors and the cross-section of stock returns," *Journal of Financial Economics* 78, pages 277-309.

Patatoukas, P.N., R. Sloan and A. Wang , 2018, "Short-Sales Constraints and Aftermarket IPO Pricing" *mimeo*, U.C. Berkeley.

Saffi, P. A.C. and K. Sigurdsson , 2011, "Price Efficiency and Short Selling" *Review of Financial Studies*, 24, pages 821-852.