NEW YORK UNIVERSITY
STERN SCHOOL OF BUSINESS
FINANCE DEPARTMENT

Working Paper Series, 1997

Are Financial Market Corners and Short Squeezes Inefficient?

Nagarajan, S.

FIN-97-5

ARE FINANCIAL MARKET CORNERS AND SHORT SQUEEZES INEFFICIENT?*

S. NAGARAJAN

FACULTY OF MANAGEMENT

MCGILL UNIVERSITY

*I thank Beth Allen, Jin Duan, Jon Faust, Larry Glosten, Ed Green, Rob Heinkel, Praveen Kumar, Pete Kyle, Suresh Sundaresan, S. Viswanathan, Neil Wallace, participants at the WFA meetings, Econometric Society Meetings, the Federal Reserve Bank of Minneapolis, and EFA meetings, and especially an anonymous referee and Franklin Allen (the Editor) for insightful comments and suggestions. Financial support of the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged. Any errors, of course, are mine.

ARE FINANCIAL MARKET CORNERS AND SHORT SQUEEZES INEFFICIENT?

ABSTRACT

This paper rigorously examines the prevalent belief that financial market corners and short squeezes reduce trading efficiency, especially when traders are privately informed. Explicit welfare criteria are proposed, and a trading model based on direct revelation mechanism design methodology is developed. It is shown that market corners and short squeezes are not *per se* inconsistent with trading efficiency if traders are sufficiently risk-averse. An impossibility theorem is then proved: The risk-aversion levels that give rise to corners in the first place can *never* be large enough to achieve efficiency starting from those corners. Increasing the supply of the traded security can restore efficiency, while neither limiting the short positions nor preventing the cornerer from squeezing the shorts improves trading welfare. These results hold for *any* set of prior beliefs, and for *any* trading game that shares the same preferences and information structure — direct or indirect, extensive-form or normal-form, with or without trading frictions.

1. Introduction

It has long been recognized that, in markets with finite or limited supply of assets, there may exist incentives for one or more parties to corner the market, in the hope of squeezing the less fortunate agents into paying exorbitant prices. Throughout history, there have been many allegations of corners and squeezes in various markets, the latest being the Salomon Brothers scandal (Gastineau and Jarrow 1991). Jegadeesh (1993) finds that the secondary market prices of two-year notes during the "Salomon squeeze" were higher than the estimated competitive prices in the four weeks after the issue. Cornell and Shapiro (1989) document an instance of mispricing of treasury bonds that strongly suggests a squeeze. According to Sundaresan (1992), however, squeezes in treasury markets have been far more common than widely believed, as reflected in the repo rates for on-the-run (roughly speaking, new) issues. Interestingly, squeezes are not necessarily illegal, except if there is collusion involved in creating a shortage of securities.

The objective of this paper is to examine rigorously the issue of whether corners and short squeezes are necessarily *inefficient* for trading welfare. The works cited above and other existing analyses assume implicitly that corners and short squeezes are generally welfare reducing. While this may be plausible, there is little theoretical analysis to support this view. Most of the existing arguments point to "artificially" higher prices during a squeeze to conclude that such a market condition must be "bad" for welfare. To be sure, abnormally high prices may indicate reduced total welfare if there is no private information (Jarrow 1992), or if there is a large number of noise traders in the market. Nevertheless, there are two reasons why such analyses are inadequate. First, as pointed out by Anderson (1984), trading prices, no matter how abnormally high, *per se* may or may not imply welfare reduction — since being side-payments, they may not always be relevant for (equally-weighted) aggregate welfare. Second, focusing on high prices alone ignores the welfare of the informed traders, and high prices may actually be a source of rents for private information.[1] Thus, any proper measure of social welfare in financial markets must take into account the welfare of *all* traders, whether they are informed or not.

This paper proposes the criterion of *incentive-constrained ex post efficiency* to evaluate the welfare implications of corners and short squeezes. This welfare criterion, common in

[1] In addition, if the welfare of informed traders is not properly accounted for, they will have less incentive to acquire costly information to startwith, resulting in reduced efficiency for the entire market (à la Grossman-Stiglitz 1980). In the interest of focus, this issue is not modeled here.

the literature on Bayesian games with asymmetric information (see Holmström and Myerson 1983), measures the equally-weighted, full-information welfare of all traders, subject to incentive compatibility and individual rationality constraints. Incentive-constrained ex post efficiency (henceforth, efficiency) implies first-best or Pareto-optimal level of allocations after all the trading is completed, even when privately informed traders behave strategically. Roughly speaking, it asks whether all trades that could have gone through under full information actually end up getting executed, even under asymmetric information. Thus, our efficiency criterion is related to the measures of information efficiency (eg. depth, tightness, etc.) originally suggested by Black (1971), and subsequently used widely in the microstructure literature. While ex post efficiency is a strong welfare measure, given our weights on traders' utilities here, it implies *interim* and *ex ante* efficiency as well. More importantly, it also implies *renegotiation-proofness*, which is a highly desirable property, since no unexploited gains to trade exist ex post, and hence traders have no incentives to retrade any further.

A model of trading based on the direct revelation mechanism approach (Fudenberg and Tirole 1992) is first developed. A key feature of this approach is that it permits *simultaneous* welfare analysis of a large family of dynamic extensive-form trading games with *multilateral* asymmetric information, by transforming them into more tractable static games called direct revelation games. This is in contrast to the standard practice of modelling only one, stylized extensive form trading game of unilateral asymmetric information, at a time. The issue then is whether ex post efficient (Pareto-optimal) allocations can result from trades involving corners and short squeezes in the Bayesian-Nash equilibrium. From a methodological viewpoint, the direct revelation mechanism design approach followed here has some advantages over the conventional rational expectations approach, in that it explicitly models the incentive constraints faced by informed traders, rather than assume that equilibrium prices convey information which the traders themselves would not voluntarily disclose in the first place.

Applying this trading model, the welfare analysis of corners and short squeezes is conducted in two stages: the first stage where trading from arbitrary endowments leads to a corner, and the second stage where trading starts from this corner. The second stage is analyzed first, and it is shown that efficiency can be obtained starting from a corner, but only if the traders' risk-aversion and the total supply of the security is sufficiently large — as the gains from risk-sharing and information-induced trading then dominate the adverse selection costs. The first stage trading is analyzed next, and the set of parameters for which corners occur in the first place is characterized. An impossibility theorem is proved, which shows that parameter values (eg. risk-aversion) for which a prior round of trading resulted in a corner are *never* large enough to achieve efficiency coming out of that corner. More generally, the direct revelation approach followed here assures that efficiency can never be achieved in *any* trading game — direct or indirect, extensive-form or normal-form, with or without trading frictions — that shares the same preferences and information structure, because the direct revelation game represents the upperbound in terms of trading efficiency. Finally, the impossibility result is robust in that it holds for any set of prior beliefs,[2] any exogenous short position limit, and also irrespective of whether the cornerer is allowed to squeeze the shorts.

Work relating to corners can be found in the mechanism design literature. For instance, Myerson and Satterthwaite (1983) showed that when two privately informed agents trade an indivisible asset (a natural corner), ex post efficiency is impossible. Extending this model to divisible assets, Cramton, Gibbons and Klemperer (1987) and McAfee (1991) show that Pareto-optimality can be achieved when traders start out with approximately equal endowments of the asset, but not when starting from corners. The Myerson-Satterthwaite and Cramton-Gibbons-Klemperer models involve risk-neutrality, *private* values, and no short-selling, and are not readily applicable to financial markets.

Related literature includes works by Allen and Gorton (1988), Allen and Gale (1991), Kyle (1984), Vila (1987), Kumar and Seppi (1989), and Seppi (1991), who study market manipulations in futures and other markets. Jarrow (1992) derives sufficient conditions on the price processes for manipulation strategies to exist, and finds that corners and squeezes always exist. Unlike the present paper, the focus in this literature is on the possibility of manipulations in rational expectations equilibria, and the existence of conditions for market manipulation strategies to be effective. The legal literature is mainly concerned with identifying and penalizing trading manipulation, and not with welfare (Johnson 1982). Finally, standard results from the auction literature do not apply, since the cornerer usually has some private information, a feature absent

[2]This property is important, since as Ledyard (1988) has shown, prior-dependence may render equilibrium predictions vacuous in Bayesian games.

in most auction models (McAfee and McMillan 1987).

The paper is organized as follows. Section 2 develops the direct revelation trading model, and characterizes the efficient trading mechanisms. This model is then used in Section 3 to study trading from a market corner. Reaching that corner is taken up in Section 4, and Section 5 concludes. Throughout, the results are illustrated with an example. Appendix A contains all the proofs, while Appendix B contains a treatment of ex ante efficiency.

2. The Trading Model

This section presents a simple model of trading between two *strategic* traders which is particularly suited for rigorous welfare analysis.[3] Each trader i starts with an initial endowment ω_i of a risky security, where $\sum_{i=1}^{2} \omega_i = \omega$ is the total supply of the risky security. To allow for market corners, the endowment ω_i could be such that $-\psi \leq \omega_i \leq \omega + \psi$, where $\psi \geq 0$ is an (exogenously set) total short position limit for the whole market.

The risky security to be traded is perfectly divisible, and its future value to the traders has two components: a value $\tilde{x}$ that is *common* to both the traders, and a value $\tilde{y}_i$ that is *private* (i.e., specific) only to trader i. While the common value aspect of a financial asset is better known, the private value component is nevertheless important. Private valuation of a trader may arise from, for instance, his/her personal consumption needs, hedging motives, portfolio positions or other (unmodeled) reasons specific to that trader. Consistent with this interpretation, assume that $\tilde{y}_1$, $\tilde{y}_2$, and $\tilde{x}$ are all *i.i.d.* It follows from the No-Trade Theorem of Milgrom and Stokey (1982) that the private values $\tilde{y}_i$ must be non-degenerate for non-zero trades to occur in equilibrium (see for instance, Glosten and Milgrom 1985, p. 77).[4] In this regard, the private valuation here plays a role similar to the heterogeneous expectations modeled by Kyle and Wang (1994) and others. For tractability, we assume that $\tilde{x} \sim \tilde{s}_1 + \tilde{s}_2 + \tilde{\epsilon}$, where $\tilde{s}_1$ and $\tilde{s}_2$ are *i.i.d.* random variables (interpreted below), and $\tilde{\epsilon}$ is such that $E[\tilde{\epsilon}|s_i] = 0$.

[3]Extending the model to include n strategic traders is straightforward but messy, and will actually increase efficiency (see also Cramton, Gibbons and Klemperer (1987)). The welfare implication of including *atomistic* traders in addition to strategic traders, however, is less clear and remains an open issue.

[4]Alternatively, introducing Kyle (1985)-type noise traders will also generate equilibrium trades. However, it then becomes difficult to make explicit welfare comparisons. Note that noise trading can be regarded as a special case of our formulation where $\tilde{y}_i = \infty$.

2.1 The Information Structure

Each trader finds out his/her *valuation* or type $\tilde{v}_i$ at the beginning of the game, where $\tilde{v}_i = \tilde{s}_i + \tilde{y}_i$. That is, each trader's private information v_i is a *combination* of two signals — one a signal s_i of the ultimate common value of the asset $\tilde{x}$, and the other his/her own private value $\tilde{y}_i$.[5] Thus, we model the more complex case of *two-sided* asymmetric information, in contrast to the one-sided asymmetric information models common in the microstructure literature. Note that, in contrast to models such as Kyle and Wang (1994) where the private value aspects (namely, the traders' heterogeneous beliefs) are common knowledge, here they are part of the traders' information set, thus allowing for richer interactions.

While v_i is known only to trader i, it is common knowledge that v_1 and v_2 are distributed i.i.d. as $G(\cdot)$, with a strictly positive, *symmetric* density function $g(\cdot)$ over the compact support $[\underline{v}, \bar{v}]$. In addition, the following hazard rate condition must be satisfied: $E(y_i|v_i) - [1 - G(v_i)]/g(v_i)$ and $E(y_i|v_i) + G(v_i)/g(v_i)$ are increasing in v_i.[6] We shall use the following expectations operators extensively: $E(\cdot)$ and $E_{v_j}(\cdot) = E_j(\cdot)$, which represent the unconditional and conditional (on v_i) expectations respectively.

