Frequent Trading and Price Impact in Thin Markets*

Marzena J. Rostek† and Marek Weretka‡

October 31, 2007

Abstract

Extensive empirical research has shown that large institutional investors have significant impact on prices and they mitigate an adverse effect through their trading strategies. We develop here a dynamic model with multiple assets wherein the key innovation is the traders' taking into account their impact on prices. The predictions match a number of empirical facts that are hard to reconcile with a competitive model: Optimal execution of trade by breaking up orders into blocks, asset price overshooting, as well as contemporaneous and dynamic relations between return, price impact, and trading volume. If a trade takes place more frequently than the dividend payments, the price impact is not stationary but instead decreases with the time to maturity. As with the presence of price impact the market and liquidation values do not coincide, we also study the valuation of assets in thin markets.

JEL Classification: D43, D53, G11, G12, L13.
Keywords: Institutional Traders, Asset Pricing, Thin Markets, Blockage Discount.

---

[*] First draft: September 2006. This paper combines the material from "Asset Valuation in Thin Financial Markets" by Marek Weretka and "The Time Structure of Price Impact" by Marzena J. Rostek and Marek Weretka. We would like to thank William Brock, Zhiwu Chen, Peter DeMarzo, Steven Durlauf, Frank Fabozzi, John Geanakoplos, Ananth Madhavan, Antonio Mello, Larry Samuelson, Elu Von Thaden, and Dimitri Vayanos for very helpful conversations. We are also grateful for comments and suggestions to seminar participants at Johns Hopkins, Mannheim (Finance), Penn State, and Wisconsin-Madison, and conference participants at the 2007 Winter Meeting of the Econometric Society in Chicago and the 2007 Summer Meeting of the Econometric Society at Duke. All errors are ours.

[†] University of Wisconsin-Madison, Department of Economics, 1180 Observatory Drive, WI 53706, Madison. E-mail: mrostek@ssc.wisc.edu.

[‡] University of Wisconsin-Madison, Department of Economics, 1180 Observatory Drive, WI 53706, Madison. E-mail: mweretka@ssc.wisc.edu.

It is now well documented that transactions by large institutional investors exert economically significant price impacts.[1] Because the stock positions of institutional traders constitute a sizable fraction of total trading, these trades often exceed the number of shares the market maker is willing to trade at the quoted bid and ask prices. As a result, the trades are likely to move prices in the direction of the trade, thereby adversely affecting the terms of trade.

In practice, extensive resources are devoted to developing trading strategies that can mitigate price impact. One common practice involves breaking up orders into *blocks*, which are then traded sequentially. The empirical finance literature has reported even in markets as deep as the NYSE, a typical institutional package will represent more than 60% of the average trading volume if traded at once (Chan and Lakonishok [1995]). Only about 20% of the value of institutional purchases is completed within a single day, while more than 50% of that volume traded takes at least four days for complete execution (Chan and Lakonishok [1995]). Techniques used to estimate market impact and facilitate trading are widespread in investment management and are available in the Market Impact Models offered by Citigroup, EQ International, ITG, MCI Barra, and OptiMark, among others.[2] At the market level, the market microstructure literature has attributed recent changes in market design towards automation of the trade execution to competition through liquidity and cost.[3] In fact, the adverse effects of market power measured in the so-called implicit trading costs dominate the explicit costs of the trades, such as commission fees (Chan and Lakonishok [1995]; Stoll [1978]; Keim and Madhavan [1995, 1996,1998]).

How important then is institutional trading in quantitative terms in contemporaneous financial markets? Over-all, institutional trading at the NYSE accounts for over 70% of the total trading volume (Schwartz and Shapiro [1990]). That trend has been sharply increasing over time; for example, total equity held by U.S. institutions at the NYSE increased from 7.2% in fifties to reach almost 49.8% in 2002.[4].

While developing tools to assess and mitigate market impact of institutional investors is currently part of practitioner cutting-edge research, the leading equilibrium asset pricing models (CAPM or Consumption CAPM) assume a price-taking behavior. Therefore such models are not suitable to address a host of questions that naturally arise given the facts about thin trading: What determines investor price impact and market illiquidity in general? If prices cannot be taken as given, what becomes the optimal strategy to execute a large order? How can one formally assess the value of an asset that is traded in thin market? This paper presents a dynamic asset pricing model where investors trade frequently and take into account their impact on prices. The model does not assume the presence of asymmetric information and exogenous costs (e.g., those associated with transactions, search or financing), and hence allows an understanding of the behavior of those markets in which such frictions are not the driving force behind a price impact.[5]

---

[1] See, e.g., Kraus and Stoll (1972); Holthausen, Leftwich and Mayers (1987); Chan and Lakonishok (1993, 1995); Keim and Madhavan (1995, 1996, 1998).

[2] Many of these are versions of the "implementation shortfall," see Perold (1988). Barclay and Holderness (1992] who discuss the legal aspects of block trades.

[3] Examples of exchanges that have adopted an electronic trading system with posted orders include the Nasdaq, NYSE, Euronext, the London Stock Exchange, the Toronto Stock Exchange, and the Vancouver Stock Exchange.

[4] `www.nysedata.com/factbook`

[5] The notion of equilibrium, based on Weretka (2006a) is closely related to the notion of equilibrium in Linear

Moreover, we model trading environments in which *all* traders have price impact. In particular, no competitive fringe (cf. Cournot) or noise traders are needed to prevent the market from collapsing. To the best of our knowledge, our approach offers the first such general-equilibrium, multiple-asset model of frequent trading in thin financial markets.

Preview of the Model and Results. The model presented here builds on the static framework proposed by Weretka (2006a), and extends that framework to a dynamic setting with non-separable utility functions. In order to delineate how the mere presence of price impact affects equilibrium, we consider an otherwise standard CAPM setting with mean-variance optimizers. The sole assumption of the CAPM from which we depart concerns the price-taking behavior of investors. The standard competitive CAPM is encompassed as a limit case of our model, which we call *Thin-Market CAPM* (*TM-CAPM*). The explicit modeling of price impact then enables us to ask new questions about the effects that are associated with thin trading. We summarize the main findings below.

- Price Impact. TM-CAPM predicts that investor's ability to affect prices depends on market depth (number of investors), risk preference and riskiness of assets. Correlation of asset payoffs induces cross-asset price impact. Our model captures the fact that the dividends are paid less frequently (typically semiannually) relative to how often the trade takes place. We show that, thus defined *frequent trading* endogenously introduces a non-stationarity of price impact: Price impact decreases with time to maturity, and hence markets are least liquid (thinnest) just prior to asset maturity. This in turn implies that liquidity will be correlated across markets, even if asset returns are independent, as long as the times in which dividends are paid coincide. [6]

- Trading Strategy. We also address the question of optimal handling of (large) orders in thin financial markets.[7] We show that rational investors break up their orders into smaller blocks and place them sequentially on the market, rather then rebalance their portfolios within one trade. The dynamics of price impact further implies that the volume of trade is the largest at the beginning of the trading period and decreases as time approaches maturity.[8]

- Temporary and Permanent Price Effects. As evidenced by the empirical literature, the exogenous shocks in asset supply such as IPOs, changes of index weights or forced liquidations result in price overshooting: Even if the shock is preannounced, on the actual event date the price drops below the new fundamental value to attain that value only in subsequent periods. The resulting price change has both a temporary and a permanent component.[9] We consider

---

Supply Functions (see Section 2.2). The manner in which we formalize the price impact allows us to circumvent the problem of indeterminacy of equilibria that is present in the deterministic models with Supply Functions.

[6] This phenomenon, often described as *liquidity spillover*, has been identified in a time series by Dufour and Engle (2000); Sarkar, Schwartz, and Wolf (2005); in a cross-section, commonality in liquidity has been documented by Hasbrouck and Seppi (2001); Huberman and Halka (2001); Chordia, Roll, and Subrahmanyam (2002).

[7] For various classes of the functional forms of price impact, see Bertsimas and Lo (1998); Almgren and Chriss (2000); Subramanian and Jarrow (2001); Dubil (2002); Almgren (2003); and Obizhaeva and Wang (2005).

[8] We discuss how this prediction relates to the much-documented pattern of intra-day trading volume, which is the largest at trade opening (Jain and Joh [1988]; Foster and Visvanathan [1990]; Gerety and Mulherin [1992]; Hamao and Hasbrouck [1995]; Aitken, Brown, and Walter [1994]; Almgren et. al. [2005]). This has been attributed to insider information and the willingness to coordinate trades with other institutional traders to ensure a thick market.

[9] The two effects where first empirically discovered by Kraus and Stoll (1972) and more recently they were

anticipated and unanticipated exogenous shocks in asset trades. Consistent with the data, TM-CAPM predicts that a temporary liquidity effect occurs whether or not the shock is anticipated. While the permanent effect always occurs immediately after the investors learn about the shock, the temporary effect does builds up to reach the maximum at the moment of trade. Moreover, in contrast to the permanent effect, whose magnitude will not depend on time, the temporary effect increases as the maturity approaches. These predictions are strongly supported by the methodology recently implemented by Citigroup to estimate price impact.[10] Finally, the model suggests that, compared to the competitive CAPM, price impact results in excess return volatility, while the non-stationarity of price impact induces volatility clustering.

- Asset Valuation in Thin Markets. In the presence of price impact, the market value of a large block of shares no-longer coincides with its liquidation value. Therefore, valuation specialists often apply the so-called *blockage discount*, defined as a "deduction from the actively traded price of a stock because the block of stock to be valued is so large relative to the volume of actual sales on the existing market that the block could not be liquidated within a reasonable time without depressing the market price" (*Handbook of Advanced Business Valuation*, p. 140).[11] In practice, blockage discounts are applied not only to stocks, but also to real estate, personal property (collections of art, antiques, and manuscripts), charitable gifts etc. These discounts have typically been estimated to range between 0% and 15%.[12] The IRS has acknowledged the concept of blockage since 1937. Yet, there are no formal guidelines as to how to determine the appropriate magnitude of the discount. We show that our results can quite directly be applied to formalize the appraisal instruments.

1 Related Models with Price Impact

Our model is related to the large literature that appeared to explain why illiquidity arises in financial markets and how that illiquidity affects the individual portfolio choices of investors. The theoretical mechanisms underlying these models can be precisely grouped into four categories. The first two are based on frictions, such as exogenous costs (transaction costs or, more recently, search costs and imperfect availability of funding to market makers)[13] or asymmetric information

---

confirmed by Harris and Gurel (1986); Holthausen, Leftwich and Mayers (1990); Chan and Lakonishok (1995); Beneish and Whaley (1996); Keim and Madhavan (1996); Lynch and Mendenhall (1997); and Greenwood (2005).

[10] In that program, the price impact is decomposed into:

(a) A permanent component ("reflects the information transmitted to the market by the buy/sell imbalance"), which is believed to be roughly independent of the trade scheduling;

(b) A temporary component ("reflects the price concession needed to attract counterparts within a specified short time interval"), which is highly sensitive to trade scheduling (Almgren et. al. [2005]).

Also, consistent with our model's predictions, Newman and Rierson (2004) found that new bond issuance in the European telecommunication sector increased yield spreads of other firms in the sector. The effect was transitory, significant and peaked on the day of issuance, not on the day of announcement. In addition, the severity of the effect was enhanced by asset riskiness and correlation, a finding that corresponds to our model's prediction.

[11] These discounts are distinct from (though sometimes confused with) *restricted stock discounts* due to difficulty in selling because of regulatory or contractual constraints.

[12] For a summary of U.S. Tax Court decisions involving blockage discounts, see Estabrook (1999, 2001); and Pratt (2001).

[13] E.g., Amihud and Mendelson (1986); Constantinides (1986); Vayanos (1998); Duffie, Garleanu and Pedersen (2005); Lagos and Rochetau (2006); Brunnermeier and Pedersen (2006); Vayanos and Wang (2007).

(regarding the fundamentals of a security, order flows, or endowments). While such models capture important features of trading environments, these mechanisms do leave room for explaining a number of empirical liquidity-related phenomena. For example, transaction-cost-based explanations favor less frequent trades, and therefore cannot account for the breaking up of orders. A search story might be consistent with the order break-up in decentralized markets, such as over-the-counter markets. Still, the phenomenon of trading in blocks has been shown to be robust to whether the market structure is centralized or not (Keim and Madhavan [1996]). Traditionally, the leading class of models with price impacts is based on asymmetric or private information (Glosten and Milgrom [1985]; Kyle [1985, 1989]; Foster and Viswanathan [1996]; Vayanos [1999]). Nonetheless, empirical studies do suggest that in many trade settings, the price impact component that is due to asymmetric information can account for a relatively small fraction of the observed magnitudes of trading volume.[14]

Rather than introducing frictions, several papers most recently have directly incorporated price impact and downward-sloping demands into asset pricing (Attari, Mello and Ruckes [2005]; Brunnermeier and Pedersen [2005]; Pritsker [2005]; and DeMarzo and Urošević [2006] extended by Urošević [2005]). These papers capture price impact by building Cournot-type models with one or N large investors and a continuum of (infinitesimally) small price-taking traders. Our model is different in the two major respects: First, it captures price impact without assuming the presence of competitive investors who form the demand function. Rather, the price impact is revealed as a price concession of a trader that is required to execute an order to all the other investors. Second, unlike a Cournot-type model, in which every large investor trades only with small investors, our model permits large investors to trade directly with one another.

