Rohit Rahi and Jean-Pierre Zigrand

A theory of strategic intermediation and endogenous liquidity

Working paper

A Theory of Strategic Intermediation and Endogenous Liquidity*

by

Rohit Rahi
Department of Finance,
Department of Economics,
and Financial Markets Group,
London School of Economics,
Houghton Street, London WC2A 2AE

and

Jean-Pierre Zigrand
Department of Finance
and Financial Markets Group,
London School of Economics,
Houghton Street, London WC2A 2AE

December 17, 2007.

*We thank seminar participants at the LSE/FMG Conference on Liquidity, the EFA meetings in Maastricht, Cass Business School, and University of Venice.

Abstract

Market liquidity is typically characterized by a number of ad hoc metrics, such as depth (or market impact), volume, intermediation costs (such as breadth) etc. No general coherent definition seems to exist, and few attempts have been made to justify the existing metrics on welfare grounds. In this paper we propose a welfare-based definition of liquidity and characterize its relationship with the usual proxies. The model on which the welfare analysis rests is an equilibrium model with multiple assets and restricted investor participation. Strategic intermediaries pursue profit opportunities by providing intermediation services (i.e. "liquidity") in exchange for an endogenous fee. Our model is well suited to study the contagion-like effects of liquidity shocks. We also consider the case in which intermediaries can optimally design securities.

Journal of Economic Literature classification numbers: G10, G20, D52, D53.
Keywords: Liquidity, intermediation, arbitrage, restricted participation, contagion, market microstructure.

1 Introduction

Liquidity has long been a puzzle to financial economists. Given the myriad connotations of "market liquidity,", it is a bit odd that more attempts have not been made to analyze and reconcile the various aspects of liquidity within one equilibrium model. Some well-known attributes of liquidity are depth (the market impact of a trade), breadth (the size of bid-ask spreads, also referred to as tightness), volume, as well as timeliness and ease of execution. Rather than define liquidity by its attributes, we define liquidity by the underlying function that gives rise to those attributes. While any one model may be too specialized to capture all, or even many, of the salient features, we believe that a general equilibrium model such as the one proposed here may help to guide our intuition as to which features of liquidity are worthy of analysis. Indeed, since most papers on market liquidity are partial equilibrium models or partial equilibrium empirical studies, it has not been obvious why the focus has been on one or the other asset or one or the other measure of liquidity. It also is not clear whether such proxies are in fact exogenous.

We believe that the study of liquidity needs to ultimately be unambiguously grounded in a general equilibrium welfare analysis. Liquidity affects trades, which then may affect depth, tightness and timeliness, which in turn affect liquidity and welfare. In order to find a metric that is at the same time intuitive and welfare-based, one does need to resort to a realistic general equilibrium model with liquidity demanders and liquidity suppliers. In this paper liquidity is provided both by investors and by financial intermediaries. We explicitly model the objectives of intermediaries and we view their profits as being part of the social welfare function. We agree with Dewatripont and Tirole (1993) when they say:

> The behavior of financial intermediaries ... largely determines the liquidity of financial markets. ... A more complete understanding of financial markets should thus explicitly integrate financial intermediation.

Within this setup, the current paper introduces and defends a particular liquidity metric and shows that there is an unambiguous relationship with the attributes of liquidity and welfare. The metric proposed is not model-dependent, but its properties of course will be. Roughly, liquidity is defined here as the gains from trade achieved in equilibrium. This liquidity metric is therefore not foremost a measure of how quickly and cheaply a particular market order is executed. It is a measure of the ease of trade perhaps of one asset or another, but mostly it measures the ease with which interpersonal gains from trade can be realized. In this sense, markets may be very liquid, allowing welfare relevant trades to happen at low cost, and yet some asset markets may have large bid-ask spreads or low depth. These particular markets may therefore be rather irrelevant from a welfare point of view, and should not affect the overall liquidity measure much. We are also concerned in this paper with the contagion effects of liquidity. Our model lends itself directly to study how in equilibrium a liquidity shock in one sector of the economy is transmitted by intermediaries to other sectors, and which markets bear the brunt of the shock. We also find a feedback effect through which a detrimental liquidity shock lowers the number of intermediaries which in turn lowers liquidity and so on.

To make room for liquidity-providing intermediaries, we model asset markets as segmented. While there are assets that each given group of investors can trade among themselves, some trades with other groups of investors require the intervention of intermediaries. As an obvious example we can cite the fact that buyers and sellers of more exotic financial instruments rely on inter-dealer brokers to facilitate liquidity by gathering pricing information and identifying counterparties with reciprocal interests. As in the real world, intermediaries are modeled as larger, strategic entities maximizing trading profits, given the trades of other intermediaries. Asset prices and bid-ask spreads are determined endogenously at a Nash equilibrium. Since we allow entry into the intermediation sector to be unrestricted, but costly, the number of arbitrageurs is endogenous within the model. In other words, liquidity depends on the number of intermediaries, and the number of intermediaries depends on the liquidity. Liquidity in our model can therefore be thought of as provided partly by the endogenous number of intermediaries ("across markets liquidity" in O'Hara (1995)) and partly through direct centralized trading in markets ("within markets liquidity" in O'Hara (1995)). This also contrasts with many market microstructure studies where all trades must pass through market makers. When linking liquidity to welfare, the endogenous transaction costs are not viewed as deadweight costs; they are rather viewed as forming the intermediaries' profits, which need to be accounted for in any welfare analysis,

In order to keep the model tractable, we abstract from some attributes of liquidity. In particular, we chose to study an economy with only two dates, so that the aspect of liquidity as the price of immediacy an investor needs to pay to an intermediary in order to transact now rather than later (as in Grossman and Miller (1988)) cannot be captured. Nor can we capture resiliency, the tendency that order flows do or do not have to induce return reversals. A second simplifying assumption imposed in this paper is that information is symmetric.

