Academic Ono College

The Business School Working Paper series

Working Paper No. 2007-2

Volume, Liquidity, and their Risks

when Supply\Demand Curves for Shares are Finitely Elastic[*]

Shmuel Hauser[**]

Haim Kedar-Levy[***]

ABSTRACT

This paper explores the effect of heterogeneous preferences on volume and liquidity in Merton's (1971) economy. Like others, Merton's asset-pricing model assumes that supply\demand curves are infinitely elastic, thereby forego volume and liquidity. We derive finitely elastic supply\demand curves, demonstrating that heterogeneity has major implications for volume, liquidity, and their associated risks. In particular, (i) volume and liquidity increase with heterogeneity; (ii) volume (liquidity) risk increases (declines) with heterogeneity; (iii) turnover-rate may exceed 100%; (iv) negative returns generate stronger illiquidity impacts; (v) firms issue (retire) shares when prices appreciate (depreciate); (vi) capital formation is low (high) in homogeneous (heterogeneous) markets.

---

[*] We are grateful to Yakov Amihud, Doron Avramov, David Feldman, Dan Galai, Arieh Gavious, Eugene Kandel, Itzhak Venezia, and Zvi Wiener for helpful comments and discussions. Previous versions benefited from comments by participants of the Finance Department Research Seminars at the Hebrew University, Jerusalem, the College for Management, the 2006 Far Eastern Meeting of the Econometric Society, Beijing, China, and the Spring 2006 Midwest Economic Theory Conference, Lansing, MI, USA. We assume full responsibility for all remaining errors.

[**] Dean, Ono Academic College, and Finance Professor at the Guilford Glazer School of Business and Management, Ben Gurion University of the Negev.

[***] The Guilford Glazer School of Business and Management, Ben Gurion University of the Negev, P.O.B. 653, Beer Sheva 84105, Israel. hlevy@som.bgu.ac.il Tel. (972) 8-6472-569, Fax (972) 8-6477 697. Corresponding author.

Key words: Volume; Liquidity risk; Heterogeneity; Contrarian; Trend-chasing.

JEL Codes: G11, G12

Volume and liquidity are key measures of efficiency and functionality in financial markets. However, they are typically absent from classic frictionless asset-pricing theories, such as the Capital Assets Pricing Model (CAPM) of Sharpe (1964) and Lintner (1965) as well as the Arbitrage Pricing Theory of Ross (1976). Empirical evidence suggests that microeconomic and microstructure measures do affect asset prices. For example, it is documented that illiquidity is priced, the structure of the book, order size, and trading halts affect the price discovery process, and elasticities of supply and demand for shares are not infinite.[1] These important effects motivate the incorporation of volume and liquidity, and their risks, in asset pricing models (see Acharya and Pedersen, 2005; Johnson, 2006, 2008; Lo and Wang, 2006; Novy-Marx, 2008).

This paper develops a framework in which volume, liquidity, and elasticities, and their associated risks, are embedded in a Merton (1971)-based economy. In our one riskless bond one risky stock economy, we show that intertemporal asset allocation rules yield either an instantaneous upward-sloping "supply" or a downward-sloping "demand" function for shares. The finite elasticities are attributable to differences in investors' attitudes toward risk, and their existence reveal several microeconomic measures of the asset pricing model. Our construction of supply and demand schedules for shares is newly developed here, and is best illustrated numerically. In particular, assume an investor has Constant Relative Risk Aversion (CRRA)

[1] A partial list of this literature includes Amihud and Mendelson (1987, 1989), Loderer, Cooney, and van Drunen (1991), Bagwell (1992), Kandel, Sarig, and Wohl (1999), Liaw, Liu, and Wei (2001), Pastor and Stambaugh (2001), Chordia, Roll, and Subrahmanyam (2002), Levin and Wright (2002), Kalay, Sade, and Wohl (2004), Avramov, Chordia, and Goyal (2006), and Hauser et al. (2006).

preferences, holding 60% of her portfolio in the riskless asset and 40% in the risky asset. Further, assume that the initial portfolio value is $100 and the stock price increases by 10%, then the total portfolio value increases to $104 (ignoring riskless interest). To maintain optimal bond-stock proportions, this new value must be allocated $62.40 and $41.60, respectively. If the investor does not trade, her current asset allocation ($60 and $44) deviates from the optimal rule, and she will not maximize expected utility of wealth. To maintain an optimal allocation the investor must sell $44-$41.60=$2.40 worth of the stock and buy $2.40 worth of the bond. If the stock price is $0.10, our investor's optimal rebalancing implies selling 24 shares. We show that the functional relationship between price changes and implied optimal trades for CRRA investors yields a downward sloping, monotone, and convex Marginal Trade Schedule (MTS).

If CRRA investors sell (buy) shares upon price increases (declines), acting like Contrarians, other investors must buy (sell) those shares for the market to clear each period, and hence act like Trend-chasers (buy upon price increases and sell when prices decline). These trades must comply with the latter investors' optimal asset allocation strategy and be consistent with the equilibrium asset price and share issues\repurchases by the firm. Our formal derivation of the specific risk preferences that imply optimal Trend-chasing reveals that Trend-chasers exhibit a lower coefficient of relative risk aversion (RRA) than the Contrarian investors. The functional relationship between price changes and trades in units of shares for Trend-chasing investors yields an upward sloping, monotone, and convex "supply-like" MTS that complies with Decreasing Relative Risk Aversion (DRRA) preferences. We show that the market price for variance risk is the boundary that distinguishes one group's RRA from the other's. Indeed, the microeconomics in this model emerges from the specific way we break down the single representative investor into two heterogeneous strategists. Because the differences in rebalancing rules stem from heterogeneous risk preferences, we assure that the assumptions in Constantinides (1982) hold and a composite investor is constructed.

Because they have relatively simple functional forms, both instantaneous supply and demand schedules lend themselves easily to the formal analysis of instantaneous volume, liquidity, and elasticities. To conduct the analysis on a time-series basis, where integration is either impossible or tedious, we simulate a benchmark financial market based on NYSE post-war statistics, and replicate it across 16 "markets". The first is the most homogeneous (the dispersion of RRAs is minimal, nearing Merton's single investor), and each subsequent market is populated by increasingly heterogeneous investors. Variables in each market are estimated across 100 sample paths of the standardized diffusion stock price processes, each made of 250 "days", i.e., 25,000 observations.

Equilibrium is supported in our most homogeneous market by about 90% Contrarian investors and only 10% Trend investors, but this composition changes linearly with heterogeneity such that Contrarians are about 10% and Trend-chasers approximately 90% in our most heterogeneous market. Weighted average RRA varies from 2.29 (most homogeneous) to 3.13 (most heterogeneous). Our major findings are: 1) Absolute volume generated by both investors increases linearly with heterogeneity, but its components are not linear: volume by Contrarian investors increases in a concave schedule to a maximum and then declines, while volume by Trend investors increases in a convex schedule. As a result, most volume is generated by Contrarians in relatively homogeneous markets, but Trend investors dominate in heterogeneous markets. Instantaneous share issues\repurchases by the firm complement excess demand\supply. 2) The correlation coefficient between the stock price and the number of shares outstanding is negative $\pm 8\%$ in relatively homogeneous markets, but positive $\pm 10\%$ in heterogeneous markets. This result has important implications for the process of capital formation. Since stock prices increase on average, and because Contrarian investors dominate homogeneous markets, their excess supply requires the firm to repurchase some of its shares, thereby reducing the pace of capital formation, and in some cases turning it negative. Conversely, since Trend investors dominate heterogeneous markets, their excess demand induces the firm to issue new shares, increasing the pace of capital formation. 3) Turnover-rate and its standard deviation increase with heterogeneity, obtaining levels about 100% in moderately heterogeneous markets. 4) The standard deviation of volume increases linearly with the absolute sum of volume, yet its components are not linear; most of the volume risk stems from Contrarian investors in homogeneous markets, but from Trend-chasers and the firm in heterogeneous markets. 5) Market liquidity, defined as the absolute proportional volume over the absolute proportional change in price, is the sum of liquidity that both the investors and the firm generate. Generally, it increases non-linearly with heterogeneity. 6) Liquidity risk, which primarily stems from Trend investors, declines non-linearly with heterogeneity; the standard deviation of liquidity in relatively homogeneous markets is about five times its magnitude in heterogeneous markets. 7) The measure of illiquidity, *ILLIQ* (Amihud, 2002), declines about 20-fold with heterogeneity in a convex schedule, together with its standard deviation. 8) The elasticity of supply is always smaller than the elasticity of demand, yet its standard deviation is smaller than that of the elasticity of demand in homogeneous markets, and exceeds it in heterogeneous markets. 9) The calibrated model generates empirically comparable measures of turnover-rates, illiquidity, liquidity risk, and supply\demand elasticities at the opening of trading days. It is further consistent with the asymmetric illiquidity impact of negative returns, and with the empirical finding that firms issue (retire) shares when their share prices are overvalued (undervalued).

Incorporating volume and liquidity in equilibrium asset pricing models is a timely issue that has been addressed recently in different ways. Closest to our characterization is Johnson's (2006) model, demonstrating that even in the simplest frictionless endowment economies one can define the sensitivity of prices to trades. Acharya and Pedersen (2005) incorporate illiquidity as an exogenous cost of selling an asset in a CAPM setup, whereas Johnson (2008) explores the relationship between volume, liquidity, and liquidity risk as they emerge from changes to the risk bearing capacity of the economy due to random entry and exit of short-lived investors. Entry and exit are exogenous in Johnson's (2008) model, but they are comparable to the endogenous trades in our model. Still, both models generate a few comparable results. The model developed by Lo and Wang (2006) is methodologically close to ours as it explores volumes that stem from allocation changes in a frictionless economy. Novy-Marx (2008) simulates an illiquidity premium in an economy where investors have no explicit demand for liquidity, demonstrating that it stems technically in any mis-specified model.

The economic setting is described in Section 1; in Section 2 we derive intertemporal trade in units of shares and define supply and demand schedules for the shares. In Section 3 we derive volume, liquidity, and elasticities, calibrate the model for NYSE and Tel Aviv Stock Exchange data, and discuss the model predictions based on the simulated economies. Section 4 concludes.

1. The economic setting

Securities

Assume a single bond and a single stock trade in frictionless financial markets. The riskless bond yields a given rate of return r and the stock earns a constant expected rate of return μ. The stock price process is given by

$$dP_t \ / P_t = \mu dt + \sigma dz_t . \tag{1}$$

Merton (1971) shows that with a price-taking representative investor and log-normality of the price processes of a finite number of risky securities, the dynamic model nests the static

CAPM. Yet the dynamics in Merton's model must be cast in terms of quantities and prices if one desires to reveal volume, as do Lo and Wang (2006) and Johnson (2006, 2008). If one further wishes to distinguish buyers from sellers on the premise of preferences, the single representative investor must be replaced with heterogeneous investors. Our goal is to reveal the intertemporal functional relationship between trade and price changes, and to analyze volume, liquidity, and elasticities, and their risks based on the resulting intertemporal supply and demand schedules. To do so we replace Merton's single representative investor with two different investor groups, assuring on the one hand that they have an incentive to trade, and on the other hand that their composite preferences are consistent with equilibrium. Constantinides (1982) demonstrates that under certain assumptions, particularly complete markets, the single-period CAPM can be derived without demand aggregation. This approach is beneficial as it does not require making strong assumptions on preferences for the purpose of demand aggregation, and avoids introducing state variables that stem from changes in the relative wealth of heterogeneous investors. [2]

Market participants

There are two investor groups, dubbed *Trend* and *Contrarian*, each consisting of homogeneous price-taking investors having infinite horizon. While the specific segmentation will be derived formally in the next section, we show here that investors whose coefficient of RRA is

---

[2] These assumptions are: 1) investors have homogeneous beliefs; 2) investors have time-additive von Neumann-Morgenstern utility functions; 3) states at *t* are independent of events at *t-dt*; 4) aggregate endowments at *t* are independent of events at *t*; 5) uncertainty in production technologies is resolved by the states, and is only dependent on events at *t-dt* through the state of the input at *t-dt*; 6) production technology has constant returns to scale; and 7) the set of technological variables and every subset thereof exhibit separability. The economic dynamics in Merton (1971) is consistent with assumptions 1, 3, 5, 6, and 7. To comply with assumption 2, we restrict all the utility functions investors have in our model to the Hyperbolic Absolute Risk Aversion (HARA) type, such that $-U'(W_k)/U''(W_k) = a_k + b_k W_k$ $(U' > 0; U'' < 0)$ for each individual investor *k*. Assumption 4 holds through the budget constraint (2). The resulting economy is dynamically complete and a composite investor may be constructed.

smaller than the market price for variance risk would optimally trade in *units of shares* in the same direction as the price change. Because such investors buy when the stock price increases and sell when it declines, we dub them the *Trend* group (interchangeably "type *T*"; $k=\tau$ for the individual or $K=T$ for the group). Additionally, we prove that investors whose RRA is *higher* than the market price for variance risk would optimally trade in an opposite direction to the price change, and thus they are denoted the *Contrarian* group (interchangeably "type *C*"; $k=c$ or $K=C$).

All investors have similar beliefs (μ,σ), and investors within a given group have similar tastes ($U(W_k)$) and endowments ($\overline{W}_k$). There are J_τ investors of type τ and J_c investors of type *c*, thus the aggregate wealth each group holds is $W_{T,t}=J_\tau \overline{W}_{\tau,t}$ and $W_{C,t}=J_c \overline{W}_{c,t}$, respectively. The aggregate wealth in the economy is $W_t = W_{T,t}+W_{C,t}$ (capital subscripts represent aggregate values). Because beliefs, tastes, and endowments are similar within each group, a representative investor can be defined for each of the groups through demand aggregation, based on Rubinstein (1974).