To complete the picture of the modeling approaches, several non-equilibrium models with price impact have been proposed, mostly by practitioners.[15] These models assume exogenous price impact functions for every trader, and given these functions, the models analyze market dynamics without solving for equilibrium.

Our model is the closest to Vayanos (1999),[16] who develops a model with asymmetric information concerning aggregate asset holdings, in which traders submit demand functions. Our model features no private information. It allows for multiple risky assets, which permits us to study asset-specific as well as aggregate liquidity and a cross-asset relation between price impacts and returns. The key difference is that in our model the trading takes place more frequently compared,

---

[14] Indeed, it has been well documented that pre-announced changes of weights in indices, such as the S&P, have a significant price effect on the day of the inclusion. Such natural experiments that allow controlling for the informational component of the price change were studied by Kaul, Mehrotra and Morck (2000), Hau and Rey (2004) Loderer, Cooney and Van Drunen (1991) and Hau, Massa and Peress (2005) for stocks and currencies or foreign equity.

Furthermore, large institutional investors do not outperform fixed benchmark portfolios, which would likely be the case if they had superior information about asset fundamentals. Finally, Seppi (1990) shows that upstairs markets that are used by traders who can credibly signal their trades are not motivated by information advantage. Madhavan and Cheng (1997) show that for the average value of a trade, the price impacts in the downstairs markets do not differ significantly from those in the upstairs markets. This suggests that in both markets, the price impact is not mainly driven by asymmetry of information.

[15] e.g., Bertsimas and Lo (1998); Almgren and Chriss (2000); Subramanian and Jarrow (2001); Dubil (2002); Almgren (2003); Huberman and Stanzl (2004).

[16] Our model is the closest to the version of Vayanos model with perfect information (the third part of the paper)

to dividend payments, which allows us to characterize the dynamics of price impact and study new effects associated with the non-stationarity of such price impact.

2 A Model of Frequent Trading in Thin Financial Markets

2.1 Market Microstructure

There are I investors, also called *liquidity providers*, where I can be a small number. Investment opportunities include N risky assets and one riskless asset (e.g., a treasury bill). Investors can trade for T periods, after which assets mature, and dividends are paid in period $T+1$. The dividends from risky assets, A, are normally distributed, $A \sim \mathcal{N}(\bar{A}, \bar{\mathcal{V}})$, where $\bar{A}$ is the vector of the expected asset payoffs and $\bar{\mathcal{V}}$ is the symmetric and positive definite variance-covariance matrix of A. For notational convenience, we assume that the APR on the riskless asset is zero.

Institutional investors enter each period $t \in \{1, 2, \ldots, T\}$ with stocks of risky assets $\theta_t^i \in \mathbb{R}^N$ and bonds $\theta_{b,t}^i \in \mathbb{R}$. After they trade $\Delta\theta_t^i$ and $\Delta\theta_{b,t}^i$ in stocks and bonds, respectively, their holdings in the next period are $\theta_{t+1}^i = \theta_t^i + \Delta\theta_t^i$ and $\theta_{b,t+1}^i = \theta_{b,t}^i + \Delta\theta_{b,t}^i$. $(\theta_0^i, \theta_{b,t}^i)$ denotes exogenously given initial portfolio of trader i, which for all investors is assumed to lie in the asset span.[17] Investors chose their trades to maximize the expected CARA utility functions. By the standard argument, such assumptions are equivalent to the ones that investors are mean-variance optimizers. That is, investor's i indirect utility function, expressed in terms of after-trade portfolios, is linear in bond holdings and quadratic in risky assets:

$$U(\theta_{T+1}^i, \theta_{b,T+1}^i) = \theta_{b,T+1}^i + \bar{A}\theta_{T+1}^i - \frac{\alpha}{2}\theta_{T+1}^i \bar{\mathcal{V}} \theta_{T+1}^i. \tag{1}$$

It is useful to define an index of a market depth as

$$\gamma \equiv 1 - \frac{1}{I-1}. \tag{2}$$

The closer γ is to one, the deeper the market is and the more competitive the market interaction. With only two traders γ is equal to zero. We argue later that equilibrium does not exist in this case. Intuitively, traders' price impacts are shown to be mutually reinforcing, and since with bilateral trade the reinforcement occurs without any discounting, it leads to infinite price impacts.[18] Therefore, we assume there are at least three liquidity providers, and hence γ ranges between one-half and one. Finally, θ^{Av} denotes an *average portfolio*, defined as a portfolio held by all institutional traders, evaluated in per capita terms:

$$\theta^{Av} \equiv \frac{1}{I}\sum_{i \in I} \theta_0^i. \tag{3}$$

Each asset market is centralized and a market maker facilitates the trade.

---

[17] As in the competitive model, this assumption can be shown to effectively imply complete markets.

[18] The non-existence of equilibrium with two traders is also obtained in a closely related model by Kyle (1989). It is shown in Weretka (2006a) that with an outside trade option (e.g., a market or a trader who can absorb an arbitrarily small fraction of trade), equilibrium exists even with two investors.

2.2 Equilibrium

In TM-CAPM, trades of the liquidity providers are large relative to the market size and can, therefore, exert a non-negligible impact on prices.The price impact of trader i is formalized as an $N \times N$ matrix $\mathcal{M}^i_t$, with a typical element (n, m) characterizing the price effect on asset m resulting from the marginal increase in demand for asset n. (In case $N = 1$, the matrix becomes a scalar, a slope of a standard demand.) As long as the *price impact matrix* $\mathcal{M}^i_t$ is non-zero, the asset demands on which investor i operates are not perfectly elastic.

In this paper, the price impacts of each trader are not exogenous, but are determined in equilibrium jointly with trades and prices. With equilibrium in every period t being now a triple $(\bar{p}_t, \Delta\bar{\theta}_t, \bar{\mathcal{M}}_t)$, the standard conditions of traders' optimization and market clearing are supplemented by a consistency condition on price impacts.

Given the price observed in the market, p_t, a portfolio traded at this price, $\Delta\bar{\theta}^i_t$, and price impact matrix, $\mathcal{M}^i_t$, the investor faces a demand function,

$$p_{p_t,\Delta\bar{\theta}^i_t,\mathcal{M}^i_t}(\Delta\theta^i) = p_t + \mathcal{M}^i_t(\Delta\theta^i_t - \Delta\bar{\theta}^i_t). \tag{4}$$

In equilibrium, the conjectured price impacts should correspond to the traders' true price impacts. This condition defines a consistency condition on matrices $\mathcal{M}^i_t$, formally stated below. For the sake of transparency, we first define an equilibrium, assuming the notion of consistency; then the definition of consistency follows.

Definition 1 *A vector* $(\bar{p}_t, \Delta\bar{\theta}_t, \bar{\mathcal{M}}_t)_{t\in\{1,\ldots,T\}}$ *is a dynamic equilibrium if in each period t*

1) asset markets clear, $\sum_i \Delta\bar{\theta}^i_t = 0$*;*

2) for any i, the trade $\Delta\bar{\theta}^i_t$ *is (subgame-perfect) optimal, given demand functions* $p_{\bar{p}_t,\Delta\bar{\theta}^i_t\bar{\mathcal{M}}^i_t}(\cdot)$*;*

3) price impact matrices $\bar{\mathcal{M}}_t$ *are mutually consistent.*

In each period, when choosing trade $\Delta\bar{\theta}^i_t$, investors anticipate the effects of their choices on tomorrow's equilibrium values of state variables, and hence, similar to a competitive model, in period t, they maximize the value function determined by tomorrow's policy functions (hence the subgame perfection condition in the definition of an equilibrium). To conceptualize the consistency of $\bar{\mathcal{M}}_t$, we consider how financial markets react to investors' deviation from equilibrium trade $\Delta\bar{\theta}^i_t$ to any $\Delta\theta^i_t > \Delta\bar{\theta}^i_t$. Competitive models assume that the effects of such deviation on prices are at most negligible, and hence they can be ignored. Here by contrast, the price decreases sufficiently to encourage other traders to absorb the additional asset supply. As a result, all markets clear, and other investors respond optimally to prices in equilibrium, also after a unilateral deviation in which one investor i is trading a suboptimal quantity of shares, $\Delta\theta^i_t$. In other words, any deviation $\Delta\theta^i_t$ by trader i triggers a subequilibrium defined as follows:

Definition 2 *Given* $\bar{\mathcal{M}}^{-i}_t$*, vector* $(p^*_t, \Delta\theta^{-i*}_t, \bar{\mathcal{M}}^{-i}_t)$ *is a subequilibrium triggered by trade* $\Delta\theta^{i*}_t$ *if:*

1) markets clear with deviation, $\Delta\theta^{i*}_t + \sum_{j\neq i} \Delta\theta^{j*}_t = 0$*;*

2) for any $j \neq i$*,, trade* $\Delta\theta^{j*}_t$ *is optimal given demand functions* $p_{p^*_t,\Delta\theta^{j*}_t,\bar{\mathcal{M}}^j_t}(\cdot)$.

Consistent price impact reflects the price change needed to clear the market for any possible deviation $\Delta\theta_t^{i*}$, given that the other traders respond rationally to market prices.

Definition 3 *$\bar{\mathcal{M}}_t^i$ is consistent with $\bar{\mathcal{M}}_t^{-i}$ if for any deviation $\Delta\theta_t^i$ of trader i*

$$p_t^* - \bar{p}_t \equiv \bar{\mathcal{M}}_t^i(\Delta\theta_t^i - \Delta\bar{\theta}_t^i), \tag{5}$$

where $\bar{p}_t$ is an equilibrium price, and p_t^* is the price in the subequilibrium triggered by deviation $\Delta\theta_t^i$.

INTERPRETATION AND PROPERTIES OF AN EQUILIBRIUM. One way to see how the equilibrium considered in this paper generalizes the competitive concept is to note that the traders are modeled as slope-takers rather than price-takers. With infinitely elastic demands, the competitive model obtains.

Within the CARA-Normal framework, the equilibrium notion is outcome-equivalent to a unique symmetric Bayesian-Nash equilibrium in a game in which traders submit Linear Supply Functions (LSF). Such a game was introduced to the financial literature by Kyle (1989) and to the industrial organization literature by Klemperer and Meyer (1989).

Our alternative formulation of equilibrium facilitates a direct comparison with the standard CAPM. More importantly our approach also offers new insights about the determinants of investors' market power in the game already studied in the literature. It shows that the only information a trader needs to respond optimally to market conditions is the slope of his own residual demand. In particular, no information about the number, let alone the utility functions or trading strategies of the trading partners, is required. In addition Weretka (2006b), considers the model in which traders enter the market with arbitrary beliefs about their price impacts, collect price-quantity data to independently estimate their true impact. If traders respond optimally to prices, given updated estimates, these estimates will converge to the consistent price impacts in our model, and the outcome will then converge to the non-competitive equilibrium studied in this paper, and also, by equivalence, results in a linear Bayesian-Nash equilibrium in the LSF game.[19] Our assumptions thus fit particularly well in anonymous markets in which investors have no other information but their past trades and market prices; therefore the traders discover their market power through statistical inference.

3 Derivation of Equilibrium

We now derive an equilibrium in our model. Throughout, we highlight the differences with the competitive benchmark. We first conjecture the functional form of the value function, find an equilibrium in given t and then, recursively determine the coefficients of the value function.

It is convenient to introduce the following two statistics. The fundamental value vector $\bar{v} \in \mathbb{R}^N$

---

[19] More precisely, the equilibrium from this paper is outcome equivalent to a NSF equilibrium, which is refined by adding noise traders and letting their distribution converge to zero.

is defined as a vector of prices predicted by the competitive CAPM

$$\bar{v} \equiv \bar{A} - \alpha \bar{\mathcal{V}} \theta^{Av}. \tag{6}$$

For each investor i, we also define an auxiliary portfolio $\tilde{\theta}_t^i$ as a convex combination of the current and the average portfolio

$$\tilde{\theta}_t^i \equiv (1 - \lambda_t)\theta^{Av} + \lambda_t \theta_t^i, \tag{7}$$

where the scalar λ_t is given by $(1-\gamma)^{T-t}$. Observe that $\tilde{\theta}_t^i$ is predetermined at t. For any period t, we propose the following value function

$$V_t^i(\Delta\theta_t^i, \Delta\theta_{t,b}^i) = \Delta\bar{\theta}_{t,b}^i + ((1-\lambda_t)\bar{v} + \lambda_t \bar{A}) \cdot \Delta\bar{\theta}_t^i - \frac{\alpha}{2}(\tilde{\theta}_t^i + \lambda_t \Delta\theta_t^i)\bar{\mathcal{V}} \cdot (\tilde{\theta}_t^i + \lambda_t \Delta\theta_t^i) + c_t, \tag{8}$$

where constant c_t does not depend on trades ($\Delta\theta_t^i$, $\Delta\theta_{t,b}^i$), and the policy function

$$\Delta\bar{\theta}_t^i = \gamma(\theta^{Av} - \theta_t^i). \tag{9}$$

As a check, observe that in the last trading period, T, the value function coincides with the utility function (1), given that $\lambda_T = 1$ and $c_T = \bar{\theta}_{T,b}^i$. In hindsight, for $t < T$, the functional form in (8) can be motivated as follows: As we show, in any given trading period, it is optimal for an investor to partially replace his risky holdings with the average portfolio. From the perspective of period $t < T$, only a fraction λ_t of the current trade $\Delta\theta_t^i$ survives to maturity in $T+1$. Hence, only λ_t of the current trade adds to the riskiness of the ultimate portfolio $\bar{\theta}_{T+1}^i$. The remaining part, $1 - \lambda_t$, is liquidated in subsequent periods at market prices equal to $\bar{v}$. We suggestively then refer to λ_t as a *survival rate*.