Related Literature. There are many categories of papers studying market liquidity directly or indirectly, with different foci. Hodrick and Moulton (2003) illustrate some of these categories in a simple reduced form model. However, we are not aware of any papers that define the "extent of liquidity" via an explicit metric that itself has a clear welfare meaning, or that contrast this definition with the different attributes of "liquidity," such as depth, bid-ask spreads, transaction costs, immediacy services and the like. Grossman and Stiglitz (1980) directly assume exogenous liquidity *trades*, rather than liquidity shocks (idiosyncratic shocks to endowments or preferences) that may give rise to optimal liquidity trades. This is also the case in the models of Kyle (1985) and Glosten and Milgrom (1985) and in much of the ensuing literature on market microstructure. A number of papers, following Diamond and Verecchia (1981), have taken this further and study how different specifications of "liquidity shocks" translate into optimal trades and equilibrium outcomes. Financial contagion has been studied by Allen and Gale (2000), Freixas et al. (2000), Fernando (2003), and Gromb and Vayanos (2007), among others.

Traditionally, liquidity has been studied mostly in single-asset models, with little attention given to multi-asset liquidity, common factors, liquidity substitutes and so forth (Chordia et al. (2000) can be consulted for the relevant references on single-asset microstructure papers). Recently, however, a few empirical market microstructure papers have started to address this omission, among them Chordia et al. (2000), Hasbrouck and Seppi (2001) and Huberman and Halka (2001). As far as theoretical modeling of multi-asset liquidity is concerned, less work has been done, be it in market microstructure or otherwise. Fernando (2003) models "liquidity shocks" as non-informative additive shocks that affect investors' marginal valuations of risky assets. Investors are subject to deadweight transaction costs that benefit nobody. There are no intermediaries in Fernando's model, so liquidity is supplied directly by the investors with low marginal valuations. While "liquidity shocks" are specified, no definition or metric of "liquidity" is proposed. Fernando's main interests are the price effects of idiosyncratic versus systematic liquidity shocks as well as how liquidity shocks to one asset affect prices of other assets.

The paper is organized as follows. In the next section we introduce our definition of liquidity and outline some of its general properties. We then proceed to analyze it in detail in the segmented markets setup of Rahi and Zigrand (2007b,a). The results we need from these papers are summarized in Section 3. Section 4 elaborates on the role played by intermediaries in the provision of liquidity. In the next few sections we relate our liquidity measure to depth, bid-ask spreads, individual asset liquidity, volume, and welfare. In Section 9 we show how our setup can be used to study contagion. In Section 10 we allow intermediaries to introduce new securities and analyze the impact on liquidity. Section 11 is devoted to extensions of our main results. Proofs are collected in the Appendix.

2 Liquidity as Realized Gains from Trade

In this section we shall define what we mean by *market liquidity*. While there are many different meanings of liquidity in general, market liquidity is usually associated with the following attributes:

1. Depth
2. Volume of trade
3. Bid-ask spreads
4. Intermediation and transaction costs (e.g. brokerage fees and the degree of competition in market making)

These are meant to be proxies for the ease with which agents can execute desirable trades. We capture this notion of ease of trade directly by defining liquidity as "realized gains from trade." We argue that this metric captures the overall *economic meaning* of liquidity, as reflected in welfare. Insofar as the attributes listed above are not a good proxy for the liquidity measure, they may not be truly economically relevant.

At the most fundamental level, markets are more liquid the more diverse are the valuations of agents in the absence of trading. In the extreme case where all agents have the same no-trade valuations, there are no gains from trade to be realized—trading volume is zero and markets can be deemed to be completely illiquid. For example, a situation in which all agents want to be on the same side of a trade, so that these trades cannot be consummated, is often referred to as a "drying up of liquidity."

We formalize this idea in a two-period economy in which assets are traded at date 0 and pay off at date 1. Our measure of liquidity involves a comparison of state-price deflators. Given a collection of J assets with random payoffs $d := (d_1, \ldots, d_J)$ and prices $q := (q_1, \ldots, q_J)$, a random variable p is called a state-price deflator[1] if $q_j = E[d_j p]$ for every asset j, or more compactly, $q = E[dp]$.

