Large investors and liquidity: a review of the literature

Matthew Pritsker[1]

Abstract

A growing share of financial assets are held by large institutional investors whose desired trades are large enough to move prices in markets. Because large investors' trades have "price impact", asset markets are not perfectly liquid from their perspective. This illiquidity is likely to influence their decisions of which assets to hold and which assets to trade, and may influence how assets are priced. These insights on illiquidity and large investors motivated Pritsker's (2002) modelling of liquidity in a market with large investors. This article is a companion piece to Pritsker (2002) which reviews the literature on asset liquidity and on large investors and suggests ways in which these research areas can be combined.

1. Introduction

The standard competitive asset pricing paradigm assumes that individual investors' desired trades are sufficiently small that each investor can take prices as given and hence choose their asset holdings while ignoring the price impact of their trades. The price-taking assumption is reasonable when applied to the trades of most individual investors, but it is less tenable when applied to the trades of institutional investors. The observed behaviour of many institutional investors - breaking apart a large trade into several smaller trades, or building up or selling a position over days - suggests that their desired trades have price impact, and that large institutions account for price impact when selecting their trading strategy (Chan and Lakonishok (1995)).

One notion of a perfectly liquid asset is an asset for which individuals can buy and sell all that they want at current prices. This notion of liquidity suggests that many markets are essentially perfectly liquid from the perspective of small investors since prices do not change much, if at all, in response to their desired trades. However, many markets are not perfectly liquid from the perspective of large investors. Because large investors are faced with imperfect market liquidity, the lack of liquidity may influence their investment decisions. For example, large investors who anticipate a potential future need to sell off assets quickly at some unexpected future date to meet cash flow obligations may desire holdings of relatively liquid assets in order to minimise the transaction costs associated with future forced sales. This desire for relatively liquid asset holdings should be reflected in equilibrium asset prices and returns.

The above observations suggest that large investors and asset market liquidity are related topics, and that whether liquidity risk is priced by the market may depend on the trading behaviour of large investors. The purpose of this paper is to review the literature on asset market liquidity and on large investors, and then suggest directions of research which synthesise the two topics. Motivated by the notion that large investors and liquidity are related, Pritsker (2002) studies asset market liquidity in a setting where there are many large and small investors who trade multiple risky assets over a large but finite number of time periods. The analysis in Pritsker builds on other models of large investors. The most closely related research is DeMarzo and Urosevic (2000), Vayanos (2001) and Urosevic (2001). The basic underlying framework in DeMarzo and Urosevic and in Vayanos is nearly identical. Both consider the behaviour of a single large investor and many small investors when the investors

[1] Board of Governors of the Federal Reserve System. The views expressed in this paper are those of the author but not necessarily those of the Board of Governors of the Federal Reserve System, or other members of its staff. Address correspondence to Matt Pritsker, The Federal Reserve Board, Mail Stop 91, Washington DC 20551. Matt may be reached by telephone on (202) 452-3534, fax (202) 452-3819, or by e-mail at mpritsker@frb.gov.

trade a single risky and risk-free asset over many time periods, and where all investors have CARA utility of consumption. The models differ in how they depart from this framework. DeMarzo and Urosevic consider a moral hazard setting in which the single large investor also expends costly effort in overseeing the activity of the firm. They model the investors' optimal oversight and portfolio choices together and then examine the implications of these decisions for asset prices. Vayanos modifies the basic setting to instead examine how asymmetric information about the large investor's holdings of risky assets influences the equilibrium behaviour of market prices. The basic framework in Urosevic (2001) extends the basic framework in DeMarzo and Urosevic to allow for multiple large investors and multiple risky assets. Urosevic then proceeds to examine the moral hazard setting when many large investors choose their portfolio holdings and the amount of effort to expend in monitoring the activity of the firms.

The basic modelling framework in Urosevic (2001) is, minus the moral hazard, essentially the same as that in Pritsker (2002).[2] Pritsker uses the basic framework to examine how market liquidity differs across large investors. He also examines how shocks to investors' endowments or to their cash flow needs affect equilibrium asset holdings and prices. Pritsker also examines how rumours or news about potential future financial distress by one large investor affects asset prices, trades and asset market liquidity. This latter exercise is of special interest in the light of the increases in asset market illiquidity that occurred following rumours of financial distress at Long-Term Capital Management, one large investor in financial markets. Pritsker's results to date are as follows:

1. *Asset pricing*: When investors hold Pareto-optimal asset allocations, then asset prices are the same as those in a competitive setting and the CAPM is satisfied. If, instead, investors' asset holdings are not Pareto-optimal, then assets' excess returns over the riskless rate satisfy a multifactor model where one factor is the market portfolio and the other factors correspond to each large investor's endowment.

2. *Market liquidity*: The presence of large investors affects market liquidity. When all investors are small, markets are perfectly liquid: no single investor's order flow has price impact. When large investors are present, then their order flow has price impact and thus they face illiquid markets. Interestingly, the amount of liquidity that is available to different large investors differs with their risk aversion: the lower a large investor's risk aversion, the greater the price impact of his/her trades.[3] Because all information in the model is public, these differences in liquidity across large investors are not related to information asymmetry; instead they are a purely strategic reflection of the imperfect competition features of the model.

