ARBITRAGE AND SHORT SELLING: A POLITICAL ECONOMY APPROACH

Paola Robotti

CSGR Working Paper 205/06

April 2006
Arbitrage and Short Selling: A Political Economy Approach*
Paola Robotti
CSGR Working Paper 205/06
April 2006

Abstract

Arbitrage and short selling are investment strategies that epitomise a particular mode of investment: smart, sophisticated, and efficiency-enhancing (as described within mainstream finance). Yet many market events of recent financial history point to a less benign face of this practice. Short selling in particular can be associated with manipulative strategies that do not add much to the efficiency of the market but rather hinder the interests of various categories of market actors, notably long-term investors and corporate actors. Despite pointing to important aspects of the capitalist relationships created by contemporary financial markets, these topics have gone unnoticed in international political economy. This article shows that a political-economy approach to arbitrage would provide important insights into the role that financial theories play in the formulation of regulatory decisions and into the impact that these decisions have on different classes of market actors.

Keywords: arbitrage; short selling; efficient market hypothesis; financial market regulation; Regulation SHO

Contact details:

Paola Robotti
University of California San Diego (UCSD)
International Relations/Pacific Studies
Email. probotti@ucsd.edu

* Financial support from the Marie Curie Programme of the European Commission is gratefully acknowledged. I thank James Angel, John Finnerty, Robert O’Brien and the Division of Market Regulation of the Securities and Exchange Commission for sharing with me their thoughts on Reg. SHO. I also thank Miles Kahler for his support and comments. All errors are my sole responsibility.
Introduction

In Issue 32(3) of *Economy and Society* Donald MacKenzie explores the sociology of arbitrage. His work touches on one of the most important principles of modern finance and a milestone of contemporary investing. MacKenzie’s idea is that a sociological approach to arbitrage helps unveil the contradictions in the argument that markets can remain efficient, intended as isolated from the turbulences of the social world.\(^1\) While becoming an increasingly familiar topic and the symbol of a particular mode of investing, arbitrage has remained neglected in political science and the subfield of international political economy. This article asks whether there is any rationale for a political economy of arbitrage and what such an approach would look like.

According to a textbook definition, arbitrage is the act of discerning and profiting from current price discrepancies that the market does not incorporate into prices. Such an opportunity arises, for instance, when the same security trade at two different prices in two different markets. An arbitrageur would buy the cheaper security and sell short the more expensive one. Short-selling means selling without the need of owning the shares to be sold short. Shares will be borrowed later on to be delivered. A classic example would be that of shares of General Motors (GM) selling at $54 per share on the New York Stock Exchange and at $56 on the London Stock exchange: arbitrageurs would sell short shares in London and buy them in New York, as a result making the stock prices in the two locations converge. By so doing, they would ‘earn sure profit without investing a penny of [their] own money.’\(^2\) This is why arbitrage is said to involve no risk and to earn a guaranteed profit.
In day-to-day market operations, the word ‘arbitrage’ is associated with very diverse investment strategies that might be only slightly related to the textbook example. There are a myriad of funds performing arbitrage strategies in securities, commodities, and bonds markets. The majority of these funds build long positions in assets that are expected to grow in value and build short positions in assets that are expected to lose value. For instance, at the end of 2004 Marks & Spencer was said to have had a difficult take-over due to the operation of risk arbitrageurs betting on the success of the deal. In this case, risk arbitrageurs performed a ‘merger arbitrage’: they bought the shares of the acquired company whose price was supposed to rise as a consequence of the deal – and which was in this sense ‘undervalued’ – and sold short the shares of the acquiring company whose price was supposed to fall and which was for the same logic ‘overvalued’. Yet another arbitrage strategy is the one performed by ‘vulture funds’, so called because they buy the debt of distressed companies or countries perhaps at half of its nominal value and sell it back later on at or above its full price. The fund that in 1996 paid $11 million to buy $20 million of Peru’s sovereign debt and that later on sued the Peruvian government for full payment plus capitalised interest was a vulture fund. Considering the attention devoted to arbitrage in investor websites, online pundits and manuals, the practice is on the increase.

Arbitrage is also a core principle of neoclassical finance since it is essential to its assumption of market efficiency. Since arbitrageurs detect and close price discrepancies, they are said to drive prices back to their equilibrium level and markets back to efficiency, even when the mass of investors is trading in an irrational way. Given that arbitrageurs are rational and compete to discover arbitrage opportunities, these opportunities are fast disappearing so that prices quickly converge to their fundamental value.
Several non-mainstream finance approaches have questioned the theory of arbitrage. Their argument is that arbitrage is unlikely to bring markets back to efficiency because investors – even the sophisticated ones or arbitrageurs – do not behave according to the model of neoclassical finance. They either herd and imitate or find it more convenient to join rather than contrast other investors’ irrationality. These approaches have been concerned with disproving the rationality assumption of mainstream finance – especially as formulated in the theory of arbitrage – and showing other sociology- and psychology-based models of investor behaviour.

This article does not focus on the behaviour and rationality of arbitrageurs. Drawing on a definition of political economy as the discipline that studies the interactions between political and economic processes, the article analyzes the regulatory treatment of arbitrage and its impact on different societal interests. These two aspects are interrelated: market practices are the result of regulatory decisions and non-decisions. This regulatory structure in turn hinders certain interests and benefits others. The theory of arbitrage is still very central since regulators have more than often taken on the efficiency argument to justify their decisions. In other words, while in day-to-day operations arbitrage might not work according to the theory – using MacKenzie’s words, it has a low performativity ratio – its ability to influence regulatory decisions is still very high. This has consequences for the outcome of regulation. In particular, it can favour certain categories of investors (e.g. short sellers, market makers) at the expenses of others (e.g. small corporate, long term investors).

A caveat is necessary. This article starts with a particular definition of arbitrage. It draws on Fama’s seminal work on the theory of arbitrage, which defines arbitrageurs by their capacity to sell short, and makes short selling the focus of the analysis. In “The Behaviour
of Stock-Market Prices’, published in the *Journal of Business* in 1965, Fama contends that the role of arbitrageurs or sophisticated traders is that of offsetting a noise generated process (e.g. a price bubble) by acting as market contrarians and that selling short is the main instrument of this offsetting strategy. Short selling, in other words, is attributed the main efficiency-enhancing role in opposing the exuberant expectations of the buying side of the market. This argument is dominant among regulators, investors and the press. Examples can be found in the commentaries following the bursting of the Internet bubble in March 2000, as section four will show.

In addition, amidst the differences in their investment strategies, arbitrageurs are all characterised by the (legal) capacity to sell short. While everyone can buy, not everyone can sell short. In many countries an investor protection rationale forbids from selling short those institutional investors that collect money from the broader public, notably mutual and pension funds. The reason is that the maximum amount one can lose in buying a stock is the price of the stock, while in the case of short selling losses can be theoretically limitless.

Who benefits from this legal capacity? Hedge funds are among the few investors with unrestricted access to short selling. This, coupled with the essence of their trade, which is very close to an arbitrage philosophy, has made of them the typical short sellers. In the late 1990s, many regulatory reforms have relaxed restrictions on short selling for mass investors, while instruments of short selling have become increasingly available to the retail player (e.g. single-stock futures). Yet the equation hedge funds=arbitrageurs is usually maintained. The SEC for instance estimates that most of the short selling is done by market makers for hedge funds’ accounts.
Now that the topic and focus of the article are outlined, it is necessary to determine the interest, from a political economy point of view, of this analysis of arbitrage/short selling. In the last 5 to 10 years short selling has become a controversial topic, not least because it was at the centre of many episodes of market unrest or manipulation (e.g. post-9/11 speculation; naked short-selling abuses). Calls for its regulation or re-regulation often emerged after a particular market crisis or lawsuit and some regulatory initiatives were launched in major financial centres (e.g. US, UK and Japan). Most practitioners and regulators, however, have opposed regulation on the ground that short selling is a tool of efficiency. If on the one hand there might be reasons to address short selling for its (increasingly) manipulative character, on the other hand short selling is seen as crucial to an efficient market. The dilemma for regulators lies in this tension between the need to protect investors and the need to promote efficiency, which regulators find equally constraining. In solving this dilemma, regulators interact with and affect various interest groups: in this lies the political interest of the analysis.