2.2 Traders' Preferences

The traders are risk-averse towards holding the risky asset but are risk-neutral with respect to money, and their utility function is given by

$$u_i[q_i, p_i|\tilde{y}_i, \tilde{x}] \quad = \quad (\tilde{x} + \tilde{y}_i)q_i - \frac{\rho}{2}q_i^2 + p_i \tag{2.1}$$

where q_i is the quantity (volume) of the risky asset holdings, p_i is any (cash) payment received by trader i, ρ is an index of risk aversion, $\tilde{x}$ is the common value, and $\tilde{y}_i$ is the private value to trader i. The risk-aversion is the same for both traders and is assumed to be common knowledge.[7]

---

[5]If the private and common valued components are observed separately by the traders, it follows from Milgrom and Stokey (1982) that welfare can be further improved by trading on the private values alone.

[6]This is a bit stronger than the usual hazard rate condition, but weaker than the requirement that $[1 - G(v_b)]/g(v_b)$ and $G(v_s)/g(v_s)$ are monotone. It is satisfied, for instance, if $\tilde{y}_i$ and $\tilde{s}_i$ are uniformly or normally distributed.

[7]If risk-aversion is also private information, then the type-space becomes multi-dimensional and the mechanism design problem becomes quite complex. See McAfee and McMillan (1988).

Non-satiation requires that $\rho \leq 2$, a condition that is assumed throughout. The traders trade directly with each other, and hence $\sum p_i = 0$ and $\sum q_i = \omega$.[8]

The expected utility of trader i, conditional on private information v_i, can be written as

$$U_i[q_i, p_i | v_i] = E[v_i + s_j + \epsilon | v_i] q_i - \frac{\rho}{2} q_i^2 + p_i. \tag{2.2}$$

Such a utility function, sometimes referred to as the *Marshallian* utility function, is common in the speculation literature (Stein 1987). It is also in the spirit of the mean-variance value functions used in many models (see Duffie and Jackson (1989) for instance). For expositional ease, the quadratic functional form presented in (2.2) is used throughout the paper. However, all the results remain valid for more general *quasi-linear* utility functions of the form, $U_i[q_i, p_i | v_i] = V_i(q_i, v_i) + p_i$, where $V_i(q_i, v_i)$ is concave in q_i, and satisfies a generalized single crossing property.[9] A key feature of such utility functions, including the quadratic version shown in (2.2), is that they imply risk-aversion towards holding the risky security, but risk-neutrality with respect to money. The chief advantage of quasi-linear utility is that the traders' welfare can be readily aggregated, thus allowing for explicit welfare comparisons.[10] Furthermore, the utility function (2.2) is separable in information and risk-aversion, resulting in separation between risk-sharing and information-induced trades. Extending the model to include more general preferences is discussed in Section 5.

2.3 The Direct Revelation Mechanism Methodology

The traders in this model can play any game among a large family of dynamic trading games. These games may possibly involve many complex rules, intricate sequence of moves, several rounds of bidding and asking, and various types of trading strategies — some of which may even include cornering the market and squeezing the shorts. Our aim here is to avoid assuming a specific extensive form trading game, lest the results become too sensitive to the assumptions

[8]This would also be true on average even if the traders trade through a competitive, risk-neutral market maker. We abstract from market making issues that have been extensively analyzed elsewhere in the literature.

[9]See Fudenberg and Tirole (1992), p. 263, for details. Additional assumptions on higher order derivatives may also be needed.

[10]A similar specification can be obtained from the usual combination of exponential utility and normal distributions, but by using a product (vs. additive) form of the welfare function.

of the extensive form game itself. On the other hand, analyzing all possible dynamic trading games will be a formidable task. Fortunately, in the absence of message space restrictions, the *revelation principle* (Dasgupta, Hammond, and Maskin 1979 and Myerson 1979) enables us to collapse all these dynamic, extensive-form games involving various trading strategies into a single static game called the *direct revelation* game, and consider without loss of generality only incentive compatible (truthful) and individually rational (voluntary participation) *direct revelation* trading mechanisms. These mechanisms effectively duplicate (for analytical purposes) the equilibrium outcomes of the original extensive-form trading games, and they work as follows.[11]

Each trader confidentially reports his/her valuation to a disinterested coordinator, who then determines the outcome of the direct trading game, called the mechanism, which consists of the quantity and the payment terms of the trade as functions of the traders' reports.[12] The insight of the revelation principle is that, subject to resource constraints, traders can be bribed to tell the truth, thus offsetting their incentives to misrepresent and profit from their private information. If a particular allocation obtains in the equilibrium of the original extensive form game, then exactly the same allocation can be duplicated in the equilibrium of a corresponding direct revelation game, by offering the traders appropriate informational rents.

To put our methodology in perspective, the direct revelation mechanism design approach offers a number of advantages over the conventional ones followed in the literature. First, in contrast to rational expectations models, our approach explicitly incorporates the traders' incentive constraints, rather than presume that equilibrium prices reveal information which the traders themselves may not voluntarily disclose in the first place. This distinction is especially important when traders may not be "informationally small" (in the sense of Gul and Postlewaite 1992) — as is the case with a trader cornering the market. Second, while the extensive form games of Kyle (1985), Glosten and Milgrom (1985), Easley and O'Hara (1987), Diamond and Verrecchia (1984) and others have shed valuable insight into price formation issues (such as the rate at which traders' private information is incorporated into security prices), they were not

[11]When there are restrictions on the space of messages that traders can communicate with — for instance, those induced by the trading mechanism itself — the revelation principle may not be directly applicable, and the analysis must take these restrictions into account (Reichelstein and Reiter 1988).

[12]This coordinator merely serves as a communication device, and has no economic role. In particular, he need not be a market maker, as his function can also be done by a machine.

specifically designed to study the welfare properties of different trading mechanisms, as they usually take the extensive form game as given. The power of the revelation principle is that it narrows the search for efficient mechanisms among countless dynamic extensive form games of varying (and possibly infinite) dimensions, to the space of static direct games of uniform dimension. This effectively permits *simultaneous* analyses of a broad class of trading games, in contrast to the standard practice of analyzing only one extensive-form game at a time.

Third, although we do not explicitly model "real-world" extensive-form trading games, the direct revelation approach followed here assures that the equilibrium allocations of direct revelation games represent the *upperbound* of all allocations that can be achieved in these indirect trading games (Fudenberg and Tirole 1992). In particular, if full efficiency cannot be achieved in direct revelation games, then it can *never* be achieved in any trading game that shares the same preferences and information structure — direct or indirect, extensive-form or normal-form, with or without frictions. Finally, in contrast to the literature, the direct revelation approach also enables us to model *multi-lateral* asymmetric information trading problems, which are virtually intractable when modeled as extensive-form signaling games (Mailath 1989). In short, while differing in its focus, the direct revelation mechanism design approach can be viewed as a useful complement to the important extensive-form approaches followed in the microstructure literature.[13]

2.4 Incentive Compatible & Individually Rational Trading Mechanisms

Applying the revelation principle, we now focus on direct revelation games. Recall that a direct revelation game is a trading game in which each trader confidentially reports his/her valuation v_i to a disinterested coordinator. The coordinator then determines the trading mechanism μ, which consists of the quantity allocation q_i and the payment p_i made to each trader i — as functions of the traders' reports. Formally, for any exogenous short position limit ψ, a direct trading mechanism $\mu = \langle q, p|\psi\rangle$ can be defined as

- $q_i(v_1, v_2)$ is the quantity or volume of the asset allocated to trader i, where $\sum_{i=1}^{2} q_i(v_1, v_2) = \omega$. The trading volume is then $|\omega_i - q_i(v_1, v_2)|$, and

[13] A drawback of the direct revelation approach is that dynamic issues (eg. price formation) cannot be easily addressed. Moreover, once an optimal direct revelation mechanism is identified, finding the equivalent indirect mechanism may be difficult.

- $p_i(v_1, v_2)$ is the total payment made to trader i, s.t. $\sum_{i=1}^{2} p_i(v_1, v_2) = 0$. This payment will be negative if the trader ends up buying instead of selling. The (per share) price is $p_i(v_1, v_2)/[\omega_i - q_i(v_1, v_2)]$. Throughout, we will find it easier to use the total payment $p_i(v_1, v_2)$, rather than the share price.

The following redefinitions will be useful:

$$\begin{aligned} Q_i(v_i) &= E_j[q_i(v_i, v_j)] \\ S_i(v_i) &= E_j[(s_j + \epsilon) q_i(v_i, v_j)] \\ R_i(v_i) &= E_j[\frac{\rho}{2} q_i^2(v_i, v_j)] \\ P_i(v_i) &= E_j[p_i(v_i, v_j)] \end{aligned} \tag{2.3}$$

forall $v_i \in [\underline{v}, \bar{v}]$ where $i \in \{1, 2\}$. The expected utility to trader i of type v_i, who reports $\hat{v}_i$ instead, can then be written as

$$U_i[\hat{v}_i | v_i] \quad = \quad v_i Q_i(\hat{v}_i) + S_i(\hat{v}_i) - R_i(\hat{v}_i) + P_i(\hat{v}_i)$$

where $Q_i(\cdot)$ is the expected quantity of the asset holdings, $S_i(\cdot)$ is the expected value of the holdings from the unknown components of the common value, $R_i(\cdot)$ is the expected risk-premium, and $P_i(\cdot)$ is the expected transfer payment to trader i — all conditional on his/her valuation v_i, and that he/she reports $\hat{v}_i$.

Since the traders may not voluntarily reveal their private information truthfully, the mechanism must take this incentive into account. A direct trading mechanism is (Bayesian) *incentive compatible* if each trader truthfully reports his/her valuation in a Bayesian-Nash equilibrium. Such a mechanism elicits truth-telling by offering the traders appropriate rents for their private information, offsetting their incentives to lie and misrepresent their valuations. As will be seen later, this is the source of adverse selection costs in trading.

Formally, the trading mechanism $\mu = \langle q, p | \psi \rangle$ is incentive compatible (IC) iff

$$U_i[v_i] \quad \equiv \quad U_i[v_i | v_i] \quad \geq \quad v_i Q_i(\hat{v}_i) + S_i(\hat{v}_i) - R_i(\hat{v}_i) + P_i(\hat{v}_i) \tag{2.4}$$

$\forall v_i, \hat{v}_i \in [\underline{v}, \bar{v}]$, where v_i is the true valuation and $\hat{v}_i$ is the reported valuation. Using standard techniques [see Fudenberg and Tirole (1992)], the incentive compatible mechanisms can be characterized as follows.

LEMMA 2.1. *The trading mechanism $\mu = \langle q, p|\psi\rangle$ is incentive compatible iff $Q_i(\cdot)$ is non-decreasing, $U_i[\cdot]$ is convex with $U_i'[\cdot] = Q_i(\cdot)$, a.e., and*

$$U_i[v_i] = U_i[v_i^m] + E_{u_i}\left[\frac{Q_i(u_i)}{g(u_i)} 1_{\{v_i^m \leq u_i \leq v_i\}}\right], \tag{2.5}$$

where $v_i^m \in [\underline{v}, \bar{v}]$ is the type of the trader i who receives the lowest net utility.

In order to induce each trader to participate in the trading game, *individual rationality* or voluntary participation conditions must also be satisfied. To this end, note that if trader i does not participate in the trading game, his/her expected payoff will be $\omega_i v_i + \omega_i E(s_j + \epsilon) - \frac{\rho}{2}\omega_i^2$. Thus, individual rationality (IR) is satisfied if

$$U_i[v_i] \geq \omega_i v_i + \omega_i E(s_j + \epsilon) - \frac{\rho}{2}\omega_i^2 \qquad \forall\ v_i \in [\underline{v}, \bar{v}].$$

The IR conditions will be binding only for some minimum net utility type $v_i^m(q_i)$, however. To see that there exists a $v_i^m(q_i)$, note from Lemma 2.1 that $U_i[\cdot]$ is convex and non-decreasing. Hence, $U_i[v_i] - \omega_i v_i - \omega_i E(s_j + \epsilon) + \frac{\rho}{2}\omega_i^2$ is also convex in v_i, and has a minimum at some $v_i^m(q_i) \in [\underline{v}, \bar{v}]$. Since $U_i[\cdot]$ is monotone, the IR conditions reduce to

$$U_i[v_i^m] \geq \omega_i v_i^m(q_i) + \omega_i E(s_j + \epsilon) - \frac{\rho}{2}\omega_i^2 \qquad \forall\ v_i \in [\underline{v}, \bar{v}]. \tag{2.6}$$

The minimum utility conditions have the following interpretation (Cramton, Gibbons and Klemperer 1987). The trader with valuation $v_i < v_i^m(q_i)$ will act like a potential seller, whereas the one with $v_i > v_i^m(q_i)$ will act like a buyer. At $v_i = v_i^m(q_i)$, he/she will be indifferent between buying and selling and does not expect to trade.[14] Note that the minimum utility itself will depend on the quantity allocation schedule $q_i(\cdot)$.