Consider first the benchmark case of a frictionless economy with complete markets. Let p^i be the no-trade valuation of agent i, i.e. the state-price deflator at which the agent chooses not to trade. Let p^W be a Walrasian state-price deflator. Then we measure the gains from trade of agent i in the equilibrium under consideration by[2] $\nu^i E[(p^i - p^W)^2]$, where ν^i is a preference parameter.[3] The corresponding liquidity measure is

$$\mathcal{L} := \sum_i \nu^i E[(p^i - p^W)^2]. \tag{1}$$

Liquidity thus defined is a measure of the gains from trade realized in equilibrium. However, there is no sense in which it can reflect bid-ask spreads, transaction costs, or intermediation costs, as these are absent in a complete-markets economy with no frictions. Accordingly, we introduce two kinds of market imperfections: incomplete markets and market segmentation. Agents can trade only a limited number of assets, and different agents (or groups of agents) have access to different sets of assets. This provides a role for intermediaries to exploit price differentials across market segments and in the process to provide liquidity.

We formalize market segmentation as follows. Assets are traded in several locations or "exchanges." There are K such exchanges, with I^k investors on exchange k. We also use K and I^k to denote the set of exchanges and the set of investors on exchange k,[4] i.e. $K := \{1, \ldots, K\}$ and $I^k := \{1, \ldots, I^k\}$. There are J^k assets available to agents on k, with the random payoff of a typical asset j denoted by d^k_j.

[1] Other terms used in the literature for "state-price deflator" are "state-price density," "stochastic discount factor," and "pricing kernel."

[2] We restrict all random variables to lie in the linear space L^2 of square-integrable random variables.

[3] For simplicity, we restrict ourselves to a single-parameter family of preferences. The precise meaning of ν^i will depend on the way preferences are modeled. We do this in the next section.

[4] Following standard convention, we use the same symbol to denote a set and its cardinality.

Asset payoffs on exchange k can then be summarized by the random payoff vector $d^k := (d_1^k, \ldots, d_{J^k}^k)$.

In general, markets are incomplete on exchange k. The set of marketable payoffs on exchange k is $M^k := \{x : x = d^k \cdot \theta, \text{ for some portfolio } \theta \in \mathbb{R}^{J^k}\}$. Given the linear space M^k, a random variable x can be split into a marketable component x_{M^k} and a non-marketable component ϵ in such a way that the mean-square distance between x and x_{M^k} is minimal. This marketable component x_{M^k} is given by the least-squares regression of x on d^k. In other words, there is a unique decomposition $x = x_{M^k} + \epsilon$, with $E[\epsilon d^k] = 0$, and x_{M^k} the payoff of a portfolio of the assets d^k, such that $E[(x_{M^k} - x)^2]$ is minimal. If markets are incomplete on exchange k there are multiple state price deflators consistent with the same asset prices and payoffs on k. However, the *traded* state-price deflator p_{M^k} is unique (see Lemma 2.1 below).

In a segmented economy agents can trade among themselves within each segment, and they can also trade across segments via intermediaries. Regardless of how intermediation is modeled, there is a natural generalization of the liquidity measure (1). Consider an equilibrium of the intermediated economy in which a state-price deflator for exchange k is given by $\hat{p}^k$, $k \in K$. Except in an ideal world of perfect intermediation, the $\hat{p}^k$'s will typically be different across exchanges. Let $p^{k,i}$ be a no-trade state-price deflator of the i'th agent on exchange k. Then we define the liquidity metric for exchange k as

$$\mathcal{L}^k := \sum_{i \in I^k} \nu^{k,i} E[(p_{M^k}^{k,i} - \hat{p}_{M^k}^k)^2], \tag{2}$$

and the aggregate liquidity measure as

$$\mathcal{L} := \sum_{k \in K} \mathcal{L}^k. \tag{3}$$

In the complete-markets frictionless case, all payoffs are marketable, and $\hat{p}^k = p^W$ for all k, so that (3) reduces to (1).

A complete characterization of this liquidity measure, and an analysis of its relationship to attributes such as trading volume and bid-ask spreads, must await a full description of the model. At this stage we motivate and describe some of its general properties that do not depend on the particular way in which equilibrium prices (the $\hat{p}^k$'s) are determined.

The term $E[(p_{M^k}^{k,i} - \hat{p}_{M^k}^k)^2]$ in the definition of liquidity is the mean-square distance between agent (k,i)'s (traded) valuation $p_{M^k}^{k,i}$ and the equilibrium (traded) valuation of exchange k, $\hat{p}_{M^k}^k$. This has the interpretation of gains from trade reaped by agent (k,i) constrained by the assets available for trade on k (in particular, if there are no markets on k, these gains are zero). More generally, we can rely on the work of Chen and Knez (1995) on market integration to provide a characterization of mean-square distance between state-price deflators:

Lemma 2.1 *Given random variables p and p', and a marketed subspace M for some collection of assets, we have:*

1. $p_M = p'_M$ *if and only if* $E[dp] = E[dp']$, *for all payoffs* $d \in M$.

2.

$$E[(p_M - p'_M)^2] = \max_{d:E[(d_M)^2]=1} \left[E(dp_M) - E(dp'_M)\right]^2$$

i.e. $E[(p_M - p'_M)^2]$ *is the maximal squared pricing error induced by* p_M *and* p'_M *among payoffs* d *with* $E[(d_M)^2] = 1$.