Optimal asset allocation

Let the stock value held by the price-taking individual investor *k* at *t* be $S_{k,t}=N_{k,t}P_t, \forall t$, where $N_{k,t}$ is number of shares and P_t their price. Bond value held by investor *k* is denoted by $D_{k,t}=Q_{k,t}B_t,\ \forall t$, where $Q_{k,t}$ is quantity of bonds and B_t their price. Total wealth of investor *k* at *t*, $\overline{W}_{k,t}$, is allocated proportionally to the stock and bond components of the portfolio, $\alpha_{k,t}=S_{k,t}/\overline{W}_{k,t}$ and $1\text{-}\alpha_{k,t}=D_{k,t}/\overline{W}_{k,t}$, respectively. Hence, given the price process (1), investor *k*'s next period wealth *prior to portfolio rebalancing* is

$$\begin{aligned}\hat{W}_{k,t+dt} &= N_{k,t}P_{t+dt}+Q_{k,t}B_{t+dt}\\ &= \overline{W}_{k,t}\big(1+(r+\alpha_{k,t}(\mu-r))dt\big)+\alpha_{k,t}\overline{W}_{k,t}\sigma dz\end{aligned}, \qquad (2)$$

where $\hat{W}_{k,t+dt}$ is wealth before rebalancing, $P_{t+dt} = P_t(1+\mu dt + \sigma dz_{t+dt})$ and $B_{t+dt} = B_t(1+rdt)$. Assuming all investors have a twice differentiable utility function $U_k(W_k)$, we apply Itô's Lemma about $U(\hat{W}_{k,t+dt})$, as well as the expectation operator

$$E\left(U(\hat{W}_{k,t+dt})\right) = U(\overline{W}_{k,t}) + U'(\overline{W}_{k,t})\overline{W}_{k,t}\left(r + \alpha_{k,t}(\mu - r)\right)dt + \frac{U''(\overline{W}_{k,t})}{2}\overline{W}_{k,t}^2\alpha_{k,t}^2\sigma^2 dt \quad . \qquad (3)$$

The first order optimality condition of (3) with respect to $\alpha_{k,t}$ yields [3]

$$\frac{dE(U)}{d\alpha_{k,t}} = U'(\overline{W}_{k,t})\overline{W}_{k,t}(\mu - r) + U''(\overline{W}_{k,t})\overline{W}_{k,t}^2\alpha_{k,t}\sigma^2 = 0 .$$

Solving for the optimal proportion $\alpha_{k,t}^*$ to be invested in the risky asset, we have

$$\alpha_{k,t}^* = -\frac{U'(\overline{W}_{k,t})}{U''(\overline{W}_{k,t})\overline{W}_{k,t}}\frac{\mu - r}{\sigma^2},$$

and by fixing the Arrow-Pratt measure of relative risk aversion $R_{W,k,t} \equiv -\frac{U''(\overline{W}_{k,t})\overline{W}_{k,t}}{U'(\overline{W}_{k,t})}$ we obtain

$$\alpha_{k,t}^* = \lambda / R_{W,k,t}, \qquad (4)$$

where $\lambda \equiv (\mu - r)/\sigma^2$ is the market price for variance risk. Assume all investors have utility functions of the following HARA type

$$U(W,t) = e^{-\rho t}\frac{\delta_k}{1-\delta_k}\left(\frac{W_{k,t}}{\delta_k} + \eta_k\right)^{1-\delta_k} . \qquad (5)$$

This function can produce relative or absolute measures of risk aversion, and both can be constant, decreasing, or increasing in wealth (Merton, 1971, fn. 19). Fixing the measure of relative risk aversion $R_{W,k,t}$ of (5) in (4) yields the optimal asset allocation rule at *t* for investor *k*

$$\alpha^*_{k,t}\overline{W}_{k,t} = \frac{\lambda}{\delta_k}\left(\overline{W}_{k,t} + \eta_k\delta_k\right), \tag{6}$$

which is similar to equation 49 in Merton's (1971) optimal investment rule with infinite horizon.

To derive the optimum change in share holdings between two periods, one must specify the asset allocation rule at $t + dt$ and distinguish between price and quantity effects. Incrementing the time index from t to $t + dt$, the optimal equity position at $t + dt$ is $\alpha^*_{k,t+dt}\overline{W}_{k,t+dt} = \lambda / \delta_k\left(\overline{W}_{k,t+dt} + \eta_k\delta_k\right)$. The property of budget conservation assures that $\overline{W}_{k,t+dt} = \hat{W}_{k,t+dt}$.[4] Hence, the optimal investment in the risky asset at $t + dt$ can be expressed as

$$\alpha^*_{k,t+dt}\overline{W}_{k,t+dt} = \lambda / \delta_k\left(\hat{W}_{k,t+dt} + \eta_k\delta_k\right). \tag{7}$$

Therefore, (7) links the optimal wealth allocations of two consecutive periods facilitating the solution for the optimal number of shares traded as investors maintain optimal portfolio rebalancing.

2. Implied supply and demand functions for shares

From (7) we derive two mutually-exclusive optimal rebalancing strategies that imply conditional buy or sell orders for shares; conditional on the direction of price change. These schedules, representing instantaneous intertemporal supply and demand for shares, are derived by rewriting (7) in terms of quantities and prices

$$N_{k,t+dt}P^H_{t+dt} = \lambda / \delta_k\left(N_{k,t}P^H_{t+dt} + Q_{k,t}B_{t+dt} + \eta_k\delta_k\right), \tag{8}$$

where P^H_{t+dt} is a variable that obtains a continuum of hypothetical prices for which we construct the

---

[3] Note that because asset prices are lognormal and the investment opportunity set is constant, the myopic solution technique (3) and a stochastic dynamic programming approach yield similar asset allocation rules. Because these are our fundamental tools of analysis we resort to the simpler technique.

[4] This corresponds to the numerical example presented in the introduction, whereby $\hat{W}_{k,t+dt} = 60 + 44 = 104$ and $\overline{W}_{k,t+dt} = 62.4 + 41.6 = 104$.

implied supply\demand schedules, and in equilibrium $P_{t+dt}^{H} := P_{t+dt}$. Defining the number of shares held by investor *k* at $t + dt$ by the number of shares held at *t* plus an optimal addition (subtraction) through trade over *dt*, we may replace $N_{k,t+dt} = N_{k,t} + dN_{k,t+dt}$ on the left side of (8) and solve for the functional relationship between P_{t+dt}^{H} and $dN_{k,t+dt}$,

$$P_{t+dt}^{H} = \frac{\lambda / \delta_k \left(\tilde{D}_{k,t+dt} + \eta_k \delta_k\right)}{N_{k,t}\left(1 - \lambda / \delta_k\right) + dN_{k,t+dt}}, \tag{9}$$

where $\tilde{D}_{k,t+dt} \equiv Q_{k,t} B_{t+dt}$. Investor *k*'s risk preferences δ_k and η_k determine whether the slope of (9) in the hypothetical-price marginal-trade plane is positive or negative, i.e., whether investors would buy or sell given P_{t+dt}. Because all within-group individual investors are homogeneous, the group can be represented by an aggregate investor, and (8) can be rewritten in its aggregate form (capital letter subscripts)

$$N_{K,t+dt} P_{t+dt}^{H} = \frac{\lambda}{\delta_K}\left(N_{K,t} P_{t+dt}^{H} + Q_{K,t} B_{t+dt} + \eta_K \delta_K\right), \tag{10}$$

based on which we have the aggregate version of (9),

$$P_{t+dt}^{H} = \frac{\lambda / \delta_K \left(\tilde{D}_{K,t+dt} + \eta_K \delta_K\right)}{N_{K,t}\left(1 - \lambda / \delta_K\right) + dN_{K,t+dt}}. \tag{11}$$

The separation of investors into two mutually exclusive groups is given by Proposition 1:

Proposition 1 (Optimal Intertemporal Trade): *If utility parameters δ and η are bounded as presented in Table 1 and investors hold long positions in both assets, (11) implies either a positive or a negative slope of the optimal intertemporal trade schedules. In both cases the relevant trade schedule is convex and passes through* P_t. Proof: Appendix A

[Table 1]

Equation (11) yields a monotonically increasing trade schedule for Trend investors and a monotonically decreasing schedule for Contrarian investors, both are convex. Utility parameters generate the distinction between both schedules as $0 < \delta_T < \lambda < \delta_C$, and the displacement parameters satisfy $\eta_T < 0$ (DRRA preferences) for Trend investors; we assume henceforth $\eta_C = 0$ (CRRA preferences) for the Contrarian investors. These schedules are not solely "supply" or "demand" functions for shares, since both supply and demand are present along each curve, depending on the sign of price changes. For example, Proposition 1 states that it is optimal for Trend investors to buy shares when the stock price increases, but to sell shares when the price declines. Therefore, the upward sloping schedule determines the intertemporal demand for shares in periods of price increases, but it determines the supply when prices decline. Since the opposite holds for the downward sloping schedule of Contrarian investors, the term Marginal Trade Schedules (MTS) appears more appropriate, although we use both terms below to simplify the discussion. We illustrate this notion graphically in Figure 1.

[Figure 1]

Summing $N_{K,t+dt}$ on the left-hand side of (10) across both groups *K*=(*T*,*C*) equals the aggregate stock value in the economy $N_{t+dt}P_{t+dt} = (N_{T,t+dt} + N_{C,t+dt})P_{t+dt}$. This value is determined by the sum of optimal wealth allocations to the risky asset across groups. It induces change in the capital stock through instantaneous share issue or repurchase by the firm whenever $N_{t+dt} \neq N_t$,

$$\begin{aligned} N_{t+dt}P_{t+dt} &= \lambda\left(\frac{N_{T,t}P_{t+dt} + Q_{T,t}B_{t+dt} + \eta_T\delta_T}{\delta_T} + \frac{N_{C,t}P_{t+dt} + Q_{C,t}B_{t+dt} + \eta_C\delta_C}{\delta_C}\right) \\ &= \lambda\left(\frac{\hat{W}_{T,t+dt} + \eta_T\delta_T}{\delta_T} + \frac{\hat{W}_{C,t+dt} + \eta_C\delta_C}{\delta_C}\right) \end{aligned}. \qquad (12)$$

3. Volume, liquidity, and finite elasticities

The model now contains all the building blocks necessary to explore volume, liquidity, and the elasticities of supply and demand within Merton's economy. Because volume is key to the definition of both liquidity and elasticities we start by providing its closed-form equations. Volume stems from three sources: the optimal rebalancing trade in units of shares between *t* and *t+dt* by each investor type, *C* and *T*, and instantaneous share issue or repurchase by the firm.

Closed-form Solutions for Volume

Proposition 2 (The Sources of Volume): *Instantaneous volume by investor group T is given by*

$$VOL(T_{t+dt}) \equiv \Delta N_{T,t+dt} = \frac{a_T N_{T,t} + b_T N_{C,t} + c_T N_{t+dt} + d/P_{t+dt}}{1/\delta_C - 1/\delta_T}, \tag{13a}$$

instantaneous volume by investor group C is

$$VOL(C_{t+dt}) \equiv \Delta N_{C,t+dt} = \frac{a_C N_{C,t} + b_C N_{T,t} + c_C N_{t+dt} + d/P_{t+dt}}{1/\delta_T - 1/\delta_C}, \tag{13b}$$

and volume generated by the firm is defined by

$$VOL(F_{t+dt}) \equiv \Delta N_{t+dt} = N_{t+dt} - N_t, \tag{14}$$

where:

$$a_T \equiv \frac{2}{\delta_T} - \frac{1}{\delta_C} - \frac{\lambda}{\delta_T^2}, \qquad a_C \equiv \frac{2}{\delta_C} - \frac{1}{\delta_T} - \frac{\lambda}{\delta_C^2},$$

$$b_T \equiv \frac{1}{\delta_C} - \frac{\lambda}{\delta_C^2}, \qquad b_C \equiv \frac{1}{\delta_T} - \frac{\lambda}{\delta_T^2},$$

$$c_T \equiv \frac{1}{\delta_C} - \frac{1}{\lambda}, \qquad c_C \equiv \frac{1}{\delta_T} - \frac{1}{\lambda},$$

$$d \equiv Q_{T,t} B_{t+dt} \frac{\delta_T - \lambda}{\delta_T^2} + Q_{C,t} B_{t+dt} \frac{\delta_C - \lambda}{\delta_C^2} + \eta_T \left(1 - \frac{\lambda}{\delta_T}\right) + \eta_C \left(1 - \frac{\lambda}{\delta_C}\right).$$

Proof: Appendix B.

Instantaneous volume depends on stock and bond allocations at the beginning of the period, on the number of shares available at the end of the period, and most importantly, on the dispersion of both RRAs from λ, henceforth referred to as "*heterogeneity*". A key result of the model, which affects all our microeconomic measures of market activity, is that volume increases with heterogeneity. Instantaneous volume will normally not be equal between the two investor types, as both investors may purchase (sell) different quantities from the instantaneous share issue (repurchase) by the firm. This latter source of volume deserves closer attention: In Merton's economy the capital stock N_{t+dt} is implied by the model as the exogenous price P_{t+dt} and the optimum demand for risky assets determine its value to satisfy equilibrium. The firm can be thought of as a risk-neutral agent that adjusts periodically the number of shares outstanding so as to meet the aggregate demand for shares. Absent transaction costs the firm would trade each period, supporting the exogenous equilibrium price. However, in the presence of transaction costs Constantinides (1986) shows that investors' rebalancing trades substantially decline, a result that may well hold for the firm's trades.[5] Alternatively, a (not-necessarily risk-neutral) market maker may provide this service in lieu of the firm by holding some inventory, and trading with the relevant investor. In what follows we demonstrate the important role that the firm's generated volume plays in the market, primarily as a moderator of volume risk, and the its implications for the capital formation process.