We will find it easier to substitute out $q_2(v)$ by noting that $q_2(v) = \omega - q_1(v)$. The following lemma characterizes the set of IC and IR trading mechanisms.

[14] Of course, not expecting to trade is riskier than staying out of the trading game. This risk is taken care of in the pricing mechanism, by offering a risk-premium. I am grateful to Larry Glosten for pointing this out.

LEMMA 2.2. *The trading mechanism $\mu = \langle q, p|\psi\rangle$ is incentive compatible and individually rational iff $Q_i(\cdot)$ is non-decreasing and*

$$\begin{aligned}\mathcal{W}[\rho,\omega|G,\psi] \quad &\equiv \quad E\Bigg\{ y_1 q_1(v) + y_2[\omega - q_1(v)] - \frac{\rho}{2}q_1^2(v) - \frac{\rho}{2}[\omega - q_1(v)]^2 - \frac{[1-G(v_1)]}{g(v_1)}q_1(v) \\ &\quad - \frac{[1-G(v_2)]}{g(v_2)}[\omega - q_1(v)] + \frac{q_1(v)}{g(v_1)}1_{\{v_1 \le v_1^m(q_1)\}} + \frac{[\omega - q_1(v)]}{g(v_2)}1_{\{v_2 \le v_2^m(q_2)\}}\Bigg\} \\ &\quad - \omega_1 v_1^m(q_1) - \omega_2 v_2^m(q_2) + \frac{\rho}{2}\omega_1^2 + \frac{\rho}{2}\omega_2^2 \\ &\geq \quad 0 \end{aligned} \tag{2.7}$$

and if there exists a payment rule $p_i(\cdot,\cdot)$ s.t.

$$\begin{aligned} p_i(v_1,v_2) = \kappa_i - \int_{\underline{v}}^{v_i} u_i dQ_i(u_i) + \int_{\underline{v}}^{v_j} u_j dQ_j(u_j) - \int_{\underline{v}}^{v_i} dS_i(u_i) + \int_{\underline{v}}^{v_j} dS_j(u_j) \\ + \int_{\underline{v}}^{v_i} dR_i(u_i) - \int_{\underline{v}}^{v_j} dR_j(u_j) \end{aligned} \tag{2.8}$$

for appropriately chosen constants κ_i.

The first two terms in (2.7) represent the expected value of the allocations from the mechanism. The next two terms represent the risk-premia for the traders. The next four terms represent the adverse selection costs of implementing incentive compatible mechanisms. They imply that in the absence of incentive compatibility, the buyer tends to shade his/her bid downwards and the seller tends to shade his/her ask upwards, in an attempt to capture informational rents. As will be seen later, this incentive to misreport drives a wedge between the bids and asks, that may restrict trades even when the buyer's true valuation exceeds the seller's true valuation. As implied by Milgrom-Stokey, the common value terms s_i do not generate any gains, and drop out of (2.7). Indirectly, however, they play an important informational role, since $v_i = y_i + s_i$. The last four terms represent the status quo utility for the traders. To summarize, condition (2.7) says that for the trading mechanism to be incentive compatible and individually rational, the cumulative welfare of the traders under the mechanism net of the informational rents, must be at least as great as the status quo utility of the "worst" type of both the traders.

As mentioned earlier, the equilibrium price is a side-payment between the traders, and hence will not be unique in general.[15] However, the *functional* form of the payment must satisfy (2.8) for the trading mechanism to be IC and IR. Furthermore, it is important to note from (2.8) that the volume of asset holdings q_i is a sufficient statistic for the payment rule $p_i(\cdot,\cdot)$ in equilibrium.

2.5 Incentive-Constrained Efficiency

An efficient mechanism maximizes the weighted sum of the traders' expected utilities. A trading mechanism can be *ex ante*, *interim* or *ex post* efficient, depending on the information used to condition the traders' expected utilities (Holmström and Myerson 1983). Accordingly, an ex ante efficient mechanism maximizes the weighted sum of traders' expected utilities, computed before they find out their private information; an interim efficient mechanism, when each trader knows only his/her private information; and an ex post efficient mechanism, when all the information is common knowledge.

In general, ex post efficiency is a strong requirement, followed by interim and ex ante efficiency. Nevertheless, we will be interested in the criterion of ex post efficiency in this paper, for five reasons. First, for symmetry, we weight the traders' expected utilities equally. Since the relative weights on the traders' utilities are the same across each vector of types, it can be shown that every ex post efficient mechanism here is also interim and ex ante efficient (Green 1995). This is important, because it ensures that the ex post efficient mechanism will be preferred by the traders not only ex post, but also at the ex ante and interim stages as well. Second and perhaps more important, ex post efficient mechanisms are also strongly *renegotiation-proof*, in the following sense: Even if the traders have the option of retrading ex post, they will not do so, as there will be no more gains to trade left to be exploited.[16] Third, ex post efficiency, if achievable, is also consistent with Coase's (1960) conjecture. Fourth, as Appendix B shows, the set of ex ante efficient, incentive feasible mechanisms is non-empty, and hence the question of

[15] This is the reason behind Jarrow's (1992) result that restrictions on price processes can not rule out corners.

[16] Of course, there are weaker renegotiation-proof concepts available (Maskin and Tirole 1992), but these tend to be very sensitive not only to the distributional assumptions made but also to the type of equilibrium refinement used (eg. Cho and Kreps 1987). They are less suitable for our purpose, as we are interested in proving robust results here.

whether ex ante efficiency can be achieved is not very interesting. Finally, ex post efficiency is consistent with the focus in the finance literature.

Formally, an ex post efficient (a.k.a. Pareto optimal or first-best or full-information) trading mechanism maximizes the total ex post gains to trade for every $\{v_1, v_2\}$.[17] That is, the ex post efficient mechanism maximizes the equally-weighted ex post utilities of the traders, or

$$\max_{\{\mu\}} \quad E(y_1|v_1)q_1(v) + E(y_2|v_2)[\omega - q_1(v)] - \frac{\rho}{2}q_1^2(v) - \frac{\rho}{2}[\omega - q_1(v)]^2 - \sum_{i=1}^{2}\omega_i v_i + \frac{\rho}{2}\sum_{i=1}^{2}\omega_i^2.$$

Using the first-order condition, ex post efficient mechanism is given by[18]

$$\begin{aligned} q_1^{FB}(v) &= \frac{\omega}{2} + q_1^+(v) + q_1^-(v), \qquad \text{where} \\ q_1^+(v) &= \min\left\{\frac{\omega}{2}, \frac{1}{2\rho}[E(y_1|v_1) - E(y_2|v_2)]^+\right\} \qquad \text{and} \\ q_1^-(v) &= \max\left\{-\frac{\omega}{2}, \frac{1}{2\rho}[E(y_1|v_1) - E(y_2|v_2)]^-\right\}, \end{aligned} \tag{2.9}$$

along with the pricing rule (2.8). Note that since we are interested in resolving the corners at the end of the game, the security allocations at that time can not involve short positions, and hence $q_i \leq \omega$. Solution (2.9) identifies two *separate* components to the first-best allocation of asset holdings — namely, the risk-sharing component $\omega/2$, and the information-induced holdings $q_1^+(v)$ and $q_1^-(v)$. The separation between risk-sharing and information-induced holdings is a consequence of the separable preference structure in (2.1).

To see how the valuations affect trading under the first-best solution, note from (2.9) that if $E(y_1|v_1) \geq E(y_2|v_2)$, then $q_1^-(v) = 0$ and $q_1^+(v) > 0$, and trader 1 will end up with more of the risky asset than what is optimal from a purely risk-sharing viewpoint (i.e., $\omega/2$), upto a maximum of ω. On the other hand, if $E(y_1|v_1) < E(y_1|v_2)$, then $q_1^+(v) = 0$ and $q_1^-(v) < 0$, and trader 1 will hold less than the pure risk-sharing component $\omega/2$, down to a minimum of zero. Observe also that, the information-induced trading volume is piecewise *linear* and non-decreasing in the differences in conditional expectation of the valuations $|E(y_1|v_1) - E(y_2|v_2)|$,

[17] Ex post efficiency here is defined for all $\{v_i\}$ instead of $\{y_i, s_i\}$, since no one observes y_i and s_i separately, even ex post.

[18] It is easy to verify that the second-order conditions are satisfied.

and non-increasing in the risk-aversion parameter ρ. The following example illustrates the first-best mechanism.

2.6 A Uniform Prior Example: Ex Post Efficient Trading Mechanisms

Assume that the traders' valuations are i.i.d. uniformly over the interval $[0,1]$. For simplicity, since the common values do not generate any trade, assume that they are zero. The ex post efficient mechanism is given by

$$\begin{aligned} q_1^{FB}(v) &= \frac{\omega}{2} + q_1^+(v) + q_1^-(v), \qquad \text{where} \\ q_1^+(v) &= \min\left\{\frac{\omega}{2}, \frac{1}{2\rho}[v_1 - v_2]^+\right\} \qquad \text{and} \\ q_1^-(v) &= \max\left\{-\frac{\omega}{2}, \frac{1}{2\rho}[v_1 - v_2]^-\right\}, \\ p_i(v_1, v_2) &= \kappa_i - \frac{1}{8\rho}\left[v_2^2 - v_1^2\right] + \left[\omega - \frac{1}{2\rho}\right]\frac{[v_1 - v_2]}{4}. \ \blacksquare \end{aligned} \tag{2.10}$$

Incentive-constrained ex post efficiency is defined to be achieved by an ex post efficient trading mechanism if it satisfies the IC and IR constraints. Thus, in order to verify whether incentive-constrained ex post efficiency (henceforth, simply efficiency) can be achieved, one must verify whether the first-best mechanism in (2.9) is incentive feasible, i.e., whether it satisfies condition (2.7). This is, in fact, the technique used in proving the main results in this paper.

If, on the other hand, ex post efficiency can not be obtained, then a useful welfare concept is incentive-constrained *ex ante* efficiency. A full treatment of ex ante efficiency is given in Appendix B. The following proposition characterizes the conditions under which ex post efficiency can be achieved.

PROPOSITION 2.1. *First-best level of allocations can be achieved, as long as the initial endowments are centered around $\omega/2$.*

The proof is a straightforward extension of that of Proposition 1 (pp. 628-629) in Cramton, Gibbons and Klemperer (1987) and is omitted. Intuitively, if a trader does not know for sure if he/she is going to be a buyer or a seller, his/her incentive to underbid if he/she wishes to buy (overbid if he/she wishes to sell) is significantly reduced, since shading his/her bid may cause him/her to end up selling (buying) instead of buying (selling). This works as long as the initial endowments are approximately evenly distributed as each trader is approximately equally likely to buy or sell, insuring that both traders have roughly equal market power.

While extremely uneven distribution of *initial* endowments may be rare, the efficiency result of Proposition 2.1 does not shed much light on corners and squeezes, however. To see this, recall that our direct revelation game is a collapsed version of some indirect extensive form game(s). Thus, Proposition 2.1 implies that efficiency can eventually be achieved starting with roughly even initial endowments in indirect games involving possibly many rounds of trading. The difficulty is, there is no way of knowing whether these intermediate trading rounds involve corners or short squeezes.

The welfare effects of corners and short squeezes can be directly addressed by modeling a two-stage game as follows. In the first stage, traders start with arbitrary initial endowments of the security, and in the process of trading, possibly arrive at a corner. The second-stage trading game will then proceed from the corner. Section 3 analyzes the second stage trading first, followed by Section 4, which examines the first stage.