3.

$$E[(p_M - p'_M)^2] = \max_{d \in M:E[(d)^2]=1} \left[E(dp) - E(dp')\right]^2$$

i.e. $E[(p_M - p'_M)^2]$ *is the maximal squared pricing error induced by* p *and* p' *among marketed payoffs* d *with* $E[(d)^2] = 1$.

The first statement says that two random variables are valid state-price deflators for a given collection of assets if and only if their marketed components are the same. Thus our liquidity measure does not depend on which state-price representation is chosen (i.e. $p^{k,i}$ could be any no-trade state-price deflator for agent (k,i) and $\hat{p}^k$ could be any equilibrium state-price deflator for exchange k). The last two statements characterize the mean-square distance between the traded state-price deflators p_M and p'_M as a bound on the difference in asset valuations induced by them. More precisely, it is the maximal squared pricing error using p and p' to price (normalized) payoffs in M, or alternatively it is the maximal squared pricing error using the traded state-price deflators themselves to price all (normalized) payoffs, whether marketed or not.

Liquidity in our setting is provided by both investors and intermediaries. We can isolate the first component as follows. Let p^k be an autarky state-price deflator for exchange k. In the absence of intermediaries, $\hat{p}^k = p^k$, so that liquidity on k is

$$\mathcal{L}^k\big|_{N=0} = \sum_{i \in I^k} \nu^{k,i} E[(p^{k,i}_{M^k} - p^k_{M^k})^2], \tag{4}$$

This is the liquidity generated from the batch auction on exchange k, without any intervention of the intermediaries. It reflects the realized gains from trade of investors on k from trading among themselves.

For much of this paper we will be studying the case in which intra-exchange liquidity, given by (4), is zero, so that all liquidity is intermediated. This is a special case of our setup in which all investors within an exchange have the same no-trade valuations, but where endowments, preferences and asset spans differ across exchanges. We call this economy a *clientele economy.* In a clientele economy, $p^{k,i} = p^k$, for all k. Then liquidity on exchange k is

$$\mathcal{L}^k = \nu^k E[(p^k_{M^k} - \hat{p}^k_{M^k})^2],$$

where $\nu^k := \sum_{i \in I^k} \nu^{k,i}$ is the aggregate preference parameter for exchange k. In the absence of intermediation $\hat{p}^k = p^k$ and $\mathcal{L}^k = 0$, for all k: markets are completely illiquid, as there are no liquidity providers. Liquidity on each exchange is also zero if the p^k's are all the same, so that there is no reason to trade across exchanges to begin with (in this case $\hat{p}^k$ must be equal to p^k for all k, as there are no profit opportunities for intermediaries).

It has been usual in the literature on liquidity, especially in applied work, to focus on depth and on spreads. While we will be more precise later on the relationship between our measure of liquidity and these proxies, a few general remarks are in order.

Analysis of depths and spreads in individual assets, as has been typical in the literature, suffers from the usual pitfalls of partial equilibrium analysis. For example, a particular asset may not be liquid, but substitutes may be liquid enough to make the overall market liquid. This calls for a *global* point of view that considers multiple assets traded in multiple markets. This is what we do in the present paper.

While depths will play a role in our measure of liquidity, we do not equate liquidity with depth. The liquidity aggregate $\mathcal{L}$ is intended to be a measure of *inside liquidity*. This is in contrast to what could be called *outside liquidity*, which is the ease of trade by agents coming from outside the Marshallian demand side, e.g. market orders submitted by noise traders. In other words, when we consider the cost related to the execution of an exogenous market order, we are thinking of liquidity as perceived by an outside trader whose utility function, endowments and so forth are not incorporated in the equilibrium demand function. Such outside liquidity is a direct unambiguous function of depths.

Spreads have been analyzed in the literature by picking a few assets and then arguing that the spread in these assets is representative of the economy as a whole. For instance, refer to the excellent monograph by Marston (1995) where the integration of various national financial markets is measured by the degree of closeness with which these markets price various money market and fixed-income securities. Our liquidity metric, on the other hand, leads naturally to a measure of spreads that is a function of the mean-square distances between the state-price deflators $\{\hat{p}^k\}_{k\in K}$. The advantage of such a measure is that it considers willingness to pay directly, rather than indirectly through proxies computed from a limited number of securities. In the latter procedure, a judgement must be made as to the most relevant assets or asset classes to compare. Furthermore, since identical assets, or more generally payoffs, may not exist on multiple exchanges, one would need to compare substitute assets. Both points raise a Pandora's box of judgmental issues which can be avoided entirely by using state prices instead. As shown in Lemma 2.1, the mean-square distance between the traded state-price deflators on two exchanges is equal to the bound on the squared pricing errors in using these state-price deflators to price any (normalized) payoff, whether marketed or not. In other words, it exactly represents what one is looking for when computing price differentials, and has the virtue of using and representing all available information.