3. *Market manipulation*: Risk-sharing and shock absorption are affected by the presence of large investors because large investors behave strategically. In equilibrium, some large investors respond to shocks by following trading strategies which appear to be like front-running. By contrast, when markets are competitive, large investors do not engage in such behaviour. More specifically, when markets are competitive and one investor is hit with a shock which substantially increases his/her risky asset holdings, his/her excess risky asset holdings are rapidly purchased by other large investors, and the market returns to optimal risk-sharing within a single period of trade. By contrast, when there are large investors in the market, if one large investor is hit with a shock which increases his/her supply of risky asset holdings, in equilibrium the other large investors respond by initially selling (not purchasing) risky assets, and then later purchasing them back. The large investors' trades are optimal because they anticipate that future sales by the large investor who was hit with the shock will

---

[2] Urosevic (2001) does not examine a setting with many large investors and many risky assets. He instead examines a setting with one large investor and many risky assets, or one risky asset and many large investors. In conversations with DeMarzo and Urosevic I learned that Urosevic has solved the multi-asset, multi-large investor case in his PhD thesis. I solved the general multi-asset, multi-large investor case independently by extending the three-period, single large investor, single asset model of Kihlstrom (2001) to allow for multiple time periods, large investors and assets. My earliest work (based on Kihlstrom) only considered investors who live for a large but finite number of time periods. After reading DeMarzo and Urosevic, I modified my model to allow for infinitely lived investors who trade risky assets for a finite number of time periods. My results after these modifications are much more elegant than those that were derived in my earlier analysis.

[3] If liquidity is measured as the slope of the price function with respect to one large investor's trades while holding the trades of other large investors fixed, then investors who are less risk-averse receive less liquidity by this measure. Alternatively, if liquidity is measured as the magnitude of the price impact associated with an investor's selling assets to raise cash, then by this alternative measure, less risk-averse investors continue to receive less liquidity from other investors.

eventually cause prices to decline. They exploit the expected future price decline by front-running the sales, selling just after the shock (selling high) and then purchasing back shares as prices decline (buying low).

4. *Shock propagation*: Shocks to participants' positions in one market affect asset prices in other markets. The price deviations in other markets due to a shock in one market depend on the covariances and variances of the assets' dividend payments. Assets whose dividends do not covary are not susceptible to endowment shocks.[4] Cash flow shocks, ie shocks which cause an individual market participant to sell assets to meet a particular cash need, cause shocks to propagate by another route. In particular, cash flow shocks can cause co-movement between the prices of assets whose dividends are not correlated.[5]

The next two sections review the literature on liquidity and on large investors. The conclusion provides suggestions for further research.

2. Market liquidity

The purpose of this section is to review the literature on market liquidity. The review is structured to cover the definition of liquidity, the sources of illiquidity, the measurement of liquidity and whether liquidity is priced in asset markets.

2.1 Market liquidity defined

Imperfect market liquidity is synonymous with the notion that there are costs associated with transacting. These costs can be explicit, such as the spread between bid and ask prices in securities markets, or they can be implicit, such as the search costs associated with matching buyers with sellers. Liquidity costs are important to the market participants that expect to bear them. These include for example broker-dealers in options markets since those dealers need to dynamically trade through time to hedge their options book.[6] Liquidity also matters to investors who may not expect to trade frequently, but might need to sell assets to meet cash needs in unforeseen circumstances.[7]

Although imperfect market liquidity is synonymous with transaction costs, it is generally impossible to define or to precisely measure the amount of market liquidity associated with a particular asset because liquidity encompasses many different attributes of the implicit and explicit structure of transaction costs. For example, Kyle (1985) describes market liquidity in terms of three attributes of transaction costs: the tightness, depth and resilience of the market, where tightness measures the cost of quickly buying and then selling a position, depth refers to the size of a transaction that is required to change prices, and resilience measures the speed at which prices recover to fundamentals after a non-informational trade. Using these attributes, it should be clear that comparing individual assets' liquidities is problematic because one asset could be more liquid along one dimension of transaction costs while the other is more liquid in a different dimension.

[4] The deviation of asset prices in market *j* due to a shock which reshuffles asset holdings in market *k* is equal to $\beta_{j,k}$ times the deviation of asset prices in market *k* where $\beta_{j,k}$ measures the covariance between dividends in markets *j* and *k* divided by the variance of dividends in market *k*.

[5] There is a large literature on shock transmission within the contagion literature. Models of how contagion occurs through financial markets include Kodres and Pritsker (2002) and Kyle and Xiong (2001).

[6] When markets are not statically complete, some market participants will find it optimal to follow a dynamic trading strategy. The standard example is a broker-dealer in a securities market who needs to dynamically hedge an option position.

[7] For small investors such circumstances might include medical emergencies or loss of a job. For large institutional investors such circumstances might include paying large insurance claims or mutual fund redemptions.

2.2 The sources of market illiquidity

Despite the lack of a precise definition of illiquidity, it is possible to model why markets may not be liquid. Three sources of illiquidity are commonly used in the academic literature. The first is exogenous transaction costs (Constantinides (1986), Heaton and Lucas (1996), Vayanos (1998), Vayanos and Vila (1999) and Huang (2002)) such as those that might arise from order processing costs, or the costs of commissions. The costs associated with search are another source of exogenous transaction costs (Duffie et al (2001)). The second major source is asymmetric information about asset payoffs (Glosten and Milgrom (1985), Kyle (1985), Kyle (1989) and Eisfeldt (2001)) or about market participants' endowments (Cao and Lyons (1999), Vayanos (1999) and Vayanos (2001)). When there is asymmetric information about asset payoffs, prices change in response to trades because of the information that the trades might convey about asset fundamentals. The resulting price response to trades is an additional cost of transacting. Similarly, if some market participants (such as broker-dealers) have private knowledge of other investors' endowments, they might be able to predict future price movements; and they might trade on this knowledge. As a result, prices will respond to the potential information content of these trades. The third major source of illiquidity is imperfect competition in asset markets due to the presence of large traders (Lindenberg (1979), Kyle (1985), Kyle (1989), Basak (1997), Cao and Lyons (1999), Vayanos (1999), DeMarzo and Urosevic (2000), Vayanos (2001), Kihlstrom (2001) and Pritsker (2002)). As noted above, a trader is large relative to the size of the market if the scale of their desired trading activity would have the effect of causing prices to change.