For the candidate value function (8), we next derive an equilibrium in period t, and determine the corresponding parameters λ_t and c_t in the value function.

PORTFOLIO CHOICE. To see how the individual portfolio choice is affected by the presence of price impact, consider the budget set of investor i who chooses trades $\Delta\theta_t^i$ and $\Delta\theta_{t,b}^i$ in period t, for fixed prices p_t and price impacts $\mathcal{M}_t^i$:

$$p_{p_t, \Delta\bar{\theta}_t^i, \mathcal{M}_t^i}(\Delta\theta_t^i) \cdot \Delta\theta_t^i + \Delta\theta_{t,b}^i \leq 0. \tag{10}$$

Because the prices of risky assets are now functions of the quantities demanded $\Delta\theta_t^i$, the budget constraint (10) is quadratic rather than linear in $\Delta\theta_t^i$. The optimal trades in assets are characterized by the condition of equality between marginal utilities and marginal revenues, which, unlike in the competitive market, depends on the quantity traded.[20] Specifically, at the optimum, the

---

[20] The marginal utility and the marginal revenue of the riskless asset are both equal to one. The Marginal Rate of Substitution and the Marginal Revenue respectively are, hence, determined by the derivatives of the utility function and the budget set with respect to the risky assets. Geometrically, condition (11) corresponds to the familiar tangency between an indifference curve and the budget set, which is represented by a parabola and the slope of which is given by the vector of the - now quantity-dependent - marginal revenues. The strict convexity of the objective function and of the budget set assures that this condition is sufficient as well as necessary for optimality.

marginal revenues from selling each asset exceed the prices by $\mathcal{M}^i_t\Delta\bar{\theta}^i_t$:

$$(1-\lambda_t)\bar{v} + \lambda_t\bar{A} - \lambda_t\alpha\bar{\mathcal{V}}\cdot(\tilde{\theta}^i_t + \lambda_t\Delta\bar{\theta}^i_t) = p_t + \mathcal{M}^i_t\Delta\bar{\theta}^i_t. \tag{11}$$

With a positive definite $\mathcal{M}^i$, we solve (11) for the individual asset demand of investor i as a function of price and price impact

$$\Delta\bar{\theta}^i_t(p_t, \mathcal{M}^i_t) = (\lambda_t^2\alpha\bar{\mathcal{V}} + \mathcal{M}^i_t)^{-1}((1-\lambda_t)\bar{v} + \lambda_t\bar{A} - \alpha\lambda_t\bar{\mathcal{V}}\cdot\tilde{\theta}^i_t - p_t). \tag{12}$$

Non-zero price impacts $\mathcal{M}^i_t$ make the demand steeper, which induces the investor to trade less aggressively. Note in passing that at the equilibrium prices $\bar{p}_t$ and price impacts $\bar{\mathcal{M}}^i_t$, the net trades of all traders must sum to zero: $\sum_{i\in I}\Delta\bar{\theta}^i_t(\bar{p}_t, \bar{\mathcal{M}}^i_t) = 0$.

Price impacts. So far, we have considered investors' price impacts as given. We now endogenize $\bar{M}$. Let the price impacts of all other traders but i be fixed, and suppose that investor i offers to sell an extra block of shares $\Delta\theta^i_t - \Delta\bar{\theta}^i_t$ above his equilibrium trade $\Delta\bar{\theta}^i_t$. For markets to clear, the price must decrease, so the other investors are willing to purchase the additional shares. By market clearing and optimization of other traders, the required price change satisfies

$$\Delta\theta^i_t + \sum_{j\neq j}\Delta\bar{\theta}^j_t(p_t, \bar{\mathcal{M}}^j_t) = 0, \tag{13}$$

The condition (13) determines the price that assures market clearing for any possible deviation $\Delta\theta^i_t$ and, thus, it implicitly defines the inverse demand function faced by trader i in period t. Substituting i's individual demand (12) into (13), solving for price p_t, and using the fact of market clearing in equilibrium yields

$$p_{\bar{p}_t, \Delta\bar{\theta}^i_t, \bar{\mathcal{M}}^i_t}(\Delta\theta^i) = \bar{p}_t + \underbrace{(\sum_{j\neq i}(\bar{\mathcal{M}}^j_t + \lambda_t^2\alpha\bar{\mathcal{V}})^{-1})^{-1}}_{\bar{\mathcal{M}}^i_t}(\Delta\theta^i_t - \Delta\bar{\theta}^i_t). \tag{14}$$

Thus, the equilibrium price impact of investor i is given by

$$\begin{aligned} \bar{\mathcal{M}}^i_t &= (\sum_{j\neq i}(\bar{\mathcal{M}}^j_t + \lambda_t\alpha\bar{\mathcal{V}})^{-1})^{-1} = \\ &= (1-\gamma)\mathcal{H}(\bar{\mathcal{M}}^j_t + \lambda_t^2\alpha\bar{\mathcal{V}} | j\neq i). \end{aligned} \tag{15}$$

Equation (15) reveals several central and novel predictions of the model. It first uncovers an interesting mathematical structure about equilibrium market power. Namely, the price impact of investor i is characterized as a harmonic average of the convexities of other investors' value functions, $\lambda_t^2\alpha\bar{\mathcal{V}}$, augmented by their price impact, $\bar{\mathcal{M}}^j_t$, and discounted by factor $1-\gamma$. To better understand why (15) holds, recall that price impact represents price concessions following the unilateral deviations of trader i from his equilibrium trade, which are sufficient for the asset markets to clear. The greater the price impact assumed by other traders, the larger the price concession that is needed to make them be willing to absorb the additional risky assets. This reasoning explains how market power mutually reinforces across investors.[21]

The relation between investor price impacts in equation (15) takes the impacts of traders other than i as given. To pin down the endogenous price impact for each trader, we find a symmetric solution to the system of consistent matrices described by I equations (15) each for one trader. In the unique symmetric solution ($\bar{M}^i = \bar{M}^j$), the consistent price impact for any $i \in I$ is equal to

$$\bar{\mathcal{M}}_t^i = \frac{(1-\gamma)}{\gamma}\lambda_t^2 \alpha \bar{\mathcal{V}}. \tag{16}$$

Investor's i price impact depends on the convexity of traders' value functions. Specifically, the TM-CAPM predicts that, what is less apparent in a symmetric solution, is other investors' risk aversion that enters trader's i price impact: More risk averse trading partners are more reluctant to increase their holdings of risky assets, which implies the larger price concessions in trading. This effect is (proportionally) enhanced by the return riskiness amplified or weakened by cross-market impact, as specified by the variance-covariance matrix $\bar{\mathcal{V}}$. As expected, price impact is partially mitigated by market depth, which reflects the number of potential liquidity providers. The effect on prices is further weakened when the survival rate λ_t is smaller; when the current trade has a small effect on the ultimate portfolio, the liquidity providers are more willing to absorb buy or sell orders.

The competitive prediction obtains if and only if investors are risk neutral ($\alpha \sim 0$), markets are infinitely deep ($\gamma \sim 1$). If the time horizon is sufficiently long, so that by the time assets mature, investors holdings can be almost entirely replaced with the average market portfolio ($\lambda_t \sim 0$), the markets are almost perfectly liquid.

ASSET PRICES.The equilibrium asset prices obtain from the market clearing condition by substituting demands (12)

$$0 = \frac{1}{I}\sum_{i \in I} \Delta\bar{\theta}^i(\bar{p}_t, \bar{\mathcal{M}}_t^i) = \bar{v} - \bar{p}_t. \tag{17}$$

where we used the symmetry of $\lambda_t^2 \alpha \bar{\mathcal{V}} + \bar{\mathcal{M}}_t^i$, the definition of $\bar{v}$, (6) and the fact that $\frac{1}{I}\sum_{i \in I} \tilde{\bar{\theta}}_t^i$ coincides with the average portfolio. The equilibrium prices are constant over time, and they coincide with the competitive prices

$$\bar{p}_t = \bar{v} \equiv \bar{A} - \alpha \bar{\mathcal{V}} \theta^{Av}. \tag{18}$$

Here we postpone the explanation and the discussion about the robustness of this surprising result until the next section.

OPTIMAL PORTFOLIOS. To see how investors rebalance their portfolios in each period, we substitute (18) into (12)

$$\Delta\bar{\theta}_t^i = \gamma(\theta^{Av} - \theta_t^i). \tag{19}$$

---

[21] That the reinforcement is governed by a harmonic average implies that traders with the flatter individual demands (with the smallest $M^j + \lambda_t^2 \alpha \bar{\mathcal{V}}$) are relatively more important in determining i's market power. Indeed, the more competitive traders will require lower price concessions to purchase a given amount of the risky assets.

Investor i sells fraction γ of a portfolio with which he entered the period and replaces it with γ of the average portfolio. Interestingly, the fraction is determined solely by the depth of the market; in particular, it is independent of risk aversion α. Given the equilibrium trades of risky assets (19), the budget constraint in (10) determines the changes of bond holdings in period t

$$\Delta\theta^i_{t,b} = -\bar{v} \cdot \Delta\bar{\theta}^i_t. \tag{20}$$

Survival rate λ_t. To show that the candidate value function (8) is valid in arbitrary period $t < T$, we use a standard inductive argument. We have already verified that (8) holds in T. If (8) is satisfied for any $t' > t$, the policy function (19) implies that for any trade $\Delta\theta^i_t$, the ultimate risky portfolio in period $T+1$ is given by

$$\bar{\theta}^i_{T+1} = (1 - (1-\gamma)^{T-t})\theta^{Av} + (1-\gamma)^{T-t}(\bar{\theta}^i_t + \Delta\theta^i_t). \tag{21}$$

With λ_t given by

$$\lambda_t \equiv (1-\gamma)^{T-t}, \tag{22}$$

portfolio $\bar{\theta}^i_{T+1}$ can be written as $\bar{\theta}^i_{T+1} = \tilde{\theta}^i_t + \lambda_t \Delta\theta^i_t$. By equation (20), the bond holdings are given by

$$\bar{\theta}^i_{T+1,b} = \Delta\bar{\theta}^i_{t,b} + \bar{v}(1-\lambda_t)\Delta\bar{\theta}^i_t + c_t, \tag{23}$$

where the constant is equal to

$$c_t = \bar{\theta}^i_{t,b} - \bar{v}(1-\lambda_t)(\theta^{Av} - \theta^i_t). \tag{24}$$

Applying (21), (23) and (24) in the utility function (1) establishes (8) as the functional form of the value function in period t.

4 Model Predictions

This section offers the main results of the paper.

The optimal trading strategies. The first result characterizes how market thinness affects the optimal execution of trade. The competitive CAPM predicts that investors instantaneously sell their initial portfolios and rebalance their holdings within one period; and that they invest in a combination of the market portfolio and the riskless asset (Two-Fund Separation). No trade takes place in subsequent periods (in a deterministic setting), as the investors' risky holdings became efficient already in the first trading period. By contrast, TM-CAPM predicts that in thin markets ($\gamma < 1$), every time they trade, investors will sell only a fraction of (the remaining part of) their initial portfolios to invest between the market portfolio and the riskless asset. Thus, a *Three-Fund Separation* obtains.

Proposition 1 (Three-Fund Separation) *For every trading period* $t = 1, ..., T$*, the risky part of the optimal portfolio is a convex combination of the initial and average portfolios,* θ^i_0 *and* θ^{Av}*,* *with the weight on* θ^{Av} *monotonically increasing over time:*

$$(1-\gamma)^t\theta_0^i + (1-(1-\gamma)^t)\theta^{Av}. \tag{25}$$

The remaining wealth is invested in the riskless asset, θ_b^i.

The optimal handling of large orders thus involves trading in blocks: The adverse effects of price impact induce the investors to break up their orders into smaller blocks and then place them sequentially on the market. Order break-up is common practice among large investors. Table 1 reports a summary of findings by Chan and Lakonishok (1995), who characterized the time structure of the institutional orders placed for 29 months on the NYSE. Only 20% of the total volume of all institutional purchases and sales is then completed within one day, and more than 30% of the orders takes at least 6 days to execute.

TABLE 1. ORDER SPLITTING

	1 Day	2-3 Days	4-6 Days	> 6 Days
Buy	20.1%	26.7%	21.7%	31.5%
Sell	22.1%	27.2%	20.5%	30.2%

(26)

Data: All trades of NYSE and AMEX stocks by 37 investment management firms from July 1, 1986, to December 30, 1988 (October 1987 excluded). A buy/sell package is defined as successive purchases/sales of a stock with at most a 5-day break between consecutive trades. The numbers are percentages of the total volume of trade measured in \$. *Source*: Chan and Lakonishok (1995, Table 1).