3. Efficiency of Trading Starting From a Corner

Assume that trader 1 has cornered the market to startwith, i.e., $\omega \leq \omega_1 \leq \omega + \psi$, and that trader 2 starts with a short position, i.e., $0 \geq \omega_2 \geq -\psi$. Without loss of generality, the analysis is carried out for $\omega_1 = \omega + \psi$ and $\omega_2 = -\psi$. In this section, it is assumed (unrealistically) that the cornerer is not allowed to call in the shorts by charging arbitrary (high) prices. That is, we assume that both incentive compatibility and individual rationality must be satisfied for the cornered trader to close out his short position. Later, Section 3.1 considers the more realistic case of the cornerer calling in the shorts. There are two reasons for taking this approach. First, the realistic case turns out to be a special case of the analysis in this section. Second, such an approach can answer questions such as: Is squeezing the shorts a source of inefficiency? What is the welfare implication of a regulatory policy preventing the cornerer from calling in the shorts?

Before proceeding formally, some intuition is in order. Recall from (2.7) that the overall welfare in a trading game depends on three factors: a) welfare improvement from pure risk-sharing, b) the gains from trading due to differences in valuations between the traders, and c) the cost of informational rents extracted by the traders. The separation result in (2.9) would seem to suggest that pure risk-sharing incentive would level out the endowments even if one of the traders started out with a corner position. Presumably, one can then apply the risk-neutrality result of Cramton, Gibbons and Klemperer (1987) (Proposition 1, pp. 628-629) to conclude that first-best can always be achieved. However, as (2.9) shows, the valuation-induced trading itself will *decrease* under risk-aversion, and adverse selection costs may limit risk-sharing. Hence, it is not obvious which of the three effects dominate. As mentioned earlier, in order to verify whether ex post efficiency can be achieved, one can substitute the first-best solution $q_1^{FB}(v)$ into the constraint (2.7) and check whether the latter is still valid. The following lemma characterizes the conditions under which trading can lead to efficient allocations when starting from a corner.

LEMMA 3.1. *Starting from a corner, ex post efficiency can be achieved iff the expected net trading gains $\mathcal{W}[\rho,\omega|G,\psi]$ are non-negative:*

$$\begin{aligned}\mathcal{W}[\rho,\omega|G,\psi] &= \frac{\omega}{2}E[V(v)] + E[V(v)(q_1^+(v) + q_1^-(v))] + \frac{\rho\omega^2}{4} - E\left[\frac{\omega}{g(v_2)}\right] \\ &\quad - \rho E[q_1^+(v) + q_1^-(v)]^2 - \omega E[s_2] - \psi[\bar{v} - \underline{v}] + \rho\psi(\omega + \psi) \\ &\geq 0, \end{aligned} \tag{3.1}$$

where

$$V(v) = E(y_1|v_1) - E(y_2|v_2) + \frac{G(v_1)}{g(v_1)} + \frac{1 - G(v_2)}{g(v_2)}.$$

It is straightforward to show that $\mathcal{W}[\rho,\omega|G,\psi]$ is continuous and differentiable in ρ everywhere except possibly at $\hat{\rho} = [E(y_1|\bar{v}) - E(y_2|\underline{v})]/\omega$, which is a set of measure-zero. A simple extension of Myerson and Satterthwaite (1983) implies that $\mathcal{W}[0,\omega|G,\psi] < 0$. We are now ready to prove the following.

THEOREM 3.1. *For every trading game starting from a corner, there exists a critical risk-aversion index $\rho^*(\psi)$, such that ex post efficiency can be achieved for all $\rho \geq \rho^*(\psi)$.*

A discussion of this result is postponed until after the analysis of the more realistic case of the cornerer calling in the shorts.

3.1 Calling in the Shorts

We now assume that the cornerer can call in the shorts. The trading process starting from a corner can then be broken up into two parts. In the first part, the cornerer calls in the shorts.

Given the existing institutional setup for short-calls, the price charged is arbitrary, subject to wealth constraints, and other constraints (if any) imposed by the exchange. Importantly, this price need not (and typically will not) be either incentive compatible or individually rational for the cornered trader. Fortunately, the short-calling part of trading need not be modeled, since such prices are sidepayments and hence irrelevant for our aggregate welfare. Moreover, the trade in the first part, i.e., the closing out of the short positions, is by fiat, and does not affect efficiency.

In the second part, the traders play the trading game described in Section 3 — in which the cornerer, having brought her position down from ω_1 ($\geq \omega$) to ω, now starts with ω, and the cornered trader, having closed out his short position, now starts with zero quantity of the security. Thus, without loss of generality, trader 1 can be assumed to start with ω, and trader 2 assumed to start with no security in the ensuing trading game. Since no information is released in the first part of trading, the prior beliefs for the second part remain the same as before. Any feasible trading mechanism must satisfy incentive compatibility and individual rationality for both traders in this second part. The question then is whether ex post efficiency can be achieved. The answer is obtained simply by noting that the second part of this game is a special case of the game analyzed in Section 3, with $\psi = 0$. Thus, the following corollary is immediate from Theorem 3.1.

COROLLARY 3.2. *If the cornerer is allowed to call in the shorts, starting from a corner, all traders with risk-aversion levels greater than $\rho^*(0)$ will achieve ex post efficiency. In particular, short squeezes have no welfare effect for these ranges of risk-aversion.*

Thus, there is an efficient way out of corners if the risk-aversion of the traders is sufficiently large. But Theorem 3.1 and Corollary 3.2 say something stronger: If traders with a given level of risk-aversion can achieve effiency starting from a corner situation, then more risk-averse traders will also achieve ex post efficiency. The intuition for this result is as follows. Generally speaking, corners can lead to ex post efficiency as long as the gains from risk-sharing and information-induced trading dominate the adverse selection costs. High risk-aversion levels induce more risk-sharing relative to the adverse selection costs, and help achieve ex post efficiency. This, in turn, depends on the parameters of the market, such as the risk-aversion of the traders ρ, the distributional characteristics of the traders' private information $G(\cdot)$, the total endowment of the asset ω, and when the cornerer is not allowed to call in the shorts, also on ψ. Note that in the realistic case when the cornerer can call in the shorts, the short position limit ψ is irrelevant for efficiency, since the 'real' game starts only after the short position is closed out by fiat.

For a given distribution G and the supply of the security ω, the set of all risk-aversion levels for which trading from a corner leads to ex post efficiency can be written as

$$\mathcal{C} = \{\rho \mid \mathcal{W}[\rho,\omega|G,\psi] \geq 0\}.$$

The set of efficient risk-aversion levels starting from a corner $\mathcal{C}$ can be shown to be convex and bounded. The properties of this efficient set will become clear below, where we discuss an example. Roughly speaking, starting from corners, ex post efficiency can be obtained if the combination of traders' risk-aversion ρ and the total endowment of the security ω is large enough.

3.2 The Uniform Prior Example: Efficiency Starting from a Corner

Consider the uniform prior example of Section 2.1. Substituting the first-best mechanism (2.10) into (3.1), the net expected trading gains from Proposition 3.1 reduce to

$$\mathcal{W}[\rho,\omega] = -\frac{\omega}{6} + \frac{1}{8\rho}\min\{\omega^4\rho^4, 1\} - \frac{\omega}{3}\min\{\omega^3\rho^3, 1\} + \frac{\omega^2\rho}{4}\min\{\omega^2\rho^2, 1\} - \psi[1 - \rho(\omega + \psi)].$$

For a given ψ and low risk-aversion levels ρ the net trading welfare remains negative, but becomes positive for larger values of ρ.[19] A similar effect is seen for the total endowment of the security ω. The set of risk-aversion levels that can achieve efficiency starting from a corner in the realistic case is obtained by setting $\psi = 0$:

$$\mathcal{C} = \{\rho \mid \omega\rho \geq 1 + (1/2)^{1/2}\},$$

and is displayed in Figure 1 as a function of ω. As mentioned earlier, it is convex and bounded. Hence, ex post efficiency can be achieved starting from a corner, except for small levels of risk-aversion and the supply of the security.

[19] The relevant interval for ρ is $[0;2]$, given the non-satiation constraint.

Finally, the result of Myerson and Satterthwaite (1983) and Cramton, Gibbons and Klemperer (1987), *viz.*, the inefficiency trading from corners under *risk-neutrality*, can be obtained as a special case of our model, by setting $\rho = 0, \quad \psi = 0$, and $\omega = 1$. The net welfare in this case is

$$\mathcal{W}[\rho = 0, \omega = 1] \quad = \quad -\frac{1}{6} \quad < \quad 0. \quad \blacksquare$$

4. Reaching a Corner

In the last section, it was demonstrated that traders can achieve ex post efficiency starting from a corner, as long as their risk-aversion is large enough. In this section, we ask whether a prior round of trading will give rise to a corner in the first place, for those levels of risk-aversion. Intuition suggests that highly risk-averse traders are not likely to corner the market in the first place, but is it possible to have moderate levels of risk-aversion that not only result in a corner but also help achieve efficiency starting from one?

This issue can be addressed by explicitly modeling the two-stage game described earlier. In the first stage, traders start with arbitrary initial positions, and in the process of trading, possibly arrive at a corner. Then the second stage game starting from the corner will proceed exactly as in Section 3. Such an analysis will in general be complicated, involving many Perfect Bayesian Equilibria. However, it can be shown that the equilibrium of the first-stage may generally involve pooling equilibria (Fudenberg and Tirole 1992). If, on the other hand, a separating equilibrium results at the end of the first stage, then fresh (private) information may arrive at the beginning of the second stage. In either case, the prior beliefs for both stages will be on the support $[\underline{v}, \bar{v}]$. Since our objective here is to prove an impossibility result, this observation leads to the following simpler approach: Rather than modeling the first stage game explicitly in this section, we simply characterize the *upperbound* of parameter values for which a corner can be reached. We then compare this set of parameters to the set of parameters derived earlier in Theorem 3.1 and Corollary 3.2 for which efficiency can be achieved starting from this corner.

To this end, consider the following modification of the first-best solution $q_1^{FB}(v)$ in (2.9) to allow for corners to be reached. Starting from any arbitrary initial position $\{\omega_1, \omega_2\}$, the first-best solution involves the corner, $q_1^{FB} = \omega + \psi$, iff

$$\begin{aligned} \omega + \psi &\leq \frac{\omega}{2} + \frac{1}{2\rho}[E(y_1|v_1) - E(y_2|v_2)] \quad \text{or} \\ \rho &\leq \frac{[E(y_1|v_1) - E(y_2|v_2)]}{\omega + 2\psi}. \end{aligned} \tag{4.1}$$

for some $\{v_1, v_2\}$. On the other hand, the outcome of a first-best solution can never be a corner, if

$$\rho \quad > \quad \hat{\rho}(\psi) \quad \equiv \quad \frac{[E(y_1|\bar{v}) - E(y_2|\underline{v})]}{\omega + 2\psi}. \tag{4.2}$$

In this case, corners will *never* be reached through trading in the first place. Denote by C' the set of risk-aversion levels ρ that satisfies (4.1), and hence can lead to a corner. Recall from Theorem 3.1 that, starting from a corner all $\rho \geq \rho^*(\psi)$ will achieve efficiency. The following result compares the risk-aversion levels ρ in (4.1) with those in Theorem 3.1 and Corollary 3.2.

THEOREM 4.1. *$C \cap C' = \phi$. That is, the set of risk-aversion levels for which a previous round of trading can lead to a corner C', and the set C which achieves efficiency starting from the corner are mutually exclusive — whether the cornerer is allowed to call in the shorts or not, and irrespective of the short position limits ψ, or the distribution of prior beliefs G.*
Thus, no two-stage trading game exists, where starting from arbitrary endowments, a corner is reached at the end of the first stage, and ex post efficiency is ultimately achieved at the end of the second stage.

REMARK: The proof technique involves using uniform distribution as the upperbound for all other distributions, and then applying Rothschild and Stiglitz's (1970) notion of increasing risk.