The heterogeneity of risk preferences drives the main results of the model in two important ways. First, the greater the (absolute) difference between a specific investor's RRA and the market price for variance risk λ, the more shares will the specific investor need to trade to maintain optimal portfolio rebalancing (given price change). The reason is that the next-period optimal

[5] We thank Yakov Amihud for highlighting this notion.

position in the risky asset is sensitive to the ratio between λ and the relevant RRA, δ_T or δ_C, as can be seen in (7) and (9). In other words, the more λ/δ differs from unity, the greater must be the change in the next-period optimal position of the risky asset versus current period, which necessitates trading a larger number of shares. Second, the model implies that the rank $0<\delta_T<\lambda<\delta_C$ must hold to distinguish intertemporal buyers from sellers for any price change. Thus, for a given λ, an increase in δ_T reduces the ratio λ/δ_T *toward* unity, and therefore *reduces* volume by Trend investors. The opposite occurs when δ_C increases, since the ratio λ/δ_C decreases, but now *away* from unity, *increasing* volume by the Contrarian investors. Because the divergence of either one of the RRAs from λ increases optimal trade, the measures of absolute volume, turnover, liquidity, and elasticities increase with heterogeneity. In a strictly homogeneous market, as in Merton (1971), volume nullifies and all other measures of market activity vanish as well. As a result, our model adds a new dimension to Merton's economy through which one can explore several microeconomic and market microstructure effects that underlie the dynamics of equilibrium asset pricing.

Simulations

Time integration of the different measures of market activity is impossible in some cases, and tedious in others. We therefore simulate the market at 16 different levels of dispersion of RRAs from λ, from almost homogeneous to highly heterogeneous markets. We are simulating the economy using NYSE postwar statistics (1/1945-12/2006) as detailed in Table 2. Because λ=2.22 in our benchmark economy, the most homogeneous market we simulate has a symmetric proportional dispersion of 4% about unity: λ/δ_T=2.22/2.13=1.04 on one hand, and λ/δ_C=2.22/2.31=0.96 on the other hand. We increase heterogeneity by widening the gap by 6% increments in each direction (e.g., 1.1 with 0.9, 1.16 with 0.84, etc.), until the most heterogeneous market is simulated where δ_T =1.14 and δ_C =37.0. In all figures below the more homogeneous markets are located on the left, and heterogeneity increases rightward. Accordingly, weighted average RRA varies from 2.29 to 3.13. Each of the 16 markets is simulated with 100 sample paths of the stock price, and each sample path is made of 250 periods, assumed to resemble daily trading intervals. The statistics reported hereunder for each of the 16 economies are therefore based on these 100 simulations, i.e., 25,000 "daily" observations.

It should be highlighted at the outset that the proportional share holdings of both investor types must vary with heterogeneity in order to adequately construct the composite investor: Type *C* investors hold about 90% of shares in the most homogeneous market, and this proportion declines linearly to about 10% in the most heterogeneous market. In other words, equilibrium pricing is supported with as little as 10% Trend investors if the market is relatively homogeneous, but the market can accommodate about 90% and more with the divergence of RRAs. This result is related to Dumas (1989), who calculates the composite investor's preferences numerically for non-HARA utilities, and to Wang (1996) who aggregates two investors with different RRA coefficients but with CRRA preferences.

The analysis of volume and turnover

While instantaneous volume is given by (13a), (13b), and (14), the time-series sum of absolute volume (in unit terms) is the empirically comparable measure of volume in the model, therefore analyzed across heterogeneity. The most homogeneous markets in Figure 2, Panel A are characterized by the lowest absolute volume, and one can readily notice that volume generally increases with heterogeneity. Volume is zero *at* the origin, where Merton's economy is populated by a single representative investor. [6]

[Figure 2]

Panel A of Figure 2 shows that the increase of absolute volume by type *T* investors is monotonic and convex, while volume by type *C* investors increases in a concave schedule to a maximum and then declines. Between the two investor types, volume is generated mainly by type *C* investors in relatively homogeneous markets while type *T* investors generate most of the volume in heterogeneous markets. The sum of volume generated by these investors increases linearly with heterogeneity. The firm generates volume as required to close the gap between the demand\supply of both investor types. As a result, it generates no volume when absolute supply and demand by both investors are equal (RRAs of about 1.52-4.11 in Panel A of Figure 2). Absolute firm volume is concave in homogeneous markets up to this point, but it increases in a convex schedule in more heterogeneous markets. Panel B of Figure 2 shows the non-absolute sum of unit volume generated by type *T* and *C* investors, and by the firm. It demonstrates that the average pace of capital formation has a minimum in relatively homogeneous markets, where type *C* investors dominate, but it is increasingly positive as heterogeneity increases and type *T* investors dominate. These findings motivate further studies of the relationship between capital market activity and the pace of capital formation toward sustainable economic growth, an issue which is beyond the scope of this paper.

---

[6] The linear pattern is a result of the proportional expansion of the ratios λ / δ_K about unity. Increasing δ_C and reducing δ_T in fixed increments, for example, would result in a convex increase of total volume, yet this would necessitate changing η_T, which is not needed under proportional expansion.

Two reasons drive the different volume patterns of our investors. First, type *C* investors hold the majority of shares in homogeneous markets, as required to construct the composite investor. The second reason is rooted in equations (13a) and (13b), and is presented in Figure 3.

[Figure 3]

This figure plots the instantaneous volume of both investor types in an arbitrary trading day where the stock price declined, therefore type *T* investors are generally selling shares (they purchase from the new issue if $\delta_T \to \lambda$, and more so as δ_C diverges from λ) while type *C* are buying shares (positive values). The most important difference between the instantaneous volume functions of our two investor types stems from the way their optimal trade varies with heterogeneity: volume by type *T* investors increases with heterogeneity whether it originates from an increase in δ_C or from a decline in δ_T. In contrast, while instantaneous volume by type *C* investors also increases with an increase in δ_C, it *declines* when δ_T declines. As a result of both these effects type *C* investors generate most of volume in relatively homogeneous markets, while type *T* investors are the main generators of volume in relatively heterogeneous markets.

Combining the two notions, the pace of capital formation and the determinants of investors' instantaneous volume, we address the correlation between the price level and capital formation. Because stock prices increase on average, the average (across sample paths) correlation coefficient between the number of shares outstanding and the stock price level is negative in relatively homogeneous markets. In such markets the dominance of type *C* investors generates excess supply, which the firm must repurchase. Conversely, the correlation coefficient is positive in heterogeneous markets where type *T* investors dominate and generate excess demand, which implies share issues by the firm. The correlation coefficient obtains values that range from -0.08 to +0.1, and the pattern by which this correlation changes across levels of heterogeneity is presented in Figure 4. This correlation pattern is consistent with the empirical evidence whereby firms tend to issue or retire shares under specific market states. Among others, Ikenberry, Lakonishok, and Vermaelen (1995, 2000), Loughran and Ritter (1995), and Pontiff and Woodgate (2008) show that firms issue shares when their prices are overvalued and buy back when they are undervalued. These states are correlated with levels of heterogeneity, since a price appreciation (depreciation) process increases the proportion of Trend-chasers (Contrarians) among shareholders, thereby increasing (reducing) the level of heterogeneity.

[Figure 4]

With the understanding of the components of (unit) volume and its overall level, we proceed to explore the standard deviation of volume (volume risk), and its sources. Figure 5 shows that the standard deviation of total not-absolute volume in the market (generated by type *C* and type *T* investors and by the firm), increases linearly with absolute volume. This linear schedule conceals non-linear components of the overall standard deviation of volume. The standard deviation of type *C* investors' trades increases in a concave manner to a maximum, and then declines; conversely, the standard deviation of type *T*'s volume increases in a convex schedule. The standard deviation of each of these components surpasses in some cases the standard deviation of overall volume risk in the market. Yet the standard deviation of volume generated by the firm offsets significant portions of the investors' volume risk, due to its negative correlations with either one of them. Therefore, overall market volume risk increases linearly with absolute volume, which in turn is positively correlated with heterogeneity.

[Figure 5]

The average of "daily" absolute volume divided by contemporaneous outstanding shares, multiplied by 250, yields the annual turnover rate, with or without the firm's volume

$$Turnover\%(T+C+F)=\frac{250}{n}\sum_{t=1}^{n}\left(\left|VOL(C)_t\right|+\left|VOL(T)_t\right|+\left|N_t-N_{t-dt}\right|\right)/N_t\,, \qquad (15a)$$

$$Turnover\%(T+C)=\frac{250}{n}\sum_{t=1}^{n}\left(\left|VOL(C)_t\right|+\left|VOL(T)_t\right|\right)/N_t\,. \qquad (15b)$$

Over the past several years annual turnover rates (empirically measured without the firm's generated volume, as in (15b)) increased in many exchanges worldwide[7] commonly exceeding 100%, and attracting much research attention. The fundamental academic paradigm is that existing asset pricing models cannot explain observed levels of volume.[8] As Panel A of Figure 6 shows, total annual turnover rate defined in (15a) increases according to the non-linear pattern of total absolute volume, as in Figure 2. While (15a) may be relevant as a market-wide measure of turnover accounting for aggregate firms' trades, individual firms do not trade often, and the relevant measure appears to be (15b). This measure increases linearly with heterogeneity, obtaining values in the neighborhood of 100% when δ_T is between 1.52-1.59 and δ_C between 3.70-4.11. Turnover rates exceeding 200% are also supported by the model in our most heterogeneous markets, where the weighted average RRA is at most 3.13. The figure also shows the boundaries of ± 2 standard deviations of turnover rate, demonstrating that turnover risk starts low in relatively homogeneous markets but increases with heterogeneity.

[Figure 6]

[7] The World Federation of Exchanges provides in its 2007 annual report turnover data (p. 88), including: Nasdaq 304%, NYSE 167%, Australian SE 102%, Hong Kong 94%, Korea 193%, Osaka 128%, Shanghai SE 211%, Shenzhen SE 389%, Tokyo SE 138%, Borsa Italiana 204%, Deutche Borse 208%, Swiss Exchange 134% and more.

[8] Grossman and Stiglitz (1980) present a rational model for trading volume, while Caballé and Sákovics (1998), and Odean (1998) explore behavioral explanations. Barber and Odean (2000) provide empirical support that individual investor's annual turnover rate exceeds 100%, and is welfare-reducing.

Liquidity and illiquidity

A few empirical measures of liquidity (or its reciprocal, illiquidity) are generally related to the notion of elasticities, where the absolute proportional quantity traded is divided by the absolute rate of return of the stock (see Johnson, 2006). Adopting this perspective, we define the measures of liquidity that type C or T investors are generating as

$$L_C = \frac{1}{n}\sum_{t=1}^{n} \frac{\left|\Delta N_{C,t+dt}\right| / N_{t+dt}}{\left|dP_{t+dt}\right| / P_{t+dt}}, \tag{16a}$$

$$L_T = \frac{1}{n}\sum_{t=1}^{n} \frac{\left|\Delta N_{T,t+dt}\right| / N_{t+dt}}{\left|dP_{t+dt}\right| / P_{t+dt}}, \tag{16b}$$

whereas liquidity generated by the firm is given by

$$L_F = \frac{1}{n}\sum_{t=1}^{n} \frac{\left|N_{t+dt} - N_t\right| / N_{t+dt}}{\left|dP_{t+dt}\right| / P_{t+dt}}. \tag{16c}$$

When both investors' measures (16a) and (16b) are added, we obtain one of the common empirical definitions for liquidity as it applies the sum of all absolute trades that do not stem from the firm in the numerator. This third measure of liquidity is denoted henceforth as "Liquidity T+C":[9]

$$L_{T+C} = \frac{1}{n}\sum_{t=1}^{n} \frac{\left(\left|\Delta N_{T,t+dt}\right| + \left|\Delta N_{C,t+dt}\right|\right) / N_{t+dt}}{\left|dP_{t+dt}\right| / P_{t+dt}}. \tag{17}$$

We define "Market liquidity" as the sum of (16c) and (17), which includes the firm's trades,

$$L_M = \frac{1}{n}\sum_{t=1}^{n} \frac{\left(\left|\Delta N_{T,t+dt}\right| + \left|\Delta N_{C,t+dt}\right| + \left|N_{t+dt} - N_t\right|\right) / N_{t+dt}}{\left|dP_{t+dt}\right| / P_{t+dt}}. \tag{18}$$

The reciprocal of (17) can be interpreted as the percentage price impact of order flow rate

---

[9] An exact valuation of (16a,b) and (17) is obtained by replacing (13a) and\or (13b) in their appropriate numerators, while an approximation of the relevant measures of liquidity can be calculated by partial derivatives. Our simulations reveal that differences between these measures are negligible.

throughout a period (Amihud, 2002), which stems from Kyle's (1985) definition of illiquidity. It is closely related to other measures, such as Silber (1975), where excess quantity demand replaces volume. We scale Amihud's (2002) measure of *ILLIQ* by the number of shares outstanding, and modify (17) to account for value rather than unit volume. The measure, given in (19), excludes the liquidity component generated by the firm in order to allow comparison with Amihud's measure.

$$ILLIQ = \frac{1}{n}\sum_{t=1}^{n}\frac{\left|dP_{t+dt}\right|/P_{t+dt}}{\left(\left|\Delta N_{T,t+dt}P_{t+dt}\right|+\left|\Delta N_{C,t+dt}P_{t+dt}\right|\right)/N_{t+dt}}. \qquad (19)$$

Our five measures of liquidity (16a,b,c, 17, and 18) depend on the amount of funds managed by each strategy, by the dispersion of RRAs about λ, and the other parameters and state variables given in (13a) and (13b). Figure 7 shows that the components of liquidity vary with heterogeneity in the same way that volume does (see the preceding sub-section). As a result, liquidity generated by Contrarian investors exceeds the liquidity level generated by Trend-chasers in relatively homogeneous markets (less than 1.52-4.11), while the opposite holds true in more heterogeneous markets. The firm repurchases (issues) shares to balance excess supply (demand), and by doing so contributes to overall market liquidity.