In addition to the three most common sources, market illiquidity has also been modelled as resulting from Knightian uncertainty. This is the type of uncertainty that might occur if traders in financial markets confront circumstances that are completely unanticipated, and for which it is not clear how to proceed. Securities dealers when confronted with such circumstances may follow very risk-averse strategies which minimise their losses should anyone wish to trade with them. The resulting spreads can be so wide, and hence the market so illiquid, that trading does not take place at the equilibrium spreads (Cherubini and Della Lunga (2001) and Routledge and Zin (2001)). Imperfect market liquidity has also been modelled as resulting from optimal securities design since a firm which sells liquid and illiquid securities can use the differences between the securities' characteristics to price discriminate between investors who care about liquidity and those who do not (Boudoukh and Whitelaw (1993) and DeMarzo and Duffie (1999)).

2.3 How is market liquidity measured?

Liquidity is important to investors because it affects the costs at which they can trade assets. The goal of liquidity measurement is to identify the cost structure which confronts investors, and hence influences their decisions on which assets to hold and when they should be traded.

Because there are many dimensions of the relevant cost structure, there is no single method for measuring market liquidity. Measures which are typically used in the empirical literature on liquidity and asset pricing include bid-ask spreads, various measures of the price impact of order flow, and various measures of order flow.[8] Of these measures, the price impact of order flow is perhaps the one that is used most widely. The advantage of this measure is that it is based on the observed price changes associated with trades. The bid-ask spread is in some sense a more limited measure since it indicates the prices for standardised relatively small trades; as a result many transactions take place at prices other than the bid or the ask. Measures that are solely based on trading volume are also limited, but for a different reason. Trading volume-based measures do not measure the transaction costs associated with trading activity; high volume is typically associated with liquidity, yet it is clear that trading volume could be high and markets could be very illiquid.[9] Despite the advantages of using the price impact of order flow as a measure of liquidity, tricky econometric issues are involved when

---

8 Measures of the price impact of order flow include price changes regressed on signed volume, or absolute price changes regressed on absolute volume, or daily price changes regressed on daily volume. Measures of volume include numbers of trades and daily volume measured in dollars.

9 The day of the October 1987 stock market crash involved high volume because many participants wanted to sell stock, but liquidity was reportedly very poor.

using the approach to uncover the cost structure faced by investors when they decide to make a trade. The tricky issues are measurement error, selection bias and simultaneity bias.

Measurement error arises from two sources. The first is that in some specifications of the relationship between asset prices and trades, the appropriate measure of trades indicates whether each trade was initiated by a participant that wanted to buy or sell an asset. Because most data sources do not indicate which side initiated a trade, the designation of the side that initiated a trade is one source of measurement error. A second source of measurement error is due to the price and quote data; often the data which is available to the econometrician is not the same as that which is available to market participants. Differences between the true and observed prices and quotes arise because for some infrequently traded assets, although traders' perceptions of prices may be updated frequently, the publicly observed prices and quotes may only be updated after a trade takes place. If trades are spaced far enough apart in time, then because the notional prices and quotes before a trade are not publicly available, it is difficult to measure the price impact or quote revision associated with the trade. Hence, measures of the price impact of trades will be only imperfectly measured.

Selection biases arise in liquidity measurement because the trades that are observed within a sample are dictated by the amount of market liquidity. To take an extreme example, suppose that markets are highly liquid at some times, and not liquid at others. If trades only occur at the liquid times then measures of liquidity which are based on the price impact of the observed trades will tend to overstate liquidity because they are based only on the select sample of times in which the markets were liquid. Another way in which sample selection biases manifest themselves in this area is that some assets may be so illiquid at all times that investors who tend to do trades above a particular size simply will not take positions in that asset. As a result the illiquidity of those assets for large trades is not identified in the data.

The final source of bias is simultaneity bias. This bias arises when trades and prices are both determined by some other difficult to control for factor such as economic news. When both variables are driven by additional factors it can appear that trades move prices, suggesting a level of market illiquidity, even when there is no relationship between trades and prices.

Most attempts to measure market liquidity using trade and quote data are carried out in the context of the market microstructure literature.[10] The potential sources of noise and bias in liquidity measurement are no doubt well known in this literature. However, there is not a generally accepted methodology to control for these biases. The significance of the potential biases that cannot be controlled for is unknown, and remains an important topic for future research especially since these estimated measures of asset illiquidity are often used to determine whether illiquidity is priced by asset markets. It is to that subject that I now turn.

2.4 Is asset illiquidity priced in asset returns?

If investors care about liquidity risk, and it influences their trading behaviour, then perhaps it should be priced into asset returns. This section reviews some of theoretical and empirical literature on whether liquidity risk is priced.

Theory

One of the earliest theoretical contributions which relates market liquidity and equilibrium expected rates of return is the model of Amihud and Mendelson (1986). Amihud and Mendelson consider a setting with risk neutral investors who differ in the time horizons over which they wish to hold risky assets. The assets in this model vary in their liquidity, where liquidity is modelled as a fixed bid-ask spread. Their principal theoretical result is that there are clientele effects in asset holdings in which investors with short horizons prefer to hold assets with small bid-ask spreads and investors with long horizons prefer to hold assets with larger spreads. As a result of the clientele effects, assets with larger transaction costs are shown to earn larger gross returns, suggesting that asset illiquidity is priced. It is important to stress that the transaction costs in the Amihud and Mendelson model are

[10] For example, Huang and Stoll (1997) provide a market microstructure model of the determinants of bid-ask spreads. Stoll (2001) provides a recent review of the empirical and theoretical market microstructure literature.

deterministic, not stochastic. To examine whether there are systematic components to liquidity, and whether these components are priced, a model with stochastic liquidity is required.