That asset-by-asset depth or asset-by-asset bid-ask spreads may have no relationship with liquidity is easy to illustrate. For instance when all exchanges have the same autarky valuations, liquidity as defined is zero, no matter how deep the respective markets are. An outside market order may have barely any effect on prices, and yet markets are illiquid since there is no reason to trade within the model. This is reflected also in zero volume of trade across the exchanges.[5] Equally, it is easy to see that the level of mispricing, e.g. the size of bid-ask spreads, for individual securities need not have any relationship with the level of overall liquidity. Consider, for the sake of illustration, an asset with payoff d, $E[d] = 0$, that is traded on two exchanges, 1 and 2. The mispricing of this asset, given by $E[(\hat{p}^1 - \hat{p}^2)d]$, may be very low. For instance it is zero if the covariance between d and $\hat{p}^1 - \hat{p}^2$ is zero. Yet markets may be very illiquid, for instance if there are no intermediaries or if the potential gains from trade are insignificant. And the same applies to the converse: liquidity may be relatively high and yet bid-ask spreads for some asset may be large. In other words, the bid-ask spread for one particular asset may not necessarily provide a reliable indication as to the level of liquidity in the markets. All information impounded into the pricing relationships and gathered from the equilibrium actions of all agents needs to be taken into consideration, as is the case when using state prices.

In summary, market liquidity as we see it is a general snapshot spread, properly aggregated across all payoffs and all market segments. The apparent drawback of our definition is that it involves terms, such as autarky state-price deflators, which are hard to estimate. We shall show in the next few sections, however, that this criticism is not warranted as a number of observed variables can serve as a proxy for these unobservable terms.

3 Equilibrium

The definition of liquidity proposed in this paper does not crucially depend on any particular choice of timing, preferences or endowments, and is therefore of universal application. However, in order to derive closed form solutions and to relate liquidity to welfare, a modeling choice must be made. The major difficulty that needs to be overcome is the fact that strategic intermediaries play a game whose payoffs are functions of the outcomes of a general equilibrium.

The setup is as in Rahi and Zigrand (2007b,a). In this section we provide a brief synopsis of the model and the characterization of equilibrium. The reader is referred to the original papers for proofs and interpretive discussions.

Investor $i \in I^k$ on exchange $k \in K$ has endowments $(\omega_0^{k,i}, \omega^{k,i})$, where $\omega_0^{k,i} \in \mathbb{R}$ is his endowment at date 0, and $\omega^{k,i}$, a random variable, is his endowment at date 1. His preferences are given by quasilinear quadratic expected utility

$$U^{k,i}(x_0^{k,i}, x^{k,i}) = x_0^{k,i} + E\left[x^{k,i} - \frac{1}{2}\beta^{k,i}(x^{k,i})^2\right],$$

[5]If the market order from the outside liquidity trader does reflect gains from trade, then the initial model was not a complete characterization of the economy and liquidity in equilibrium is not zero.

where $x_0^{k,i} \in \mathbb{R}$ is consumption at date 0, and $x^{k,i}$ is a random variable representing consumption at date 1. The coefficient $\beta^{k,i}$ is positive. Investors are price-taking and can trade only on their own exchange. It will be useful later to characterize exchange k in terms of its aggregate preference parameter $\beta^k := [\sum_i (\beta^{k,i})^{-1}]^{-1}$, and its aggregate date 1 endowment $\omega^k := \sum_i \omega^{k,i}$. Similarly we define the corresponding parameters for the entire economy: $\beta := [\sum_k (\beta^k)^{-1}]^{-1}$ and $\omega := \sum_k \omega^k$. We assume that $1-\beta\omega \geq 0$, which says that the representative investor with aggregate preference parameter β is weakly nonsatiated at the aggregate endowment ω.

In addition to the price-taking investors, there are N arbitrageurs (with the set of arbitrageurs also denoted by N) who possess the trading technology which allows them to also trade *across* exchanges, or in other words, which allows them to act as intermediaries if they so wish. For simplicity, we assume that arbitrageurs only care about time zero consumption. They are assumed to be imperfectly competitive in our model, as they clearly are in actual financial markets. We also assume they have no endowments, so they can be interpreted as pure intermediaries.

We assume that all random variables (asset payoffs and endowments) have finite support. Then we can describe the uncertainty by a finite state space $S := \{1, \dots, S\}$.

The interaction between price-taking investors and strategic arbitrageurs involves a Nash equilibrium concept with a Walrasian fringe, pioneered by Gabszewicz and Vial (1972). Let $y^{k,n}$ be the supply of assets on exchange k by arbitrageur n, and $y^k := \sum_{n \in N} y^{k,n}$ the aggregate arbitrageur supply on exchange k. For given y^k, $q^k(y^k)$ is the market-clearing asset price vector on exchange k, with the asset demand of investor i on exchange k denoted by $\theta^{k,i}(q^k)$.