Thus, there is both good news and bad news about market corners. The good news, according to Theorem 3.1and Corollary 3.2, is that corners *per se* are not bad for welfare, and if the corner positions were due to exogenous reasons (as opposed to resulting from a previous round of trading), they can lead to ex post efficient allocations, as long as the traders' risk-aversion is sufficiently large. These exogenous factors could include, for instance, the case of an entrepreneur making an initial public offering (IPO), a government agency/firm being privatized, the treasury department auctioning a new debt instrument, or a firm inheriting a large block of securities in another company as a result of a merger or an acquisition. The bad news conveyed by Theorem 4.1 is that if the corner resulted from a previous round of trading, then its very occurrence may indicate that traders were not sufficiently risk-averse in the first place, and that the cornerer is now more likely to be interested in extracting rents for his information than in risk-sharing. This may lead to reduced trading in the second stage, resulting in less than full efficiency. However, it is important to note from Theorems 3.1 and 4.1 that the real source of inefficiency is not the squeezing of the shorts, as the problem remains even in the absence of short positions. Rather, the inefficiency results from the combination of large positions and low risk- aversion levels that enables the cornerer to restrict trades in the second stage.[20]

The mechanism design approach followed here assures that since efficiency cannot be achieved starting from a trading-induced corner in direct revelation games, it can *never* be achieved in any trading game that shares the same preferences and information structure — direct or indirect, extensive-form or normal-form, with or without frictions (Fudenberg and Tirole 1992). This is because the direct revelation games represent the upperbound of efficiency for a given environment (eg. preferences, information structure), and frictions, if anything, reduce efficiency. Theorem 4.1 is also independent of the prior beliefs of the traders or any short-position limits set by the regulator. In fact, the result is also independent of whether or not the cornerer is allowed to call in the shorts. This is illustrated below by the uniform prior example discussed earlier.

4.1 The Uniform Prior Example: Efficiency To and From a Corner

For $\omega = 1$, Figure 2 shows the sets of risk-aversion levels (as functions of short-position limits) that help achieve efficiency when trading towards as well as away from corners — when the cornerer is not allowed to call in the shorts. Figure 3 illustrates the corresponding sets when the cornerer can call in the shorts. In both figures, the bottom hatched portion represents the set of risk-aversion levels for which a corner can be reached through trading, whereas the top hatched portion represents those risk-aversion levels that allow efficiency to be obtained starting from a corner. If the cornerer is not allowed to call in the shorts (Figure 2), the minimum (critical) risk-aversion for achieving efficiency starting from a corner decreases with short position limits.

[20]Note that it is not necessary to consider second-best solutions to the first-stage trading problem for the following reason. If, as shown, the first-best solution does not result in corners for the range of risk-aversion levels of interest, then cornering is even less likely in the second-best solution, as trading is always more restricted in the latter case.

By contrast, if she is allowed to call in the shorts, this minimum risk-aversion remains the same (Figure 3). Nevertheless, this does not result in an overlap of C and C' in either case — although they come closer for less restrictive short position limits in the former case — since the maximum risk-aversion for which a corner can be reached in the first place also decreases with short position limits. In both cases, the "gap" between the two hatched portions denotes risk-aversion levels for which a corner can never be reached through trading, and even if trading had been started (exogenously) from a corner, efficiency can never be obtained. ∎

5. Conclusion

This paper has presented a simple, but rigorous, model of trading in financial markets based on the direct revelation mechanism design methodology. Using incentive-constrained ex post efficiency as the welfare criteria, it has been shown that corners and short squeezes *per se* are not necessarily inefficient. However, if the corners were reached through previous rounds of trading, then trading from these corners and any short squeezes will *never* be efficient. Given the mechanism design methodology, this impossibility result holds for *any* type of trading game that shares the same preferences and information structure — direct or indirect, extensive form or normal form, with or without frictions. The results are also robust in the sense that they are independent of the distributional assumptions, and more importantly, of any exogenous short sale or short-calling restrictions.

Since trading-induced corners are inefficient, the likelihood of occurence of corners in our model can be reduced *ex ante*, by starting out with a larger supply of the security.[21] As seen from equation (4.2), a large ex ante supply reduces the set of risk-aversion levels for which corners can be reached in a previous round of trading, *ceteris paribus*. Increasing the supply *ex post*, i.e., after a corner has been reached, can improve welfare when the cornerer is prevented from calling in the shorts (cf. equation (3.1)). In the more realistic case, offering the cornered trader (ex post) an additional supply exceeding his short positions may achieve the same result.

The result that short position limits do not improve trading welfare suggests that such measures may be of limited value. Moreover, a comparison of Theorem 3.1 and Corollary 3.2 suggests that efficiency can not be improved by preventing the cornerer from calling in the

[21] This argument applies only to cases where the increase in supply does not affect the "intrinsic" value of the security.

shorts. Finally, the importance of risk-aversion in reaching and getting out of corners efficiently suggests that when-issued markets, where dealers can pre-hedge their positions, may have the unintended effect of reducing the dealers' *effective* risk-aversion when bidding in the primary auction. A lower effective risk-aversion may give the dealers an incentive to corner the primary issue, and to engage in short squeezes in the secondary markets.

As mentioned earlier, the model presented in this paper contains simplifying assumptions for the sake of tractability and exposition. These suggest a number of directions in which the model can be extended. First is to introduce more general preferences which may result in the traders' welfare not being easily aggregated. One solution then is to follow Mirlees' (1971) and Wilson's (1992) indirect approach in applying Green's theorem. The qualitative results derived in this paper are likely to continue to hold under more general preferences as well, although, being an indirect approach, the role of risk-aversion will be less transparent. A second possible extension could relax the assumption that the utility function is separable in risk-aversion and information. More generally, one would expect the risk-aversion to decrease as the information becomes more favorable.[22] Thus, the likelihood of getting into corners may increase with favorable information, but decrease with unfavorable information. Once the corner is reached, however, the lower risk-aversion would make it more difficult to eventually obtain efficiency. Hence, the impossibility result is likely to continue to hold. A third extension could include multiple traders, some of whom may be strategic and others atomistic and non-strategic. Finally, the criteria of incentive-constrained efficiency and the mechanism design trading model developed in this paper are well-suited for analyzing welfare issues in a host of specific markets (eg. treasury, futures) as well.

[22] I am grateful to an anonymous referee for pointing this out.

Appendix A

Proof of Lemma 2.2. The total expected utility to both the traders is,

$$
\begin{aligned}
E\Big[E(y_2|v_2)q_2(v) &- \frac{\rho}{2}q_1^2(v) - \frac{\rho}{2}q_2^2(v) + x + E(y_1|v_1)q_1(v)\Big] \\
&= E_2[U_2(v_2)] + E_1[U_1(v_1)] \\
&= U_2[v_2^m] + E_{v_2}\left[E_{u_2}\left[\frac{Q_2(u_2)}{g(u_2)}1_{\{v_2^m \le u_2 \le v_2\}}\right]\right] + U_1[v_1^m] + E_{v_1}\left[E_{u_1}\left[\frac{Q_1(u_1)}{g(u_1)}1_{\{v_1^m \le u_1 \le v_1\}}\right]\right] \\
&\quad \text{(from (2.5), since } \mu \text{ is incentive compatible)} \\
&= U_2[v_2^m] + U_1[v_1^m] + E_{u_2}\left[\frac{[1-G(u_2)]}{g(u_2)}Q_2(u_2)\right] + E_{u_1}\left[\frac{[1-G(u_1)]}{g(u_1)}Q_1(u_1)\right] \\
&\quad - E_{u_2}\left[\frac{Q_2(u_2)}{g(u_2)}1_{\{\underline{v} \le u_2 \le v_2^m\}}\right] - E_{u_1}\left[\frac{Q_1(u_1)}{g(u_1)}1_{\{\underline{v} \le u_1 \le v_1^m\}}\right] \\
&\quad \text{(by changing the order of expectations)} \\
&= U_2[v_2^m] + U_1[v_1^m] + E_2\left[\frac{[1-G(v_2)]}{g(v_2)}E_1[q_2(v)]\right] + E_1\left[\frac{[1-G(v_1)]}{g(v_1)}E_2[q_1(v)]\right] \\
&\quad - E_2\left[\frac{E_1[q_2(v)]}{g(u_2)}1_{\{\underline{v} \le u_2 \le v_2^m\}}\right] - E_1\left[\frac{E_2(q_1(v)]}{g(u_1)}1_{\{\underline{v} \le u_1 \le v_1^m\}}\right] \\
&= U_2[v_2^m] + U_1[v_1^m] + E\left[\frac{[1-G(v_2)]}{g(v_2)}q_2(v)\right] + E\left[\frac{[1-G(v_1)]}{g(v_1)}q_1(v)\right] \\
&\quad - E\left[\frac{q_2(v)}{g(u_2)}1_{\{\underline{v} \le u_2 \le v_2^m\}}\right] - E\left[\frac{q_1(v)}{g(u_1)}1_{\{\underline{v} \le u_1 \le v_1^m\}}\right] \qquad \text{(A1)}
\end{aligned}
$$

The last equality follows from the law of iterated expectations. Rearranging the left hand side and the right hand side of (A1) and imposing the IR constraints (2.6) yields (2.7).

To prove the if part, let $Q_i(\cdot)$ be non-decreasing, and let condition (2.7) hold. We can construct a pricing rule, such that μ is IC and IR, as follows.

$$
\begin{aligned}
p_i(v_1, v_2) = \kappa_i - \int_{\underline{v}}^{v_i} u_i dQ_i(u_i) + \int_{\underline{v}}^{v_j} u_j dQ_j(u_j) - \int_{\underline{v}}^{v_i} dS_i(u_i) + \int_{\underline{v}}^{v_j} dS_j(u_j) \\
+ \int_{\underline{v}}^{v_i} dR_i(u_i) - \int_{\underline{v}}^{v_j} dR_j(u_j) \qquad \text{(A2)}
\end{aligned}
$$

where κ_i are appropriately chosen constants. It is straightforward to verify that this mechanism is IC and IR using standard techniques (see MS Theorem 1, pp. 270-271), and we omit this part here. ∎

Proof of Lemma 3.1: First, note that $\omega_1 = \omega + \psi$ and $\omega_2 = -\psi$ for a corner imply that $v_1^m(q_1) = \bar{v}$ and $v_2^m(q_2) = \underline{v}$ respectively. Rewriting (2.7),

$$\begin{aligned}
\mathcal{W}[\rho,\omega|G] &= \omega E[E(y_2|v_2)] + E\Big\{\Big[E(y_1|v_1) - E(y_2|v_2) + \rho[\omega - q_1(v)]\Big]q_1(v)\Big\} \\
&\quad + E\left[\frac{G(v_1)}{g(v_1)}q_1(v)\right] - \omega E\left[\frac{[1-G(v_2)]}{g(v_2)}\right] + E\left[\frac{[1-G(v_2)]}{g(v_2)}q_1(v)\right] - \omega\bar{v} \\
&\quad - \psi[\bar{v} - \underline{v}] + \psi\rho(\omega+\psi) \\
&= E\Big\{\Big[E(y_1|v_1) - E(y_2|v_2) + \rho(\omega - q_1(v))\Big]q_1(v)\Big\} + E\left[\frac{G(v_1)}{g(v_1)}q_1(v)\right] \\
&\quad - \omega E\left[\frac{1}{g(v_2)}\right] + E\left[\frac{[1-G(v_2)]}{g(v_2)}q_1(v)\right] - \omega E[s_2] \\
&\quad - \psi[\bar{v} - \underline{v}] + \psi\rho(\omega+\psi) \\
&\quad \left(\text{since} \quad \omega E[E(y_2|v_2)] + \omega E\left[\frac{G(v_2)}{g(v_2)}\right] - \omega\bar{v} = -\omega E[s_2]\right) \\
&= E\Big\{\Big[E(y_1|v_1) - E(y_2|v_2) + \rho[\omega - q_1(v)] + \frac{G(v_1)}{g(v_1)} + \frac{[1-G(v_2)]}{g(v_2)}\Big]q_1(v)\Big\} \\
&\quad - \omega E\left[\frac{1}{g(v_2)}\right] - \omega E[s_2] - \psi[\bar{v} - \underline{v}] + \psi\rho(\omega+\psi).
\end{aligned}$$