Acharya and Pedersen (2002) present a model in which liquidity is stochastic. The model contains small investors who have CARA utility and face stochastic dividends and an exogenous stochastic transaction cost associated with selling assets. Their main insight is that returns net of transaction costs should satisfy the CAPM in this framework. They use this insight to solve for asset prices in an overlapping generations model framework in which each generation of investors lives for two periods. They show that within this framework, asset returns (not net of transaction costs) have a conditional four-factor structure with non-zero alpha:

$$E_{t-1}(r_t^i - r_t^f) = E_{t-1}c_t^i + \lambda_{t-1}\, Cov_{t-1}(r_t^i, r_t^m) + \lambda_{t-1}\, Cov_{t-1}(c_t^i, c_t^m)$$
$$- \lambda_{t-1}\, Cov_{t-1}(r_t^i, c_t^m) - \lambda_{t-1}\, Cov_{t-1}(c_t^i, r_t^m) \qquad (1)$$

where c measures transaction costs, and i and m denote asset i and the market portfolio respectively. This equation makes several contributions to the literature on liquidity. First, it shows how stochastic transaction costs fit into the general asset pricing framework. Second, it shows that when estimating asset pricing models using returns which do not net out transaction costs, then asset returns have a four-factor structure and a non-zero alpha which is related to expected transaction costs.

It is important to stress that the Acharya and Pedersen framework is not truly a four-factor model; the only true factor is the market portfolio. However, the model appears to have four factors because it is written in terms of gross returns (which are irrelevant to investors) instead of returns net of transaction costs (which investors care about).[11] This raises a second issue; since all investors in Acharya and Pedersen's framework have CARA utility and since the only sources of risk are traded asset risk, asset markets in their framework are effectively complete both statically and dynamically.[12] As a result, agents in this framework do not hold liquid assets for their insurance value in meeting unforeseen future cash needs. Further, market participants in Acharya and Pedersen's framework do not have incentives to hedge against changes in future market liquidity. This suggests that in a more realistic setting, asset liquidity may affect asset prices in ways which are not accounted for in this framework.[13]

An issue related to how illiquidity affects asset returns is how it affects asset prices. Duffie et al (2001) address this issue in the context of the prices of durable goods such as houses (or stocks). The source of illiquidity in the model is repeated adverse selection which arises because the seller of the asset knows more about its quality than the buyer.[14] A consequence of adverse selection is that a seller of a house may sometimes choose to forgo some favourable moving opportunities (such as career change) because the adverse selection problem prevents him/her getting a high enough price for the house. Duffie et al show that the discounted expected value of these future missed opportunities is built into the price of the house. A similar mechanism appears to be operating in financial markets with imperfect competition. In Pritsker (2002), imperfect competition in the asset markets causes traders to adjust their asset positions slowly towards Pareto-optimal asset allocations. The discounted future deviations from Pareto-optimal asset allocations are one determinant of the current price of the asset.

11 Another way to see that there is only one factor is to note that the market price of risk of all the factors in the four-factor model are identical, indicating there is really only one factor.

12 That is, asset prices are the same as they would be if a full set of Arrow-Debreu contingent securities was allowed to be traded in the economy.

13 Holmstrom and Tirole (2001) differ from Acharya and Pedersen in that they examine how binding, borrowing constraints on corporate borrowers generate a desire for corporations to hold liquid assets to hedge against the market incompleteness generated by the borrowing constraint. Holmstrom and Tirole's analysis is related to corporations' need for liquidity, but it is not related to asset market liquidity where liquidity is measured as a transaction cost.

14 Their model assumes that the purchaser of an asset may be imperfectly informed about asset quality at the time of purchase, but better informed at the time of asset sale. This is reasonable for houses, but less reasonable for financial assets such as stocks.

Empirical evidence

The early literature on liquidity and asset pricing was motivated by the framework in Amihud and Mendelson (1986), and thus studied whether stocks earn higher returns if they are less liquid, where liquidity is measured by the stock's bid-ask spread as a proportion of asset price. In their analysis, they regressed stocks' excess returns over the riskless rate on estimated market β's and on the proportional bid-ask spread. Their analysis suggested that assets with higher transaction costs, as measured by the spread, earn higher returns. Later work by Brennan and Subrahmanyam (1996) was unable to find reliable evidence that bid-ask spreads were priced. The sources of the differences from Amihud and Mendelson's earlier results are not resolved in their paper, but obvious candidates for explaining the differences are that Brennan and Subrahmanyam used a different econometric testing approach, and additionally they controlled for the factors that Fama and French (1993) showed appear to have power for pricing assets.[15] While Brennan and Subrahmanyam did not find evidence that bid-ask spreads were priced, they found evidence that the market rewarded stocks for which the price impact of trades was higher, where the price impact of trades was estimated based on a market microstructure methodology.

Several new empirical papers have been written on liquidity and asset pricing. These papers are motivated by recent empirical evidence that the liquidities of many assets tend to move together through time, suggesting that there are common factors which determine assets' market liquidity.[16] Pastor and Stambaugh (2001) create data series which measure time variation in the liquidity of individual stocks. They then use market-wide averages of these data series as a proxy for a systematic liquidity factor. The liquidity measure for individual firms is based on the tendency of a firm's excess returns over a market index to experience negative autocorrelation in returns over a two-day period given high trading volume on the first day. This approach builds on the notion that price changes that are due to illiquidity are mean-reverting, and that liquidity measures should be based on the magnitude of price changes relative to volume, with greater price change for a given volume interpreted as evidence of illiquidity.[17,18] They then test whether their measure of market illiquidity is priced by asset markets. They find strong evidence that it is priced even after controlling for the Fama and French (1993) factors. In my view, the Pastor and Stambaugh results are intriguing, but because it is not entirely clear whether they have found a proxy for liquidity or for something else, more work needs to be done on the properties of their proxies before their results can be viewed as entirely convincing.