PARTIAL DIVERSIFICATION OF RISK. TM-CAPM further predicts that when the number of trading periods is sufficiently large, the portfolios will converge to the competitive holdings, but for any finite number of periods, the portfolios will be distinct. Consequently, at any point in time, the idiosyncratic risk will not be perfectly hedged. On the other hand, the allocation can be arbitrarily close to efficiency, provided that the time to maturity is sufficiently long.

In deeper markets, individual risky holdings converge faster to the competitive outcome, that is, the average portfolio held by all investors (see Figure 1). Somewhat surprisingly, not only does γ increase the speed of trade, but it actually fully determines it. In particular, the speed of trade does *not* depend on risk aversion α, as long as $\alpha > 0$. Intuitively, higher α is associated with greater gains to trade, and hence encourages more aggressive hedging through faster trading. But it also amplifies the price impacts of all traders, making the interactions less competitive and reducing the trade. In a quadratic model, in symmetric equilibrium, the two effects of risk aversion offset each other. Thus, even if large institutional traders are almost risk-neutral, as is often assumed in the finance literature, they will choose to trade slowly.

The breaking up of orders further induces an interesting dynamics of the trading volume, which is largest at the beginning of the trading period and exponentially decreases over time. Empirical evidence on intra-day trading shows that the volume of trade is largest at the market opening (e.g., Jain and Joh [1988], Foster and Visvanathan [1990], Gerety and Mulherin [1992], Hamao and Hasbrouck [1995], and Aitken, Brown, and Walter [1994]). Our model is consistent with that piece of evidence when $T+1$ is interpreted as the end of the trading day rather than the maturity period.[22] The pattern found in the data is often U-shaped and the model does not capture the end-of-day increase in the trading volume. In the data, the latter effect is smaller and has been shown to depend on whether the market structure is centralized or not - the feature taken as given in the centralized market studied in this paper.[23]

NONSTATIONARITY OF PRICE IMPACT. Proposition 1 raises the following question: Why is the trade clustered in earlier trading periods, rather than being, for example, equally spread over time? TM-CAPM offers an explanation through the dynamics of price impact, derived in Proposition 2, which is the key dynamic result of the model. The noncompetitive CAPM predicts that the price impact is not constant across the trading periods $\{1, ..., T\}$, but instead increases as the time approaches maturity: The further away from maturity, the more opportunities to diversify and re-trade, the less costly it is for the investors to depart from their current holdings. That mechanism is apparent in the value function (8), which becomes more concave over time to reflect the traders' increasing effective risk aversion, and decreasing willingness to trade risky assets at given price concessions. In short, our model predicts that the apparent coordination of trade in earlier periods occurs because the investors choose to trade when their price impacts are smallest.[24] To the extent that liquidity can proxy the level of market competitiveness, TM-CAPM predicts that markets are least competitive just prior to maturity.

Proposition 2 (Time Structure of Price Impact) *In the TM-CAPM, the price impact exponentially decreases with time-to-maturity:*

$$\bar{\mathcal{M}}_t^i = \frac{(1-\gamma)^{2(T-t)+1}}{\gamma} \alpha \bar{\mathcal{V}}. \tag{27}$$

Remarkably, the derived schedule of price impact matrices, $\bar{\mathcal{M}}_t^i$, is directly proportional to the variance-covariance matrix of returns, $\bar{\mathcal{V}}$. Thus, assets with more risky dividend payoffs are less liquid (in a cross section as well as a time series). In addition, formula (27) shows that there are cross-market price-impact effects. When the payoffs of two stocks are positively correlated, the sale of one asset inflicts a downward pressure on the price of other assets. In addition, how asset riskiness ($\bar{\mathcal{V}}$) and risk aversion (α) – the essential determinants of convexity of the investors'

---

[22] Two qualifications are in order here. Recall that in our deterministic model, T is defined as the final trading period followed by asset maturity. It follows from our analysis in Section 5 then that introducing information revelation to the model would generate the observed intra-day patterns, in which case T would be interpreted as a period prior to which information about asset returns is partially revealed. One possible explanation is that after the markets close, information accumulates overnight only to be revealed at the next markets' opening.

[23] There is also evidence that demonstrates that the volume of trade is the smallest on Fridays or before major holidays.

[24] Incidentally, common wisdom among financial professionals explains trade concentration by the need to assure sufficiently deep markets.

Sarkar, Schwarz, and Wolf (2005) found clustering of trade in equity markets and interpreted the evidence as supportive of asymmetric information or beliefs as a trade motive. The model presented in this paper shows that trade clustering might occur even without the asymmetries.

preferences – affect price impact in (27) is strengthened by time-to-maturity, $T-t$, but weakened by market depth, γ.

Crucially, the derived time structure of price impact reveals a nonstationarity that appears endogenous in the model. The mere fact that trade takes place more frequently than dividend payments (frequent trading) leads to nonstationarity of price impact in thin markets. This aspect uncovers an important channel through which systemic risk can destabilize markets. For instance, markets can endogenously become illiquid prior to a crisis in which a large amount of information about asset returns is expected to be revealed. This mechanism also may lead to the commonality in liquidity across assets, already widely documented in the empirical literature (see, e.g. the survey by Amihud, Mendelson and Pedersen [2005]).

Similarly, Proposition 2 suggests that in thin markets asset maturity becomes an active instrument in stabilizing markets. We discuss further implications of such nonstationarity in Section 5.3.

SECURITY MARKET LINE. On of the most celebrated and controversial results in the standard CAPM is the Security Market Line, which asserts that the return of an asset traded in a thin market can be explained solely by the covariance of its return with the return of the market portfolio. Analyzing the tradeoff between risk and return is much harder in the non-competitive model, as the asset prices no longer coincide with the marginal utilities of traders, and, moreover, marginal utilities typically differ across agents.

However, we demonstrated that, under our assumptions, the equilibrium asset prices derived in (18) coincide with asset fundamental values and are, therefore, identical to the competitive prices in every period. It follows that asset returns are the same random variables as in the competitive model and their expectations lie on the Security Market Line, spanned by the riskless return and the return on the average portfolio. Let $\bar{R}^{Av}$ denote the expected return of an average portfolio, let $\beta_n = \bar{\mathcal{V}}_{Average,n}/\mathcal{V}_{Average}$ be the beta of asset n, and let $\bar{R}_n$ be its expected return.

Proposition 3 (Security Market Line) *In thin markets with I liquidity providers the expected returns of individual assets at any t are located on the Security Market Line*

$$\bar{R}_n - R = \beta_n(\bar{R}^{Av} - R), \tag{28}$$

The result herein implies that thin trading *per se* does not distort asset returns. The fact that the market power of the traders does not affect equilibrium prices compared to the competitive CAPM is, nonetheless, somewhat surprising. This result, however, is non-generic, as it critically relies on the joint assumption of (i) quadratic and (ii) homogenous utility functions, and (iii) the deterministic structure of the model. Intuitively, with quadratic-homogenous utilities, symmetric price impacts for buyers and sellers make them reduce their demand and supply for each asset by the same factor γ. As a result, thin markets clear at the competitive prices, even though the trades are not competitive. Relaxing one of the three assumptions thus introduces systematic price effects. We briefly discuss these effects in Section 6.

We should stress here also that in our model, the average portfolio is defined as the risky portfolio held by a possibly small group of liquidity providers trading in specific asset markets.

Therefore, the standard approach to testing the predictions of CAPM, based on a market portfolio (defined, e.g. as all assets traded on the NYSE) should not be applied in this instance. To empirically test TM-CAPM, one should first properly identify a thin market.

5 Price Effects in Thin Markets

In this section, we examine new implications for asset pricing that arise from the presence of price impact and its nonstationarity. We demonstrate that thin markets significantly differ from the perfectly competitive setting in how they respond to exogenous shocks in asset supply. We argue that incorporating price impact into asset pricing can help us understand several empirical phenomena that are hard to reconcile within the competitive model: Asset price overshooting (Section 5.1); stylized facts about return volatility (Section 5.3); existence of instruments for asset valuation that try to account for price impact (Section 5.4).

Examples of liquidity shocks in financial markets would be the inclusion of an asset into the S&P index, a change of index weights,[25] forced liquidation, issuance of new debt, or selling Initial Public Offerings (IPOs). It has been much documented that such shocks lead to a significant price change followed by a partial reversal of the price change in subsequent periods. Crucially, in the data, the significant price change occurs on the date of the shock even if the shock was *pre-announced.* The observed price effects thus cannot be attributed to any revelation of information, which should have been incorporated on the day of the announcement. This phenomenon, often referred to as *price overshooting*, should not occur in competitive markets. What should be observed, instead, is that the prices adjust to the new fundamental value immediately following the shock announcement.

To examine how thin markets react to exogenous shocks in asset supply (or demand), we consider an unanticipated (Section 5.1) as well as an anticipated (Section 5.5) exogenous sale of a large block of shares by a trader other than N liquidity providers. Hereafter, portfolio θ_t^{Av} will denote the expected average portfolio in period $T+1$, given the information in period t. Vector $\bar{v}_t$ will stand for the fundamental value of an asset determined by (6) and θ_t^{Av}.

5.1 Price Overshooting

Consider an unanticipated one-time shock in asset supply, that is, in period t^*, a portfolio $\hat{\theta}$ is being liquidated along with the trade by liquidity providers. Since all institutional investors know about the shock in period t^*, the average portfolio $\theta_{t^*}^{Av}$ attains a new level

$$\theta_{t^*}^{Av} \equiv \theta_{t^*-1}^{Av} + \hat{\theta}/I. \tag{29}$$

In the competitive model, the risky holdings of liquidity providers would instantaneously change from θ_{t-1}^{Av} to $\theta_{t^*}^{Av}$, and the asset prices would adjust to the new fundamental value,

$$p_{t^*} = \bar{v}_{t^*} \equiv \bar{A} - \alpha \bar{\mathcal{V}} \theta_{t^*}^{Av}, \tag{30}$$

---

[25] Index funds invest a constant fraction of wealth in companies that are included in an index, regardless of the performance of an asset, and, therefore a change of index weights can induce a demand shock.

and no further price change would be observed till maturity (Figure 2). In a thin market, until period $t^* - 1$, investor i follows the trading strategy described in Proposition 1 and, hence, enters period t^* with the risky holdings of

$$\bar{\theta}^i_{t^*} = (1 - (1-\gamma)^{t^*})\theta_0^{Av} + (1-\gamma)^{t^*}\bar{\theta}^i_0. \tag{31}$$

In period t^*, traders learn about the shock and the market clearing condition becomes

$$0 = \sum_{i \in I} \Delta\bar{\theta}^i_{t^*}(\bar{p}_{t^*}, \bar{\mathcal{M}}^i_{t^*}) + \hat{\theta}. \tag{32}$$

Market clearing (32) implies that the prices, which have remained constant at $\bar{v}_0$ up to $t^* - 1$, are equal to

$$\bar{p}_{t^*} = \bar{v}_{t^*} - \frac{(1-\gamma)^{2(T-t^*)+1}}{\gamma I}\alpha\bar{\mathcal{V}}\hat{\theta}, \tag{33}$$

and the individual trades are given by

$$\Delta\bar{\theta}^i_{t^*} = \gamma(\bar{\theta}^{Av}_{t^*} - \bar{\theta}^i_{t^*}) + \hat{\theta}/I. \tag{34}$$

Since no other shocks follow after t^*, in all periods, $t > t^*$, prices return to $\bar{p}_t = \bar{v}_{t^*}$ and trades take place according to $\Delta\bar{\theta}^i_t = \gamma(\bar{\theta}^{Av}_{t^*} - \bar{\theta}^i_t)$.

The cash obtained from liquidating portfolio $\bar{p}_{t^*} \cdot \hat{\theta}$ is lower than its pre-liquidation market value $\bar{p}_{t^*-1} \cdot \hat{\theta}$ due to two effects. These effects, we call the *fundamental effect*, Δ^f, and the *liquidity effect*, Δ^l, and they are characterized in Proposition 35.

Proposition 4 (Asset Price Overshooting) *Following an unanticipated liquidity shock, $\hat{\theta}$, in period t^*, the change of equilibrium prices on the equilibrium path is given by*

$$-\Delta\bar{p}_t = \underbrace{\frac{\alpha}{I}\bar{\mathcal{V}}\hat{\theta}}_{\equiv\Delta^f} + \underbrace{\frac{(1-\gamma)^{2(T-t^*)+1}}{\gamma}\frac{\alpha}{I}\bar{\mathcal{V}}\hat{\theta}}_{\equiv\Delta^l} \tag{35}$$

The liquidity effect Δ^l disappears in all periods $t > t^*$.