The article analyses some short-selling-related events, which have an international resonance, and focuses more specifically on the case of the US, where a new rule on short selling was introduced in 2004. Since 1999 the Securities and Exchange Commission (SEC) has been involved in a major revision of short sale regulation in the US cash equity markets. A consultation process was launched in which all actors either engaged in shorting the market or at the receiving end of short selling were invited to contribute. This debate is an exceptional laboratory to assess the beliefs, claims and arguments on short selling and its impact in relation to different categories of market actors. Methodologically, the article draws on academic contributions, regulatory documents, and interviews with both practitioners and regulators.
The article is organised in four sections. Sections one to three review the current approaches to the study of arbitrage in neoclassical finance, behavioural finance and sociology respectively. Section four identifies a political economy approach by focusing on the regulatory attitude towards short selling at both the domestic and international level. Section five draws some concluding remarks and outlines the agenda for future research on the political economy of short selling.

1. The mainstream finance approach to arbitrage

Arbitrage is one of the major tenets of modern finance, if not ‘the most fundamental valuation principle in finance’. This section reviews the mainstream financial analysis of arbitrage and shows how arbitrage is a core component of the theory of efficient market and of the consequent idea of fair market. In order to do so, it is essential to first locate this theory within the project of modern finance.

This project started in the 1950s at the University of Chicago and gave rise to modern academic finance. This does not mean that there was no academic interest in finance before that time, but that the work done at the University of Chicago provided ‘a rigorous theory to a large body of empirical results’. The Efficient Market Hypothesis (EMH) and the Theory of Arbitrage were among the most prominent results of that work. Their most comprehensive enunciation is in the writings of Eugene Fama and Milton Friedman.

These scholars were confronted with a problem of resource allocation. Fama writes: ‘The primary role of the capital market is allocation of ownership of the economy’s capital stock.’ Given this problem, they theorised how the market could achieve the optimal
distribution of that ownership. As Fama writes, ‘the ideal is a market in which prices provide accurate signals for resource allocation’, that is, a market in which the price at ‘any time fully reflects all available information’. Fama called such a market ‘efficient’. Moreover, because in this situation market signals appear in the same way to each individual, an efficient market also becomes a ‘fair’ one. Hence the allocation of the economy’s capital stock is optimal. Fama’s definition of efficiency differs from any common-sense understanding or previous usage of the word. In particular, it is the first time that the adjective efficient is used with reference to markets and not to individual portfolios and that becomes synonymous with optimality in the allocation of societal resources.

But which are the conditions that make markets efficient according to Fama? In order for markets to be efficient, investors need to be rational, which in the framework of the theory means that they look at information concerning fundamental values of financial assets in making their investment decisions. The fundamental value of an asset is determined by what the asset will earn in the future. As much as efficiency, rationality is given a specific meaning – that of looking at fundamental values – which cannot be reduced to its common-sense usage.

Why is rationality crucial for market efficiency? If investors are rational, they will compete to discover all fundamental information on financial assets and by so doing they will help information to become immediately incorporated in asset prices. This in turn means that prices will only change in response to new information, which is by definition unpredictable. If this is the case, price changes are unpredictable too. If prices are unpredictable, no one has privileged information or a viewpoint from which to make investment decisions. This conveys the idea that the market is not only efficient, but also
fair. It is an extremely powerful conclusion given that Fama’s problem was the ‘allocation of ownership of the economy’s capital stock’.  

More controversial was the proof of whether the theory stood the test of empirical evidence. Soon after the publication of Fama’s article in 1970, the EMH was questioned on many grounds, especially the rationality assumption. This was also the time when much of the work on behaviour finance began.

Another theory that complements the EMH, however, partly overcomes the problem with the rationality assumption. This is when arbitrage starts to matter. Friedman first conceptualised it in ‘The Case for Flexible Exchange Rates’, which appeared in his Essays in Positive Economics in 1953. In 1965 Fama elaborated the same concept in his article ‘The Behaviour of Stock-Market Prices’ published in the Journal of Business.

Whereas in the EMH the precondition for markets to be efficient is that investors are rational, the theory of arbitrage says that markets are still efficient even if some investors are irrational. Suppose, Fama says, that a noise generating process gets started – e.g. investors trade on noise rather than on fundamental news, that is, they act irrationally. Suppose also that this process is subject to herding behaviour: these irrational investors follow each other and end up trading in the same way. Prices would tend to run well above their intrinsic value, that is, to create ‘bubbles’. Even in this case, however, the end result will still be an efficient market, as irrational investors will be met in the market by sophisticated traders (arbitrageurs) who, by acting as market contrarians, offset and neutralise price bubbles.

If there are many sophisticated traders […] they will be able to recognise situations where the price of a common stock is beginning to run up above
 Arbitrageurs are not only perfectly rational and able to recognise the fundamental value of an asset, but are also able to correct other investors’ irrationality. By so doing, they will earn a net profit for themselves and bring prices in line with their fundamental values. Arbitrageurs thus become for the EMH what the invisible hand is for the theory of perfect competition.

This concept of arbitrage is central to models such as Sharpe’s Capital Asset Pricing Model and Ross’s Asset Pricing Theory. In these cases, and in much of the subsequent literature, arbitrage has been defined more specifically as the simultaneous buying and short selling of the same or essentially similar security that trades at different prices in two different markets. In other words, it has been defined to include only those situations in which there is no risk and which entail a sure profit. This paper, however, argues that Fama’s example points to a more general characteristic of arbitrageurs: that of being able to offset bubbles each time they get created in a noise generating process. As it will be explained later, this conceptualisation has a far bigger explanatory and legitimating power than the narrower one.

2. The behavioural finance approach to arbitrage

Among the critiques moved to the Efficient Market Hypothesis, behavioural finance is certainly the most popular one. Behavioural financial economists question the rationality assumption and in particular the assumption that arbitrage can bring prices back to their equilibrium level. They formulate the ‘theory of limited arbitrage’ and the ‘theory of investor sentiment’ to explain why arbitrageurs might not work as expected to.
The theory of investor sentiment says that investors do not behave rationally but over- or under-react to news. This noise generating process gives rise to bubbles or momentum in market prices. This is no different from the process described by Fama. What is different in behavioural finance is that arbitrageurs or sophisticated traders might be unwilling to offset this irrational behaviour.

Suppose that arbitrageurs know “the model”, i.e. exactly the form of behaviour that the noise traders follow, and trade with the noise traders to take advantage of their misperceptions. Would such arbitrageurs bring the price of an asset closer to its fundamental value? At least in principle, it is possible that in some situations it pays arbitrageurs to anticipate future noise trader demand, and to “jump on the bandwagon” when prices are high rather than sell the asset.

Instead of having an incentive to sell the asset short – as in Fama’s theory – they might fuel and increase the noise generating process. Shleifer takes as example the Internet bubble of 2000 and many other financial bubbles documented in the history of finance.

The theory of limited arbitrage says that arbitrageurs might be unable to close price discrepancies. This is due to the fact that prices can move in opposite directions to what arbitrageurs expect. Whereas the theory of arbitrage assumes that arbitrageurs can buy the undervalued of two essentially similar securities and sell short the overvalued one, behavioural finance points out that ‘many securities do not have perfect or even good substitutes’, so that the very idea of two essentially similar securities is in question. If securities are not essentially the same, they will be affected by different fundamental news or the same news will affect them differently. It might happen that news on the overvalued stocks will be particularly good and news on the undervalued stocks will be particularly bad, so that prices further diverge from their intrinsic values. In this case arbitrageurs will register a net loss instead of a net profit and will not close out price discrepancies (interviewed practitioners speak of ‘asymmetric positions’). If both
mechanisms – the theory of limited arbitrage and the theory of investor sentiment – are active, arbitrageurs are unlikely to perform the role of rational stabilisers foreseen by the EMH. In this scenario, markets are no longer expected to be efficient.