Acharya and Pedersen (2002) also examine whether liquidity risk is priced, but they use a different measure from that used by Pastor and Stambaugh. Acharya and Pedersen's measures of liquidity are based on daily absolute price changes normalised by daily trading volume. This measure of liquidity is similar to that of Pastor and Stambaugh in that both account for the relationship of volume and price movement; however, the Acharya and Pedersen measure does not condition on the tendency for prices to reverse themselves. Acharya and Pedersen create their proxies of liquidity for individual stocks and for a proxy for the market portfolio. They find that the estimated coefficients on the liquidity variables tend to have the correct sign, and to be economically significant, but usually they are not estimated precisely enough to be statistically significant.

Chordia, Subrahmanyam and Anshuman (2001) also analyse the role of liquidity in asset pricing, but unlike most other analyses, they focus on variability in liquidity as proxied for by variability in measures of trading volume. They hypothesise that risk-averse investors should dislike variability in liquidity and thus stocks with more variability in liquidity should have lower prices and hence earn higher expected returns. In fact, their results are of a somewhat puzzling nature because they find strong evidence that the opposite of their hypothesis is true. My view is that the evidence in Chordia, Subrahmanyam and Anshuman may not be as puzzling as it seems, but that it instead points towards a need for more

---

[15] The Fama and French factors had not been discovered at the time that Amihud and Mendelson wrote their paper.

[16] Chordia et al (2000) document liquidity commonality for stocks. Chordia, Sarkar and Subrahmanyam (2001) show that there are common components which drive liquidity in the stock and bond markets.

[17] Recall that as noted earlier, if trades and prices are both responding to some other factor, such as economic news, then measures of liquidity which are based on the relationship between trades and prices can be misleading.

[18] Pastor and Stambaugh do not carefully justify why their liquidity measure for individual stocks is based on the autocorrelation of excess returns over a market index, as opposed to autocorrelation of the firms' returns.

theoretical research of how liquidity matters in portfolio choice. I discuss this point further in the conclusions.

The next section examines the literature on large investors, which is one source of market illiquidity.

3. Large investors

There has been enormous growth in the share of asset trades that are done by institutional investors. Since these investors often take large positions relative to the size of the markets in which they trade, even in the absence of other transaction costs, some markets may not be liquid from their perspective. Pritsker (2002) examines the role of institutional investors in determining market liquidity. The related literature on large investors in markets can be broadly broken down into three areas: why there are large investors; how they affect equilibrium asset pricing; and whether they stabilise or destabilise asset markets. Most of my discussion is related to the first two of these subjects. A comprehensive review of the literature on the third subject might require a separate paper.

3.1 Why are there large investors?

Many of the models of large investors and asset pricing take as primitives the set of large and small investors in the economy, their preferences and trading mechanisms, and then given this setup solve for the behaviour of asset prices. The contribution of some of the related literature on financial intermediation is that it derives the structure of the participants in financial markets; in essence it establishes why some investors in financial markets are large while others remain small and it establishes why small and large investors can coexist. Ideally, this literature can also be extended to model the behaviour of financial markets when there is intermediation. The full set of related financial intermediation literature is too large to review here. But I will discuss it briefly and highlight a few recent articles that are of interest.

Theories of financial intermediation provide a natural explanation for why there are large investors in financial markets since most financial intermediaries are large. Financial intermediaries such as insurance companies, banks, pension funds and mutual funds issue liabilities to small investors and then purchase assets to back up those liabilities. The traditional basis for why small investors enter into contracts with financial intermediaries includes pooling of risk (insurance companies), pooling of risk from liquidity needs (banks and mutual funds) and economising on the costs of monitoring borrowers (banks).[19] The growth of large institutions which is due to these sources has led to a deepening of markets, and to ever more complex financial products. This trend toward complexity is a self-reinforcing contributor to the increasing role of large institutional investors in markets and to the shrinking role of small investors, since large institutions are the only investors that can afford to pay the high fixed information costs associated with pricing and trading complex products (Allen and Gale (1999)).[20]

Although the amount of direct participation in markets by small investors is shrinking, there may be room for large institutional investors and small investors to both interact in markets. Two recent articles derive roles for small and large investors in the context of theories of mutual funds. The first, by Nanda and Singh (1998), emphasises the liquidity services provided by mutual funds. Their model has two types of small investors: one which is vulnerable to idiosyncratic future liquidity shocks à la Diamond and Dybvig (1983) and the other which is less vulnerable. The vulnerable investors invest in a mutual fund which holds sufficient liquid assets to meet their joint liquidity needs. They pay the mutual funds a

[19] Diamond and Dybvig (1983) model how banks pool liquidity risk. Nanda and Singh (1998) model how mutual funds economise on liquidity risk; they also analyse the relationship between the structure of mutual fund pricing and liquidity. Diamond (1984) models how financial intermediaries economise on monitoring costs.

[20] The high fixed information costs associated with learning to price and trade these products makes it prohibitively expensive for small investors to do so on their own. Instead small investors use large financial intermediaries to trade these products on their behalf. Since these intermediaries can spread the fixed costs of information over many small investors, it becomes economical for small investors to benefit from these products when intermediation is available.

fee for these liquidity services which is reflected in the funds providing performance which is not as good as the market return. Investors who are less vulnerable to liquidity shocks do not invest in the mutual fund and thus remain small. Mamaysky and Spiegel (2002) emphasise a different aspect of mutual funds, namely their ability to reduce the informational costs of following dynamic trading strategies. When markets are statically incomplete, it may be optimal for some investors to follow dynamic trading strategies. Because such strategies may involve closely monitoring market developments, the monitoring costs of a single investor implementing the strategy may be prohibitive. On the other hand, paying an institutional investor to follow a strategy which is customised to the optimal dynamic strategy of each small investor customer is also prohibitively expensive. In the light of these costs, Mamaysky and Spiegel argue that if families of mutual funds advertise funds which follow different dynamic strategies, then an investor who splits his/her wealth among different funds within the same fund family essentially creates a dynamic strategy which may come reasonably close to the optimal dynamic strategy and which may produce a better outcome than paying the costs of continually monitoring market developments. Of course, while there may be some small investors who use mutual funds for the purposes of dynamic strategies, the logic of their model suggests that some small investors for whom markets are "complete enough" will choose to remain small. The contribution of these models to modelling the behaviour of large investors is that they go some distance towards highlighting the differences in objectives of large and small investors. Ideally the differences in these objectives should be reflected in models of large investors and asset pricing. Unfortunately, for reasons of tractability, almost all large investor models assume that market participants have CARA utility or are risk neutral.[21] It is to these asset pricing models that I now turn.