The fundamental effect arises because the exogenous supply of risky assets $\hat{\theta}$ increases the average risky holdings held by the liquidity providers, which in turn reduces their average marginal utility from these assets to $\bar{v}_{t^*}$. This effect is permanent. Notice that the fundamental effect would be observed in a market with price-taking liquidity providers, as long as the number of such providers remained small (so that the market demand were downward-sloping). It is the temporary effect that lowers the price at t^* below $\bar{v}_{t^*}$, due to the non-competitive nature of trade. Why the price moves beyond the change in the fundamental value on the day of the shock can be explained as follows: With a positive net supply of risky assets at t^*, $\hat{\theta}$, in equilibrium, investors demand, on average, positive amounts of such assets; the investors' average marginal revenue exceeds the market price by $\bar{\mathcal{M}}^i_{t^*}\hat{\theta}/I > 0$. It follows then from optimality that the price must be below the average marginal utility.[26] Why is the liquidity effect temporary? In subsequent periods, no exogenous shocks occur, and the average trade of the liquidity providers becomes equal to zero (by market clearing); hence, the price is equal to the average marginal utility, and the price attains its new fundamental value. Observe that the fundamental effect Δ^f does not depend on the timing of the shock t^*, while the magnitude of overshooting Δ^l is greater when the time to maturity is shorter. Figure 2 depicts both effects for the exogenous shock $\hat{\theta}$ in period t^*. Panel A shows the path for the trade of an asset, and panel B, the price of an asset $p_{t,n}$.

The derived permanent and temporary price impacts are linear in block size. The linearity of the permanent effect has been confirmed by Almgren et al. (2005).[27] Their study estimates temporary price impact as a concave function of block size, which by now is a robust and well-documented result (e.g. Kempf and Korn [1999]; Plerou et. al. [2002]). However, much of the empirical evidence has been established, assuming that price impact is time independent, and our model predicts instead that price impact increases over the trading period. If price impact is indeed not constant (the findings by Madhavan, Richardson and Roomans [1997] are supportive), and large blocks are liquidated when the price impact is small, while small ones when it is large, then the estimation that assumes stationary price impact would lead to a spurious concavity of the temporary price effect, even if that effect were linear in every period.

In the literature to date, the permanent price effect is often attributed to the revelation of insider information about the future dividends thorough price mechanism (e.g., the literature initiated by Kyle [1985, 1989]).[28] In this paper, the permanent price change occurs even though investors do not infer any new information about future payoffs per share. Instead, it results from demand-supply balancing. Following a positive (negative) exogenous shock, the average holdings of risky assets of liquidity providers permanently increase (decrease); this increase permanently lowers their marginal utility, and the price falls. Such a price change is temporarily enhanced by the liquidity effect, which in our model arises because non-competitive investors reduce their demands, as explained above.

When the asset returns are correlated, exogenous sales of one asset affect the values of other assets. The fundamental effect of a supply shock $\hat{\theta}$ on the value of an arbitrary portfolio θ is given by

$$\theta \cdot \Delta^f = \frac{\alpha}{I} Cov(A\hat{\theta}, A\theta). \tag{36}$$

---

[26] This can easily be seen by taking the average of equation (11) across all traders, and observing the left-hand side (the average marginal utility) is equal to $\bar{v}_{t^*}$ in all periods $t \geq t^*$, while the right hand side (the average marginal revenue) is equal to $p_{t^*} + \bar{\mathcal{M}}^i_{t^*}\hat{\theta}/I$ in t^* and p_t from $t^* + 1$ on.

[27] It is also consistent with the findings of Loeb (1983), and Chan and Lakonishok (1995), who show that order size is critical for the price effect.

[28] In a recent paper, Brunnermeier and Pedersen (2005) explained price overshooting by "predatory trading": When a trader needs to quickly liquidate the portfolio, other investors sell and subsequently buy back the asset. This strategy lowers the price at which they can obtain the liquidated portfolio. The mechanism arises due to the presence of long-run investors who define a downward-sloping demand, buying assets when they are expensive and selling when assets are cheap. Our explanation of overshooting is complementary in that predatory trading does not occur in our model, since all traders maximize their preferences.

In addition, the liquidity effect temporarily (at t^*) reduces the market value of a portfolio θ

$$\theta \cdot \Delta^l = \frac{(1-\gamma)^{2(T-t^*)+1}}{\gamma} \frac{\alpha}{I} Cov(A\hat{\theta}, A\theta). \tag{37}$$

Interestingly, when the correlation between the payoffs of two portfolios is negative, the supply shock $\hat{\theta}$ permanently increases the value θ. This happens because the shock increases the attractiveness of θ as a hedge against the risk of $\hat{\theta}$.

5.2 Limits to Arbitrage

The overshooting result from Proposition 4 may seem puzzling at first. Since the liquidity providers know at t^* that in the next period the prices will revert, why do they not arbitrage the temporary price differential between t^* and $t^* + 1$? (In the competitive model, the no-arbitrage condition prevents price overshooting.) Note that, when in equilibrium, the liquidity providers are on average net buyers as they absorb the shock. In order to exploit the price change, they would have to make their risky positions in t^* even longer, which would result in higher asset prices. This, in turn would exert a negative externality (first order) by increasing the cost of their equilibrium purchases.[29] In a non-competitive equilibrium, such a marginal cost equals the marginal benefit from inter-temporal arbitrage. TM-CAPM thus predicts that overshooting is not a friction, but rather an equilibrium phenomenon in thin markets.

The argument above also illustrates another new feature of the model. According to the competitive theory of asset pricing, in the presence of arbitrage opportunities, traders can make infinite profits by taking unbounded positions. This implies that, even if the number of traders in a given market is small, price overshooting is ruled out by the presence of potential entrants. By contrast, in the model of thin markets, price impact limits the benefits from arbitrage, reducing the incentives to enter the market. Suppose that, at t^*, an entrant purchases a block of assets to be sold it the next period. The benefit from such a market operation is bounded for the following two reasons: Due to price impact at t^*, buying more shares than $\hat{\theta}$ would result in a negative profit, as the buy would drive the price above the fundamental value in this period; The profit is further limited by the fact that selling the block of shares in period $t^* + 1$ would also have adverse effects on the price, further magnified by the nonstationarity of price impact. Therefore, unlike in a competitive model, the incentives to enter the market are bounded, and even small fixed entry costs[30] may prevent outsiders from arbitraging the price overshooting. On the other hand, with zero transaction costs, free entry will prevent price overshooting even if markets are thin.

---

[29] Suppose there is only one risky asset. If in t^* the trader increases the trade by ε and sells the same amount the next period, the benefit from arbitraging the price differential is $\varepsilon \times \bar{\mathcal{M}}\hat{\theta}/I$. On the other hand the additional demand increases the price in t^* by $\bar{M} \times \varepsilon$ which negatively affects additional trade per capita $\hat{\theta}$. This implies that on an equilibrium path the marginal benefit is exactly equal to the marginal negative externality inflicted on today's trade.

[30] Such costs include explicit trading costs, such as transactions costs, but also the cost associated with learning the characteristics of the stocks.

5.3 Unanticipated Shocks and Price Volatility

The following robust findings about return volatility have long been documented in empirical finance: (1) The magnitude of return volatility is not justified by the volatility of asset fundamentals; (2) Volatility exhibits persistence, i.e., clustering of large and small moves; (3) Changes in volatility are largely unrelated to fundamentals; (4) Unconditional distribution of asset returns have heavy tails. We now offer a heuristic argument to illustrate how the sheer presence of price impact, coupled with its nonstationarity, may affect the moments of the returns in a manner consistent with these stylized facts.

Suppose that instead of once-and-for-all shock $\hat{\theta}$, we observe a sequence of i.i.d. random shocks in asset supply $\hat{\theta}_t$ (generated, for example, by noise traders), and suppose that such shocks are unanticipated by the traders.[31] If markets are perfectly deep, $\gamma = 1$, and traders are price takers, each trader absorbs only an infinitesimal fraction of the shock, and hence, the average marginal utility is not affected. It follows then that the asset prices are constant throughout the trading period. When $\gamma < 1$ and the liquidity providers still behave as price takers, the price follows a random walk. The prices are affected by supply shocks because positive realizations of $\hat{\theta}_t$ increase the number of shares held by the liquidity providers, and therefore, they permanently affect traders' average marginal utility (equal to price). The limited ability to absorb exogenous supply shocks makes the price more volatile compared to the perfectly deep markets. The permanent price effect depends on the size of the realization, and price volatility, is constant in all trading periods. This changes when traders recognize their impact on prices. Then, the price becomes even more volatile, as the fundamental component is reinforced by the non-stationary liquidity effect. Due to the latter effect, the unconditional distribution of $\Delta\bar{p}_t$ evolves over time, and its variance increases exponentially.

In sum, the model of frequent trading in thin markets predicts that the high (low) volatility of price, and hence return, is clustered at the end (in the beginning) of the trading horizon, and the changes of volatility are independent of the asset fundamentals. Finally, observe that given the presence of an exponential component in the variance, the price change in the last periods puts a large mass on realizations that are far away from the typical variability over the trading horizon. If the distribution of $\Delta\bar{p}_t$ is estimated, assuming that a price change is i.i.d., such approach might lead to an empirical distribution that has heavy tails, even if the kurtosis of the distribution in every trading period is less than 3. This outcome is yet another consequence of the nonstationarity of price impact.

5.4 Market Value and Blockage Discount

In perfectly competitive markets, the value of a block of shares, $\hat{\theta}$, is simply equal to the quantity of shares times the corresponding prices currently observed on the market $\bar{p} \cdot \hat{\theta}$. When markets are thin, selling a large block of shares exerts a downward pressure on prices, and the market

---

[31] This is only a heuristic argument, as the proper formulation of the problem would require that agents take the volatility of the supply into account when trading. The extra risk would result in a more concave value function (for a model with price uncertainty see, for example, Vayanos [1999]). This would not qualitatively change the effects of supply shocks on prices.

value no longer reflects the actual amount of cash that would be obtained by selling block $\hat{\theta}$. The problem of apprizing assets traded in thin markets has been recognized by valuation specialists, who apply the so called *blockage discount*. The blockage discount is defined as a "deduction from the actively traded price of a stock because the block of stock to be valued is so large relative to the volume of actual sales on the existing market that the block could not be liquidated within a reasonable time without depressing the market price" (*Handbook of Advanced Business Valuation*, p. 140).[32] In practice, blockage discounts are applied not only to stocks, but also to real estate, personal property (collections of art, antiques and manuscripts), charitable gifts, etc. The discounts have typically been estimated to range between 0 and 15 percent.[33] The IRS has acknowledged the concept of blockage discount since 1937. According to Federal Tax Regulations, the burden of demonstrating that a blockage discount is justified lies on the taxpayer. Yet, there is no objective guidance on how to assess the cash value of assets and the appropriate amount of blockage discounts. Practitioners have developed a range of heuristic methods for how to adjust the values of assets (see, e.g., Estabrook [1999, 2001]), and these methods have been adopted in appraisal businesses and in valuation consulting. Our results can quite directly be applied to address asset valuation in thin markets formally and thus derive blockage discounts.

The challenge in formalizing appraisals when markets are thin does arise because assets are often transferred outside of the market. For example, a typical instance where blockage discounts are applied involves a transfer of a property in a case of a divorce. It is in the interest of the divorcees to claim a large price impact which implies a large tax discount. The relevant question is: What would be the value of the property if it were sold on the market (even though it will *not* be)? This reasoning with how the price impact is calculated in our model. Let $\bar{p}_t$ be the observed market price and $\hat{p}_t$ be the hypothetical price that would be obtained if the block were offered on the market. The blockage discount is equal to $BD \equiv \hat{\theta} \cdot (\bar{p}_t - \hat{p}_t) = -\hat{\theta} \cdot \Delta\bar{p}_t$, where $\Delta\bar{p}_t$ is as in Proposition 4. Consequently, the blockage discount becomes

$$BD = \underbrace{\frac{\alpha}{I}\hat{\mathcal{V}}}_{\hat{\theta}\cdot\Delta^f} + \underbrace{\frac{(1-\gamma)^{2(T-t^*)+1}}{\gamma}\frac{\alpha}{I}\hat{\mathcal{V}}}_{\hat{\theta}\cdot\Delta^l}, \tag{38}$$

where $\hat{\mathcal{V}} \equiv Var(A \cdot \hat{\theta})$ measures the risk of the block. The discount is positively correlated with the riskiness of the block and the risk aversion of the liquidity providers. The higher the variance of an asset payoff, or the more risk averse the liquidity providers, the harder it will be to liquidate the asset on the market. Market depth γ reduces the blockage discount, and at the extreme, in perfectly competitive markets ($I = \infty$), the blockage discount is equal to zero. Finally, blockage discount depends linearly on the variance of the payoff, and therefore, it is quadratic in the size of a block.

In the derivation of formula (38), we made two implicit assumptions: That the block is being sold all at once and that the owner does not have any other assets but the considered block. In

---

[32] These are distinct from (though sometimes confused with) *restricted stock discounts* due to the difficulty in selling because of regulatory or contractual constraints.

[33] For a summary of U.S. Tax Court decisions involving blockage discounts, see Estabrook (1999, 2001); and Pratt (2001).

practice, traders break up large packages into smaller blocks and sell them slowly over time to mitigate partially the adverse effects resulting from market thinness. Therefore, formula (38) is likely to overestimate the value of a blockage discount, and hence it should be interpreted as an upper bound on the discount. The lower bound for the blockage discount is the fundamental effect, as that effect is present even if the trade is spread over time. If a trader has other assets which are not included in $\hat{\theta}$ and whose payoffs are positively (negatively) correlated with the liquidated portfolio, then the liquidation also affects the values of these assets. The blockage discount should then be adjusted upwards (downwards) accordingly.