3. The sociology approach to arbitrage

The ‘sociology of arbitrage’ owes its name to the work of Donald MacKenzie. Like behavioural financial economists, MacKenzie questions the capacity of arbitrage to offset investors’ irrationality and thus to bring prices back to their equilibrium level. Contrary to behavioural finance, however, for MacKenzie the flaws in the arbitrage mechanism lay less with psychological motives (responsiveness, conservatism, etc.) than with sociological reasons (various forms of social connectivities: e.g. imitation).

His analysis of arbitrage coincides with his departure from the ‘Parsons’ Pact’, that is, the belief that the economy is a separate system from the social and that its study is the business of economists, not of sociologists. He sees arbitrage as defined in neoclassical finance as the tool for this separation between economics and sociology:

There is a sense in which arbitrageurs are the border guards, in economic practice, of the Parsonian boundary between economics and sociology.

If indeed arbitrage does not guarantee the efficiency effect, the Parson’s Pact is undermined and sociological explanations become quintessential to market outcomes. He develops his argument by looking at the case of Long-Term Capital Management (LTCM), the hedge fund that went close to bankruptcy in August 1998. Of all the explanations of LTCM collapse that have been advanced, he opts for a process of a more sociological kind: imitation.
LTMC’s success led to widespread imitation […], and the imitation led to a ‘superportfolio’ of partially overlapping arbitrage positions. Sales by some holders of the superportfolio moved prices against others, leading to a cascade of self-reinforcing adverse price movements.40

McKenzie draws three conclusions from the LTMC case. First, despite the global reach of its activity, LTMC evidenced how interactions in financial markets are of a ‘local’41 and ‘Granovetterian’ nature,42 that is, they are ‘interactions between relatively limited numbers of people who are in some sense known to each other’. Second, LTMC evidenced that such forms of interactions – imitation, reciprocal monitoring of each other trading strategies, etc. – can lead to economic consequences – a liquidity crisis in this case. ‘The key risks may be ‘social’ risks from patterns of interaction within the financial markets, rather than shocks from the real economy or from events outside the markets’.43 Third, and as a consequence of the first two explanations, LTMC showed the inability of arbitrage to insulate the economic from the social.

In a response to MacKenzie, Hardie argues that the failure of arbitrage to bring prices back to equilibrium very much depends on our definition of it.44 The failure, in other words, is a definitional one. Hardie argues that the arbitrage MacKenzie and behavioural financial economists refer to (‘what the real world defines as arbitrage’,45) is not really arbitrage. He proposes to narrow its definition to include only those trades that involve no risk and have a guaranteed profit.46 This narrower definition – which goes back to Ross’s theory as reported above – ‘still serves the role of arbitrage in financial economics, in closing mispricing between assets’.47 As for what is commonly – and according to Hardie – wrongly termed arbitrage, Hardie argues that it should be treated like most other investor activity: they both work according to the sociological processes (e.g.
imitation) evidenced by MacKenzie. If properly defined, arbitrageurs can still be the border guards of the Parsonian Pact.

4. A political economy approach to arbitrage

Investors’ behaviour – how they think and consequently trade – is central to all the approaches to arbitrage analysed so far. Neoclassical finance sees investors’ – and especially arbitrageurs’ – rationality as the critical assumption in the theory of efficient markets. Behavioural finance sees investors’ psychological motives/irrationality as the crucial reason to reject the efficiency paradigm, especially since arbitrageurs might not be willing to behave as the theory expects them to. In MacKenzie’s sociology of arbitrage, sociological processes such as imitation prevent the achievement of efficiency and more generally prevent the economy from remaining insulated from the social. Someone might expect that a political economy approach to arbitrage will identify the political processes that influence the behaviour of investors and prevent them from preserving the efficiency of the market. Indeed such an approach already exists within mainstream finance. According to mainstream approaches in finance, in fact, the process of price formation can be inefficient because of regulatory and legal constraints. These constraints prevent investors from driving prices back to their equilibrium level. (This is for instance the argument of the ‘Coasians’, who interpret the Coase theorem to signify that government actions, including regulation, cannot produce a better result than relying on negotiations between individuals in the market.48)

This is not, however, the main purpose of this article, though the focus is still on regulatory policies and attitude. The political economy approach proposed in this article shifts the focus away from investors’ rationality and towards the interests that revolves
around the practice of arbitrage. The efficiency-argument is still crucial, since it has influenced the work of regulators and legislators alike and has informed the debate between those who benefit and those who claim to lose from arbitrage. Yet the article aims at looking not at the efficiency hypothesis *per se*, but rather at its impact on and interrelations with the practice of financial regulation.

Before proceeding with the analysis, however, it is important to emphasise once again what is meant by arbitrage in this context and why it differs from the approaches analysed so far. Hardie rightly points out that the definition of arbitrage is crucial for its conceptualization. While he advocates for a narrower and more ‘purist’ definition according to which arbitrage is still able to perform its efficiency-enhancing role, this paper argues for a rather opposite position. ‘Real world’ understanding of arbitrage includes a variety of mechanisms and instruments that are perceived to be able to offset a price bubble thanks to superior technique or knowledge, independently of whether they are risk-free or profit-guaranteed. Fama wrote in his example that if the price of a stock runs above its intrinsic value, sophisticated traders have an incentive to sell this stock short, offsetting irrational behaviour and bringing efficiency back to the market. This paper argues that in regulatory and policy circles short selling is attributed the same efficiency-enhancing role that is commonly ascribed to arbitrage. In addition to shifting the focus away from investors’ behaviour and towards their interests, this article stresses the need to look more specifically at the practice of short selling.

More specifically, the article looks at the way short selling is allowed to operate in the market and the role of the efficiency explanation within this. Regulatory decisions and non-decisions determine whether a strategy is permitted and the range of choices open to investors. This article shows that even if the *performativity* of the EMH has been
questioned in both theoretical analyses and empirical tests, its power to legitimise regulatory and market practice is as strong as ever.\textsuperscript{51} Besides, this legitimisation is not without consequences for the distribution of gains and losses from the practice of short selling. These aspects will be analysed through a series of market events of the recent financial history and, as a particular case study, in relation to the reform of the short-sale rule in US equity markets.

The role of short selling has been increasingly debated. On the one hand, short selling is said to be able to offset price bubbles and push prices towards their equilibrium level (Fama’s argument). In particular, it is said to perform three positive functions: to incorporate information into prices thus making the process of price formation more efficient; to provide liquidity;\textsuperscript{52} and, in the case of derivatives arbitrage, to keep derivatives linked to the cash market.\textsuperscript{53} On the other hand, short selling played a dubious role in many events of the recent financial history – e.g. speculative pressures in the aftermath of 9/11; naked short selling abuses. More generally, short selling is said to be able to exaggerate share price declines through the use of manipulative strategies; to exacerbate a bear market; and to increase share price volatility. The argument of the benign role of short selling is dominant throughout the financial community – practitioners, academics and regulators alike. This explains why any call for its regulation is measured against the possibility of hindering its efficiency-enhancing role. The argument of the manipulative role of short selling is voiced by long-term shareholders and issuers, especially small capitalization companies, which fear the aggressive practices of short sellers – e.g. the practice of spreading negative information about a company to push its stock price down. It is true that in principle nobody is in favour of manipulative short selling and that manipulation is condemned as an aberration of an otherwise optimal strategy. Yet the distinction of what is manipulative and what is not is hard to
maintain. In addition, manipulative short selling has been on the increase and in many markets it has been the rule and not the exception/aberration. The various interests linked to the practice of short selling are at the centre of this paper and will be examined in the remainder of the section. Before doing so, however, it is essential to understand who is selling short and the relevance of this mode of trading.

To begin with, there is a lack of data on short selling. Short and long positions in equity markets are not subject to the same disclosure requirements. In the UK, for instance, anyone buying more than 3 percent of the shares of a company is required to disclose, but this does not apply if someone is selling short more than 3 percent of the same shares. Similarly, if these transactions are carried out through put options, which are the economic equivalent of short selling, they are not subject to any disclosure requirement, while call options (the buying side) are. In the U.S., short selling in cash equity markets – in terms of aggregate short position in a share – is reported, but this leaves out short selling carried out through derivatives markets, which is the bulk of it. In addition, the identity of those who sell short remains unknown. Regulators estimate that most of the short selling is done by hedge funds (through market makers) or similar vehicles. This can partially answer the question of the relevance of this practice in financial markets: the growing importance of short selling is directly proportional to the huge amounts of money flowing into the hedge fund industry – an industry which passed from 300 billion in capital under management in 1998 to 1 trillion in 2005.