Definition 1 *Given an asset structure $\{d^k\}_{k \in K}$, a Cournot-Walras equilibrium (CWE) of the economy is an array of asset price functions, asset demand functions, and arbitrageur supplies, $\{q^k : \mathbb{R}^{J^k} \to \mathbb{R}^{J^k}, \theta^{k,i} : \mathbb{R}^{J^k} \to \mathbb{R}^{J^k}, y^{k,n} \in \mathbb{R}^{J^k}\}_{k \in K, i \in I^k, n \in N}$, such that*

1. *Investor optimization: For given q^k, $\theta^{k,i}(q^k)$ solves*

$$\max_{\theta^{k,i} \in \mathbb{R}^{J^k}} x_0^{k,i} + E\Big[x^{k,i} - \frac{\beta^{k,i}}{2}(x^{k,i})^2\Big]$$

subject to the budget constraints:

$$\begin{aligned} x_0^{k,i} &= \omega_0^{k,i} - q^k \cdot \theta^{k,i} \\ x^{k,i} &= \omega^{k,i} + d^k \cdot \theta^{k,i}. \end{aligned}$$

2. *Arbitrageur optimization: For given $\{q^k(y^k), \{y^{k,n'}\}_{n' \neq n}\}_{k \in K}$, $y^{k,n}$ solves*

$$\max_{y^{k,n} \in \mathbb{R}^{J^k}} \sum_{k \in K} y^{k,n} \cdot q^k\Big(y^{k,n} + \sum_{n' \neq n} y^{k,n'}\Big)$$

$$s.t. \quad \sum_{k \in K} d^k \cdot y^{k,n} \leq 0.$$

3. Market clearing: $\left\{q^k(y^k)\right\}_{k\in K}$ *solves*

$$\sum_{i\in I^k}\theta^{k,i}(q^k(y^k)) = y^k, \qquad \forall k \in K.$$

Let $p^{k,i} := 1 - \beta^{k,i}\omega^{k,i}$ and $p^k := 1 - \beta^k\omega^k$. This is consistent with our usage of $p^{k,i}$ and p^k in Section 2, as it can be shown that $p^{k,i}$ is a no-trade state-price deflator for agent (k,i) and p^k is an autarky state-price deflator for exchange k. Indeed, for given arbitrageur supply y^k,

$$q^k(y^k) = E\left[d^k[p^k - \beta^k(d^k\cdot y^k)]\right]. \tag{5}$$

Thus $p^k - \beta^k(d^k\cdot y^k)$ is a state-price deflator for exchange k. The autarky state-price deflator p^k is obtained by setting $y^k = 0$.

Proposition 3.1 (Cournot-Walras equilibrium: Rahi and Zigrand (2007b)**)** *There is a unique CWE.*[6]

1. *Equilibrium arbitrageur supplies are given by*

$$d^k\cdot y^{k,n} = \frac{1}{(1+N)\beta^k}\left(p^k_{M^k} - p^A_{M^k}\right), \qquad k\in K \tag{6}$$

   where $p^A \geq 0$ *is a state-price deflator for the arbitrageurs.*

2. *Equilibrium asset prices on exchange* k *are given by* $\hat{q}^k := E[d^k\hat{p}^k]$*, where*

$$\hat{p}^k := \frac{1}{1+N}p^k + \frac{N}{1+N}p^A. \tag{7}$$

   Thus $\hat{p}^k$ *is an equilibrium state-price deflator for exchange* k*.*

3. *Aggregate arbitrageur profits originating from exchange* k *are given by*

$$\Phi^k := \hat{q}^k\cdot y^k = \frac{N}{(1+N)^2\beta^k}\,E[(p^k_{M^k} - p^A_{M^k})^2]. \tag{8}$$

4. *The equilibrium demands of investors are given by*

$$d^k\cdot\theta^{k,i} = \frac{1}{\beta^{k,i}}(p^{k,i}_{M^k} - \hat{p}^k_{M^k}), \qquad i\in I^k, k\in K. \tag{9}$$

[6] We should point out some notational differences relative to Rahi and Zigrand (2007b). Here we denote equilibrium asset prices on exchange k by $\hat{q}^k$ instead of q^k; the latter notation is reserved for asset prices in autarky (see Section 6).

5. *The equilibrium utilities of investors are given by*

$$U^{k,i} = \bar{U}^{k,i} + \frac{1}{2}\beta^{k,i} E[(d^k \cdot \theta^{k,i})^2], \qquad i \in I^k, k \in K, \tag{10}$$

where $\bar{U}^{k,i}$ is a constant that does not depend on the asset structure or investor portfolios.

The random variable p^A is a state-price deflator for the arbitrageurs in the sense that it is a state-price deflator, i.e. $\hat{q}^k = E[d^k p^A]$ for all k, and moreover $p^A(s)$ is the arbitrageurs' marginal shadow value of consumption in state s (formally, $p^A(s)$ is the Lagrange multiplier attached to the arbitrageurs' no-default constraint in state s). Note that p^A can be chosen not to depend on N.

Given the centrality of the arbitrageur valuation p^A, it is important to provide an explicit characterization of it. To this end, we define a Walrasian equilibrium with restricted consumption as an equilibrium in which agents can trade any asset on a centralized exchange, facing a common state-price deflator p^{RC}, but agents on exchange k can consume claims in M^k only.[7] There are no arbitrageurs.

Proposition 3.2 (Arbitrageur valuations: Rahi and Zigrand (2007a) **)** *Arbitrageur valuations in the CWE coincide with valuations in the Walrasian equilibrium with restricted consumption, i.e. $p^A_{M^k} = p^{RC}_{M^k}$, for all k. Consequently $\lim_{N\to\infty} \hat{q}^k = E[d^k p^{RC}]$.*

Thus asset prices in the arbitraged economy converge to asset prices in the restricted-consumption Walrasian equilibrium, as the number of arbitrageurs goes to infinity (note, that this is an immediate consequence of (7) once it is established that $p^A_{M^k} = p^{RC}_{M^k}$).[8] It is in this sense that arbitrageurs serve to integrate markets.