5.5 Anticipated Shocks

So far, we have considered the supply shocks that are announced only at the moment of trade. Many shocks in financial markets are announced long before they actually occur (e.g., changing the weights of a stock market index, selling IPOs). Extensive empirical evidence shows that such future shocks have strong effects on today's price. In this section, we discuss how prices are affected by the timing between announcement and of the shock occurrence. We assume that in period t^* liquidity providers learn that an extra supply of assets will become available in period t^{**}. A formal derivation of such an equilibrium with anticipated (sequence of) shocks is presented in the Appendix.

The evolution of price response to an announced once-and-for-all shock is depicted in Figure 3. For all $t \geq t^*$, the fundamental value $\bar{v}_t$ is determined by formula (35), where θ_t^{Av} is an expectation of the ultimate average portfolio θ_{T+1}^{Av}, given the information in period t. The information about the future sales of an asset is instantaneously incorporated into expectation $\theta_{t^*}^{Av}$, and the fundamental value decreases during the period of the announcement, t^*. Since no additional information is released after this period, $\bar{v}_t$ remains at the new level till T. Similar to a competitive model, this price pattern is observed even though the actual sales take place only in t^{**}. We conclude then that the permanent, fundamental effect Δ^f occurs at the moment of the announcement *and not* at the moment of trade. The liquidity effect t^{**}, takes place only in the period of the actual shock, Δ^l, as the latter is caused by the change in the order of non-competitive traders and not learning about the stocks of sharers. In t^{**}, large liquidity providers (on average) buy risky assets; hence, their average marginal revenue exceeds the observed price. When the shock is anticipated, in addition to both the fundamental and liquidity effects, we also observe a third effect, that occurs between t^* and $t^{**}-1$. Namely, in all periods, the price is lowered by $\gamma\Delta^l$. The third effect arises because the anticipation of the depressed price in t^{**} discourages traders from buying assets in all periods prior to t^*, and hence, buyers require price discounts while sellers are willing to accept it. In short, at t^*, the equilibrium price $\bar{p}_t$, overshoots its new fundamental value $\bar{v}_{t^*}$ by $\gamma\Delta^l$. This additional overshooting persists until t^{**}, when the full liquidity effect is realized, and the price attains the new fundamental value in all the following periods.

The equilibrium trading strategies are similar to those following unanticipated shocks. After t^*, in each period, the traders hedge their idiosyncratic risk by replacing faction γ of their current portfolios with γ of the new average portfolio, and in t^{**}, in addition, they absorb the shock.

Given perfect foresight, in period t^*, traders learn that at t^{**} price will be temporarily depressed (strictly) below the current price $\bar{v}^i_{t^*} - \gamma\Delta^l$. Yet, they do not sell the assets prior to the shock to rebuy them again in t^{**}. The intuition for why traders do not take advantage of the arbitrage opportunity is similar to the case of unanticipated shocks. At the moment of the shock, the traders buy on average positive number of shares, which implies that the average marginal revenue, equal to $\bar{v}^i_{t^*}$, is higher than the price and is equal to a fundamental value. Buying more in t^{**} would, thus, not be profitable.

The empirical studies by Newman and Rierson [2004] and Citygroup [2005] support our findings.

5.6 Anticipated Multiple Blocks

In practice, liquidated portfolios often exceed the average volume of trade on the market, and cannot be liquidated all at once without significantly depressing the price. In such cases, Liquidators break up the portfolios into smaller blocks placed sequentially on the market. In Section 5.3, we studied the price effects resulting from sales of multiple blocks. We assumed that, in each period, new shocks came as a surprise. In this section, we examine the effects of sequential trading when the entire sequence of trades is announced before $t = 1$. We assume that the announcement is credible in that the liquidator can commit to such sales.

Without loss of generality, we consider portfolio $\hat{\theta}$ that is liquidated between $t = 1$ and T by selling small blocks $\{\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_T\}$, such that $\hat{\theta} = \sum_{t=1}^T \hat{\theta}_t$.

The price impacts over time are not affected by the sales of the multiple blocks, and their values are given by Proposition 2. The next proposition characterizes prices on the equilibrium path, given the anticipated sales of multiple blocks. The fundamental value is given by

$$\bar{v} = \bar{A} - \alpha\mathcal{V}\theta^{Av}_{T+1} = \bar{A} - \alpha\mathcal{V}\theta^{Av}_0 + \underbrace{\frac{\alpha}{I}\bar{\mathcal{V}}\hat{\theta}}_{\equiv\Delta^f}. \tag{39}$$

and it is equal to the average marginal utility in the last period of the trade, after absorbing all shocks. The last term in (39) is a cumulative fundamental effect equal to the sum of the fundamental effects in all periods $\Delta^f = \sum_{t=1}^T \Delta^f_t$. Observe that the permanent effect on $\bar{v}$ is independent from how portfolio $\hat{\theta}$ is partitioned. The price path, however, and the cash obtained by selling the portfolio do depend on the block sizes.

Proposition 5 (Anticipated Multiple Blocks) *Consider a vector $\{\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_T\}$ of anticipated sales. In a trading period $t = 1, ..., T$ the price is equal to*

$$\bar{p}_t = \bar{v} - \Delta^l_t - \gamma\sum_{l=1}^{T-t}\Delta^l_{t+l}, \tag{40}$$

where the liquidity effect in period t, Δ^l_t is given by

$$\Delta^l_t = \frac{(1-\gamma)^{2(T-t^*)+1}}{\gamma}\frac{\alpha}{I}\bar{\mathcal{V}}\hat{\theta}_t. \tag{41}$$

In any period, the price departs from the fundamental value by the current liquidity effect, reinforced by fraction γ of the cumulative effect of all future liquidity effects. The current liquidity effect in (40), results from the non-competitive trading of the investors, as discussed in Section 5.1. The cumulative effect of the multiple anticipated block is new: The anticipation of the depressed price in the future reduces the incentives to buy and strengthens the incentives to sell today, which lowers the price. Interestingly, irrespective of how far in the future the liquidity shock occurs, it always affects today's price with weight γ. The farther away the liquidity shock is from today, the smaller will be the fraction of today's trade that survives till the shock period with lower price. On the other hand, the cumulative effects affects all prices between today and the period of the shock, which increases the weight.

In sum, similar to an unanticipated sequence of sales, long run asset prices are not affected by how the portfolio is divided into smaller blocks or the time at which the trade takes place. Right afterwards the liquidation prices converge to new fundamental value $\bar{v}$. By contrast, the price path is sensitive to how the portfolio is partitioned. This occurs because future sales depress the price during the whole period between the announcement and the shock, and the effects on prices are cumulative. As a result, the liquidator should be expected to concentrate most of the trade in the first period.

Equation (40) suggests that the liquidator has strong incentives not to announce the liquidation. If the sequence to-be-placed on the market is announced in advance, and the whole portfolio $\hat{\theta}$ is being liquidated, the fundamental value instantaneously drops by Δ^f. Furthermore, the ability to smooth the current liquidity effect is reduced, as future liquidity effects adversely affect prices today. If instead the sequence is not announced the fundamental value of the portfolio decreases only slowly in each period in which the blocks are traded. This benefits the liquidator, who receives a better price for the initial blocks. Also the cumulative liquidity effects are not present, and hence the sequence can be made arbitrary long, making the current liquidity effects negligible. Therefore, the liquidator is strictly better off by not announcing the liquidation, as he benefits from a slower adjustment of the fundamental value and the absence of cumulative liquidity effect.

6 Robustness of Model Predictions

6.1 Beyond Homogeneity of Risk Aversion

In Section 4, we argued that non-competitive trading does not alter asset prices (and returns) and, hence, the Security Market Line holds. One of the key assumptions maintained in the analysis was that all investors are equally risk averse. In this section, in a simple example[34] with $T = 1$ and one risky asset, we illustrate how the asset returns are affected by the heterogeneous risk attitudes of the traders α^i. We begin by observing that if investors can be strictly ordered according to the degree of their risk aversion, so can their price impacts.

Lemma 1 *Suppose* $\alpha^i > \alpha^j$. *Then, in equilibrium,* $\bar{\mathcal{M}}^i < \bar{\mathcal{M}}^j$.

---

[34] The example is borrowed from Weretka(2006), where it was given in a more abstract context.

Lemma 1 asserts that a relatively more risk averse trader has a lower price impact. Focusing on the highest-α^i investor, he trades with partners whose marginal utility is flatter, and who thus requires only modest price concessions to absorb potential deviations.

Suppose there are two types of traders, buyers, and sellers, such that their initial portfolios satisfy $\theta_0^s > \theta_0^b$.

The market clearing can be written as

$$\begin{aligned}\sum_{i\in I}\Delta\bar{\theta}^i(\bar{p},\bar{\mathcal{M}}^i) &= \sum_{i\in I}\frac{\bar{A}-\alpha^i\bar{\mathcal{V}}\theta_0^i-\bar{p}}{\bar{\mathcal{M}}^i+\alpha^i\bar{\mathcal{V}}} = \\ &= \sum_{i\in I}\frac{\alpha^i\bar{\mathcal{V}}}{\bar{\mathcal{M}}^i+\alpha^i\bar{\mathcal{V}}}\frac{\bar{A}-\alpha^i\bar{\mathcal{M}}^i\theta_0^i-\bar{p}}{\alpha^i\bar{\mathcal{V}}} \\ &= \sum_{i\in I}\frac{\alpha^i\bar{\mathcal{V}}}{\bar{\mathcal{M}}^i+\alpha^i\bar{\mathcal{V}}}\Delta\bar{\theta}^i(\bar{p},0).\end{aligned} \tag{42}$$

If traders are equally risk averse, then their price impacts $\mathcal{M}^i$, and hence ratios, $\alpha^i\bar{\mathcal{V}}/(\bar{\mathcal{M}}^i+\alpha^i\bar{\mathcal{V}})$ also coincide, and in the last line of (42) the multipliers can be factored out. This fact allowed us in Section 4 to concluded that the price that clears the market in the TM-CAPM also clears the perfectly competitive model. Now, suppose that the buyers are more risk averse than the sellers, $\alpha^b > \alpha^s$. By Lemma 1, the price impact of the buyers is smaller than that of the sellers, $\mathcal{M}^b < \mathcal{M}^s$, and hence the ratios in (42) can be ranked as follows

$$\frac{\alpha^b\bar{\mathcal{V}}}{\bar{\mathcal{M}}^b+\alpha^b\bar{\mathcal{V}}} > \frac{\alpha^s\bar{\mathcal{V}}}{\bar{\mathcal{M}}^s+\alpha^s\bar{\mathcal{V}}}, \tag{43}$$

Given that, in the last line of (42), the sellers' competitive trades are negative and have smaller weights, replacing the ratios with weights equal to one for all traders would make the sum negative

$$\sum_{i\in I}\Delta\bar{\theta}^i(\bar{p},0) < 0. \tag{44}$$

It follows then that at a (non-competitive) equilibrium price, $\bar{p}$, the competitive excess demand is strictly negative. Since the competitive aggregate demand is strictly decreasing in price, the price that clears the competitive market $\bar{p}^{Competitive}$ must be smaller than $\bar{p}$. This results in a negative non-competitive bias in the asset return (for any possible realization of A)

$$R_n < R_n^{Competitive}. \tag{45}$$

Intuitively, in thin markets with $\alpha^b > \alpha^s$ the asset supply is affected more than the demand for any p, and hence price must go up to clear the market. By the symmetric argument, the sign of the price bias would revert when the sellers are more risk averse than the buyers, and so would the asset returns. The example demonstrates that in the TM-CAPM, the price can be greater, the same, or below the competitive one, depending on which side of the market is more risk averse. Consequently, the expected asset returns may be located below or above the Security Market Line. We next discuss how additional assumptions about the heterogeneity of risk aversion identify systematic biases in TM-CAPM with the CRRA utility functions.

6.2 Beyond CARA-Normal Assumption

In Carvajal and Weretka (2007) we examine the effects of thin trading on asset returns, without making any specific assumptions on the functional form of the utility function. We find that when investors have identical, but not necessarily quadratic, utility functions, the derived prices would systematically differ from the competitive prices. The sign of the bias depends on the convexity of the marginal utility. A positive (negative) third derivative of the utility function is associated with a higher (lower) price than the competitive price. In this paper, the indirect utility function is quadratic, and hence, the third derivative is equal to zero. Therefore, the TM-CAPM represents a knife-edge case for which price effects are not observed.

Allowing for utility functions characterized by constant relative risk aversion (CRRA) would yield a positive price bias. In the financial markets, this translates into a spread between the risky and riskless return that is greater than that in the C-CAPM (the equity premium puzzle). In addition, the endogenously determined, risk-free interest rate is lower than that predicted by the competitive model (the interest rate puzzle). Thus, thin trading is associated with the mechanism that has suggests an explanation for the empirically observed biases in asset returns.

Appendix

Proof. (Proposition 1: Three Fund Separation) The result follows from (19). ■

Proof. (Proposition 2: Time Structure of Price Impact) The result follows from (16) and (22). ■

Proof. (Proposition 3: Security Market Line) The result follows from the two observations: By (18), the prices, and hence the asset returns, are as in the competitive model; Second, formula (28) holds in the competitive model. ■

We now give an auxiliary result that is helpful in proving the remaining Propositions of the paper. Let θ_t^{Av} be the average portfolio in period t defined in (29), $\bar{v}_t$ be the average utility on the last day of trade (called a fundamental value)

$$\bar{v} = \bar{A} - \alpha \mathcal{V} \theta_{T+1}^{Av}. \tag{46}$$

and λ_t and Δ_t^l are defined as in (22) and (41), respectively.