In the last decade the benign role of short selling was mentioned many times during currency crises, especially when the currency was said to be pegged in an unsustainable way (currency was overvalued) and shorting it was seen as a way to bring it back to its fundamental value. In equity markets, it was the Internet bubble that brought short
selling to the headlines. In 2000, when the internet bubble burst, one of the reasons for the hyper-inflated valuations of dot.com companies was said to be the lack of mechanisms to sell short or the existence of constraints in the use of short selling. Legislation such as the Commodity Futures Modernization Act (CFMA) of 2000, which re-introduced single-stock futures (a way to sell stocks short) into the US system, was welcomed as a way to make short selling easier and cheaper. Most of the Finance literature on the topic focuses on the constraints to short selling that prevents it from performing its efficiency-enhancing role. Once again the concept of arbitrage is further extended: it is not only arbitrageurs that bring efficiency by spotting price discrepancies, but every tool of short selling can counterbalance the mass of investors that buy long, thus becoming efficiency-enhancing.

Alongside this positive attitude towards short selling, several market events brought up other aspects of the practice and questioned its benign character. Short selling became particularly contested in the aftermath of 9/11. After the terrorist attacks of 9/11, both the financial and non-specialised press repeatedly reported news of hedge funds ‘profiteering’ from the terrorist attacks and especially from the collapse in the share prices of airlines and insurance companies. The activity of shorting the market was identified as being to blame. Emphasis was placed on the gains hedge funds and short sellers in general could make by selling short the stocks of airline and insurance companies. ‘A typical hedge fund short selling transaction would net “tens of millions of dollars” in profit on a trade that could be valued at hundreds of millions of dollars’. The Wall Street Journal reported that Pequot Capital Management, one of the biggest hedge funds in the US, ‘in the seven trading days since the market reopened […] made a profit of about $700 million […]’.
Discussions followed on the effect and desirability of short selling and calls were made to restrict this practice. Since in order to sell short hedge funds need to borrow stocks from institutional investors or banks, one proposal was to stop or limit the lending of stocks to hedge funds for speculative purposes. It is interesting to analyse the reasons that were given to defend the practice. A hedge fund manager said, ‘[Banning short selling] would be foolish. It would be taking out a standard and useful practice, which keeps markets efficient. It’s our job to make the markets more efficient’.\(^6\) This view was sustained not only in the hedge fund industry, but also among regulators. Regulators in both London and Washington stated that they felt no need to intervene, considering short selling to be beneficial to the market.\(^6\) Howard Davies, former Chairman of the Financial Services Authority (FSA) in London, said that monetary authorities had no intention to somehow limit the use of short selling and the lending of stocks to short sellers. He said: ‘In normal circumstances short selling is a natural and important feature of the market. Many would argue it improves “price discovery”’.\(^5\)

Short selling was also addressed at the Financial Stability Forum in 2002.\(^6\) The FSF felt compelled to debate the issue after the allegations of massive short selling in the aftermath of September 11 and in the Japanese stock market in the spring of 2002. The FSF report concluded that these allegations were not verified.\(^6\) FSF members found that short selling was on average providing a useful service to the market and that any attempt to regulate it would have been counterproductive. In this context short selling was also targeted for its role in foreign exchange markets. (Short selling in foreign exchange markets means selling a currency short in the expectation of its devaluation.) If regulation of short selling is scant in securities markets, the situation is even worse in currency markets, where no data on either short or long positions taken by market participants is available. Indeed some discussions on disclosing short selling in foreign exchange
markets did take place in Basel, but no decision was taken. The debate also revealed
divisions among countries. German and French regulators were much more inclined to
regulate it than the UK and US ones. In 2002 Hans Eichel – Germany's finance minister
– proposed banning short-selling in times of crisis.\textsuperscript{68} UK and US regulators instead
defended the practice. Howard Davies, at the time Chairman of the Financial Services
Authority (FSA), showed at best scepticism about any attempt to regulate short selling, a
practice – he says – that

\begin{quote}
Is a normal activity, which, indeed, has some pluses associated with it in terms of speeding the necessary process of price adjustment. There are those, notably in Japan and Germany who want to control short-selling directly. Good luck to them. But a financial centre like London which prides itself on its openness and flexibility should think long and hard before imposing such restrictions. [...] I hope that at the FSA we can hold the line and not be pushed into more restrictive measures [...]\\textsuperscript{69}
\end{quote}

\textit{The Economist} and the financial press in general have been sanguine too in defending short
selling against any possible regulation. When in 2003 Eliot Spitzer, New York state's
attorney-general, included short-selling in its campaign against Wall Street, \textit{The Economist}
did not only take the view that ‘in bull markets short-sellers can help to put a brake on
irrational exuberance’, but also that ‘those who call for restrictions are often trying to run
companies in trouble, or entire economies that are on the ropes’. This was like saying
that companies or countries calling for a regulation of short selling have something to
hide. ‘Last year, Hans Eichel, Germany's finance minister, proposed banning short-selling
in times of crisis. In 2002 the German stock market fell by 44%.’\textsuperscript{70}

These examples show how the efficiency-role of short selling is endorsed by regulators,
practitioners and market observers in major financial centres. They also show, however,
that agreement is not universal and that a divide can be spotted between those against
any regulation of short selling and those in favour of some control and restrictions. The
focus on a specific instance of short selling regulation – regulation SHO in the US – will allow completion of the political economy analysis by highlighting the interests that are attached to each claim. Which interest do short sellers represent? Are those willing to limit the practice of short selling promoting an inefficient market?

The short sale rule in US equity markets was originally formulated in 1938s at a time when the impact of short selling was far from being seen as benign. In the late 1930s-early 1940s many in the US blamed the 1929 stock market crash on the strategy of selling securities short in an attempt to drive down their price. It was indeed a core issue in the post-Depression debate that brought about the reform of the investment companies and securities laws in the US (Investment Company Act of 1940, Securities Act of 1933, and Securities Exchange Act of 1934). Congress delegated authority to regulate short selling to the Securities and Exchange Commission. The SEC issued a rule that was going to restrict the short sales of any security registered on a national securities exchange in the US (Rule 10a-1 or up-tick rule). Rule 10a-1 allowed selling short a listed security only ‘at a price above the price at which the immediately preceding sale was realized (plus tick) or at the last sale price if it is higher than the last different price (zero plus tick)’. Beyond the technical jargon, this provision precluded short sales of a listed security at successively lower prices so as to prevent investors from driving down the market.

The debate in Congress and at the SEC, however, did not stop there. In 1963 Congress required the SEC to carry out a study of the relationships between changes in short positions and subsequent price trend. The request of such a study suggests that Congress was concerned that the up-tick rule might not stop the process of price manipulation for which it was designed. In effect the study proved that ‘the ratio of short sales to total volume increases in a declining market’, which means that a downward
trend could be triggered despite the up-tick rule. In 1976 the SEC proposed the elimination of the up-tick rule, but market participants, including the major exchanges, NYSE and AMEX, strongly opposed the suspension of the up-tick rule, arguing that it provided important protection for investors that should not be removed. In 1991, the House Committee on Government Operations released a report on short selling that once more questioned the effectiveness of the up-tick rule and asked whether a similar rule should be implemented for NASDAQ trading. Following this report, the NASD proposed a short sale rule covering NASDAQ National Market System (NMS) securities, which are those securities that have similar characteristics to listed securities but are traded on the NASDAQ. The SEC approved this rule (Rule 3350) in 1994.