We obtain a sharper characterization of p^A under some restrictions on the asset structure $\{d^k\}_{k\in K}$. Let p^* denote the complete-markets Walrasian state-price deflator of the entire integrated economy with no participation constraints. It can be shown that

$$p^* = \sum_{k\in K} \lambda^k p^k,$$

where

$$\lambda^k := \frac{\frac{1}{\beta^k}}{\sum_{j=1}^K \frac{1}{\beta^j}}, \qquad k \in K.$$

The state-price deflator p^* reflects the investors' economy-wide average willingness to pay, with the willingness to pay on each exchange weighted by its relative depth.

Now consider the following spanning condition:

[7] See Rahi and Zigrand (2007a) for a formal definition, and also for a discussion of the subtle difference between this notion of equilibrium and Walrasian equilibrium with restricted participation. In the latter, agents face a common state-price deflator, but agents on exchange k can trade claims in M^k only.

[8] The equilibrium allocation (for investors) in the arbitraged economy also converges to the restricted-consumption Walrasian equilibrium allocation.

(S) Either (a) $M^k = M$, $k \in K$, or (b) $p^k - p^* \in M^k$, $k \in K$.

Under **S**(a) we have an standard incomplete markets economy in which all investors can trade the same set of payoffs. **S**(b) is the condition that characterizes an equilibrium security design (see Rahi and Zigrand (2007b)). We have the following analogue of Proposition 3.2:

Proposition 3.3 (Arbitrageur valuations II: Rahi and Zigrand (2007b) **)** *Suppose condition* **S** *holds. Then, arbitrageur valuations in the CWE coincide with valuations in the complete-markets Walrasian equilibrium, i.e.* $p^A_{M^k} = p^*_{M^k}$, *for all* k. *Consequently* $\lim_{N\to\infty} \hat{q}^k = E[d^k p^*]$.

4 Intermediation and Liquidity

Now that we are armed with a model with a closed form solution of the unique equilibrium, we can explicitly characterize the properties of the liquidity measure defined in Section 2. In our model, the natural choice of the preference parameter $\nu^{k,i}$ is $1/\beta^{k,i}$. Then the liquidity measure for exchange k is

$$\mathcal{L}^k = \sum_{i \in I^k} \frac{1}{\beta^{k,i}} E[(p^{k,i}_{M^k} - \hat{p}^k_{M^k})^2], \tag{11}$$

with economy-wide liquidity

$$\mathcal{L} = \sum_{k \in K} \mathcal{L}^k. \tag{12}$$

For a clientele economy ($p^{k,i} = p^k$, all k), liquidity on exchange k is

$$\mathcal{L}^k = \frac{1}{\beta^k} E[(p^k_{M^k} - \hat{p}^k_{M^k})^2]. \tag{13}$$

We shall henceforth restrict ourselves to a clientele economy. The intuition for the general case where investors can trade among themselves on their own exchange and also across exchanges via intermediaries is very similar. The only difference is that, in general, liquidity has two components, one coming from the batch auctions on individual exchanges and the other one coming from intermediation, as explained in Section 2. Our primary focus in this paper is on intermediation. The relevant extensions of the main results to the general case are gathered in Section 11.

So how does intermediation create liquidity? Intermediation does not affect the spans $\{M^k\}_{k \in K}$, as there is no asset with a new dimension of spanning that becomes available due to pure intermediation.[9] What is achieved through intermediation is that the existing assets can be used more fruitfully. Intermediaries provide liquidity in the very direct sense of being the counterparties to trades made possible due to their

[9]The case where intermediaries can issue assets to optimally intermediate is studied in Section 10.

diverse customer base that reaches across various clienteles. Without intermediaries, no gains from trade can be reaped.

Thanks to intermediation investors can trade on better terms. Suppose, for example, there are two exchanges, 1 and 2, with the same asset structure. Suppose there is an asset with payoff d for which the autarky price on exchange 1, $q^1 := E[dp^1]$, is lower than the autarky price on exchange 2, $q^2 := E[dp^2]$. Investors on 1 want to short the asset, while investors on 2 want to go long. By Proposition 3.3, we can choose $p^A = p^*$, which is a convex combination of p^1 and p^2. Hence the arbitrageurs' valuation of this asset, $q^A := E[dp^A]$, lies between p^1 and p^2. In the intermediated equilibrium, q^1 is pushed up and q^2 is pulled down (due to (7), $\hat{p}^k$ is closer to p^A than is p^k, for both exchanges). Intermediaries allow investors on exchange 1 to sell on better terms, while investors on exchange 2 can buy on better terms, with the spread narrowing. The welfare of investors increases even though intermediaries take home some profits.

Notice that liquidity for clientele k is scaled by $1/\beta^k$. From (5) it is clear that β^k is the price impact of a unit of arbitrageur trading on exchange k: the state s value of the state-price deflator $p^k - \beta^k(d^k \cdot y^k)$ falls by β^k for a unit increase in arbitrageur supply of s-contingent consumption. Thus $1/\beta^k$ is the depth of exchange k.