Proposition 6 (Equilibrium with Supply Shocks) *Let $\{\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_T\}$, be an anticipated sequence of liquidated blocks: In the TM-CAPM, the equilibrium prices are given by*

$$p_t = \bar{v} - \Delta_t^l - \gamma \sum_{l=1}^{T-t} \Delta_{t+l}^l \tag{47}$$

and the policy functions are

$$\Delta\theta_t = \gamma\bar{\theta}_t^{Av} - \gamma\bar{\theta}_t + \frac{\hat{\theta}_t}{I} \tag{48}$$

Proof. (Proposition 6: Equilibrium with Supply Shocks) Let $(\hat{\theta}_1, ..., \hat{\theta}_T)$ be a sequence of anticipated liquidity shocks.

We proceed by induction. In the last trading period, T, the value function is given by

$$v_T(\Delta\theta^i_T, \Delta\theta^i_{b,T}) = \bar{\theta}^i_{b,T} + \Delta\bar{\theta}^i_{b,T} + \bar{A}(\bar{\theta}^i_T + \Delta\bar{\theta}^i_T) - \frac{\alpha}{2}(\bar{\theta}^i_T + \Delta\bar{\theta}^i_T)\mathcal{V}(\bar{\theta}^i_T + \Delta\bar{\theta}^i_T) \tag{49}$$

Setting equal marginal utility and marginal revenue yields

$$\bar{A} - \alpha(\bar{\theta}^i_T + \Delta\bar{\theta}^i_T)\mathcal{V} = \bar{p}_T + \mathcal{M}^i_T\Delta\bar{\theta}^i_T \tag{50}$$

Solving for the portfolio and summing horizontally for all other traders gives price impact equal to $\bar{\mathcal{M}}^i_T = \alpha\mathcal{V}(1-\gamma)/\gamma$. Averaging F.O.C. (50) across all trades and using $\sum_{i\in I}\Delta\bar{\theta}^i_T = \hat{\theta}_T$ gives

$$\bar{A} - \alpha\mathcal{V}(\bar{\theta}^{Av}_T + \frac{\hat{\theta}_T}{I}) = \bar{p}_T + \frac{1-\gamma}{\gamma}\alpha\mathcal{V}\frac{\hat{\theta}_T}{I}, \tag{51}$$

from which we derive the equilibrium prices in period T,

$$\bar{p}_T = \bar{v} - \frac{1-\gamma}{\gamma}\alpha\mathcal{V}\frac{\hat{\theta}_T}{I} = \bar{v} - \Delta^l_T. \tag{52}$$

Substituting the derived prices (52) back into the F.O.C. (50), we obtain the

$$\Delta\bar{\theta}^i_T = \gamma\left(\bar{\theta}^{Av}_T - \bar{\theta}^i_T\right) + \frac{\hat{\theta}_T}{I} \tag{53}$$

$$\begin{aligned} \bar{\theta}^i_{T+1} &= \bar{\theta}^i_T + \Delta\bar{\theta}^i_T = \bar{\theta}^i_T + \gamma\left(\bar{\theta}^{Av}_T - \bar{\theta}^i_T\right) + \frac{\hat{\theta}_T}{I} = \gamma\bar{\theta}^{Av}_T + (1-\gamma)\,\bar{\theta}^i_T + \frac{\hat{\theta}_T}{I} = \\ &= \gamma\bar{\theta}^{Av}_T + (1-\gamma)\left(\bar{\theta}^i_{T-1} + \Delta\theta^i_{T-1}\right) + \frac{\hat{\theta}_T}{I} \end{aligned} \tag{54}$$

$$\begin{aligned} \bar{\theta}^i_{b,T+1} &= \bar{\theta}^i_{b,T} - \bar{p}_T\Delta\theta^i_T = \bar{\theta}^i_{b,T} + \bar{p}_T\left(\gamma\Delta\theta^i_{T-1} + \gamma\bar{\theta}^i_{T-1} - \gamma\bar{\theta}^{Av}_T - \frac{\hat{\theta}_T}{I}\right) = \\ &= \bar{\theta}^i_{b,T} + \left(\bar{v} - \frac{1-\gamma}{\gamma}\alpha\mathcal{V}\frac{\hat{\theta}_T}{I}\right)\left(\gamma\Delta\theta^i_{T-1} + \gamma\bar{\theta}^i_{T-1} - \gamma\bar{\theta}^{Av}_T - \frac{\hat{\theta}_T}{I}\right) = \\ &= \bar{\theta}^i_{b,T} + \left(\bar{v} - \frac{1-\gamma}{\gamma}\alpha\mathcal{V}\frac{\hat{\theta}_T}{I}\right)\left(\gamma\bar{\theta}^i_T - \gamma\bar{\theta}^{Av}_T - \frac{\hat{\theta}_T}{I}\right) \end{aligned} \tag{55}$$

Suppose now that for all periods from $t+1$ to T, the policy functions (θ_t, p_t, M_t) are as in Section 5.1. We will show that it also holds in t. Since these are policy functions the ultimate portfolio

$\bar{\theta}_{T+1}, \bar{\theta}_{b,T+1}$, as a function of trades in period t is given by

$$\bar{\theta}_{T+1} = \gamma \bar{\theta}_{T}^{Av} + (1-\gamma)\,\bar{\theta}_T + \frac{\hat{\theta}_T}{I}. \tag{56}$$

Since

$$\Delta\theta_{t+l} = -\gamma\,(1-\gamma)^{l-1}\left(\bar{\theta}_t + \Delta\theta_t\right) + c_t \tag{57}$$

then

$$\bar{\theta}_{b,T+1} = \bar{\theta}_{b,t} + \Delta\theta_{b,t} + \left(\bar{\theta}_t + \Delta\theta_t\right)\sum_{l=1}^{T-t} p_{t+l}\gamma\,(1-\gamma)^{l-1} \tag{58}$$

Then the average F.O.C. can be written as

$$\underbrace{\lambda_t\left[\bar{A} + \alpha V\left((1-\lambda_t)\,\theta_{t+1}^{Av} + \lambda_t\left(\bar{\theta}_t + \Delta\theta_t\right) + \sum_{l=1}^{T-t}\frac{\hat{\theta}_{t+l}}{I}\right)\right]}_{\text{Term 1}} + \underbrace{\sum_{l=1}^{T-t} p_{t+l}\gamma\,(1-\gamma)^{l-1}}_{\text{Term 2}} = p_t + \tilde{\theta}_t \tag{59}$$

We average the F.O.C. on both sides. The averaged Term 1 is equal to $\lambda_t\bar{v}$ while Term 2 is a sum of prices in periods that follow period t. Since, by assumption, prices are linear functions of trades $\tilde{\theta}_k$ (for $k \geq t$) and $\bar{v}$, the term is also a linear function of those variables. Now we find coefficients that multiply them. Since $\bar{v}$ is part of any of the prices the coefficient multiplying can be found by simply applying formula for a finite sum of geometric sequence

$$\bar{v}\sum_{l=1}^{T-t}\gamma\,(1-\gamma)^{l-1} = \gamma\left(\frac{1-(1-\gamma)^{T-t}}{1-1-\gamma}\right) = \bar{v}\left(1-(1-\gamma)^{T-t}\right) = \bar{v}\,(1-\lambda_t) \tag{60}$$

An arbitrary shock from period k, $\tilde{\theta}_k$ is a component of all prices between t and $k-1$ (multiplied by a coefficient γ) and in period k (without any multiplication). Crucially, this term is not present in prices that follow period k

$$-\gamma\tilde{\theta}_k\left[\gamma\sum_{l=1}^{k-1-t}(1-\gamma)^{l-1} + (1-\gamma)^{k-t-1}\right] = \tag{61}$$

$$-\gamma\tilde{\theta}_k\left[\gamma\frac{1-(1-\gamma)^{k-t-1}}{\gamma} + (1-\gamma)^{k-t-1}\right] = -\gamma\tilde{\theta}_k \tag{62}$$

Note that this holds for any $k \in [t, T]$. Hence

$$Term2 = \bar{v}\,(1-\lambda_t) - \gamma\sum_{l=1}^{T-t}\tilde{\theta}_{t+l}. \tag{63}$$

Therefore, the averaged F.O.C. can be rewritten as

$$\underbrace{\lambda_t\bar{v}}_{\text{Term 1}} + \underbrace{\bar{v}\,(1-\lambda_t) - \gamma\sum_{l=1}^{T-t}\tilde{\theta}_{t+l}}_{\text{Term 2}} = \bar{v} - \gamma\sum_{l=1}^{T-t}\tilde{\theta}_{t+l} = p_t + \tilde{\theta}_t \tag{64}$$

so that

$$p_t = \bar{v} - \tilde{\theta}_t - \gamma \sum_{l=1}^{T-t} \tilde{\theta}_{t+l}$$

Hence, the formula for price holds also in period t. Finally, we verify that the policy functions satisfy the conjectured recursive relation.

$$\begin{aligned}
& \lambda_t \left[\bar{A} - \alpha V \left((1-\lambda_t)\, \theta^{Av}_{t+1} + \lambda_t \left(\bar{\theta}_t + \Delta\theta_t \right) + \sum_{l=1}^{T-t} \frac{\hat{\theta}_{t+l}}{I} \right) \right] + \sum_{l=1}^{T-t} p_{t+l} \gamma \left(1-\gamma\right)^{l-1} \\
= \quad & p_t + \frac{(1-\gamma)^{2(T-t)+1}}{\gamma} \alpha V \Delta\theta_t
\end{aligned}$$

$$\begin{aligned}
& \lambda_t \left[\bar{A} - \alpha V \left((1-\lambda_t)\, \theta^{Av}_{t+1} + \lambda_t \left(\bar{\theta}_t + \Delta\theta_t \right) + \sum_{l=1}^{T-t} \frac{\hat{\theta}_{t+l}}{I} \right) \right] + \bar{v}\,(1-\lambda_t) - \gamma \sum_{l=1}^{T-t} \tilde{\theta}_{t+l} \\
= \quad & \bar{v} - \tilde{\theta}_t - \gamma \sum_{l=1}^{T-t} \tilde{\theta}_{t+l} + \frac{(1-\gamma)^{2(T-t)+1}}{\gamma} \alpha V \Delta\theta_t
\end{aligned}$$

$$\begin{aligned}
& \lambda_t \left[\bar{A} - \alpha V \left((1-\lambda_t)\, \theta^{Av}_{t+1} + \lambda_t \bar{\theta}^{Av}_{t+1} + \sum_{l=1}^{T-t} \frac{\hat{\theta}_{t+l}}{I} \right) + \lambda_t \alpha V \left(\bar{\theta}^{Av}_{t+1} - \bar{\theta}_t - \Delta\theta_t \right) \right] - \bar{v}\lambda_t \\
= \quad & \frac{(1-\gamma)^{2(T-t)+1}}{\gamma} \alpha V \Delta\theta_t - \tilde{\theta}_t
\end{aligned} \tag{65}$$

$$\begin{aligned}
\left(\bar{\theta}^{Av}_{t+1} - \bar{\theta}_t - \Delta\theta_t \right) \gamma &= (1-\gamma) \left(\Delta\theta_t - \frac{\hat{\theta}_t}{I} \right) \\
\gamma \bar{\theta}^{Av}_t + \gamma \frac{\hat{\theta}_t}{I} - \gamma \bar{\theta}_t - \gamma \Delta\theta_t &= (1-\gamma)\, \Delta\theta_t - (1-\gamma)\, \frac{\hat{\theta}_t}{I} \\
\Delta\theta_t &= \gamma \bar{\theta}^{Av}_t - \gamma \bar{\theta}_t + \frac{\hat{\theta}_t}{I}
\end{aligned} \tag{66}$$

■

Proof. (LEMMA 1: PRICE IMPACT ORDER)

Suppose $\alpha^i > \alpha^j$ and $\bar{\mathcal{M}}^i \geq \bar{\mathcal{M}}^j$. Then, $\bar{\mathcal{M}}^i + \alpha^i \bar{\mathcal{V}} > \bar{\mathcal{M}}^j + \alpha^j \bar{\mathcal{V}}$. Consequently, the harmonic mean condition becomes

$$\mathcal{H}\left(\bar{\mathcal{M}}^k + \alpha^k \bar{\mathcal{V}} | k \neq i \right) < \mathcal{H}\left(\bar{\mathcal{M}}^k + \alpha^k \bar{\mathcal{V}} | k \neq j \right). \tag{67}$$

In equilibrium,

$$\begin{aligned}\bar{\mathcal{M}}^i &= (1-\gamma)\,\mathcal{H}\left(\bar{\mathcal{M}}^k + \alpha^k \bar{\mathcal{V}} | k \neq i\right), \\ \bar{\mathcal{M}}^j &= (1-\gamma)\,\mathcal{H}\left(\bar{\mathcal{M}}^k + \alpha^k \bar{\mathcal{V}} | k \neq j\right),\end{aligned} \tag{68}$$

which yields a contradiction: $\bar{\mathcal{M}}^i < \bar{\mathcal{M}}^j$. ■

References

[1] Aitken, M., Brown P. and Walter, T. (1994): "Intraday Patterns in Returns, Trading Volume, Volatility, and Trading Frequency on SEATS', University of Sydney, working paper.