No further changes to this regime were implemented until 1999, when the SEC started a process of revision. By 1999, many things had changed in the market. Transactions in listed securities now represent only a small fraction of equity trading. The bulk of short selling is executed in over-the-counter markets and through derivatives – none of which is subject to short sale restrictions (e.g. no short sale rule applies to futures exchanges). In addition, the Commodity Futures Modernization Act of 2000 (CFMA 2000) reintroduced into the US market single-stock futures (SSF), which are a fast, cheap and easy way to conduct short selling in futures exchanges. Single-stock futures – contracts by which one part promises future delivery of an asset to another – allow selling stocks more easily than by going to the cash market because instead of having to borrow the stocks to sell short it is enough to sell a SSF contract. Those willing to sell short can move to futures markets, where no up-tick rule is in place. Last but not least, the US moved from a fractional to a decimal system for pricing securities. While in a fractional system it made sense to say that short sales could not be executed at a price lower by a fraction than the preceding one, in a decimal system the difference can be so small
0.00000001) that it is no longer a problem to wait – actually it is possible to manipulate
stocks so that they move to the threshold price quicker. These events lessened the
rationale of the up-tick rule (as currently formulated).

On the other hand, a series of abuses linked to short selling caused many in the financial
community to call for more and not less regulation. The most often reported abuses are
the so-called ‘bear-raid’, ‘short selling at the box’ and ‘naked short selling’. In a bear-raid
short selling is done in an effort to drive down the price of an asset. A variation of this
strategy is short selling at the box, where traders first ride the price of the stock up in a
boxed position and, when they think the upward trend is about to turn, start selling
short. With the help of negative comments on the company posted on the internet, they
drive the price of the stock down to buy it back cheaply. In the practice of naked short
selling, shares are sold short without any arrangement or intention to borrow the stocks
for delivery by settlement date. In legal terms this means shorting without an ‘affirmative
determination’ of the existence of the shares to short against (sources close to the SEC,
2004). With this practice market makers sell short more shares than are actually issued by
the company, which has consequences for both parties. By artificially inflating the
number of shares in the market, this practice has the effect of diluting the price value of
current shareholders. By diluting the value of the stock, affected companies will find it
more expensive to raise capital. In addition, brokers can lend shares for shorting without
the prior consent of the shareholder. Suppose someone owns shares in a broker account.
The brokers are allowed to lend those shares to someone else for the purpose of selling
them short. The shareholders will have their value diluted due to a short selling activity
carried out with their very own shares. All these abuses were reported in the comment
letters sent to the SEC during its process of consultation.79
The SEC concluded the process of consultation and came up with a new rule in June 2004 (reg. SHO). But before analysing the SEC final rule, it is useful to go through the period of consultation and show how contested the issue was and the nature of the opposing interests. The SEC opened the consultation process in 1999 asking comments on whether the up-tick rule should be reformed, abolished or maintained. The opening sentences of the concept release recognise the role of arbitrage:

Arbitrageurs contribute to pricing efficiency by utilizing short sales to profit from price discrepancies [...] 81

This type of arbitrage was already exempted from any regulation in Rule 10a-1. The SEC, however, makes the same comments about short selling tout court:

Short selling can also contribute to the pricing efficiency of the equities markets. When a short seller speculates on a downward movement in a security, his transaction is a mirror image of the person who purchases the security based upon speculation that the security’s price will rise. Such short sellers add to stock pricing efficiency because their transactions inform the market of their evaluation of future stock price performance.82

Reference to the efficient market hypothesis and to the role that short sellers play in contributing to market liquidity and pricing efficiency is also made in many comment letters. These letters can be grouped into three categories: (a) letters calling for the elimination of any restriction on short selling; (b) letters in favour of maintaining or improving upon the current rule on short sales; (c) letters calling for more stringent short-sale regulation – particularly in those markets where small-capitalization stocks trade – in order to guard against short sale abuses.83

Type A letters were sent especially by brokers/dealers (e.g. Charles Schwab, Cornerstone Securities, Hill Thompson) and associations of hedge funds (Managed Funds Association). They stress ‘the important role that short sellers have traditionally played in the market by contributing to both market liquidity and pricing efficiency’84 and that any
rule on short selling – in whatever market – greatly diminishes this positive contribution. ‘The existence of the short sale rule introduces pricing inefficiencies’.85 ‘There is nothing wrong with the practice of sort selling because it may enhance market liquidity’.86 Of course all these letters are against the use of short selling as a manipulative device, but they see manipulation as an aberration, which should not shed any negative light on the overall practice.

A few words are needed to analyse the interests that these letters reflect. The interest of hedge funds is fairly straightforward: hedge funds are major short sellers and legally the ones with the highest capacity to sell short. As for brokers/dealers, they make large profits from servicing the hedge fund industry and providing them with the instruments to sell short. In order to sell short, in fact, investors need to first borrow the securities from a broker/dealer, who will earn an interest on the proceeds of the short selling, trading and clearing commissions, and possibly income from derivative transactions.87 The more unrestricted short selling is the more investors short and the higher the fees and commissions for the brokers are.

Most type B letter were sent by: (1) major national stock exchanges, such as the New York Stock Exchange (NYSE) and the Boston Stock Exchange (BSE); by self-regulatory organizations such as the National Association of Securities Dealers (NASD), which watches over the NASDAQ exchange;88 by voluntary associations such as the Association of Publicly Traded Companies (APTC) and the North America Securities Administrators Association (NAASA). They were all in favour of retaining the short-sale rule or improving upon it. NASD pointed out that today’s markets are ‘just as vulnerable to bear raids and piling on as they were at the inception of the Rule’.89 No ‘empirical data exists that would change the need for short sale regulation since 1938 when the rule was
adopted. NYSE argued that the up-tick rule still serves to protect individual investors, prevent fraud and manipulation, and promotes just and equitable principles of trade. The mission of some of these organizations is to protect investors (e.g. NAASA) or represent the interests of issuers (e.g. APTC), so that their view can be seen as representative of a larger group of market actors. Corporate and long-term investors in particular are concerned with ‘short sellers depleting the value of their securities’. Even if they recognise that theoretically short selling can bring efficiency and liquidity to the market, they see short sellers as aggressively trying to find and spread negative information on their stocks in order to push (and keep) the price down. In this type of letter, the efficiency-enhancing role of short selling is not the primary consideration. Focus is on the protection of issuers’ and shareholders’ rights, which is partly seen in opposition to those of short sellers.

Most type C letters were sent by small capitalization companies, also called ‘penny stocks’, or by their shareholders. These are stocks traded in the NASDAQ Smallcap, OTC:BB, or Pink Sheets markets. Type C letters advocate the extension of short sale regulation to these markets, where most short sale abuses have been reported (most of the lawsuits involve these stocks; none involved NYSE stocks) and where no rule on short selling has ever been implemented. They argue that the lack of regulation of short selling in the OTC:BB, Pink Sheet and SmallCap markets produces high level of volatility in stock prices. For instance, Edwin Marshall, Chairman and CEO of Medizone International, said that the price of his company’s shares were recursively distorted and manipulated by short sellers. According to Marshall’s account,

On the 11th of June 2001 Medizone opened at $0.39. The stock then rose from the previous close of $0.38 to $0.47 in the first 12 minutes of trading. On heavy buy volume the market makers then walked the stock down to $0.28 over the next few hours. They short sold the stock down about 40% in two hours on heavy buying! One market maker would not be capable of that kind
of play. In my opinion, there has to be manipulation and collusion… The basic law of supply and demand was circumvented.\(^9\)

Small capitalisation companies with a high degree of vulnerability – that is, high risk of failure – are short sellers’ favourite target. Short sellers can bet against these companies’ stocks and even push them into bankruptcy. Once bankruptcy is declared, short sales will be automatically covered, so that this will sort out even the delivery problem. Mr Marshall refers to ‘market makers’, which are mainly brokers/dealers acting for their own and their clients’ (especially hedge funds) accounts. Market makers are supposed to maintain an orderly market by ‘acting as the counterweight to public exuberance, helping to temper volatile price rises and subsequent drops’,\(^9\) like the sophisticated traders in Fama’s example. This might be true in many cases, but, given the number of complaints received on this issue – type C was the largest group of letters, accounting for over 2000 of the comments received by the SEC\(^9\) – it can be concluded that sophisticated traders are not always doing their job of bringing prices back to their equilibrium level. They are often manipulating the market. Type C letters reported several cases of abuses (naked short selling, bear raid, short selling at the box). In these letters the very concept of short sale is often seen as immoral and unethical and its efficiency-enhancing role rarely mentioned (‘it remains a bastion of immorality in the market’;\(^9\) ‘It is not fair’;\(^9\) ‘Unfair advantage of powerful short selling by large institutions’;\(^9\) ‘The concept of short selling is abominable’;\(^9\) ‘The lack of regulation encourages criminal activity by the market makers’\(^9\)).\(^9\)