3.2 Large investors and asset pricing

When investors are large enough that they do not take prices as given, then their non-price-taking behaviour is a deviation from the classical price-taking assumption. One issue which the large investor literature seeks to examine is how deviations from price-taking behaviour affect equilibrium returns in asset markets. There are two classes of model which examine how large investors affect equilibrium asset returns; the first class of models are those in which the presence of large investors is the only market imperfection; the second are those in which there are also additional sources of market imperfections. The additional market imperfections typically take the form of asymmetric information about asset payoffs or investors' asset holdings. A second source of market imperfection involves agency problems between firm management and shareholders. To begin I will address models in which the only market imperfection is the presence of non-price-taking investors.

When there are investors who do not take prices as given, markets are not competitive by definition, hence one would expect that asset prices will vary from their levels in competitive asset markets. Lindenberg (1979) shows that this intuition is correct in static one-period models of asset prices. In particular he shows that when both large and small investors who have mean-variance utility are present, asset returns have a multifactor structure where one factor is the market portfolio and the other factors correspond to the endowments of large investors. By contrast, if the investors in Lindenberg's model behave competitively, then it is well known that the CAPM holds.

It turns out that in multiperiod models markets are much more competitive than in the case considered by Lindenberg, and can produce the same asset prices as in competitive models in limiting cases. The basic intuition for why multiperiod models produce intense competition is based on Coase (1972). Coase argued that a monopolist selling durable goods today could not credibly commit to not sell the same goods in the future at a lower price, and therefore since durable goods today are close substitutes for durable goods in the future, his/her future and current sales would compete, forcing down prices. Moreover, if the time periods when the monopolist can sell are spaced arbitrarily closely together, then Coase conjectured that the competition across time periods would be so intense that the monopolist would be forced to charge the competitive price in the limit. Kihlstrom (2001) argues that the Coasian logic applies to sales of financial assets since they are also durable goods, and hence competition through time would force the sale price of stocks to be lower than in the monopolist case. DeMarzo and Urosevic (2000) consider an infinite horizon setting and show, in their model of a

[21] A notable exception is Basak (1997). However, because of intractability, Basak does not use his model to study whether the Coasian dynamics that are discussed in the next section are present in his framework, or how they would affect his results.

single large investor, that as the time between trades goes to zero, prices converge to those that would be found in a perfectly competitive model. They also find that when the time between trades is finite, then asset prices contain a risk premium term as a result of non-optimal risk-sharing among investors. In a setting with many large investors one would expect the competition to be more intense than when a single large investor is present, hence one would expect that Coasian dynamics would also be present in such a setting. Urosevic (2001) shows that Coasian dynamics are present in his model. They are also present in Pritsker (2002). Based on the literature with only a single market imperfection, it is clear that in a multiperiod setting, markets can be nearly as competitive as those in which perfect competition is present, with perfectly competitive markets as a limiting case.

When there are additional sources of market imperfections, they sometimes have the effect of reducing the competitiveness of multiperiod models. There is a large literature on the behaviour of large investors who have private information, dating back to Kyle (1985). Kyle's setting has a single large trader who is informed about the liquidation value (or end-of-day value) of a risky asset, noise traders who trade for reasons unrelated to fundamentals, and competitive market-makers who set prices equal to the expected fundamental liquidation value of the asset conditional on the information in traders' order flow. Kyle shows that the large investor's information is incorporated into prices through time, and that even in the limit as the time between trades gets small, the large investor's information is only slowly incorporated into prices, and importantly is incorporated much more slowly than it would be in a competitive framework. This suggests that competitive Coasian dynamics may not dominate the behaviour of asset prices when other sources of market imperfections, such as information asymmetries, are present. Vayanos (1999) considers a different type of information asymmetry, knowledge of large investors' private endowments. More specifically, he considers a setting in which there are several large investors who are subject to endowment shocks which only they observe. Hence they have private information about their own endowments. They trade in a multiple period setting by submitting linear demand curves. The resulting equilibrium price is that which clears markets. In this asymmetric information setting, Vayanos shows that a higher trading frequency does not cause asset prices to become competitive in the limit as the time between trades goes to zero; ie in this setting traders continue to hide their information. Vayanos also shows that if investors' endowments are public information, then the asset dynamics are the same as in a Coasian model, ie prices are competitive. Vayanos (2001) considers a different setting in which there is a single large trader, competitive market-makers and noise traders. The large trader receives endowment shocks as before, and these endowments are privately observed. In each time period, the large investor receives a private endowment shock, then the market-maker forms his/her optimal demand curve. The large trader takes this demand curve as given when choosing the quantity that he wishes to purchase. The large trader's order flow and the demands of the noise trader are submitted together at each market clearing. In each period the resulting price is set to clear the market. Then time passes and a new period starts. In the setting of Vayanos (2001), the informed investor's information is quickly revealed to the market, and asset prices quickly become competitive. It is not clear why information is revealed so quickly in Vayanos (2001) while it is revealed slowly in Kyle (1985) and Vayanos (1999).