[Figure 7]

The measure of *ILLIQ* (Figure 8, Panel *A*) obtains its highest value in our most homogeneous market, declining rapidly and in a convex schedule with heterogeneity. This result is related to our prior findings whereby volume and market liquidity increase with heterogeneity. The magnitude obtained in our second most homogeneous market, 0.284, is comparable to Amihud (2002), who reports a median of annual means for the period 1963-1996 of 0.308. Yet, while Amihud reports a mean annual standard deviation of 0.512, our simulated economy generates an order of magnitude smaller standard deviation of 0.056, probably since daily market returns deviate from the model-assumed normal distribution. The chart shows that the standard deviations of *ILLIQ* diminish rapidly with heterogeneity, as the ± 2 standard deviations plots asymptotically converge to the mean at relatively low RRAs.

[Figure 8]

Next, we plot in Panels B and C of Figure 8 the correspondence between *ILLIQ* and the associated positive and negative stock returns at different levels of heterogeneity. We calculate the average return impact on *ILLIQ* in four groups, each made up of four simulated markets, starting from the four most homogeneous markets (1-4), and ending with the four most heterogeneous markets (13-16). The scattered diamond symbols, having the legend "*ILLIQ 1-4*", show the correspondence between positive and negative returns on the vertical axis, and the average value of *ILLIQ* shown across the four most homogeneous markets on the horizontal axis. Panel B excludes volume generated by the firm, showing the pattern by which high positive and negative rates of return correspond with higher values of *ILLIQ*. Negative returns are associated with higher illiquidity, capturing the well-documented asymmetric price impact of trade (see Hameed, Kang, and Viswanathan, 2007, and the references therein). Repeating the same exercise with increasingly heterogeneous markets in the remaining three groups, Panel B reveals the diminishing effect of rates of return on *ILLIQ*, until it hardly varies in the four most liquid markets, 13-16. Panel C repeats the same exercise, but incorporates the firm's volume when calculating *ILLIQ*. One can readily notice two effects: first, the asymmetric impact has largely faded, and second, the average level of *ILLIQ* diminished. Lastly, we report (without a chart) that the average percentage change in *ILLIQ* is very high in homogeneous markets, exceeding 60%, but declines rapidly and obtains values between 9%-0.25% for moderate levels of heterogeneity. This range is consistent with the empirical results of Acharya and Pedersen (2005).

Liquidity risk

A number of papers explore the relationship between liquidity risk and different measures of volume. Acharya and Pedersen (2005) and Johnson (2008) incorporate liquidity risk in asset pricing models. Acharya and Pedersen show that a lasting negative shock to an asset's liquidity reduces contemporaneous return and increases expected return, i.e., predicting a positive instantaneous correlation between changes in liquidity and rates of return.[10] Johnson (2008) reports a number of implications, among them a zero correlation between liquidity and volume, but one of the intriguing implications is that changes in liquidity increase with turnover.[11] To compare our model's predictions with Acharya and Pedersen's empirical results, we examine the correlations between percentage changes in the measure of liquidity (17) and the contemporaneous stock return.[12] We report mixed results: We found indeed that the correlations are positive for moderate levels of heterogeneity, albeit less than 5%, but that they turn negative in highly heterogeneous markets (Figure 9). This result stems from the correlation pattern presented in Figure 4 between total number of shares outstanding and the price level. In relatively homogeneous markets this latter correlation is negative due to the dominance of Contrarian investors, but it turns positive in more heterogeneous markets, where Trend investors dominate. Because on average the price level increases, N_{t+dt} declines in homogeneous markets but increases in heterogeneous markets, thereby increasing (17) in the former case and reducing it in the latter. This yields the correlation pattern described in Figure 9.

[Figure 9]

[10] Acharya and Pedersen's results stem from the notion of commonality in liquidity, as reported by Hasbrouk and Seppi (2001), Huberman and Halka (2001), and Chordia, Roll, and Subrahmanyam (2000), and while the rationale in our model is different.

[11] Zero or negative association between liquidity and volume is reported by a number of authors, including Galati (2000), Danielsson and Payne (2002), Evans and Lyons (2002), Jones (2002), Fleming (2003), and Watanabe (2004).

[12] We use this measure since it is comparable with actual data, where firms do not trade frequently.

Johnson's (2008) model predicts a cone-shaped scatter between log-volume and changes to the measure of illiquidity (ibid, Figure 3). Replacing the absolute value of rate of return in the numerator of (19) with the rate of return, our model reveals a similar cone pattern. Yet this pattern is present in relatively homogeneous markets only; the cone flattens rapidly with the increase of heterogeneity. Figure 10 shows this plot for the average of our four most homogeneous markets in Panel *A* (Markets 1-4), and also shows that the cone-shape almost vanishes when calculated based on the four moderately homogeneous markets (Markets 5-8). The cone-shape fades with heterogeneity because our model derives liquidity from the trade of optimizing investors who exchange shares with each other and with the firm, thereby incorporating the relevant negative correlations. Conversely, in Johnson's (2008) model it is the random entry and exit from the market of short-lived investors that generates liquidity risk. The decline of liquidity risk is further evidenced in Panel *C* of Figure 10, now measured by its standard deviation. This chart shows that it is primarily generated by type *T* investors.

[Figure 10]

Finite elasticities

Elasticities of aggregate "supply" and "demand" functions for shares can be derived based on (11). We first solve (11) for $dN_{K,t+dt}$, then take the partial derivative $\partial dN_{K,t+dt} / \partial P^H_{t+dt}$ and multiply through by P^H_{t+dt} / N_{t+dt}. Applying the expectation operator, we find the expected value of the price elasticity of quantity demanded or supplied (ignoring instantaneous trades by the firm):

$$E(\rho^K_{t+dt}) \equiv \frac{E(\partial dN_{K,t+dt}) / N_{t+dt}}{E(\partial P^H_{t+dt}) / E(P^H_{t+dt})} = \frac{\frac{\lambda}{\delta_K}\left(\tilde{D}_{K,t+dt} + \eta_K \delta_K\right)}{E(P^H_{t+dt}) N_{K,t}\left(1 - \frac{\lambda}{\delta_K}\right) - \frac{\lambda}{\delta_K}\left(\tilde{D}_{K,t+dt} + \eta_K \delta_K\right)}. \qquad (20)$$

$E(\rho^{T}_{t+dt})$ is the expected value of the elasticity of supply (type *T* investors), while $E(\rho^{C}_{t+dt})$ is the expected value of the elasticity of demand (type *C* investors). Equation (20) may be considered the general equation of elasticities, as it enables the measurement of elasticities along any point on a given MTS when all other terms are assumed constant. However, the interesting empirical implication of (20) is the value of elasticities *at* the equilibrium price, where the terms that were assumed constants in the general case cannot obtain arbitrary values in this particular solution. The detailed derivation of the solution for both elasticities at the equilibrium price is given in Appendix C. Their expected values are

$$E({\rho^{T^*}}_{t+dt}) = \frac{\lambda^2 E(P_{t+dt})\left(\tilde{D}_{T,t+dt} + \eta_T \delta_T\right)\left(\frac{1}{\delta_T} - \frac{1}{\delta_C}\right)}{N_{T,t}E(P_{t+dt})(\delta_T - \lambda) - \lambda\left(\tilde{D}_{T,t+dt} + \eta_T \delta_T\right)}, \tag{21A}$$

$$E({\rho^{C^*}}_{t+dt}) = \frac{\lambda^2 E(P_{t+dt})\left(\tilde{D}_{C,t+dt} + \eta_C \delta_C\right)\left(\frac{1}{\delta_C} - \frac{1}{\delta_T}\right)}{N_{C,t}E(P_{t+dt})(\delta_C - \lambda) - \lambda\left(\tilde{D}_{C,t+dt} + \eta_C \delta_C\right)}. \tag{21B}$$

The elasticities of supply and demand differ in their sign and magnitude due to the boundaries on RRAs derived for each investor group ($\delta_T < \lambda < \delta_C, \eta_T < 0, \eta_C = 0$), implying that (21A) is strictly positive and (21B) strictly negative for all long positions in both assets.

Using these two equations, we have calculated the average (absolute) values of the elasticities of supply and demand across the simulated markets, as well as their standard deviations (Figure 11, Panels A and B). Because our simulations assume that all investors rebalance their portfolios daily, while in practice this probably occurs closer to weekly or monthly, we divided both elasticities by 5 (=250/50), roughly representing weekly rebalancing.[13] Both measures, the

[13] Empirical findings suggest that trading frequency varies across investor types. Daily portfolio rebalancing appears to be relevant for some professional investors. Kumar and Lee (2006) analyze individual investor trades and find that

elasticities and their standard deviations, near zero in homogeneous markets, but increase monotonically with heterogeneity. The elasticity of demand is higher than the elasticity of supply, with the intuitive explanation that type *C* investors trade against changes in the stock price, while type *T* investors trade in the same direction. The standard deviation of the weekly elasticity of demand is higher in relatively homogeneous markets, but it is surpassed by the standard deviation of the elasticity of supply in heterogeneous markets, where type *T* investors dominate share holdings and volume.

[Figure 11]

Finally, we compare the simulated results of our model with the empirical findings of Kalay, Sade, and Wohl (2004), who measure elasticities on a typical trading day.[14] They report four measures of elasticities (one of them is not comparable with our model); each is calculated at the price increase and price decline segments of the corresponding curve, and depends on the value assigned to the unobservable denominator (q) in the definition of elasticity $\rho = \Delta q / q / \Delta p / p$. We calibrate our model to the TASE data and compare our model estimates with their findings. The detailed parameter values, RRA combinations and results appear in Appendix D. The measures of weekly elasticities implied by our model appear to be consistent with their findings for price increases, first since the elasticities of supply are smaller than the elasticities of demand, and second since the ranges of values generally align.

---

a typical investor executes nine trades per year. Hence, trading frequency of the average investor must be less than daily but more than once every six weeks.

[14] A number of papers empirically explore the slope of demand and supply schedules for shares upon exogenous changes to aggregate supply/demand (Kandel, Sarig, and Wohl 1999, Liaw, Liu, and Wei 2001, and Wurgler and Zhuravskaya 2002,). Kim, Lee, and Morck (2004) find that the average elasticity of demand (supply) at the opening was 13.27 (13.30) before the Korean crisis but 7.08 (6.64) after the crisis.

4. Summary

This paper analyzes different measures of volume, turnover, liquidity, and elasticities, and their associated risks in the classic Merton (1971) model. We solve for the implied intertemporal supply and demand schedules for shares of two representative investors who differ in their attitude toward risk: one has CRRA preferences and the other DRRA preferences. We provide a proposition showing that the latter has a smaller coefficient of relative risk aversion (RRA), adopting Trend-chasing portfolio rebalancing rules, while the former applies Contrarian trades. The composite investor is consistent with the requirements in Constantinides (1982).

When calibrating the model for post-war NYSE statistics, we create a benchmark economy and simulate it across levels of heterogeneity, measured by the distance of both investor's RRAs from the market price for risk. Generally, market activity, measured by volume, turnover, and liquidity, increases with heterogeneity. We find that the standard deviations of volume and turnover increase with heterogeneity, while liquidity risk declines with it. Illiquidity risk increases with volume, but only in relatively homogeneous markets; the measure of illiquidity, *ILLIQ*, obtains comparable measures with the empirical findings of Amihud (2002). We find asymmetric illiquidity effects (Hameed, Kang, and Viswanathan, 2007), and reveal that measures of elasticities are consistent with the empirical findings of Kalay, Sade, and Wohl, (2004). The model implies that the dominance of Contrarian investors in homogeneous markets generates excess supply that induces the firm to repurchase some of its shares, while the firm issues shares in heterogeneous markets where the dominance of Trend-chasers generates excess demand. As a result, the pace of capital formation generally increases with heterogeneity, which have important implications for the role of financial markets in economic growth. A number of other predictions are consistent with related asset pricing models such as Acharya and Pedersen (2005), Lo and Wang (2006), and Johnson (2006, 2008).