The first consideration that can be made by looking at the three groups of comment letters is that there is a divide along the corporate/finance axis. Firms are worried that manipulative short selling can curtail the value of their stocks, which in turn would hinder their capacity to raise fund. For them the efficiency argument is less compelling,
especially if short selling is done in an aggressive way to keep prices artificially low. Short sellers and market makers instead claim that restrictions on short selling cause stock prices to inflate – in other words, they claim that restrictions cause an ‘upward manipulation’ – and draw on the efficiency-argument to back their claim. The conflict between short sellers and corporate actors is becoming increasingly fierce. One emblematic case is that of Patrick Byrne, CEO of Overstock.com, which has initiated a real battle against naked short selling, especially against failures to deliver the shares sold short. Mr Byrne argues that these failures to deliver reduce many stocks to ‘penny stocks’ (i.e. each stock is worth less than a dollar) and open the way to their bankruptcy. Large institutional investors in fact never invest in penny stocks, so that the demand for these stocks after they have been shorted never recovers. On the other hand, the literature has started documenting instances in which it is companies that – sometimes as a reaction to short-sale abuses – use a variety of methods to indirectly self-restrict short selling and create a short sale squeeze.

A second divide can be traced along the line of long/short positions: shareholders that buy stocks to keep (which are therefore ‘long’) have different interests than short-sellers, who, most of the time, do not even own the stocks. The nature of the conflict is evident if we start from the simple consideration that when prices go up, long-holders gain and short sellers lose money, while when prices go down long-holders lose and short sellers gain. Short sellers blame long-holders of ‘upward manipulation’ (creating a price bubble), while long-holders blame short sellers of ‘downward manipulation’ (depressing prices by aggressively shorting stocks). In favour of short sellers it was a chorus of commentators during the Internet boom who pointed to short selling as a remedy against the excessively-inflated buying side. In favour of long-term shareholders two considerations are usually made. The first is that downward manipulation is executed by very
sophisticated traders who are not only able to drive prices down but also to keep them
down for long period of time. Sometimes the imbalance between long and short
positions is such – short positions are much larger – that it discourages people from
buying, as it would take them too much capital to raise the price. The second
consideration is that the downward and upward manipulation might be initiated by the
same category of investors, so that one single instance of manipulation is taking place. In
the case of the so-called ‘bandwagon effect’, for instance, sophisticated investors drive
the price of stocks up only to short them when prices are high enough to allow them a
profit. After all, if short selling is supposed to offset price bubbles, where were short
sellers during the mounting of the Internet bubble? The conflict, therefore, can be re-cast
as one between sophisticated traders/short sellers on the one hand and the mass of
ordinary investors on the other.

The final consideration is that small firms are more likely to suffer from the activities of
short sellers than large ones. It is unlikely that short sellers attack a stock like Microsoft.
Indeed very few comment letters were received from large issuers. Most of the
complaints came from small capitalization companies or, more generally, from those
companies whose vulnerability is higher. As said before, short sellers have a particular
preference for vulnerable companies running risky businesses. The example of Medizone
International is very appropriate: pharmaceutical companies producing new drugs or
vaccines are considered at high risk of failures and more likely to be attacked by short
sellers.

Given this set of preferences for the regulation/non-regulation of short selling, what did
the SEC do in the end? Which set of interests – if any – prevailed in finalising Regulation
SHO? Once again it is to be said that Regulation SHO is for a good part invalidated by
the fact that other instruments of short selling are abundantly present in the market (e.g. single-stock futures), so that whichever conclusion the SEC reaches it will not change the situation in any fundamental way. After more than five years of consultation, the SEC issued a final rule in June 2004. It thought of reaching a compromise by tackling the abuses while at the same time freeing short selling of the up-tick rule. The SEC Economic Analysis Department thought that, especially because of the introduction of the decimalisation system, the up-tick rule was ineffective in preventing attempts to drive down the market (bear raids or ‘death spirals’) and that it could be lifted without major consequences. To be sure of it, the SEC decided to have a pilot period in which the up-tick rule would be suspended for about 40 percent of listed stocks and 40 percent of NASDAQ NMS stocks (plus some stocks in small-caps and OTC:BB markets). The pilot will be used as a study to assess the general usefulness of the up-tick rule. After one year (the pilot started in May 2005), a decision will be taken to either suspend the rule (most likely event) or reform it otherwise.

The SEC decided to concentrate its efforts on tackling abuses and separating the aberration from the benign, efficiency-enhancing practice. In this respect it decided to enforce two rules: a ‘locate’ requirement and a ‘delivery’ requirement. First, it established a ‘locate requirement’ for everyone willing to sell short, so as to prevent naked-short selling. If a trader wants to sell short a number of shares, they must first be located: this means that, although the trader does not have to borrow them straight away, an arrangement must be made to do it in due time (which is called ‘affirmative determination’). An exception was made for market makers and sophisticated traders for what is called ‘bona fide’ short selling. Bona fide short selling refers to those trades that are meant to provide liquidity – bona fide hedging – and do not have any speculative or investment purpose. The rationale of this exception is to maintain the capacity of short
selling to bring liquidity and efficiency to the market. Yet the exception provides a potential loophole, since it will not be terribly easy to distinguish bona fide from other types of short selling executed by the same broker/dealer. Another potential loophole in the ‘locate’ requirement concerns international arbitrage. For instance, an investor might say that short positions are covered with stocks soon to be delivered from a foreign stock exchange. Other jurisdictions, however, might not have any restriction against short selling, not even naked short selling. This is the case of Canada. Brokers and investors can sell short through Canadian institutions that sell into the US markets. It is clear that this might give rise to another potential loophole.

The second action the SEC took is a mandatory requirement to stop selling short and close positions for all those securities that do not settle within 13 days from the trading day and that represent more than 0.5 percent of the total outstanding shares issued by the company (‘close out rule’). These are called ‘threshold securities’ and Reg. SHO requires self-regulatory organizations (SROs) such as NYSE, NASDAQ, AMEX to disseminate a daily list of threshold securities. Until securities are in the list, they cannot be sold short without first borrowing the shares. Yet market observers noticed that since the publication of the lists on 3 January 2005 many securities have been in the threshold list for well beyond 13 consecutive days. Sometimes they stay for months or forever. This can be explained by the possibility for short sellers to cover a short sale opened with one broker by taking another one with a different broker. Reg. SHO prohibits such behaviour, but the evidence from one year of threshold listing shows that the prohibition can be evaded. In addition, threshold lists can only be made for securities of reporting companies, that is, companies that are registered with the SEC. Companies in the Pink Sheets, which are highly affected by short sale abuses, are not registered with the SEC – hence their securities cannot be threshold securities. Finally, the close-out rule does not
apply to open fail positions that existed prior to the effective date of Regulation SHO (3 January 2005). This way of grandfathering the existing ‘fails-to-deliver’ (FTD) might leave the abuses unpunished and, if there has been any damage to the shorted stocks, perpetuate the damage.

Overall, however, the problem seems to be one of enforcement rather than introduction of new rules. As many commentators pointed out, Reg. SHO does not add anything new to the current legal system: rules against market manipulation already existed in the US system. The problem was that the SEC was not able to enforce them. Even with Reg. SHO it is not clear which kind of enforcement mechanisms are in place and how strict the ‘punishment’ for not complying with the ‘locate’ and ‘delivery’ requirements is. The persistence of threshold securities says that enforcement and sanctioning mechanisms are not working and this furthers the impression that the distinction between benign and manipulative short selling is not straightforward.