The second type of market imperfection is models of agency problems. A standard agency problem is moral hazard resulting from firm management that cannot be perfectly monitored, and that in the absence of monitoring may choose to shirk on their duties by expending too little effort, or worse yet expropriate the shareholders' assets and instead spend them on salary and perquisites for the firm's management. If investors hold widely diversified portfolios, then an individual investor's incentive to monitor a particular firm is small since the benefits accruing to them are small (they hold few shares), but the costs of monitoring may be high. Moreover, each investor has an incentive not to monitor if they believe other investors will do it for them. This free-rider problem can result in an amount of monitoring which is socially suboptimal. If instead there is a large investor, then because their stake in the firm is relatively large, their incentives to monitor are larger as well; hence the presence of a large investor may help to overcome the free-rider problem. On the other hand, a large investor may be underdiversified, so there is a tension between optimal risk-sharing and monitoring. These issues are addressed and discussed by Admati et al (1994), DeMarzo and Urosevic (2000) and Urosevic (2001). One of the interesting findings in DeMarzo and Urosevic (2000) is that the speed of convergence of asset prices depends on whether the agency problem is "small". When it is small enough, asset prices quickly converge to their competitive values when the time between trades goes to zero, but when the agency problem is large enough, they do not, and instead the Coase conjecture does not hold in their setting.

Before proceeding, it is useful to summarise the theoretical results on large investors and asset pricing. It appears that if the time between trades is large enough, or if there are other market imperfections such as asymmetric information, then the presence of large investors may slow the rate at which participants adjust their positions towards optimal risk-sharing. As a result market prices will reflect deviations from optimal risk-sharing. On the other hand, even with some market imperfections, prices can appear to be very close to those in a competitive framework. In the end, whether large investors significantly affect equilibrium asset returns is an unsettled question.[22]

3.3 Do large investors stabilise markets?

One of the reasons that large investors receive so much attention is because they are so often blamed for speculating against currencies, or manipulating markets, causing exchange rate pegs to collapse. In addition, some empirical literature claims that large investors herd or engage in positive feedback trading, and that this activity can destabilise markets.

Some of the large investor models that were discussed in the previous subsection appear to have the feature that large investors can manipulate markets.[23] Vayanos (2001) finds circumstances in which the large investor appears to follow a market manipulation strategy in which he sells more shares to the market-makers than would be required for competitive risk-sharing, and then buys them back; this sell high, buy low strategy looks like market manipulation. Vayanos attributes the resulting price movements to the information asymmetry in his model. Pritsker (2002) also finds that large investors appear to engage in manipulative behaviour by responding to a positive endowment shock to one large investor by initially short selling stock to other price-taking investors, and then buying the stock back as prices decline. This behaviour looks similar to that in Vayanos, but there is no asymmetric information in Pritsker. This suggests information asymmetry may not be required to generate trades that appear to look like market manipulation. It may suffice to have a model with large investors.

One of the most important questions about the role of large investors in financial markets is whether their presence helps to coordinate speculative attacks on a currency. Corsetti, Dasgupta, Morris and Shin (2001) discuss this general issue. They find that the addition of a large investor into a financial market can cause other investors to attack a currency more aggressively, but the net effects of this activity can be small. However, if the large investor can signal his/her position (or trade) to small investors before they act, then other investors' ability to condition on the trades of the large investor help those investors solve a coordination problem, significantly increasing the prospects of the speculative attack's success. Corsetti, Pesenti and Roubini (2001) review the related empirical evidence on the role of large investors during the recent Asian crisis.

4. Conclusions

This paper has reviewed the literature on market liquidity and on large investors. The analysis points towards two areas where more research could be fruitful. The first is more theoretical research on how asset liquidity should affect asset returns. This theoretical research should motivate the empirical literature. To illustrate why such research might be useful, it is useful to first revisit the findings of Chordia, Subrahmanyam and Anshuman (2001). They find that variability in a proxy for assets' liquidity appears to be priced negatively, ie high variability of liquidity implies lower expected returns. One possible explanation for this finding is that variability in liquidity is valuable to investors because it has option value - that is, investors can hold stocks with high variability in liquidity in order to only trade the most liquid assets at any particular point in time. The associated reductions in transaction costs may make it very desirable to hold stocks with high variability in liquidity. As a result, these stocks

[22] Another reason why this question remains unsettled is the intractability of large investor models. The typical model assumes large investors are risk neutral or have CARA utility. These assumptions provide tractability, but they are not without loss of generality. I suspect functional forms for utility in which risk aversion depends on wealth would lead to different results - if anyone could solve a large investor model with such utility functions.

[23] For a model of market manipulation, see Jarrow (1992).

might be expected to have lower expected returns.[24] If this explanation is correct, it suggests that to properly model why liquidity matters, theoretical models need to consider the dynamics of asset trading and assets' liquidity.

An additional area where more research might be fruitful is in empirically relating liquidity premia to institutional investors' asset holdings. Since institutional investors may focus more on liquidity than small investors, careful studies of institutional investors' trading strategies, with a particular focus on the choices of the assets that they choose to hold and choose not to hold, may help to contribute to our understanding of how asset liquidity affects equilibrium stock returns.

Bibliography

Acharya, V V and L H Pedersen (2002): "Asset pricing with liquidity risk", *Working Paper*, Stern School of Business, New York University.

Admati, A, P Pfleiderer and J Zechner (1994): "Large shareholder activism, risk sharing, and financial market equilibrium", *The Journal of Political Economy*, 102, 1097-130.

Allen, F and D Gale (1999): "Innovations in financial services, relationships, and risk sharing", *Management Science*, 45, 1239-53.

Amihud, Y and H Mendelson (1986): "Asset pricing and the bid-ask spread", *Journal of Financial Economics*, 17, 223-49.

Basak, S (1997): "Consumption choice and asset pricing with a non-price taking agent", *Economic Theory*, 10, 437-62.

Boudoukh, J and R F Whitelaw (1993): "Liquidity as a choice variable: a lesson from the Japanese government bond market", *Review of Financial Studies*, 6, 265-92.

Brennan, M J and A Subrahmanyam (1996): "Market microstructure and asset pricing: on the compensation for illiquidity in stock returns", *Journal of Financial Economics*, 41, 441-64.