Appendix A

Proof of Proposition 1 (optimal intertemporal trade)

Part A: The Contrarian strategy (k=c). *The partial derivate of* P^H_{t+dt} *with respect to* $dN_{c,t+dt}$ *in (9) must be negative to comply with a Contrarian strategy for investors holding long positions in both assets,*

$$\frac{\partial P^H_{t+dt}}{\partial(dN_{c,t+dt})} = -\frac{\lambda/\delta_c\left(Q_{c,t}B_{t+dt}+\eta_c\delta_c\right)}{\left(N_{c,t}\left(1-\lambda/\delta_c\right)+dN_{c,t+dt}\right)^2} < 0, \qquad (A1)$$

implying $0 < Q_{c,t}B_{t+dt}+\eta_c\delta_c$ *(or* $-Q_{c,t}B_{t+dt}/\delta_c < \eta_c$*) must hold in the numerator. Therefore,* η_c *may be negative, zero, or positive. To find the boundaries on* δ_c*, solve (7) for* λ/δ_c

$$\frac{\lambda}{\delta_c} = \frac{N_{c,t+dt}P_{t+dt}}{N_{c,t+dt}P_{t+dt}+Q_{c,t+dt}B_{t+dt}+\eta_c\delta_c}\,. \qquad (A2)$$

While $\tilde{D}_{c,t+dt} = Q_{c,t}B_{t+dt}$ is close, but not equal, to $D_{c,t+dt} = Q_{c,t+dt}B_{t+dt}$, *we assume* η_c *is sufficiently large to satisfy* $\eta_c >> Max(-\tilde{D}_{c,t+dt}/\delta_c, -D_{c,t+dt}/\delta_c)$ *and therefore* $\lambda < \delta_c$ *holds in (A2). Hence, investors would apply Contrarian trades if their utility parameters are bounded within the inequalities* $\left\{\begin{matrix} -\tilde{D}_{c,t+dt}/\delta_c << \eta_c \\ 0 < \lambda < \delta_c \end{matrix}\right\}$ *satisfying DRRA and DARA if* $\eta_c < 0$*, DARA if* $0 < \eta_c$*, or CRRA if* $\eta_c = 0$*. To assess convexity the second partial derivative of (A1) must be positive,*

$$\frac{\partial(P^H_{t+dt})^2}{\partial(dN_{c,t+dt})^2} = \frac{2\lambda/\delta_c\left(\tilde{D}_{c,t+dt}+\eta_c\delta_c\right)}{\left(N_{c,t}\left(1-\lambda/\delta_c\right)+dN_{c,t+dt}\right)^3} > 0 \qquad (A3)$$

By (A1) the numerator is positive, and given the requirement on δ_c*, the denominator in (A3) must be positive, implying* $\delta_c > \lambda N_{c,t}/N_{c,t+dt}$*. Thus, (9) is convex for Contrarian investors.*

Part B: The Trend strategy (k= τ ***).*** *The partial derivate of* P^H_{t+dt} *with respect to* $dN_{\tau,t+dt}$ *in (9) must be positive to comply with a Trend-chasing strategy when investors hold long positions in both assets,*

$$\frac{\partial P^H_{t+dt}}{\partial(dN_{\tau,t+dt})} = -\frac{\lambda/\delta_\tau\left(Q_{\tau,t}B_{t+dt}+\eta_\tau\delta_\tau\right)}{\left(N_{\tau,t}\left(1-\lambda/\delta_\tau\right)+dN_{\tau,t+dt}\right)^2} > 0 . \qquad (A4)$$

This implies that $Q_{\tau,t}B_{t+dt}+\eta_\tau\delta_\tau < 0$ *(or* $\eta_\tau < -Q_{\tau,t}B_{t+dt}$*) must hold in the numerator. Therefore,* η_τ *must be strictly negative. The conditions on* δ_τ *are given by solving (7) for* λ/δ_τ *,*

$$\frac{\lambda}{\delta_\tau} = \frac{N_{\tau,t+dt}P_{t+dt}}{N_{\tau,t+dt}P_{t+dt}+Q_{\tau,t+dt}B_{t+dt}+\eta_\tau\delta_\tau} . \qquad (A5)$$

While $\tilde{D}_{\tau,t+dt} = Q_{\tau,t}B_{t+dt}$ is close, but not equal to $D_{\tau,t+dt} = Q_{\tau,t+dt}B_{t+dt}$*, we assume* η_τ *is sufficiently small to satisfy* $\eta_\tau << Min(-\tilde{D}_{\tau,t+dt}/\delta_\tau, -D_{\tau,t+dt}/\delta_\tau)$ *and hence* $\delta_\tau < \lambda$ *holds in (A5). As a result, Trend-chasing trades would be implemented if the investor's utility parameters are bounded within the inequalities* $\left\{\begin{matrix}\eta_\tau < -\tilde{D}_{\tau,t+dt}/\delta_\tau \\ 0 < \delta_\tau < \lambda\end{matrix}\right\}$*, satisfying DRRA and DARA. Convexity holds when the second partial derivative of (A4) is positive,*

$$\frac{\partial(P^H_{t+dt})^2}{\partial(dN_{\tau,t+dt})^2} = \frac{2\lambda/\delta_\tau\left(\tilde{D}_{\tau,t+dt}+\eta_\tau\delta_\tau\right)}{\left(N_{\tau,t}\left(1-\lambda/\delta_\tau\right)+dN_{\tau,t+dt}\right)^3} > 0 . \qquad (A6)$$

According to (A4) the numerator is negative, and given the requirement on δ_τ *, the denominator in (A6) must be negative, satisfying* $\lambda N_{\tau,t}/N_{\tau,t+dt} > \delta_\tau$*. Thus, (9) is convex for Trend-chasing investors. Q.E.D.*

Appendix B

Proof of Proposition 2: The sources of volume

Using the equality $\overline{W}_{K,t+dt} = \hat{W}_{K,t+dt}$ *and replacing* $N_{T,t+dt} = N_{T,t} + \Delta N_{T,t+dt}$ *and* $N_{C,t+dt} = N_{t+dt} - N_{T,t} - \Delta N_{T,t+dt}$*, in equation (12) we identify the optimal trade by T*

$$N_{t+dt}P_{t+dt} = \lambda \left(\frac{(N_{T,t} + \Delta N_{T,t+dt})P_{t+dt} + D_{T,t+dt} + \eta_T \delta_T}{\delta_T} + \frac{(N_{t+dt} - N_{T,t} - \Delta N_{T,t+dt})P_{t+dt} + D_{C,t+dt} + \eta_C \delta_C}{\delta_C} \right). \qquad (B1)$$

Solving for $\Delta N_{T,t+dt}$ *yields*

$$\Delta N_{T,t+dt} = \frac{N_{T,t}\left(\frac{1}{\delta_T} - \frac{1}{\delta_C}\right) + N_{t+dt}\left(\frac{1}{\delta_C} - \frac{1}{\lambda}\right) + \frac{1}{P_{t+dt}}\left(\frac{D_{T,t+dt}}{\delta_T} + \frac{D_{C,t+dt}}{\delta_C} + \eta_T + \eta_C\right)}{\frac{1}{\delta_C} - \frac{1}{\delta_T}} \,. \qquad (B2)$$

Repeating the procedure by replacing $N_{C,t+dt} = N_{C,t} + \Delta N_{C,t+dt}$ *and* $N_{T,t+dt} = N_{t+dt} - N_{C,t} - \Delta N_{C,t+dt}$ *in order to solve for the optimum trade of type C investors yields*

$$\Delta N_{C,t+dt} = \frac{N_{C,t}\left(\frac{1}{\delta_C} - \frac{1}{\delta_T}\right) + N_{t+dt}\left(\frac{1}{\delta_T} - \frac{1}{\lambda}\right) + \frac{1}{P_{t+dt}}\left(\frac{D_{T,t+dt}}{\delta_T} + \frac{D_{C,t+dt}}{\delta_C} + \eta_T + \eta_C\right)}{\frac{1}{\delta_T} - \frac{1}{\delta_C}}. \qquad (B3)$$

Yet, the optimum bond holdings $D_{T,t+dt}$, $D_{C,t+dt}$ *are determined simultaneously with the choice of stock holdings, therefore they must be replaced with*

$$D_{K,t+dt} = \overline{W}_{K,t+dt} - \lambda / \delta_K \left(\overline{W}_{K,t+dt} + \eta_K \delta_K\right) = \overline{W}_{K,t+dt}\left(1 - \lambda / \delta_K\right) - \lambda \eta_K \,, \qquad (B4)$$

being total wealth minus the optimal stock holdings. Replacing (B4) in (B2) and (B3) for K=T,C, and again using the equality $\overline{W}_{K,t+dt} = \hat{W}_{K,t+dt}$*, we obtain instantaneous trade by investor type T*

$$\Delta N_{T,t+dt} = \frac{a_T N_{T,t} + b_T N_{C,t} + c_T N_{t+dt} + d/P_{t+dt}}{1/\delta_C - 1/\delta_T}, \qquad (B5)$$

and by investor type C

$$\Delta N_{C,t+dt} = \frac{a_C N_{C,t} + b_C N_{T,t} + c_C N_{t+dt} + d/P_{t+dt}}{1/\delta_T - 1/\delta_C}, \qquad (B6)$$

where

$$a_T \equiv \frac{2}{\delta_T} - \frac{1}{\delta_C} - \frac{\lambda}{\delta_T^2}, \qquad a_C \equiv \frac{2}{\delta_C} - \frac{1}{\delta_T} - \frac{\lambda}{\delta_C^2}$$

$$b_T \equiv \frac{1}{\delta_C} - \frac{\lambda}{\delta_C^2}, \qquad b_C \equiv \frac{1}{\delta_T} - \frac{\lambda}{\delta_T^2}$$

$$c_T \equiv \frac{1}{\delta_C} - \frac{1}{\lambda}, \qquad c_C \equiv \frac{1}{\delta_T} - \frac{1}{\lambda}$$

$$d \equiv Q_{T,t} B_{t+dt} \frac{\delta_T - \lambda}{\delta_T^2} + Q_{C,t} B_{t+dt} \frac{\delta_C - \lambda}{\delta_C^2} + \eta_T \left(1 - \frac{\lambda}{\delta_T}\right) + \eta_C \left(1 - \frac{\lambda}{\delta_C}\right).$$

Instantaneous volume generated by the firm is the change in the number of shares outstanding

$$VOL(F_{t+dt}) \equiv \Delta N_{t+dt} = N_{t+dt} - N_t. \qquad (B7)$$

Q.E.D.

Appendix C

Derivation of Expected Supply and Demand Elasticities at Equilibrium

Let us rewrite the general, aggregate version of the supply\demand function (9) $P^H_{t+dt} = \frac{\lambda/\delta_K\left(\tilde{D}_{K,t+dt} + \eta_K\delta_K\right)}{N_{K,t}\left(1-\lambda/\delta_K\right) + dN_{K,t+dt}}$ as

$$p = \frac{a}{b+q}, \tag{C1}$$

where $a \equiv \lambda/\delta_K\left(\tilde{D}_{K,t+dt} + \eta_K\delta_K\right)$, $b \equiv N_{K,t}\left(1-\lambda/\delta_K\right)$ and $q \equiv dN_{K,t+dt}$, to simplify notation. At equilibrium, stock price must be consistent with wealth allocation, which implies that once P_{t+dt} and B_{t+dt} determine the available wealth at $t+dt$ for each investor, the optimal allocation of shares between both investor groups cannot be a free variable. Hence, the optimal proportional holdings $\pi^*_{K,t} \equiv N^*_{K,t}/N_t$ is a function of existing parameters and state variables, and thus the equilibrium price is denoted $p^* = p(\pi^*_{K,t})$. Rewrite (C1) at the equilibrium price where quantity traded at equilibrium is a function of the equilibrium price,

$$q^* = \frac{a}{p(\pi^*_{K,t})} - b. \tag{C2}$$

By the chain rule, the marginal change in quantity traded depends on the optimal allocation at equilibrium

$$\frac{dq^*}{d\pi^*_{K,t}} = \frac{dq^*}{dp^*}\frac{dp^*}{d\pi^*_{K,t}}. \tag{C3}$$

The change in price due to a marginal change in share allocation is derived from (12) after dividing through by N_{t+dt}, which yields the equilibrium allocation across both investor types. A partial derivative with respect to equilibrium allocation of each investor type *C* and *T* yields

$$\frac{dp^*}{d\pi^*_{T,t}} = \lambda P_{t+dt}\left(\frac{1}{\delta_T} - \frac{1}{\delta_C}\right), \tag{C4}$$

$$\frac{dp^*}{d\pi^*_{C,t}} = \lambda P_{t+dt}\left(\frac{1}{\delta_C} - \frac{1}{\delta_T}\right), \tag{C5}$$

and we note that since $\delta_T < \delta_C$, (C4) is strictly positive and (C5) is strictly negative. Since at equilibrium $p^* = p(\pi^*_{k,t})$, the definition of (i.e., the expected value of) both elasticities can be written as $E(\rho^{K^*}{}_{t+dt}) = dq^*/q\big/d\pi^*_{K,t}/p^* = (dq^*/d\pi^*_{K,t})(p^*/q)$. Replacing (C3) with $dq^*/d\pi^*_{K,t}$ we obtain

$$E(\rho^{K^*}{}_{t+dt}) = \frac{dq^*}{dp^*}\frac{dp^*}{d\pi^*_{K,t}}\frac{p^*}{q}. \tag{C6}$$

Writing back the original definitions of *a*, *b*, and p^*, multiplying numerators and denominators by λ, and reorganizing it yields (21A) and (21B) in the text.

Appendix D

Data and Model-implied Elasticities for the Tel-Aviv Stock Exchange

The average market return in the 10-year period prior to 1998 was 27%, with 27% standard deviation, while the riskless (nominal) interest rate was about 7% in 1998, hence, $\lambda = 2.74$. The ratio between the aggregate market values of stocks and bonds was 1.09, and the average price of a traded share was 3.0. (Source: different TASE annual statistical reports.) Portfolio rebalancing frequency is assumed weekly.