Substituting $q_1^{FB}(v)$ from (2.9),

$$\begin{aligned}
\mathcal{W}[\rho,\omega|G,\psi] &= E\Big\{\Big[E(y_1|v_1) - E(y_2|v_2) + \frac{G(v_1)}{g(v_1)} + \frac{[1-G(v_2)]}{g(v_2)}\Big]\frac{\omega}{2} \\
&\quad + \Big[E(y_1|v_1) - E(y_2|v_2) + \frac{G(v_1)}{g(v_1)} + \frac{[1-G(v_2)]}{g(v_2)}\Big]\Big[q_1^+(v) + q_1^-(v)\Big] \\
&\quad + \rho\left[\frac{\omega}{2} + q_1^+(v) + q_1^-(v)\right]\left[\frac{\omega}{2} - q_1^+(v) - q_1^-(v)\right] - \frac{\omega}{g(v_2)} - \omega E[s_2|v_2]\Big\} \\
&\quad - \psi[\bar{v} - \underline{v}] + \psi\rho(\omega+\psi),
\end{aligned}$$

and the result follows. ∎

Proof of Theorem 3.1: First note that $\mathcal{W}[\rho]$ is differentiable everywhere except possibly at $\hat{\rho} = [E(y_1|\bar{v}) - E(y_2|\underline{v})]/\omega$. From (3.1),

$$\begin{aligned}
\mathcal{W}[\rho,\omega|G,\psi] =& \frac{\omega}{2}E[V(v)] + E[V(v)(q_1^+(v) + q_1^-(v))] + \frac{\rho\omega^2}{4} - E\left[\frac{\omega}{g(v_2)}\right] \\
& - \rho E[q_1^+(v) + q_1^-(v)]^2 - \omega E[s_2] - \psi[\bar{v} - \underline{v}] + \psi\rho(\omega+\psi)
\end{aligned}$$

$$
\begin{aligned}
=&\frac{\omega}{2}E[V(v)] + 2E[V(v)q_1^+(v)] + \frac{\rho\omega^2}{4} - E\left[\frac{\omega}{g(v_2)}\right] - 2\rho E[q_1^+(v)]^2 - \omega E[s_2] \\
&- \psi[\bar{v} - \underline{v}] + \psi\rho(\omega + \psi)
\end{aligned}
$$

(since the distribution G is symmetric)

$$
\begin{aligned}
=&\frac{\omega}{2}[\bar{v} - \underline{v}] + 2E\Bigg\{\left[E(y_1|v_1) - E(y_2|v_2) + \frac{G(v_1)}{g(v_1)} + \frac{1 - G(v_2)}{g(v_2)}\right] \\
&\min\left[\frac{\omega}{2}, \frac{[E(y_1|v_1) - E(y_2|v_2)]^+}{2\rho}\right]\Bigg\} + \frac{\rho\omega^2}{4} - \omega[\bar{v} - \underline{v}] \\
&- 2\rho E\Bigg\{\min\left[\frac{\omega^2}{4}, \frac{[E(y_1|v_1) - E(y_2|v_2)]^{+2}}{4\rho^2}\right]\Bigg\} - \omega E[s_2] \\
&- \psi[\bar{v} - \underline{v}] + \psi\rho(\omega + \psi) \\
=& -\frac{\omega}{2}[\bar{v} - \underline{v}] + \omega E\Bigg\{\Big[E(y_1|v_1) - E(y_2|v_2)\Big].\min\left[1, \frac{[E(y_1|v_1) - E(y_2|v_2)]}{\rho\omega}\right]1_{\{v_1 > v_2\}}\Bigg\} \\
&+ \omega E\Bigg\{\left[\frac{G(v_1)}{g(v_1)} - \frac{G(v_2)}{g(v_2)}\right]\min\left[1, \frac{[E(y_1|v_1) - E(y_2|v_2)]}{\rho\omega}\right]1_{\{v_1 > v_2\}}\Bigg\} + \frac{\rho\omega^2}{4} \\
&- \frac{\rho\omega^2}{2}E\Bigg\{\min\left[1, \frac{[E(y_1|v_1) - E(y_2|v_2)]^2}{\rho^2\omega^2}\right]1_{\{v_1 > v_2\}}\Bigg\} - \omega E[s_2] \\
&- \psi[\bar{v} - \underline{v}] + \psi\rho(\omega + \psi) \\
=& -\frac{\omega}{2}[\bar{v} - \underline{v}] + \omega E\Bigg\{\Big[E(y_1|v_1) - E(y_2|v_2)\Big]\min\left[\frac{[E(y_1|v_1) - E(y_2|v_2)]}{\rho\omega} - 1, 0\right]1_{\{v_1 > v_2\}}\Bigg\} \\
&+ \omega E\Bigg\{\left[\frac{G(v_1)}{g(v_1)} - \frac{G(v_2)}{g(v_2)}\right]\min\left[\frac{[E(y_1|v_1) - E(y_2|v_2)]}{\rho\omega} - 1, 0\right]1_{\{v_1 > v_2\}}\Bigg\} + \frac{\rho\omega^2}{4} \\
&- \frac{\rho\omega^2}{2}E\Bigg\{\min\left[\frac{[E(y_1|v_1) - E(y_2|v_2)]^2}{\rho^2\omega^2} - 1, 0\right]1_{\{v_1 > v_2\}}\Bigg\} \\
&+ \omega E\Bigg\{\Big[E(y_1|v_1) - E(y_2|v_2)\Big]1_{\{v_1 > v_2\}}\Bigg\} + \omega E\Bigg\{\left[\frac{G(v_1)}{g(v_1)} - \frac{G(v_2)}{g(v_2)}\right]1_{\{v_1 > v_2\}}\Bigg\} \\
&- \frac{\rho\omega^2}{2}E\Bigg\{1_{\{v_1 > v_2\}}\Bigg\} - \omega E[s_2] - \psi[\bar{v} - \underline{v}] + \psi\rho(\omega + \psi) \\
=& -\frac{\omega}{2}[\bar{v} - \underline{v}] + \omega E\Bigg\{\Big[E(y_1|v_1) - E(y_2|v_2)\Big]\left[\frac{[E(y_1|v_1) - E(y_2|v_2)]}{\rho\omega} - 1\right]1_{\{v_1, v_2\}}\Bigg\} \\
&+ \omega E\Bigg\{\left[\frac{G(v_1)}{g(v_1)} - \frac{G(v_2)}{g(v_2)}\right]\left[\frac{[E(y_1|v_1) - E(y_2|v_2)]}{\rho\omega} - 1\right]1_{\{v_1, v_2\}}\Bigg\} + \frac{\rho\omega^2}{4} \\
&- \frac{\rho\omega^2}{2}E\Bigg\{\left[\frac{[E(y_1|v_1) - E(y_2|v_2)]^2}{\rho^2\omega^2} - 1\right]1_{\{v_1, v_2\}}\Bigg\} \\
&+ \omega E\Bigg\{\Big[E(y_1|v_1) - E(y_2|v_2)\Big]1_{\{v_1 > v_2\}}\Bigg\} + \omega E\Bigg\{\left[\frac{G(v_1)}{g(v_1)} - \frac{G(v_2)}{g(v_2)}\right]1_{\{v_1 > v_2\}}\Bigg\} \\
&- \frac{\rho\omega^2}{2}E\Bigg\{1_{\{v_1 > v_2\}}\Bigg\} - \omega E[s_2] - \psi[\bar{v} - \underline{v}] + \psi\rho(\omega + \psi) \qquad \text{(A3)}
\end{aligned}
$$

where $1_{\{v_1,v_2\}} \equiv 1_{\{\min[E(y_1|\bar{v}),E(y_2|v_2)+\omega\rho]\geq E(y_1|v_1)>E(y_2|v_2)\}}$. Differentiating with respect to ρ (every where except at $\hat{\rho}$), yields,

$$
\begin{aligned}
\frac{\partial \mathcal{W}}{\partial \rho} =& -\omega^2 E\left\{\left[\frac{[E(y_1|v_1)-E(y_2|v_2)]^2}{\rho^2\omega^2}\right]1_{\{v_1,v_2\}}\right\} \\
& -\omega^2 E\left\{\left[\frac{G(v_1)}{g(v_1)}-\frac{G(v_2)}{g(v_2)}\right]\left[\frac{[E(y_1|v_1)-E(y_2|v_2)]}{\rho^2\omega^2}\right]1_{\{v_1,v_2\}}\right\}+\frac{\omega^2}{4} \\
& +\omega^2 E\left\{\left[\frac{[E(y_1|v_1)-E(y_2|v_2)]^2}{\rho^2\omega^2}\right]1_{\{v_1,v_2\}}\right\}+\psi(\omega+\psi) \\
& -\frac{\omega^2}{2}E\left\{\left[\frac{[E(y_1|v_1)-E(y_2|v_2)]^2}{\rho^2\omega^2}-1\right]1_{\{v_1,v_2\}}\right\}-\frac{\omega^2}{2}E\left\{1_{\{v_1>v_2\}}\right\} \\
=& -E\left\{\left[\frac{G(v_1)}{g(v_1)}-\frac{G(v_2)}{g(v_2)}\right]\left[\frac{[E(y_1|v_1)-E(y_2|v_2)]}{\rho^2}\right]1_{\{v_1,v_2\}}\right\}+\frac{\omega^2}{4}+\psi(\omega+\psi) \\
& -\frac{\omega^2}{2}E\left\{\left[\frac{[E(y_1|v_1)-E(y_2|v_2)]^2}{\rho^2\omega^2}-1\right]1_{\{v_1,v_2\}}\right\}-\frac{\omega^2}{2}E\left\{1_{\{v_1>v_2\}}\right\} \qquad \text{(A4)} \\
\geq & \ 0
\end{aligned}
$$

Since $\mathcal{W}[0,\omega|G,\psi]<0$ (Myerson and Satterthwaite 1983), the result follows by continuity. ∎

Proof of Theorem 4.1: Since $\mathcal{W}[\cdot]$ is monotone, it suffices to prove that $\mathcal{W}[\hat{\rho}]<0$, where $\hat{\rho}=[E(y_1|\bar{v})-E(y_2|\underline{v})]/(\omega+2\psi)$. For the sake of clarity, the proof is given for $\psi=0$, and the corresponding proof for $\psi>0$ is a straightforward (but messy) extension of what is given below.

Substituting $\hat{\rho}=[E(y_1|\bar{v})-E(y_2|\underline{v})]/\omega$ in (A3), and rearranging yields,

$$
\begin{aligned}
\mathcal{W}[\hat{\rho}] =& -\frac{\hat{\rho}\omega^2}{4}-\frac{\omega}{2}[\bar{s}-\underline{s}]+\frac{\hat{\rho}\omega^2}{2}E\left\{\left[\frac{[E(y_1|v_1)-E(y_2|v_2)]}{\hat{\rho}\omega}-1\right]^2 1_{\{v_1,v_2\}}\right\} \\
& +\omega E\left\{\left[\frac{G(v_1)}{g(v_1)}-\frac{G(v_2)}{g(v_2)}\right]\left[\frac{[E(y_1|v_1)-E(y_2|v_2)]}{\hat{\rho}\omega}-1\right]1_{\{v_1,v_2\}}\right\} \\
& +\omega E\left\{\left[E(y_1|v_1)-E(y_2|v_2)+\frac{G(v_1)}{g(v_1)}-\frac{G(v_2)}{g(v_2)}-\frac{\hat{\rho}\omega}{2}\right]1_{\{v_1>v_2\}}\right\}-\omega E[s_2] \qquad \text{(A5)} \\
=& -\frac{\hat{\rho}\omega^2}{4}-\frac{\omega}{2}[\bar{s}-\underline{s}]+T_1+T_2+T_3-\omega E[s_2], \qquad \text{(A6)}
\end{aligned}
$$

where T_1, T_2 and T_3 are used to denote the third, fourth and fifth terms in (A5), for ease of further analysis. Simplifying the terms T_1, T_2 and T_3 separately, we get

$$
T_1=\frac{\hat{\rho}\omega^2}{2}E\left\{[1-G(v_2)]\right\}+\frac{1}{2\hat{\rho}}E\left\{E(y_1|v_1)^2G(v_1)\right\}-\frac{1}{\hat{\rho}}E\left\{E(y_1|v_1)E(y_2|v_2)1_{\{v_1>v_2\}}\right\}
$$

$$- \omega E\Big\{E(y_1|v_1)G(v_1)\Big\} + \omega E\Big\{E(y_2|v_2)[1-G(v_2)]\Big\} + \frac{1}{2\hat{\rho}}E\Big\{E(y_2|v_2)^2[1-G(v_2)]\Big\}$$

$$T_2 = -\frac{\bar{v}}{\hat{\rho}}E\Big\{E(y_2|v_2)\Big\} + \frac{1}{\hat{\rho}}E\Big\{v_1 E(y_2|v_2)1_{\{v_1>v_2\}}\Big\} - \omega E\Big\{\frac{G^2(v_1)}{g(v_1)}\Big\}$$

$$+ \frac{1}{\hat{\rho}}E\Big\{v_2 E(y_1|v_1)1_{\{v_1>v_2\}}\Big\} + \frac{1}{\hat{\rho}}E\Big\{\frac{E(y_2|v_2)G(v_2)}{g(v_2)}\Big\} + \omega E\Big\{\frac{[1-G(v_2)]G(v_2)}{g(v_2)}\Big\}$$

$$T_3 = -\omega E\Big\{\frac{[1-G(v_2)]G(v_2)}{g(v_2)}\Big\} - \omega E\Big\{[s_1 - s_2]1_{\{v_1>v_2\}}\Big\} + \omega E\Big\{\bar{v} - v_2\Big\} - \frac{\hat{\rho}\omega^2}{2}E[1-G(v_2)].$$