Cao, H H and R K Lyons (1999): *Inventory information*, mimeo, Haas School of Business, University of California at Berkeley.

Chan, L K C and J Lakonishok (1995): "The behavior of stock prices around institutional trades", *Journal of Finance*, 50, 1147-74.

Cherubini, U and G Della Lunga (2001): "Liquidity and credit risk", *Applied Mathematical Finance*, 8, no 2, 79-95.

Chordia, T, R Roll and A Subrahmanyam (2000): "Commonality in liquidity", *Journal of Financial Economics*, 56, 3-28.

Chordia, T, A Sarkar and A Subrahmanyam (2001): "Common determinants of bond and stock market liquidity: the impact of financial crises, monetary policy, and mutual fund flows", *Working Paper*, The Federal Reserve Bank of New York.

Chordia, T, A Subrahmanyam and V R Anshuman (2001): "Trading activity and expected stock returns", *Journal of Financial Economics*, 59, 3-32.

Coase, R (1972): "Durability and monopoly", *Journal of Law and Economics*, 15, 143-9.

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

Corsetti, G, A Dasgupta, S Morris and H S Shin (2001): *Does one Soros make a difference? A theory of currency crises with large and small traders*, mimeo, Yale University.

[24] Chordia, Subrahmanyam and Anshuman mentioned this optionality aspect as one possible explanation for their findings in early drafts of their paper but later removed references to it because they did not have a way of empirically testing the explanation.

Corsetti, G, P Pesenti and N Roubini (2001): *The role of large players in currency crises*, mimeo.

DeMarzo, P M and D Duffie (1999): “A liquidity-based model of security design”, *Econometrica*, 67, 65-99.

DeMarzo, P M and B Urosevic (2000): “Optimal trading by a large shareholder”, *Working Paper*, Haas School of Business, University of California at Berkeley.

Diamond, D W (1984): “Financial intermediation and delegated monitoring”, *Review of Economic Studies*, 51, 393-414.

Diamond, D W and P Dybvig (1983): “Bank runs, deposit insurance, and liquidity”, *Journal of Political Economy*, 105, 928-56.

Duffie, D, N Gârleanu and L H Pedersen (2001): “Valuation in dynamic bargaining markets”, *Working Paper*, Graduate School of Business, Stanford University.

Eisfeldt, A (2001): *Endogenous liquidity in asset markets*, mimeo, Kellog Graduate School of Management, Northwestern University, April.

Fama, E F and K R French (1993): “Common risk factors in the returns on stocks and bonds”, *Journal of Financial Economics*, 33, 3-56.

Glosten, L R and P R Milgrom (1985): “Bid, ask, and transaction prices in a specialist market with heterogeneously informed traders”, *Journal of Financial Economics*, 14, 71-100.

Heaton, J and D J Lucas (1996): “Evaluating the effects of incomplete markets on risk sharing and asset pricing”, *Journal of Political Economy*, 104, 443-87.

Holmstrom, B and J Tirole (2001): “LAPM: a liquidity-based asset pricing model”, *The Journal of Finance*, 56, 1837-67.

Huang, M (2002): “Liquidity shocks and equilibrium liquidity premia”, *Journal of Economic Theory*, forthcoming.

Huang, R D and H R Stoll (1997): “The components of the bid-ask spread: a general approach”, *Review of Financial Studies*, 10, 995-1034.

Jarrow, R A (1992): “Market manipulation, bubbles, corners, and short squeezes”, *Journal of Financial and Quantitative Analysis*, 27, no 3, 311-36.

Kihlstrom, R (2001): “Monopoly power in dynamic securities markets”, *Working Paper*, The Wharton School, University of Pennsylvania.

Kodres, L E and M Pritsker (2002): “A rational expectations model of financial contagion”, *Journal of Finance*, 57, 769-99.

Kyle, A S (1985): “Continuous auctions and insider trading”, *Econometrica*, 53, 1315-35.

——— (1989): “Informed speculation with imperfect competition”, *Review of Economic Studies*, 56, 317-56.

Kyle, A S and W Xiong (2001): “Contagion as a wealth effect”, *Journal of Finance*, 56, 1401-40.

Lindenberg, E B (1979): “Capital market equilibrium with price affecting institutional investors”, in *Portfolio theory 25 years after: essays in honor of Harry Markowitz*, Elton, E J and M J Gruber, eds, North Holland, pp 109-24.

Mamaysky, H and M Spiegel (2002): “A theory of mutual funds: optimal fund objectives and industry organization”, *Working Paper*, Yale School of Management.

Nanda, V and R Singh (1998): “Mutual fund structures and the pricing of liquidity”, *Working Paper*, University of Michigan Business School.

Pastor, L and R F Stambaugh (2001): “Liquidity risk and expected stock returns”, *Working Paper*, University of Chicago Graduate School of Business.

Pritsker, M (2002): *Large investors, implications for equilibrium asset returns, shock absorption, and liquidity*, mimeo, Federal Reserve Board.

Routledge, B R and S E Zin (2001): *Model uncertainty and liquidity*, mimeo, July.

Stoll, H R (2001): “Market microstructure”, *Working Paper* 01-16, Financial Markets Research Center, Vanderbilt University.

Urosevic, Branko (2001): “Dynamics of insider ownership: theory and evidence”, *Working Paper*, Haas School of Business.

Vayanos, D (1998): “Transaction costs and asset prices: a dynamic equilibrium model”, *Review of Financial Studies*, 11, 1-58.

——— (1999): “Strategic trading and welfare in a dynamic market”, *Review of Economic Studies*, 66, 219-54.

——— (2001): “Strategic trading in a dynamic noisy market”, *Journal of Finance*, 56, 131-71.

Vayanos, D and J L Vila (1999): “Equilibrium interest rate and liquidity premium with transaction costs”, *Economic Theory*, 13, 509-39.