Data

Parameter	Value	Parameter	Value(s)
μ	27%	δ_C	3.0, 4.0, 5.0
r	7.0%	δ_T	1.5, 2.0, 2.5
σ	27%	η_C	0.0
$W_{C,0} = W_{T,0}$	100	η_T	Varies to assure Eq./Bond=1.09
P^*	3.0	Equities/Bonds	1.09

Model-implied Elasticities (weekly rebalancing)

Elasticity of Demand

δ_T \ δ_C	3.00	4.00	5.00
1.50	-432	-549	-603
2.00	-249	-411	-504
2.50	-109	-289	-410

Elasticity of Supply

δ_T \ δ_C	3.00	4.00	5.00
1.50	94	259	351
2.00	32	130	213
2.50	4	35	76

The three relevant measures of the elasticity of demand reported by Kalay, Sade, and Wohl (2004) at the price increase segment are 415, 103, and 78 for the three relevant definitions of q, while our model predicts values of 432, 249, and 109 for $\delta_C = 3.0$ and $\delta_T = 1.5, 2.0, 2.5$, respectively. Comparing the corresponding elasticities of supply we find that their estimates are 64, 10, and 6, while our measures are 94, 32, and 4. Note that the measures of elasticities in Kalay et al. and in our calibration have different origins: our estimates are based on different RRAs, while they apply different definitions of q. If one assumes that their three definitions for q and our three measures of RRA reasonably cover each relevant range, then average elasticities may be informative.

References

Acharya V. V., Pedersen, L. H., 2005. Asset pricing with liquidity risk. Journal of Financial Economics 77, 375-410.

Amihud, Y., Mendelson, H., 1987. Trading mechanisms and stock returns: An empirical investigation. The Journal of Finance 42, 533-553.

Amihud, Y., Mendelson, H., 1989. The effect of beta, bid-ask spread, residual risk and size on stock returns. The Journal of Finance 44, 479-486.

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. Journal of Financial Markets 5, 31-56.

Avramov, D., Chordia, T, Goyal, A., 2006. Liquidity and autocorrelations in individual stock returns. The Journal of Finance 61, 2365-2394.

Bagwell, L. S., 1992. Dutch auction repurchases: An analysis of shareholder heterogeneity. The Journal of Finance 47, 71-105.

Barber, B.M. and Odean, T., 2000. Trading is hazardous to your wealth: the common stock investment performance of individual investors. Journal of Finance 55, 773–806.

Bossaerts, P., 2004. Equilibration of real financial markets: Theory and experimental evidence. Working paper, California Institute of Technology.

Caballé, J., and Sákovics, J., 2003. Speculating against an overconfident market. Journal of Financial Markets 6, 199-225.

Chordia, T., Roll, R., Subrahmanyam, A., 2000. Commonality in liquidity. Journal of Financial Economics 56, 3-28.

Chordia, T., Roll, R., Subrahmanyam, A., 2002. Order imbalance, liquidity, and market returns. Journal of Financial Economics 65, 111-130.

Cochrane, J. H., 2001. Asset Pricing. Princeton University Press, Princeton, NJ.

Constantinides, G. M., 1982. Intertemporal asset pricing with heterogeneous consumers and without demand aggregation. Journal of Business 55, 253-267.

Constantinides, G. M., 1986. Capital Market Equilibrium with Transaction Costs. The Journal of Political Economy 94, 842-862.

Constantinides, G. M., 2002. Rational asset pricing. The Journal of Finance 57, 1567-1591.

Danielsson, J., Payne, R., 2002. Liquidity determination in an order driven market. Working paper. London School of Economics.

Dumas, B., 1989. Two-person dynamic equilibrium in the capital market. Review of Financial Studies 2, 157-188.

Evans, M. D. D., Lyons, R.K., 2002. Time varying liquidity in foreign exchange. Journal of Monetary Economics 49, 1025–1051.

Federal Reserve Statistical Release, Flow of Funds Accounts of the US, 3, 2006.

Fleming, M.J., 2003. Measuring treasury market liquidity. Federal Reserve Bank of New York Economic Policy Review 9, 83–108.

Galati, G., 2000. Trading Volumes, Volatility and Spreads in Foreign Exchange Markets: Evidence from Emerging Market Countries. BIS Working Paper No. 93.

Grossman, S. J., and Stiglitz, J. E. 1980, On the impossibility of informationally efficient markets, American Economic Review 70, 393–408.

Hameed, A., Kang, W., Viswanathan, S. 2007. Stock market declines and liquidity. Working paper. Duke University.

Hasbrouck, J., Seppi, D.J., 2001. Common factors in prices, order flows and liquidity. Journal of Financial Economics 59, 383-411.

Hauser, S., Kedar-Levy, H., Pilo, B., Shurki, I., 2006. The effect of trading halts on the speed of price discovery. Journal of Financial Services Research 29, 83-99.

Huberman G., Halka, D., 2001. Systematic liquidity, The Journal of Financial Research 2, 161-178.

Ikenberry, D., Lakonishok, J., and Vermaelen, T., 1995. Market underreaction to open market share repurchases. Journal of Financial Economics 39, 181-208.

Ikenberry, D., Lakonishok, J., Vermaelen, T., 2000. Stock repurchases in Canada: performance and strategic trading.. Journal of Finance 55, 181-208.

Johnson, T. C., 2006. Dynamic liquidity in endowment economies. Journal of Financial Economics 80, 531-562.

Johnson, T. C., 2008. Volume, liquidity, and liquidity risk. Journal of Financial Economics 87, 388-417.

Jones, C.M., 2002. A century of stock market liquidity and trading costs. Working paper. Columbia University.

Kalay, A., Sade, O., Wohl, A., 2004. Measuring stock illiquidity: An investigation of the demand and supply schedules at the TASE. Journal of Financial Economics 74, 461-486.

Kandel, S., Sarig, O., Wohl, A., 1999. The demand for stocks: An analysis of IPO auctions. Review of Financial Studies 12, 227-247.

Kim, J. W., Lee, J., Morck, R., 2004. Heterogeneous investors and their changing demand and supply schedules for individual common stocks. NBER Working Paper No. 10410.

Kumar, A., Lee, C. M. C., 2006. Retail investor sentiment and return comovements. The Journal of Finance 61, 2451-2486.

Kyle, A., 1985. Continuous auctions and insider trading. Econometrica 53, 1315–1335.

Levin, E. J., Wright, R. E., 2002. Estimating the price elasticity of demand in the London Stock Market. The European Journal of Finance 8, 222-237.

Liaw, G., Liu, Y., Wei, K.C. J., 2001. On the demand elasticity of initial public offerings: An analysis of discriminatory auctions. International Review of Finance 2, 151-178.

Lintner, J. 1965. The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. Review of Economic Statistics, 47, 13-37.

Lo, A. W., Wang, J., 2006. Trading volume: Implications of an intertemporal capital asset pricing model. The Journal of Finance 61, 2805-2840.

Loderer, C., Cooney,J. W., van Drunen, L. D., 1991. The price elasticity of demand for common stock. The Journal of Finance 46, 621-651.

Loughran, T., Ritter, J., 1995. The new issues puzzle. Journal of Finance 50, 23-51.

Mehra, R., Prescott, E., 1985. The equity premium: A puzzle. Journal of Monetary Economics 15, 145-161.

Merton, R. C., 1971. Optimum consumption and portfolio rules in a continuous-time mode. Journal of Economic Theory 3, 373-413.

Novy-Marx, R., 2008. Excess returns to illiquidity. Working paper, University of Chicago Graduate School of Business.

Odean, T., 1998. Volume, volatility, price and profit when all traders are above average. Journal of Finance 53, 1887–1934.

Pastor, L., Stambaugh, R. F., 2001. Liquidity risk and expected stock returns. Journal of Political Economy 111, 642-685.

Pontiff, J. E., Woodgate, A., 2008. Shares issuance and cross-sectional returns. Journal of Finance 63, 921-945.

Ross, S., 1976. The Arbitrage Pricing Theory of Capital Asset Pricing. Journal of Economic Theory 13, 341-360.

Rubinstein M., 1974. An aggregation theorem for securities markets. Journal of Financial Economics 1, 225-244.

Sharpe, W., 1964. Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk, Journal of Finance 19, 425 - 442.

Silber, W.L., 1975. Thinness in capital markets: the case of the Tel Aviv Stock Exchange. Journal of Financial and Quantitative Analysis 10, 129–142.

Wang, J., 1996. The term structure of interest rates in a pure exchange economy with heterogeneous investors. Journal of Financial Economics 41, 75-110.

Watanabe, A., 2004. Macroeconomic Sources of Systematic Liquidity. Working paper. University of Alberta.

World Federation of Exchanges, Annual Report 2007, http://www.world-exchanges.org

Wurgler, J., Zhuravskaya, K., 2002. Does arbitrage fatten demand curves for stocks? Journal of Business 75, 583-608.

Table 1

The separation of preferences aimed to yield

convex Marginal Trade Schedules (MTSs)

MTS	δ	η	Attitude toward risk*
Increasing MTS (Trend; "supply")	$0<\delta_T<<\lambda$	$\eta_T<<-\tilde{D}_{T,t+dt}/\delta_T<0$	DRRA & DARA
Declining MTS (Contrarian; "demand")	$0<\lambda<<\delta_C$	$-\tilde{D}_{C,t+dt}/\delta_C<<\eta_C$	DRRA & DARA if $\eta_C<0$ DARA if $\eta_C>0$ or CRRA if $\eta_C=0$

* DRRA: Decreasing Relative Risk Aversion; CRRA: Constant Relative Risk Aversion; DARA: Decreasing Absolute Risk Aversion.

When utility parameters satisfy the Trend strategy preferences it would be optimal for such investors to *buy* if the price increases and *sell* if the price declines between t and $t+dt$. Due to the resulting positive slope of this MTS, it is referred to as the "supply schedule" and it complies with DRRA and DARA preferences. On the other hand, the optimal portfolio rule for a Contrarian investor, whose preferences are bounded as specified in the second row, results in *selling* shares when prices increase and *buying* shares when prices decline. It therefore obtains a negatively-sloped MTS, also denoted the "demand schedule". Both schedules are convex. The motivation to trade stems from the fact that following a price increase, the value of shares in the Contrarian (Trend) investor’s portfolio is higher (lower) than her optimal rebalancing rule requires, thus the need to sell (buy) units of shares. The opposite holds for price declines.

Table 2

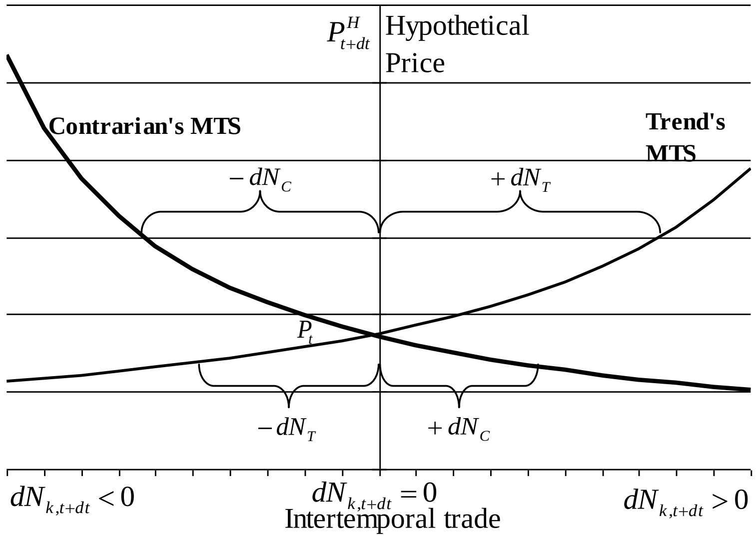
Benchmark parameters for model calibration

Parameter	Value	Parameter	Value(s)
μ	10%	δ_C	2.31-37.0
r	2.0%	δ_T	1.14-2.13
σ	19%	η_C	0.0
$W_C = W_T$	100	η_T	-42.80
$E(P)$	$31.65	Equities/Bonds	1.1

The parameters in this Table are assumed to represent relevant parameters of post-war US financial market averages between January 1945 and December 2006, with an equity premium of 8%, a riskless rate of 2.0%, and standard deviation on equities of 19%; hence, $\lambda = 2.22$ (Source: Robert Shiller's website: http://www.econ.yale.edu/~shiller/data.htm). A range of RRAs assumed for both investor groups, *C* and *T*, to represent different markets. The weighted average RRA across our 16 simulated economies ranges from 2.29 (most homogeneous) to 3.13 (most heterogeneous). The literature considers a range of 2-5 for the aggregate coefficient of RRA to be reasonable (Mehra and Prescott, 1985; Cochrane, 2001, p. 464; Constantinides, 2002). The displacement parameters are $\eta_C = 0$, representing CRRA preferences for the Contrarian investors, while η_T=-42.8, assuring that the resulting ratio between aggregate stock and bond holdings is 1.1 (Source: Federal Reserve Statistical Release, Flow of Funds Accounts of the US, 3, 2006, Table L.100), and assuring that the composite investor's preferences support the equilibrium price in all 16 markets.

Figure 1

Implied Marginal Trade Schedules (MTSs)

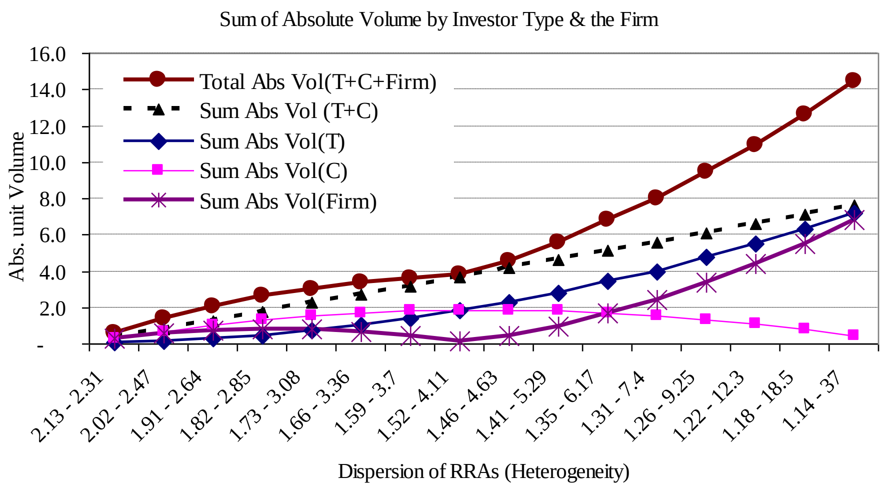


This Figure describes the MTSs for Trend (type *T*) and Contrarian (type *C*) investors. A price increase vs. P_t implies a demand for shares by group *T* ($+dN_T$) and a supply of shares by group *C* ($-dN_C$). Conversely, a price decline yields a supply of shares by group *T*, ($-dN_T$) but a demand for shares by group *C* ($+dN_C$). In both cases, trades are necessary for both investors to maintain optimal intertemporal asset allocation. Both functions are monotone and convex, but they are not necessarily symmetric. Excess supply (demand) is complemented by share repurchase (issue) by the firm.