The persistence of threshold securities also triggered further complaints against the working of the SEC. By not enforcing the rules and not sanctioning the abuses, critics say, many investors and issuers were placed at a disadvantaged position in relation to other market actors, notably broker/dealers. In addition, critics reckon that the failures to deliver are of considerable size and might lead to a systemic meltdown and a financial crisis of large proportions. Conversely, SEC officers reply that these claims are exaggerated. They argue that many of the complaints come from issuers whose management practices are equally obscure and therefore deserve to have their stocks shorted. Or they argue that those complaining against short selling abuses are ‘paranoids and lunatics’, who blow it all out of proportion and make up unrealistic conspiracy theories. In both cases the SEC is reluctant to accept that naked short selling and
strategic failures-to-deliver are as endemic as complainers report. But behind these motivations it is possible to spot once again the influence of the efficiency argument: short selling penalizes those stocks that are overpriced and those complaining have something to hide that short sellers can reveal for the benefit of the whole market.

It is undeniable that the SEC is faced with a dilemma. As a source close to the SEC said, on the one hand SEC officials do not want to restrain something that is ‘a tool of efficiency’, but on the other hand they know that many players (e.g. hedge funds) can use it for manipulation purposes. The dilemma SEC regulators face needs to be contextualised in an environment where a more positive attitude towards short selling (‘more positive’ with reference to the post-Depression years) has come to dominate market practice. This environment was described at the beginning of this section, with the reintroduction of single-stock futures and regulators’ rejection of the need to restrict short selling after 9/11 and other episodes of market unrest (example of The Economist). The efficiency-argument was implicitly or explicitly at the centre of all these decisions and opinions. Other developments can be brought to bear. For instance, US regulators relaxed some of the rules that limited the use of short selling by mutual and pension funds. One of these rules was the so-called ‘short short rule’ or ‘30 percent test’ and required that mutual funds derived less than 30 percent of their gross income from the sale of securities held for less than three months and from the use of short term strategies in general if they wanted to qualify as Regulated Investment Companies (RIC). Once qualified for RIC treatment, they could avoid corporate taxation. So the trade off was between a more liberal regime in the use of derivatives and leverage and a more accommodating tax regime. This ceased to be a trade-off in 1997, when the 106th Congress passed the ‘Taxpayer Relief Act’, which, among other things, repealed Section 851(b)(3) (the ‘short-short rule’) of the Internal Revenue Code. One of the consequences
of the repeal of this provision was that it allowed mutual funds to include short selling in their trading strategies without them losing their tax exemption. In other words, taxation has become more ‘short selling-friendly’. Another example can be found in the way short selling has come to be portrayed by online pundits. As the manager of a financial boutique observed, the tone of most articles on short selling is pedantically didactical. He said: ‘It seems to me that commentators want to inform investors that there is a new way of playing the market – shorting it – and that this is a normal, positive development and perhaps the only smart way to play the market’. In all these examples the emphasis on the profitability of short selling is backed and reinforced by the positive role short selling is said to play in keeping a market efficient.

All these trends reinforce and at the same time are reinforced by the legacy of the EMH. This section showed that this legacy is still very strong, despite the fact that its explanatory power has been questioned by different strands of the finance literature (e.g. behavioural finance, the sociology of arbitrage). Although the performativity of the EMH is far from perfect, as MacKenzie points out, its power to legitimate market practices and regulatory regimes is still very strong. This legitimating power is not without consequences for the distribution of gains from trading and investing. As the process leading to Regulation SHO showed, different categories of market actors are affected by short selling in different ways. Hence any regulatory initiative is likely to benefit certain categories of investors at the expenses of others. The efficiency-argument of short selling seems to be the key concept around which the various claims revolve.

This article provides a first attempt to conceptualise short selling in a political economy perspective. Additional research is warranted to understand the gains and losses from this
practice and to which extent a redistribution of wealth from long holder to short sellers is taking place.

**Concluding remarks**

In neoclassical finance, markets are not only efficient but fair. They do not admit the possibility of conflicting social welfares since the ethical properties attributed to markets treat all players equally. Arbitrage and short selling are interesting features of the fair and efficient market and offer particular insights into the arguments surrounding the concept. This paper has shown that conflicting social welfares exist and that a practice like short selling – generally described as useful and efficiency-enhancing – is perceived differently by corporate actors, traders, and small businesses.

Finance’s justification of short selling as efficiency-enhancing seems to be the key concept around which different claims revolve. Regulators do not want to restrain something that is ‘a tool of efficiency’, but on the other hand they know that many players (e.g. hedge funds) can use short selling for manipulative purposes. The compromise is usually achieved by tackling the abuses as if they were isolated aberrations from an otherwise optimal behaviour. However, the distinction between an aberration on the one hand, and normal and efficient behaviour on the other, is problematic, as the loopholes in Regulation SHO demonstrate.

More generally, the efficiency justification of short selling produces a divide along the lines of corporate/finance and long/short holders of capital. More research is warranted on the gains and losses from the practice of short selling for different categories of investors. Contrary to the dominant view in neo-classical finance, no financial practice is positive or negative in absolute terms but only in relative terms. There are always social
conflicts surrounding market practices and a political economy approach can help understand them by mapping ‘who gets what’ and bringing in the political, contested, nature of their management and regulation.

3 Investors build long positions when they acquire stocks, while they build short positions when they sell short stocks, according to the definition given before.
8 In talking about the ‘performativity of economic theory’, MacKenzie draws on the work of Micheal Callon, *The Laws of the Markets* (Blackwell, 1998). MacKenzie refers to a process by which economic theory does not describe a reality separated by the very theory that describes it but rather alters economic practices towards conformity with the theory. ‘[…] An aspect of economics is used in economic practice, its use has effects, and amongst those effects is to alter economic processes so as to make them more like their depiction by economics’. In this context, a low performativity ratio means that arbitrage theory has lost – or never possessed – the capacity to shape the reality that is describes. See Donald MacKenzie, ‘Is Economics Performative? Option Theory and the Construction of Derivatives Markets’, paper presented at the annual meeting of the History of Economics Society, Tacoma, WA, June 25, 2005, p. 23, www.sps.ed.ac.uk/staff/is%20economics%20performative.pdf (accessed 10 January 2006).
10 If I sell short a stock at 5 and its price rises to 10 instead of falling, I lose 5, but losses can be limitless since in principle there is no limit to the price increase.
11 When the first hedge fund was created in 1949, its inventor Alfred Winslow Jones wanted to have an instrument to go long (buy) shares that he considered ‘undervalued’ and go short on (sell short) shares that he considered ‘overvalued’. If the investor was sophisticated and smart enough to pick the right stocks, s/he could neutralise market risk for her/himself and avoid bubbles for the market. This is where the name ‘smart money’ comes from.
12 In the US, for instance, the so-called ‘short short rule’, which limited mutual funds’ ability to sell short, was lifted. See section four for a detailed analysis.
18 Ibid.
19 The Oxford Dictionary defines efficiency in different ways – the fact of being an operative agent or efficient cause; fitness or power to accomplish or success in accomplishing the purpose intended; adequate power, effectiveness, efficacy’ (Oxford English Dictionary Online 1989). Efficiency has assumed a variety of meanings in Economics – for instance marginal, economic, or technical efficiency. Markowitz (1952), the founder of modern finance, used it in the engineering sense of ‘maximising output relative to input, or minimising input relative to output’ (Chancellor: 256). It is in this sense that Markowitz referred to an ‘efficient portfolio’ of stocks.


Hardie, ‘The Sociology of Arbitrage: a Comment on Mackenzie’


See endnote 8 for an explanation of performativity.