Substituting T_1, T_2 and T_3 back into (A6) and simplifying,

$$\mathcal{W}[\hat{\rho}] = -\frac{\hat{\rho}\omega^2}{4} - \frac{\omega}{2}[\bar{s} - \underline{s}] - \omega E[s_1 + s_2] + \frac{1}{\hat{\rho}}E\Big\{\frac{G(v_1)G(v_2)}{g(v_1)g(v_2)}1_{\{v_1>v_2\}}\Big\}$$

$$- \frac{1}{\hat{\rho}}E\Big\{s_1 s_2 1_{\{v_1>v_2\}}\Big\} + 2\omega E[s_1 G(v_1)] + \frac{\bar{v}}{\hat{\rho}}E[s_2] - \frac{1}{2\hat{\rho}}E\Big\{v_2^2 - y_2^2\Big\}$$

$$- \frac{1}{\hat{\rho}}E\Big\{\frac{s_2 G(v_2)}{g(v_2)}\Big\} - \omega E\Big\{[s_1 - s_2]1_{\{v_1>v_2\}}\Big\}$$

$$\leq -\frac{\hat{\rho}\omega^2}{8} - \frac{\omega}{4}[\bar{s} - \underline{s}] + \frac{[\bar{s} - \underline{s}]^2}{8\hat{\rho}} - \frac{1}{\hat{\rho}}E\Big\{s_1 s_2 1_{\{v_1>v_2\}}\Big\} - \omega E\Big\{[s_1 + s_2]1_{\{v_2>v_1\}}\Big\}$$

$$+ \frac{\bar{v}}{\hat{\rho}}E[s_2] - \frac{1}{2\hat{\rho}}E\Big\{v_2^2 - y_2^2\Big\} - \frac{1}{\hat{\rho}}E\Big\{\frac{s_2 G(v_2)}{g(v_2)}\Big\} \qquad \text{(A7)}$$

$$< \quad 0,$$

where the inequality (A7) comes from the fact that

$$\int_{\underline{v}}^{v_1} G(v_2)dv_2 \quad < \quad \int_{\underline{v}}^{v_1} \frac{[v_2 - \underline{v}]}{[\bar{v} - \underline{v}]}dv_2$$

for any distribution $G(\cdot)$, as per Rothschild and Stiglitz (1970). ∎

Appendix B: Ex ante Efficient Trading Mechanisms

If incentive-constrained ex post efficiency is not obtained, then a useful welfare concept is ex ante efficiency (again, incentive-constrained). The trading mechanism that achieves ex- ante efficiency is given by the solution of the program

$$\mathbf{P}: \quad \max_{\{q(\cdot,\cdot)\}} E\left\{ y_1 q_1(v) + y_2[\omega - q_1(v)] - \frac{\rho}{2} q_1^2(v) - \frac{\rho}{2}[\omega - q_1(v)]^2 - \sum_{i=1}^{2} \omega_i v_i + \frac{\rho}{2} \sum_{i=1}^{2} \omega_i^2 \right\}$$

$$\text{s.t.} \quad \text{Constraint (2.7).}$$

The ex ante efficient solution to the incentive constrained program **P** is given by the following Lemma.

LEMMA B.1. *The ex ante efficient trading mechanism is*

$$\begin{aligned}
\mathbf{S}:\ q_1^*(v_1, v_2) &= \frac{\omega}{2} + q_1^{+*}(v) + q_1^{-*}(v) \quad \textit{where} \\
q_1^{+*}(v) &= \min\left\{ \frac{\omega}{2}, \frac{1}{2\rho}\left[E(y_1|v_1) - E(y_2|v_2) + \alpha\left[\frac{G(v_1)}{g(v_1)} - \frac{G(v_2)}{g(v_2)}\right] \right.\right. \\
&\qquad \left.\left. - \alpha\frac{1}{g(v_1)} 1_{\{v_1 > v_1^m\}} + \alpha \frac{1}{g(v_2)} 1_{\{v_2 > v_2^m\}} \right]^+ \right\} \quad \textit{and} \qquad \text{(B1)} \\
q_1^{-*}(v) &= \max\left\{ -\frac{\omega}{2}, \frac{1}{2\rho}\left[E(y_1|v_1) - E(y_2|v_2) + \alpha\left[\frac{G(v_1)}{g(v_1)} - \frac{G(v_2)}{g(v_2)}\right] \right.\right. \\
&\qquad \left.\left. - \alpha\frac{1}{g(v_1)} 1_{\{v_1 > v_1^m\}} + \alpha \frac{1}{g(v_2)} 1_{\{v_2 > v_2^m\}} \right]^- \right\}
\end{aligned}$$

where $\alpha = \lambda/(1+\lambda)$ is obtained by solving (2.7) as an equality, λ being the Lagrange multiplier for the constraint (2.7), and $0 \le \alpha \le 1$.

Proof of Lemma B.1: The proof of existence is similar to that in MS (Theorem 2, pp. 275-276), and hence we present only the characterization here. Ignoring the constants, rewrite Program **P** as

$$\max_{q_1(\cdot),\lambda} \quad \mathcal{L}(q_1, \lambda) = E\left\{ (1+\lambda)\left[y_1 q_1(v) + y_2[\omega - q_1(v)] - \frac{\rho}{2} q_1^2(v) - \frac{\rho}{2}[\omega - q_1(v)]^2 \right] - \lambda \left[\frac{[1 - G(v_1)]}{g(v_1)} q_1(v) \right.\right.$$

$$
- \frac{[1-G(v_2)]}{g(v_2)}[\omega - q_1(v)] + \frac{q_1(v)}{g(v_1)} 1_{\{v_1 \le v_1^m(q_1)\}} + \frac{[\omega - q_1(v)]}{g(v_2)} 1_{\{v_2 \le v_2^m(q_2)\}}
$$

$$
- \omega_1 v_1^m(q_1) - \omega_2 v_2^m(q_2) \Big] \Big] \Big\} \tag{B2}
$$

where λ is the Lagrange multiplier for the constraint (2.7). Define $\alpha = \dfrac{\lambda}{1+\lambda}$. Pointwise optimization of (B2) yields the first-order condition, which, upon rearranging yieds (B1). It is straightforward to verify that the second-order condition is satisfied (see Fudenberg and Tirole 1992, p. 263). ∎

Comparing (2.8) and (B1), note that the first-best solution corresponds to $\alpha = 0$, and the second-best solution **S** involves $\alpha \geq 0$. Note that α itself is a(n increasing) function of the *shadow price* λ of the IC & IR constraint (2.7). It depends on the parameters of the problem, including the distribution functions and ω. More importantly, α is also a measure of trading inefficiency. To see this, consider the following two cases for each trader:

(a) He/she expects to be a seller (i.e., $v_i < v_i^m$), in which case $\alpha > 0$ implies that he/she ends up holding *more* of the asset than what he/she would under the first-best solution,

(b) He/she expects to be a buyer (i.e., $v_i > v_i^m$), in which case $\alpha > 0$ implies that he/she ends up holding *less* of the asset than what he/she would under the first-best solution.

Both these cases are inefficient relative to the first-best, and hence a higher α *decreases* the total expected gains to trade.

References

Allen, F. and Gale, D. "Stock-price Manipulation," *Review of Financial Studies,* 5 (1992), 503-529.

Allen, F. and Gorton, G. "Stock price manipulation, market microstructure and asymmetric information," *European Economic Review*, 36 (1992), 624-630.

Anderson, R. "The Industrial Organization of Futures Markets: A Survey," in *The Industrial Organization of Futures Markets*, Lexington Books, Lexington, MA, 1984.

Black, F. "Towards a Fully Automated Exchange: Part I," *Financial Analysts Journal*, 27 (1971), 29-34.

Cho, I.-K. and Kreps, D. "Signalling Games and Stable Equilibria," *Quarterly Journal of Economics*, 102 (1987), 179-221.

Coase, R. "The Problem of Social Cost," *Journal of Law and Economics*, 3 (1960), 1-44.

Cornell, B. and Shapiro, A. "The Mispricing of U.S. Treasury Bonds: A Case Study," *Review of Financial Studies,* 2 (1989), 297-310.

Cramton, P., Gibbons, R. and Klemperer, P. "Dissolving a Partnership Efficiently," *Econometrica* (1987), .

Dasgupta, P., Hammond, P., and Maskin, E. "The Implementation of Social Choice Rules: Some results on Incentive Compatibility," *Review of Economic Studies*, (1979) 46, 185-216.

Diamond, D. and Verrecchia, R. "Information Aggregation in a Noisy Rational Expectations Economy," *Journal of Financial Economics*, 9, (1981), 221-235.

Duffie, D. and Jackson, M. "Optimal Innovation of Futures Contracts," *Review of Financial Studies,* 2 (1989), 275-296.

Easley, D. and O'Hara, M. "Price, Trade Size and Information in Securities Markets," *Journal of Financial Economics*, 19 (1987), 69-90.

Fudenberg, D. and Tirole, J. *Game Theory,* MIT press, Cambridge, (1992).

Gastineau, G. and Jarrow, R. "Large Trader Impact and Market Regulation," *The Financial Analysts Journal*, (1991), 47, 40-72.

Glosten, L. and Milgrom, P. (1985). "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders," *Journal of Financial Economics*, 14, 71-100.

Green, E. "Efficient Implementation of Financial Intermediation Mechanisms," Working paper, Federal Reserve Bank of Minneapolis, 1995.

Grossman, S. and Stiglitz, J. "On the Impossibility of Informationally Efficient Markets," *American Economic Review*, 70 (1980), 393-408.

Gul, F. and Postlewaite, A. "Asymptotic Efficiency in Large Exchange Economies with Asymmetric Information," *Econometrica*, 60 (1992), 1273-1292.

Jarrow, R. "Market Manipulations, Bubbles, Corners and Short Squeezes," Journal of Financial and Quantitative Analysis, (1992), 27, 311-336.

Jegadeesh, N. "Treasury Auction Bids and the Salomon Squeeze, " *Journal of Finance* (1993) 48, 1403-1419.

Johnson, P. *Commodities Regulation,* Boston: Little, Brown (1982).

Kumar, P. and Seppi, D. "Futures Manipulation with Cash Settlement," *Journal of Finance*, 1993.

Kyle, A. "A Theory of Futures Market Manipulations," in *The Industrial Organization of Futures Markets*, Lexington Books, Lexington, MA, 1984.

Kyle, A. and Wang, A. "Speculation Duopoly with Agreement to Disagree," Duke University Working paper, 1994.

Ledyard, J. "On the Scope of the Hypothesis of Bayesian Equilibrium," *Journal of Economic Theory*, 47 (1988), .

Mailath, G. "Bilateral Signalling," *Journal of Economic Theory*, 47 (1988).

Maskin, E. and Tirole, J. "Principal Agent Relationship with an Informed Principal II: Common Values," *Econometrica*, 60 (1992), 1-42.

McAfee, P. "Efficient Allocation with Continuous Quantities," *Journal of Economic Theory*, 53 (1991) 51-74.

McAfee, P. and McMillan, J. "Multidimensional Incentive Compatibility and Mechanism Design," *Journal of Economic Theory*, (1988), 46, 335-354.

McAfee, P. and McMillan, J. "Auctions and Bidding," *Journal of Economic Literature*, (1987), 25, 699-738.

Milgrom P. and Stokey, N. (1982). "Information, Trade and Common Knowledge," *Journal of Economic Theory*, 26, 17-27.