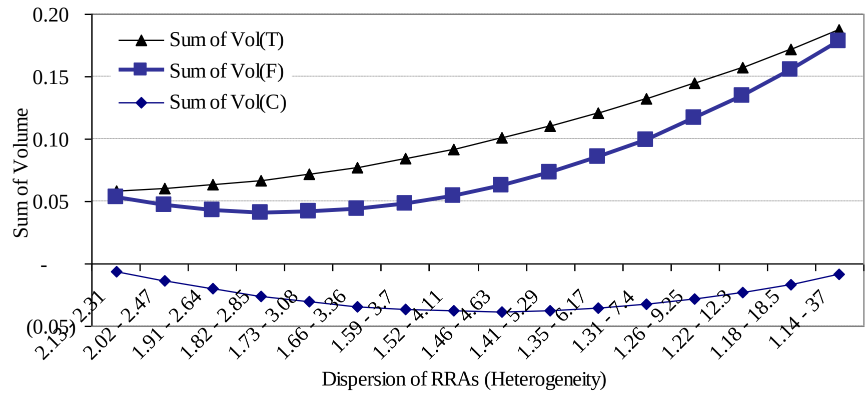
Figure 2

The generators of volume and capital formation

Panel A: Absolute unit volume by source



Panel B: Non-absolute sum of volume generated by investors *T*, *C*, and the firm



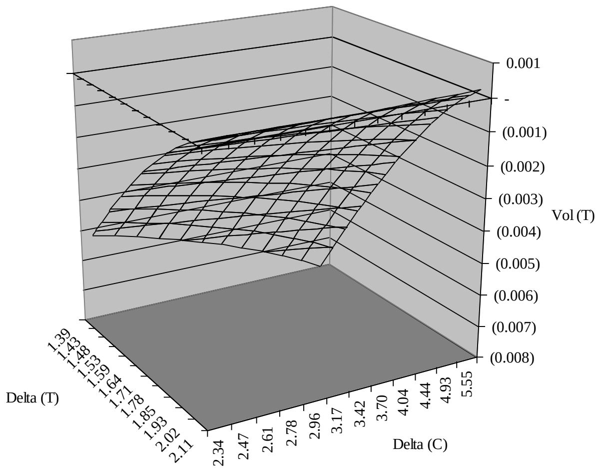
Panel A: The sum of absolute volume in the market increases non-linearly with heterogeneity. Volume by Trend-chasing investors increases monotonically and in a convex schedule, while volume by the Contrarian investors follows a concave schedule, having its maximum at moderate levels of RRA. The sum of both increases linearly with heterogeneity. Absolute volume by the firm closes the vertical gap between the trades of both investors.

Panel B: Non-absolute unit volume by the firm first declines, but then increases with heterogeneity. The decline is due to the excess demand by type *C* investors in relatively homogeneous markets, but once the level of heterogeneity increases beyond 1.52-4.11, type *T* investors dominate and generate excess demand. Hence, capital formation is low in relatively homogeneous markets, but increases with the level of heterogeneity.

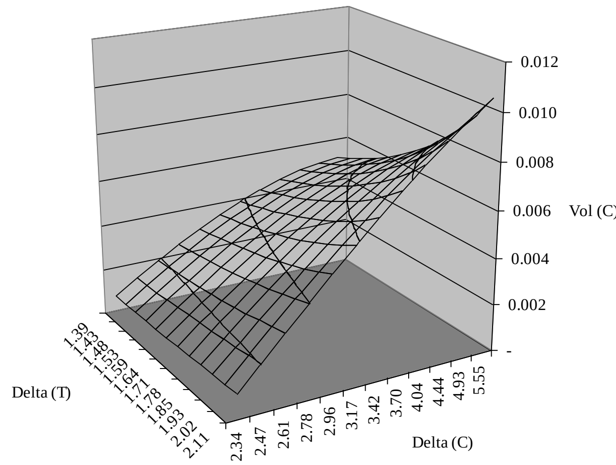
Figure 3

Instantaneous volume by both investors as a function of RRAs

Volume by a Trend investor



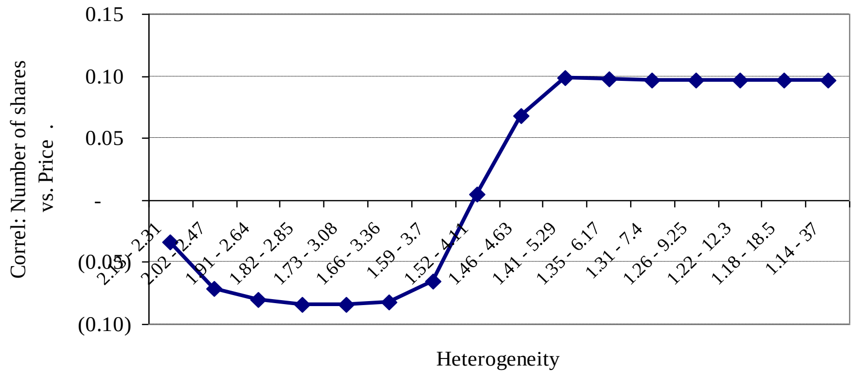
Volume by a Contrarian investor



This figure shows the different optimal trades of both investor types following a given price change as a function of the distance of both RRAs δ_T and δ_C from λ (based on eqs. 13a & 13b). This chart was derived for a price decline, thus type *T* investors are selling and type *C* are buying shares. Volume is minimal when both RRAs are close to λ=2.22, but increases with heterogeneity, i.e., declining δ_T and increasing δ_C. Volume by type *T* investors increases with heterogeneity due both to a decline of δ_T and an increase of δ_C. Conversely, volume by *C* indeed *increases* with δ_C, but *declines* as δ_T diminishes.

Figure 4

Average correlation coefficients between number of shares outstanding and price level

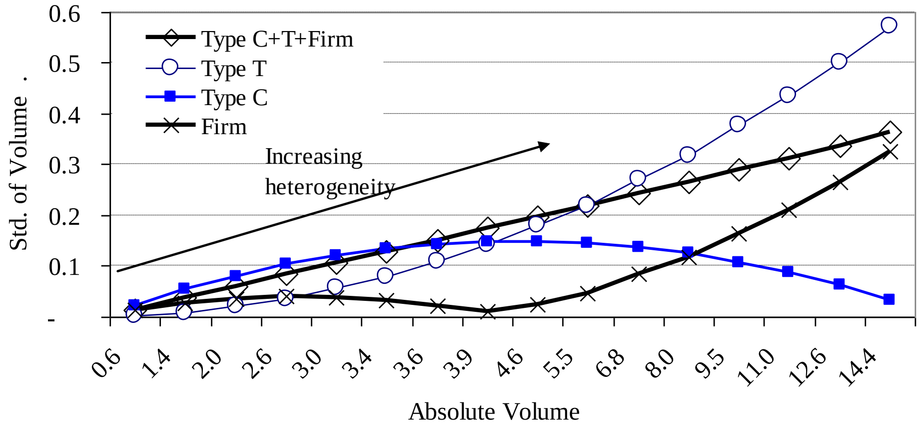


The average correlation coefficients across levels of heterogeneity is negative in homogeneous markets, but positive in heterogeneous markets. Because stock prices increase on average, type *C* (*T*) investors induce excess supply (demand) in homogeneous (heterogeneous) markets, which induces the firm to repurchase (issue) shares.

Figure 5

Standard deviation of volume vs. sum of absolute volume

Both investor types and the firm

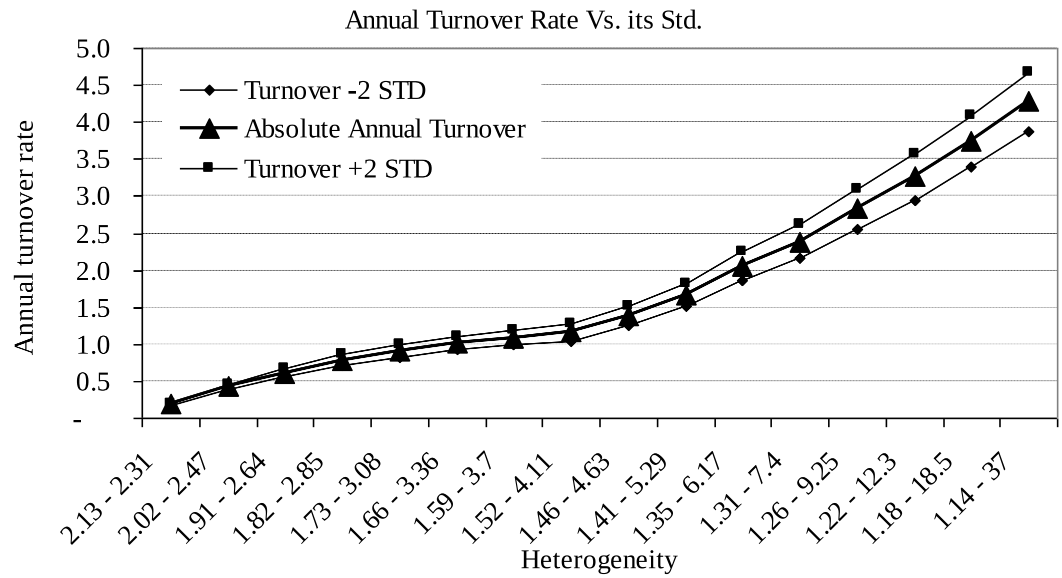


The standard deviation of volume generally increases with absolute volume at a different pace: volume risk by type *T* investors increases in a convex schedule, while volume risk by type *C* investors increases to a maximum and then declines. This maximum standard deviation occurs at a moderate level of absolute volume, which is also the point of heterogeneity where absolute volume by *C* is equal to the absolute volume generated by *T*. The firm generates volume equal to the difference between both investor types. The sum of all sources of volume risk increases linearly with absolute volume, due to the interplay of negative correlations among the three sources of trade.

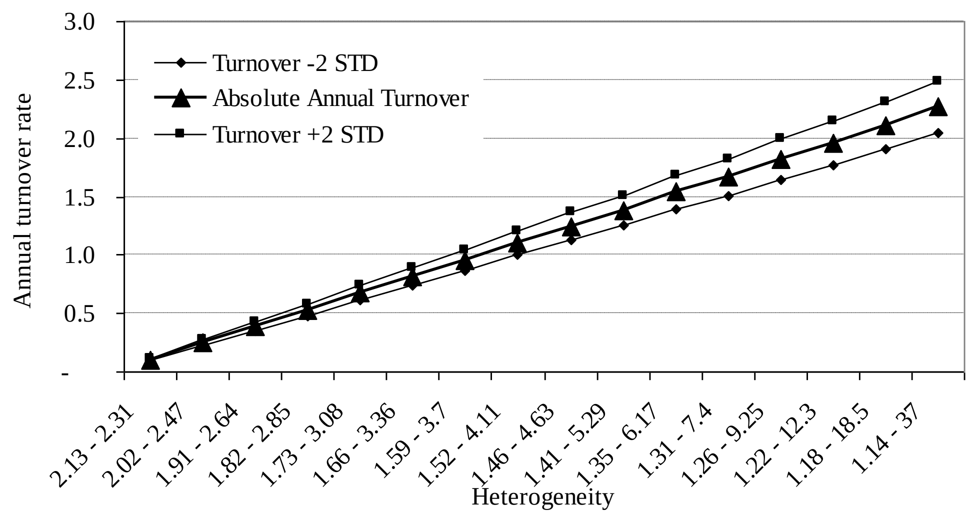
Figure 6

Annual turnover rate and its standard deviation

Panel A: Turnover generated by Contrarian, Trend-chasers, and the Firm



Panel B: Turnover generated by Contrarian and Trend-chasers only

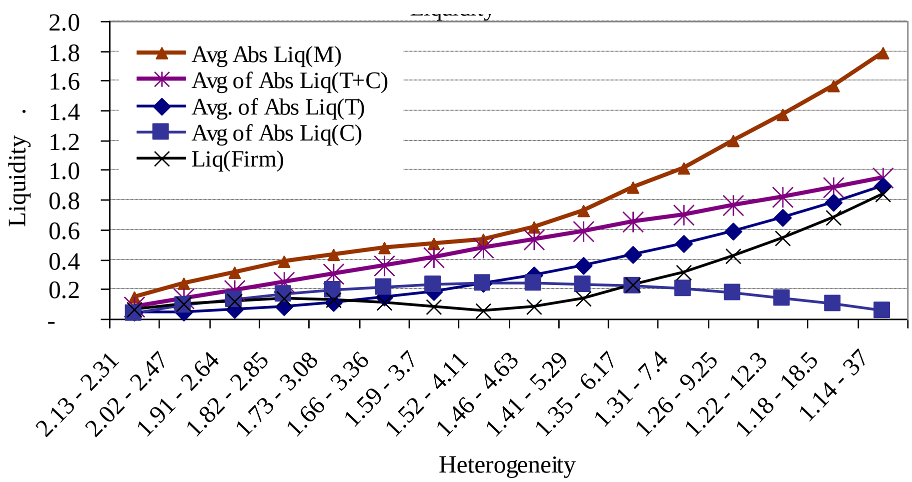


Panel A shows the non-linear increase in turnover rate when it includes volume generated by the firm.