[2] Almgren, R. (2003): 'Optimal Execution with Non-linear Impact Functions and Trading Enhanced Risk," *Applied Mathematical Finance*, 10: 1-18.

[3] Almgren, R., and Chriss, N. (2000): "Optimal Execution of Portfolio Transactions," *Journal of Risk*, 3: 5-39.

[4] Almgren, R., Thum, C., Hauptmann, E. and Li, H. (2005): "Equity market impact," *Risk.*

[5] Amihud, Y., Mendelson, H. (1986): "Asset Pricing and Bid-Ask Spread," *Journal of Financial Economics*, 17: 223-249.

[6] Amihud, Y., Mendelson, H. and Pedersen, L. (2005): 'Liquidity and Asset Prices," *Foundations and Trends in Finance*, 1: 269-364.

[7] Attari, M. Mello, A. and Ruckes, M. (2005): "Arbitraging Arbitrageurs," *Journal of Finance*, 60: 6247-2511.

[8] Barclay, M. and Holderness, C. (1992): 'The Law and Large Block Traders," *Journal of Law & Economics*, 35: 265-294.

[9] Beneish, M. and Whaley, R. (1996): "An Anatomy of the 'S&P Game, The Effects of Changing the Rules," *Journal of Finance*, 5: 1909-1930.

[10] Benston, G., and Hagerman, R. (1974): "Determinants of the Bid-ask Spreads in the Over-the-Counter Markets," *Journal of Financial Economics,*1: 353-364.

[11] Bertsimas, D. and Lo, A. (1998): "Optimal Control of Execution Costs," *Journal of Financial Markets,"* 1: 1-50.

[12] Biais, B., Glosten, L. and Spatt, C. (2002): "The microstructure of stock markets," *Journal of Financial Intermediation*, forthcoming.

[13] Brunnermeier, M. and Pedersen, L. (2005): "Predatory Trading," *Journal of Finance*, 4: 1825-1863.

[14] Brunnermeier, M. and Pedersen, L. (2006): "Market Liquidity and Funding Liquidity," working paper.

[15] Carvajal, A. and Weretka, M. (2007): "State Prices and Arbitrage in Thin Financial Markets, working paper," University of Wisconsin-Madison.

[16] Chan, L. and Lakonishok J. (1993): "Institutional Traders and Intraday Stock Price Behavior," *Journal of Financial Economics*, 33: 173-199.

[17] Chan, L. and Lakonishok, J. (1995): "The Behavior of Stock Price Around Institutional Trades," *Journal of Finance*, 50: 1147-1174.

[18] Chordia, T., Roll, R. and Subrahmanyam, A. (2001): "Market Liquidity and Trading Activity," *Journal of Finance*, 56: 501-530.

[19] Chordia, T., Roll R., and Subrahmanyam, A. (2002): "Commonality in Liquidity," *Journal of Financial Economics*, 56: 3-28.

[20] Cochrane, J. (2005): "Asset Pricing Program Review: Liquidity, Trading and Asset Prices," NBER Reporter.

[21] Constantinides, G. M. (1986): "Capital market equilibrium with transaction costs," *Journal of Political Economy*, 94: 842-62.

[22] DeMarzo, P. M. and Urošević, B. (2006): Ownership Dynamics and Asset Pricing with a Large Shareholder," *Journal of Political Economy*, 4: 145-174.

[23] Domowitz, I., Hansch, O. and Wang, X. (2005): 'Liquidity Commonality and Return Comovement," *Journal of Financial Markets*, 8: 351–376.

[24] Dubil, R. (2002): "Optimal Liquidation of Venture Capital Stakes," *Journal of Entrepreneurial Finance and Business Ventures*, 7, 2.

[25] Duffie, D., Garleanu, N. and Pedersen, L. H. (2005): "Over-the-counter markets," *Econometrica*, 73: 1815-47.

[26] Dufour, A. and Engle, R. (2000): "Time and the Price Impact of the Trade," *Journal of Finance*, 6: 2467-2498.

[27] Easley, D. and O'Hara, M. (2003): "Microstructure and Asset Pricing," in: Constantinides, G., Harris, M. and Stulz R. (eds.), Handbook of Financial Economics, B.V. North Holland, Elsevier Science Publishers.

[28] Estabrook, J. (1999): "Blockage Discounts," in: Handbook of Advanced Business Valuation, R. F. Reilly and R. P. Schweihs eds., McGraw-Hill.

[29] Estabrook, J. (2001): "Blockage Discounts: What are they and How do they Affect the Fair Market Value of Publicly Traded Stocks, Real Estate, and Other Assets," American Society of Appraisers International Appraisal Conference.

[30] Foster F. and Visvanathan, S. (1990): "A Theory of the Intraday Variation in Volume, Variance, and Trading Costs in Security Markets," *Review of Financial Studies*, 3: 539-624.

[31] Foster, F. D. and Viswanathan, S. (1996): "Strategic Trading When Agents Forecast the Forecasts of Others," *Journal of Finance*, 51(4), 1437-78.

[32] Gerety, M. and Mulherin, J. (1992): "Trading Halts and Market Activity: an Analysis of Volume at the Open and the Close," *Journal of Finance*, 5: 1765-84.

[33] Glosten, L. and Milgrom, P. (1985): "Bid, ask, and transaction prices in a specialist market with heterogeneously informed traders," *Journal of Financial Economics*, 13: 71-100.

[34] Greenwood, R. (2005): "Short- and Long-term Demand Curves for Stocks: Theory and Evidence on the Dynamics of Arbitrage," *Journal of Financial Economics*, 75, 3: 607-649.

[35] Hamao, Y. and Hasbrouck, J. (1995): "Securities Trading in the Absence of Dealers: Trades and Quotes on the Tokyo Stock Exchange," *Review of Financial Studies*, 8: 849-878.,

[36] Harris, L. and Gurel, E. (1986): "Price and Volume Effects Associated with Changes in the S&P 500: New Evidence for the Existence of Price Pressures," *Journal of Finance*, 41: 815-829.

[37] Hasbrouck, J. and Seppi, D. (2001): "Common Factors in Prices, Order Flows and Liquidity," *Journal of Financial Economics*, 59: 383–411.

[38] Hau, H. and Rey, H. (2004): "Can portfolio Rebalancing Explain the Dynamics of Equity Returns, Equity Flows, and Exchange Rates?," *American Economic Review,* 94: 126-133.

[39] Hau, H., Massa, M., and Peress, J. (2005): "Do Demand Curves for Curriencies Slope Down? Evidence from the MSCI Global Index Change," *CERP Discussion Papers*, 4862.

[40] Holthausen, R., Leftwich R. and Mayers, D. (1987): "The Effects of Large Block Transactions on Security Prices: A Cross Sectional Analyses," *Journal of Financial Economics*, 19: 237-267.

[41] Holthausen, R., Leftwich R. and Mayers, D. (1990): "Large Block Transactions The Speed of Response and Temporary and Permanent Stock-Price Effects', *Journal of Financial Economics*, 26: 71-95.

[42] Huberman, G. and Halka, D. (2001): "Systematic Liquidity," *Journal of Financial Research*, 24: 161-178.

[43] Huberman, G. and Stanzl, W. (2004) "Price Manipulation and Quasi-Arbitrage," *Econometrica*, 74(4): 1247-1276..

[44] Jain, P. and Joh, G. (1988): "The dependence between the hourly prices and trading volume," *Journal of Financial and Quantitative Analysis*, 9: 269-83.

[45] Kaul, A., Mehrotra V. and Morck R. (2000): "Demand Curves for Stocks Do Slope Down: New Evidence from an Index Weights Adjustment," *Journal of Finance*, 55, 2: 893-912.

[46] Keim, D. and Madhavan, A. (1995): "The Anatomy of the Trading Process: Empirical Evidence on the Behavior of Institutional Traders," *Journal of Financial Economics*, 37: 371-398.

[47] Keim, D., and Madhavan, A., (1996): "The Upstairs Markets for Large-Block Transactions. Analyses and Measurement of Price Effects," *Review of Financial Studies*, 9: 1-39.

[48] Keim, D. and Madhavan, A. (1998): “Execution Costs and Investment Performance: An Empirical Analysis of Institutional Equity Trades,” working paper, USC.

[49] Kempf, A. and Korn, O. (1999): “Market Depth and Order Size,” *Journal of Financial Markets*, 2: 29-48.

[50] Klemperer, P. and Meyer, M. (1989): “Supply Function Equilibria in Oligopoly under Uncertainty,” *Econometrica*, 57: 1243-1277.

[51] Kraus, A. and Stoll, H. (1972): “Price Impact on Block Trading on the New York Stock Exchange,” *Journal of Finance*, 27: 569-588.

[52] Kyle, A. (1985): “Continuous Auctions and Insider Trading’, *Econometrica*, 53: 1315-1336.

[53] Kyle, A. (1989): Informed Speculation and Imperfect Competition,” *Review of Economic Studies*, 56: 517-556.

[54] Lagos, R. and Rochetau, G. (2006): “Search in Asset Markets,” Federal Reserve Bank of Minneapolis, Research Department Staff Report 375.

[55] Loderer, C., Cooney, J., and Van Drunen, L. (1991): “The Price Elasticity of Demand for Common Stock,” *Journal of Finance*, 46: 621–651.

[56] Loeb, T. (1983): “Trading Cost: The Critical Link between Investment Information and Results,” *Financial Analysts Journal*, 39: 39-44.

[57] Lynch A. and Mendenhall, R. (1997): “New Evidence on Stock Price Effects Associated with Changes in the S&P Index,” *Journal of Business*, 70: 351-383.

[58] Madhavan, A. (2000): “Market Microstructure: A Survey,” *Journal of Financial Markets*, 3: 205-258.

[59] Madhavan, A. and Cheng, M. (1997): “The upstairs markets for large-block transactions: Analysis and measurement of price effects,” *Review of Financial Studies*, 9: 1-36.

[60] Madhavan, A., Richardson, M., and Roomans, M. (1997): “Why Do Security Prices Change? A Transaction-Level Analysis of NYSE Stocks,” *Review of Financial Studies*, 10: 1035-1064.

[61] Newman, Y. and Rierson, M. (2004): “Illiquidity Spillovers: Theory and Evidence from European Telecom Bond Issuance,” working paper.

[62] O’Hara, M. (1995): Market Microstructure Theory, Blackwell Publishers, Cambridge, MA.

[63] Obizhaeva, A. and Wang, J. (2005): “Optimal Trading Strategy and Supply/Demand Dynamics,” working paper.

[64] Perold, A. (1988): “The Implementation Shortfall: Paper versus Reality,” *Journal of Portfolio Management*, 14: 4-9.

[65] Plerou, V., Gopikrishnan, P., Gabaix, X., and Stanley, H (2002): “Qunatifying Stock-price Response to Demand Fluctuactions,” *Phisical Review E*, 66: 027104.1-4.

[66] Pratt, S. (2001): “Business Valuation Discounts and Premiums,” John Wiley & Sons Inc.New York.

[67] Pritsker, M. (2005): “Large Investors: Implications for Equilibrium Asset, Returns, Shock Absorbtion, and Liquidity,” Fincance and Economic Discussion Series 2005-36, Board of Governors of the Federal Reserve System.

[68] Sarkar, A., Schwartz, R. and Wolf A. (2005): “Inter-Temporal Trade Clustering and Two Sided Markets: Insights Into Trading Motives,” working paper.

[69] Schwartz, R. and Shapiro, J. (1992): “The challenge of institutionalization for the equity markets,” in Anthony Saunders, ed.: *Recent Developments in Finance*, (New York Salomon Center, New York, NY and Business One Irwin, Homewood, IL)

[70] Seppi, D. (1990): “Equilibrium Block Trading and Asymmetric Information,” *Journal of Finance*, 45: 73-94.

[71] Stoll, H. (1978): “The Pricing of Security Dealers Service: An Empirical Study of NASDAQ Stocks,” *Journal of Finance*, 33, 1153-1172.

[72] Subramanian, A., and Jarrow R. (2001): “The Liquidity Discount,” *Mathematical Finance*, 11, 447-474.

[73] Urošević, B., (2005): “Moral Hazard and Dynamics of Insider Ownership Stakes,” working paper, University Pompeu Fabra.

[74] Vayanos, D. (1998): “Transaction Costs and Asset Prices: A Dynamic Equilibrium Model,” *Review of Financial Studies*, 11: 1-58.

[75] Vayanos D., (1999): “Strategic Trading and Welfare in a Dynamic Market,” *Review of Economic Studies*, 66, 219-54.

[76] Weretka, M., (2006a): “Endogenous Market Power, working paper,” University of Wisconsin-Madison.

[77] Weretka, M., (2006b): “Discovery of Market Power in Bilateral Monopoly, working paper,” University of Wisconsin-Madison.

[78] Wurgler, J., and Zhuravskaya, E. (2002): “Does Arbitrage Flatten Demand Curves for Stocks,” *Journal of Bussiness* 75: 583-608.