Market makers are said to provide liquidity since they provide a two-sided quote: they match investors’ request to buy or sell

For instance, if futures are mispriced in relation to the spot market, a trader can buy the cheap future and sell short the dear stock thus keeping the two markets linked.
55 Sources close to the Securities and Exchange Commission and the Commodity Futures Trading Commission.
59 A letter published by an association of hedge funds, Albourne Associates, which contested the allegations against hedge funds, gives an idea of the terms of the discussion, Albourne Partners Ltd ‘Outrageous slur on the hedge fund community’, Albourne Village, village.albourne.com/welcome.html (published and accessed 17 September 2001).
60 Since hedge funds were described as the ones better equipped to sell short, it was inevitable to equate them with short sellers.
64 Ibid.
65 Hughes, C. ‘War on Terrorism: Market - City Watchdog Refuses to Bring in Curbs on Short Selling Activities’, The Independent, 26 September 2001.
67 Ibid.
70 The Economist, Don’t shoot the messenger, 27 February 2003.
73 Zero plus- (or up-) tick means that, if there is no change in the last trade price of the particular stock, a short sale can be executed only if the previous trade price is higher than the trade price that preceded it.
75 Ibid, p. 5.
76 Ibid, p. 5.
77 Respectively the New York Stock Exchange and the American Exchange.
The following analysis is based on the text of the letters, available at www.sec.gov, and on interviews conducted with some of those who sent the letters.


Schifrin, M. 'Nobody Wants to Kill the Golden Goose', Forbes, 14 December 1998


Castagna, P. 'To Stop the Stag', Forbes, 14 December 1998


Medizone International is a company based in California which is developing an ozone-based treatment for the cure of Hepatitis C. Interview with CEO E. Marshall was carried out in November 2001. Marshall’s comment letter to the SEC is available at sec.gov/rules/concept/s72499/marshal2.txt.

Medizone 2000, my italics

Hill Thompson, Comment Letter to the SEC.

2200 out of 2770 were type C letters, SEC, Summary of comment letters received by 1 September 2000.


This aspect needs further evidence. There were in fact reported cases in which it was the penny-stock companies that manipulated the market at the expenses of shareholders. They issued more shares than it would have been warranted by their capitalisation and by so doing they diluted the value of current shareholders (source close to the SEC).

See http://www.businessjive.com for a summary of Patrick Byrne’s public commentaries against naked short selling.


The NASDAQ system consists of two areas: the National Market System and the OTC market. The National Market System (NMS) comprises so-called Tier 1 OTC stocks, a select part of the OTC market whose issues meet higher volume and price requirements than the remainder of the OTC market.


More precisely, threshold securities are defined as those equity securities that (1) have a fail to deliver position for five consecutive settlement days, (2) total 10,000 shares or more and (3) equal to at least 0.5 per cent of the issuer’s total outstanding shares.


See endnote 18 for a definition of Pink Sheet market.

Securities and Exchange Commission, Division of Market Regulation: Key Points About Regulation SHO.


115 Anonymous interviewee managing a financial market boutique.
<table>
<thead>
<tr>
<th>Date</th>
<th>Author(s)</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>180/05 November</td>
<td>Ariel Buira</td>
<td>The Bretton Woods Institutions: Governance without Legitimacy?</td>
</tr>
<tr>
<td>181/05 November</td>
<td>Jan-Erik Lane</td>
<td>International Organisation Analysed with the Power Index Method.</td>
</tr>
<tr>
<td>182/05 November</td>
<td>Claudia M. Fabbri</td>
<td>The Constructivist Promise and Regional Integration: An Answer to ‘Old’ and ‘New’ Puzzles: The South American Case.</td>
</tr>
<tr>
<td>183/05 December</td>
<td>Heribert Dieter</td>
<td>Bilateral Trade Agreements in the Asia-Pacific: Wise or Foolish Policies?</td>
</tr>
<tr>
<td>184/05 December</td>
<td>Gero Erdmann</td>
<td>Hesitant Bedfellows: The German Stiftungen and Party Aid in Africa. Attempt at an Assessment</td>
</tr>
<tr>
<td>185/05 December</td>
<td>Nicola Maaser and Stefan Napel</td>
<td>Equal Representation in Two-tier Voting Systems</td>
</tr>
<tr>
<td>186/05 December</td>
<td>Gianluca Grimalda</td>
<td>Can Labour Market Rigidity Lead to Economic Efficiency? The Technological Change Link</td>
</tr>
<tr>
<td>187/06 January</td>
<td>Leonardo Ramos</td>
<td>Collective political agency in the XXIst century: Civil society in an age of globalization</td>
</tr>
<tr>
<td>188/06, January</td>
<td>Mustafizur Rahman and Wasel Bin Shadat</td>
<td>NAMA Negotiations in the WTO and Preference Erosion: Concerns of Bangladesh and Other Regional LDCs</td>
</tr>
<tr>
<td>189/06, January</td>
<td>Amrita Dhillon, Javier Garcia-Fronti, Sayantan Ghosal and Marcus Miller</td>
<td>Bargaining and Sustainability: The Argentine Debt Swap</td>
</tr>
<tr>
<td>190/06, January</td>
<td>Marcus Miller, Javier Garcia-Fronti and Lei Zhang</td>
<td>Contradictionary devaluation and credit crunch: Analysing Argentina.</td>
</tr>
<tr>
<td>191/06, January</td>
<td>Wyn Grant</td>
<td>Why It Won’t Be Like This All The Time: the Shift from Duopoly to Oligopoly in Agricultural Trade</td>
</tr>
<tr>
<td>192/06, January</td>
<td>Michael Keating</td>
<td>Global best practice(s) and electricity sector reform in Uganda</td>
</tr>
<tr>
<td>193/06 February</td>
<td>Natalie Chen, Paola Conconi and Carlo Perroni</td>
<td>Does migration empower married women?</td>
</tr>
<tr>
<td>194/06 February</td>
<td>Emanuel Kohlscheen</td>
<td>Why are there serial defaulters? Quasi-experimental evidence from constitutions.</td>
</tr>
<tr>
<td>195/06 March</td>
<td>Torsten Strulik</td>
<td>Knowledge politics in the field of global finance? The emergence of a cognitive approach in banking supervision</td>
</tr>
<tr>
<td>196/06 March</td>
<td>Mark Beeson and Hidetaka Yoshimatsu</td>
<td>Asia’s Odd Men Out: Australia, Japan, and the Politics of Regionalism</td>
</tr>
<tr>
<td>Date</td>
<td>Presenter(s)</td>
<td>Title</td>
</tr>
<tr>
<td>------------</td>
<td>---------------------------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>197/06 March</td>
<td>Javier Garcia Fronti and Lei Zhang</td>
<td>Political Instability and the Peso Problem</td>
</tr>
<tr>
<td>198/06 March</td>
<td>Hidetaka YOSHIMATSU</td>
<td>Collective Action Problems and Regional Integration in ASEAN</td>
</tr>
<tr>
<td>199/06 March</td>
<td>Eddy Lee and Marco Vivarelli</td>
<td>The Social Impact of Globalisation in the Developing Countries.</td>
</tr>
<tr>
<td>200/06 April</td>
<td>Jan Aart Scholte</td>
<td>Political Parties and Global Democracy</td>
</tr>
<tr>
<td>201/06 April</td>
<td>Peter Newell</td>
<td>Civil society participation in trade policy-making in Latin America: The Case of the Environmental Movement</td>
</tr>
<tr>
<td>202/06 April</td>
<td>Marcus Miller and Dania Thomas</td>
<td>Sovereign Debt Restructuring: The Judge, the Vultures and Creditor Rights</td>
</tr>
<tr>
<td>203/06 April</td>
<td>Fondo Sikod</td>
<td>Globalisation and Rural Development in Africa: The Case of the Chad-Cameroon Oil Pipeline.</td>
</tr>
<tr>
<td>204/06 April</td>
<td></td>
<td>The Translation of a Crucial Political Speech: G.W.Bush’ State of the Union Address 2003 in Le Monde</td>
</tr>
<tr>
<td>205/06 April</td>
<td>Paola Robotti</td>
<td>Arbitrage and Short Selling: A Political Economy Approach</td>
</tr>
</tbody>
</table>

**Centre for the Study of Globalisation and Regionalisation**

University of Warwick  
Coventry CV4 7AL, UK  
Tel: +44 (0)24 7657 2533  
Fax: +44 (0)24 7657 2548  
Email: csgr@warwick.ac.uk  
Web address: [http://www.csgr.org](http://www.csgr.org)