Panel B shows that if turnover rate is calculated without the firm's trades, the schedule increases linearly with heterogeneity. Yet this linear pattern stems from our symmetrically-proportional expansion of RRAs about unity (see footnote 6).

Figure 7

Market liquidity as the sum of Contrarians', Trend-chasers', and Firm's liquidities

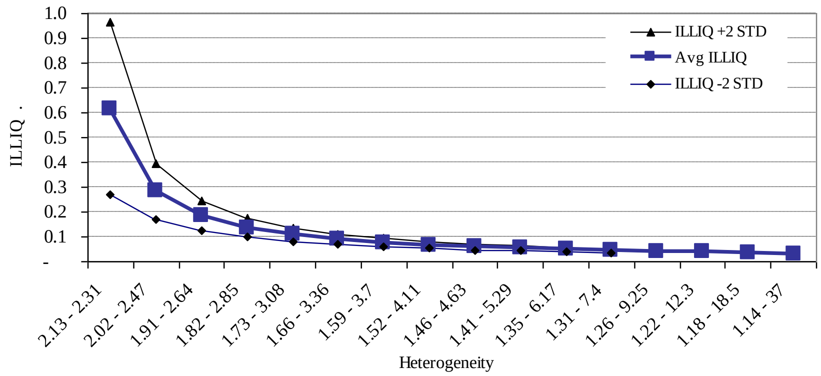


Liquidity by investor types *T* and *C* (eq. 17) is defined as the sum of liquidities provided by both investor types, and it increases linearly with heterogeneity. The liquidity schedules are determined by the optimal trading volume generated by both investor types (see Figure 2), and their sum implies a linear increase of market liquidity with heterogeneity.

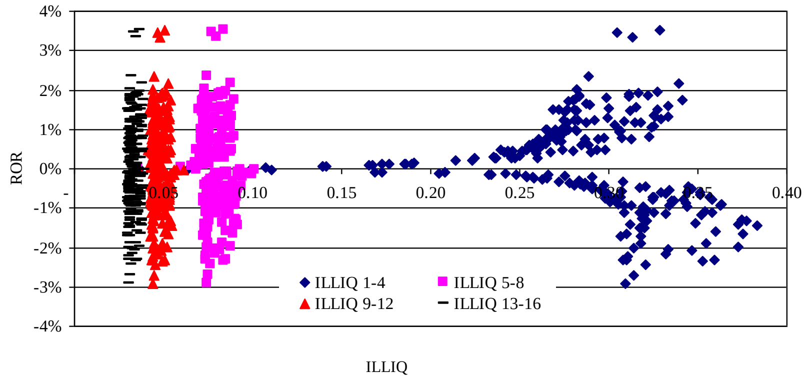
Figure 8

ILLIQ across market conditions

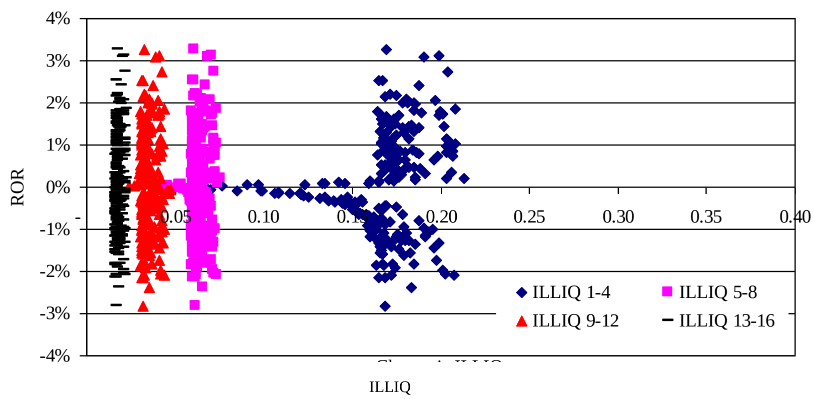
Panel *A*: Average and standard deviation of *ILLIQ* across heterogeneity



Panel *B*: Return impact on *ILLIQ*: Investors' trades only



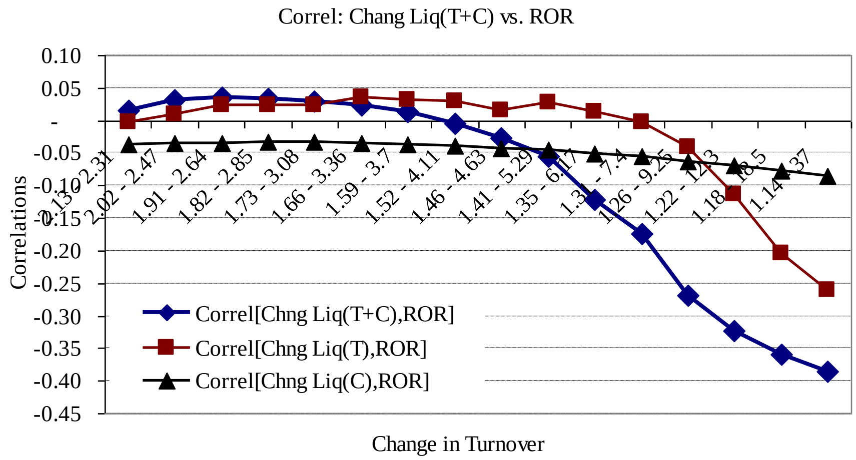
Panel *C*: Return impact on *ILLIQ*: Investors' + Firm trades



Panel A: *ILLIQ* and its standard deviation decline with heterogeneity; Panel B: If the firm's trade is excluded from the calculation of *ILLIQ* (as in Amihud, 2002), negative returns are associated with higher levels of ILLIQ than positive returns are; Panel C: Incorporating the firm's trades in the definition of *ILLIQ* both reduces its level and eliminates much of the asymmetric return impact.

Figure 9

Correlations between changes in liquidity and contemporaneous stock returns

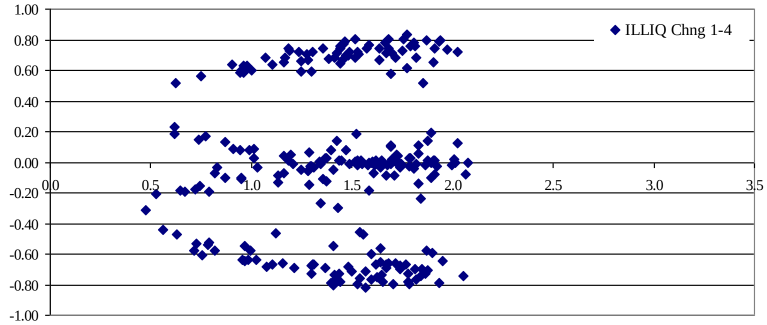


The correlation coefficient of liquidities generated by type *C* and *T* investors (without the firm) is positive in homogeneous markets, consistent with Acharya and Pedersen (2005). However, it turns negative in heterogeneous markets due to the correlation pattern between price level and the number of shares outstanding (Figure 4). The impact is inverse, since c.p., an increase (decline) in the number of shares outstanding, reduces (increases) liquidity (see eq. 17).

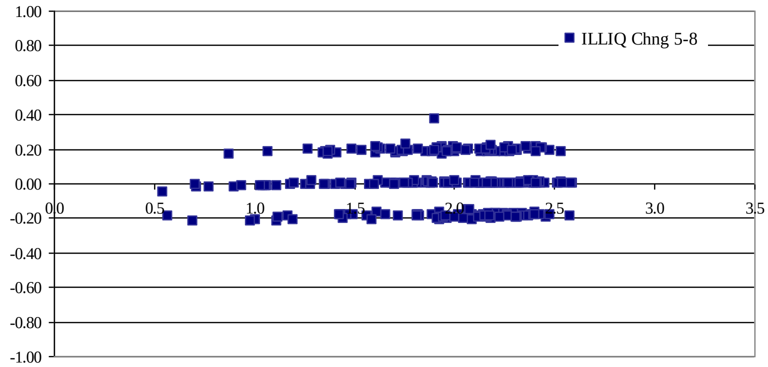
Figure 10

Illiquidity- and liquidity-risk

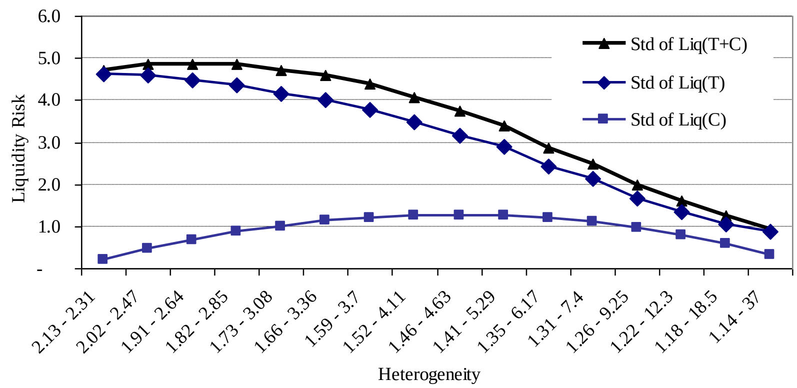
Panel *A*: Changes in non-absolute *ILLIQ* vs. Log Volume: Highly homogeneous markets



Panel *B*: Changes in non-absolute *ILLIQ* vs. Log Volume: Moderately homogeneous markets



Panel *C*: Liquidity risk declines with heterogeneity

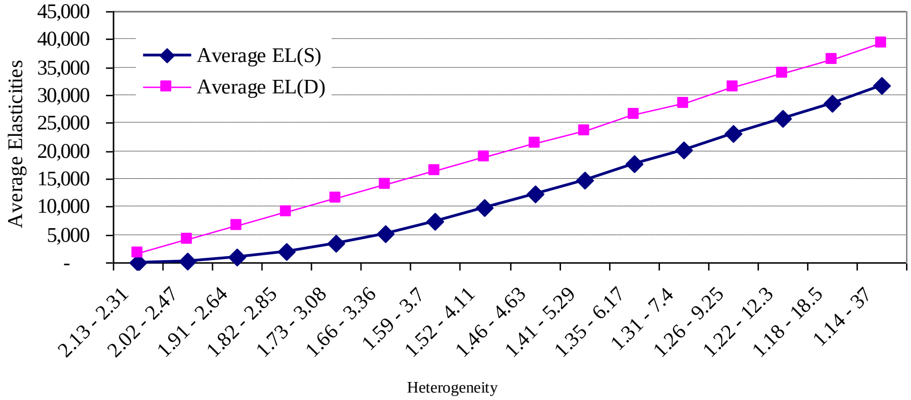


Panel A: A cone-shape scatter between log-volume ($x10^4$) and changes to *ILLIQ* (where the rate of return, in its numerator, is not absolute) implies that illiquidity risk increases with volume in the four most homogeneous markets (1-4). This figure is consistent with Johnson (2008, figure 3). Panel B: Repeating the analysis of Panel A in the four moderately homogeneous markets (5-8) shows that illiquidity risk mitigates with heterogeneity, consistent with Panel A of Figure 8. Panel C: Measuring liquidity risk as the standard deviation of the liquidity measure of both investor types (without firm-generated liquidity, eq. 17) demonstrates first, that overall liquidity risk declines with heterogeneity, and second, that most of the liquidity risk stems from type *T* investors.

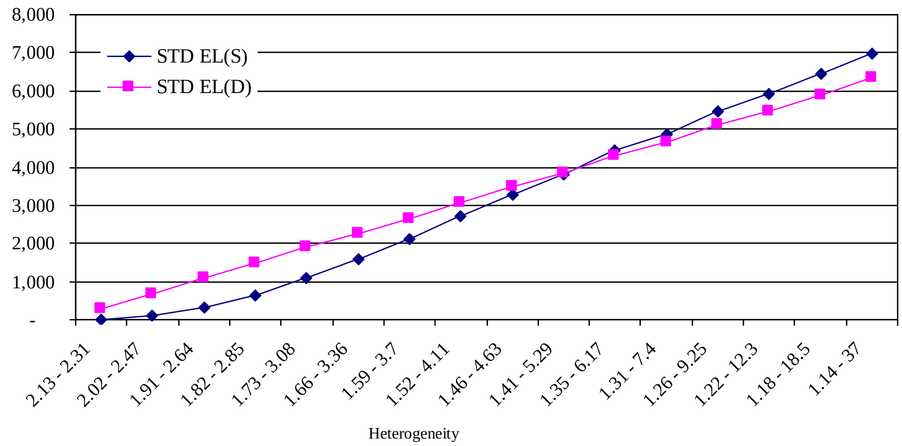
Figure 11

Average and standard deviation of elasticities

Panel *A*: Average (absolute) weekly elasticities



Panel *B*: Standard deviation of weekly elasticities



Panel A: The average values of weekly elasticities are high, and they increase with heterogeneity. Heterogeneity, in turn, is positively correlated with volume and liquidity. The elasticity of the supply schedule is lower than that of the demand schedule.

Panel B: The standard deviation of the elasticity of supply is lower than that of the elasticity of demand in homogeneous markets, but it exceeds it in heterogeneous markets. The reason is that type *T* investors, whose risk preferences determine the supply schedule, have many type *C* investors to trade with in homogeneous markets, but their presence in heterogeneous markets declines linearly.