Mirlees, J. "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic*

Studies, (1971) 38, 175-208.

Myerson, R. "Incentive Compatibility and the Bargaining Problem," *Econometrica*, 47 (1979), 61-74.

Myerson, R. and Satterthwaite, M. "Efficient Mechanisms for Bilateral Trading," *Journal of Economic Theory* (1983).

Reichelstein, S. and Reiter, S. "Game Forms with Minimal Message Spaces," *Econometrica*, 56 (1988), 661-692.

Rothschild, M. and Stiglitz, J. "Increasing Risk I: A Definition," *Journal of Economic Theory*, 2 (1970), 225-243.

Seppi, D. "Equilibrium Block Trading and Asymmetric Information," *Journal of Finance*, (1990), 45, 73-94.

Stein, J. "Informational Externalities, and Welfare-Reducing Speculation," *Journal of Political Economy*, (1987), 95, 1123-1145.

Sundaresan, S. "An Empirical Analysis of U.S. Treasury Auctions: Implications for Auction and Term Structure Theories," Working paper, 1992.

Vila, J.-L. "The Role of Information in the Manipulation of Futures Markets," Working paper, 1987.

Wilson, R. "Design of Efficient Trading Procedures," in D. Friedman and J. Rust (eds.) *The Double Auction Market*, Reading, MA: Addison Wesley.

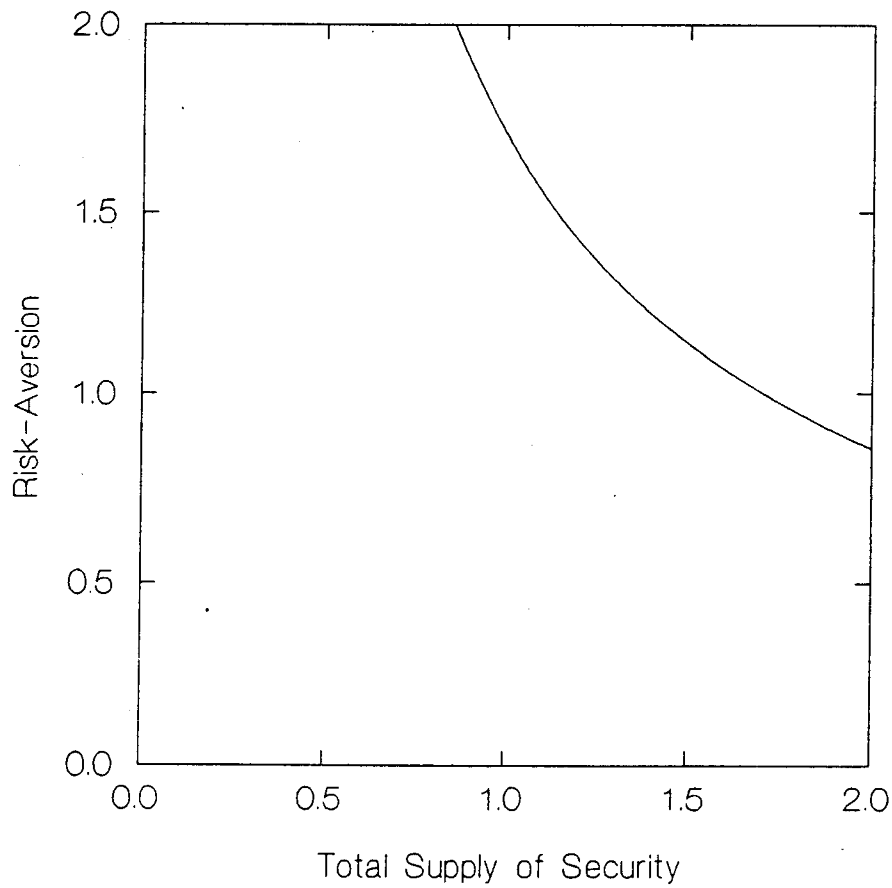

Figure 1. The minimum risk-aversion levels $\rho*$ needed to achieve efficiency starting from corners, plotted as a function of the total supply of security ω --- when the cornerer can call in the shorts.

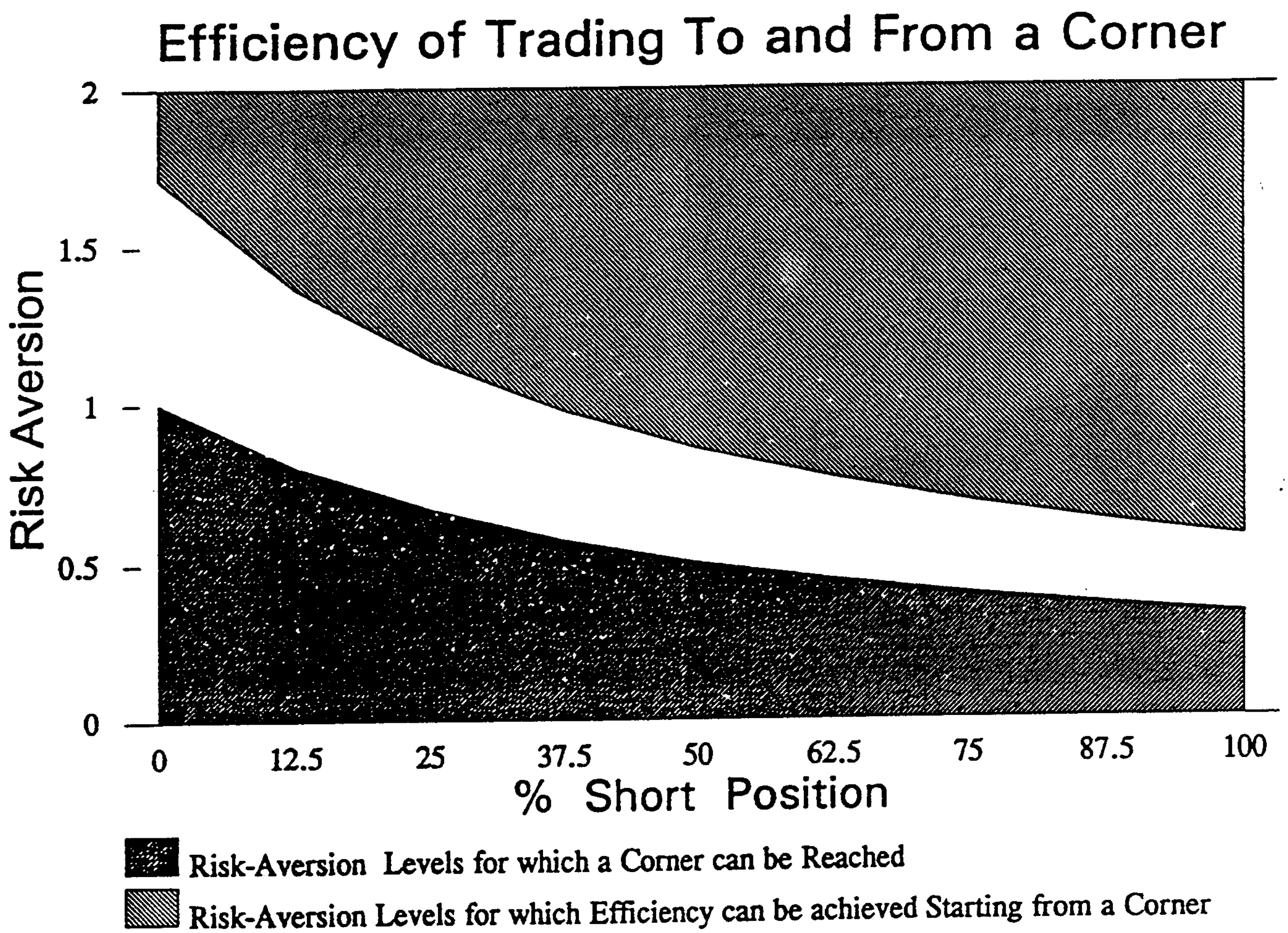

Figure 2. The bottom hatched portion represents the set of risk-aversion levels for which a corner can be reached, and the top hatched portion represents the set of risk-aversion levels for which efficiency can be achieved starting from corners --- when the cornerer *cannot* call in the shorts. Both are plotted as functions of the exogenously set short position limit for the market.

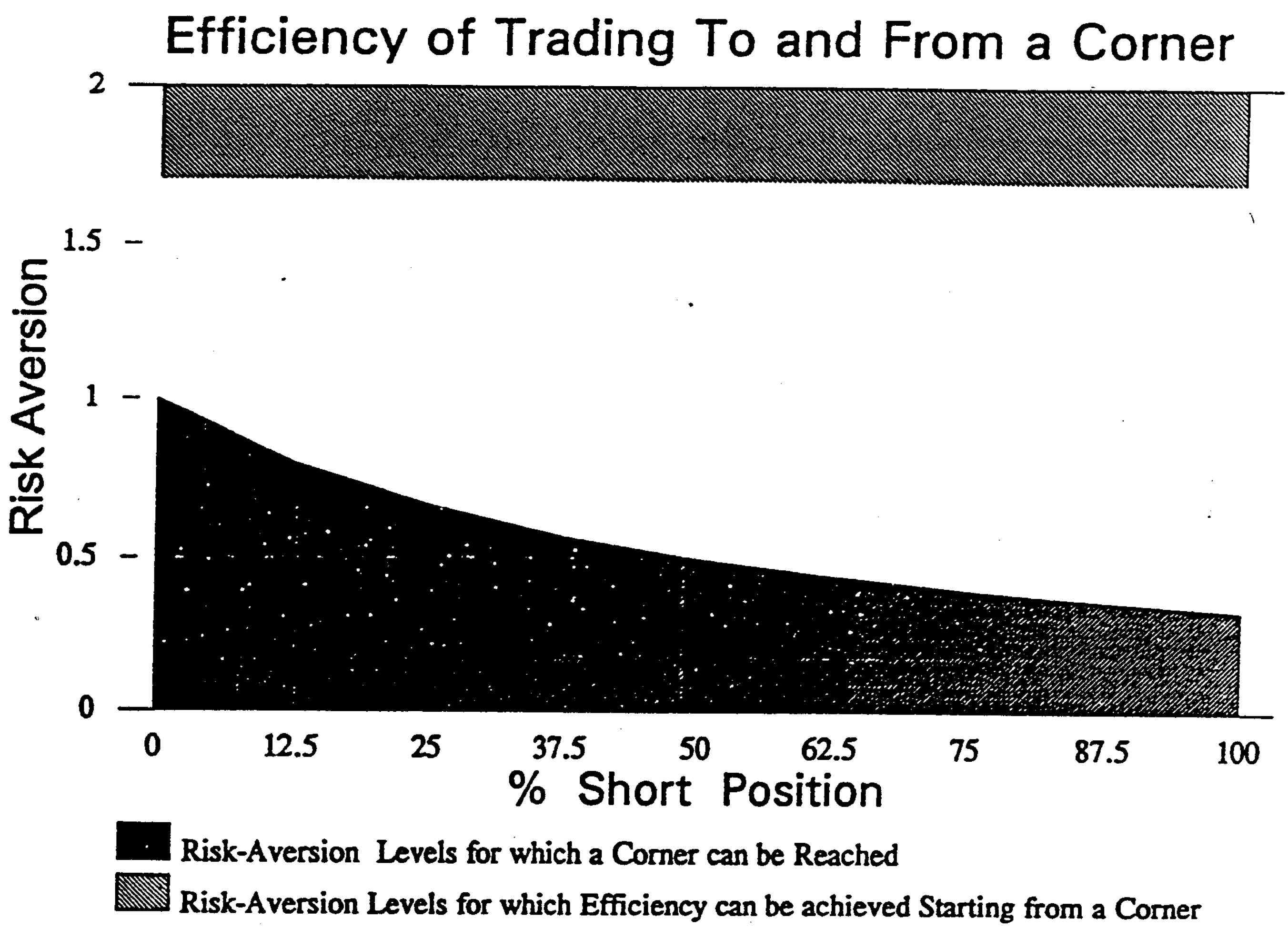

Figure 3. The bottom hatched portion represents the set of risk-aversion levels for which a corner can be reached, and the top hatched portion represents the set of risk-aversion levels for which efficiency can be achieved starting from corners --- when the cornerer can call in the shorts. Both are plotted as functions of the exogenously set short position limit for the market.