The equilibrium arbitrageur supply, given by (6), is very intuitive. Assuming for the moment that markets are complete on all exchanges, an arbitrageur supplies state s consumption to those exchanges which value it more than he does ($p^k_s - p^A_s > 0$). How much he supplies to exchange k depends on the size of the mispricing $|p^k_s - p^A_s|$, on the depth $1/\beta^k$, with more consumption supplied the deeper the exchange, and finally on the degree of competition N. If markets are incomplete, however, the difference between state prices may not be marketable. The arbitrageur would then supply state-dependent consumption as close to $p^k - p^A$ as permissible by the available assets d^k. The closest such choice turns out to be the projection $(p^k - p^A)_{M^k} = p^k_{M^k} - p^A_{M^k}$. The greater the number of arbitrageurs competing for the given opportunities, the smaller is each arbitrageur's residual demand, and so the less each one supplies. In the limiting equilibrium, as N goes to infinity, arbitrageurs virtually disappear in that individual arbitrageur trades vanish, as does their aggregate consumption, $\sum_k \Phi^k$, and they perform the reallocative job of the Walrasian auctioneer at no cost to society (as formalized in Proposition 3.2).

Another way to see this is to compare realized and potential gains from trade. Since arbitrageur valuations are Walrasian (Proposition 3.2), we can define the potentially achievable or total gains from trade as

$$\overline{\mathcal{L}} := \sum_{k \in K} \overline{\mathcal{L}}^k, \tag{14}$$

where

$$\overline{\mathcal{L}}^k := \frac{1}{\beta^k} E[(p^k_{M^k} - p^A_{M^k})^2]. \tag{15}$$

$\overline{\mathcal{L}}$ measures the gains from trade that can be reaped if the economy moves from the autarky equilibrium to a perfectly intermediated, Walrasian, equilibrium, with the asset spans remaining unchanged. $\overline{\mathcal{L}}^k$ measures the total gains from trade between k and the rest of the economy. These gains ultimately arise from differences in preferences (e.g. risk aversion) and endowments. In that sense, one can interpret date zero as the time when investors learn about their preferences and endowments, i.e. about their idiosyncratic "liquidity shocks."

Proposition 4.1 (Competition and liquidity) *In a clientele economy,*

$$\mathcal{L}^k = \left(\frac{N}{1+N}\right)^2 \overline{\mathcal{L}}^k, \qquad k \in K. \tag{16}$$

In particular, $\mathcal{L}^k$ is strictly increasing in N, $\mathcal{L}^k = 0$ at $N = 0$, and $\lim_{N\to\infty} \mathcal{L}^k = \overline{\mathcal{L}}^k$. Consequently, aggregate liquidity $\mathcal{L}$ is increasing in N, $\mathcal{L} = 0$ at $N = 0$, and $\lim_{N\to\infty} \mathcal{L} = \overline{\mathcal{L}}$.

This result follows from the fact that $p^k - \hat{p}^k = \frac{N}{1+N}(p^k - p^A)$, due to (7). The expression (16) shows how our liquidity measure captures the general costs of trading due to the noncompetitive nature of the intermediation business. More competition improves upon the extent of gains from trade realized in the markets. In the limit as competition becomes perfect, all potential gains from trade are exploited.

One of the advantages of our setup is that it is straightforward to endogenize the number of intermediaries as a function of the cost of entry into the intermediation business. While there are a number of related concepts of entry, the following is simple and sensible. Suppose each arbitrageur must bear a fixed cost c in order to set up shop and intermediate across all markets. First we determine the number of arbitrageurs N', not necessarily a natural number, so that each one of the N' arbitrageurs makes a profit of 0 after having borne the fixed costs. Using (8), (14) and (15), N' solves

$$c = \frac{1}{N'} \sum_k \Phi^k(N') = \frac{\overline{\mathcal{L}}}{(1+N')^2}.$$

Second, this number is rounded down to the nearest natural number:

$$N = \mathrm{rd}\left(\sqrt{c^{-1}\overline{\mathcal{L}}} - 1\right),$$

where the operator "rd" rounds the real number in parenthesis down to the next natural number. In particular, arbitrageurs are allowed to make profits in equilibrium, but not enough to attract one further arbitrageur. We must have $c \le \overline{\mathcal{L}}/4$ in order for intermediation to arise. N increases as c falls, with $\lim_{c\to 0} N = \infty$.

The assumption of unrestricted but costly entry provides us with a simple proxy for liquidity:

Proposition 4.2 *In a clientele economy,*

$$\mathcal{L} = \left(\frac{N}{1+N}\right)^2 (1+N')^2 c, \qquad N = rd(N').$$

This is immediate from Proposition 4.1. If we ignore integer constraints, then

$$\mathcal{L} = cN^2. \tag{17}$$

With estimates of c and N, an estimate of liquidity is then simply the cost of entry times the square of the number of intermediaries, or equivalently the total cost borne by the intermediation sector times the number of intermediaries. Notice that even though depth is a crucial ingredient in liquidity, it appears only in as far as depth affects the endogenous number of intermediaries N. An added bonus is that N is a variable which can in principle be observed directly rather than having to be estimated.

5 Depth and Spreads

Depth, $1/\beta^k$, enters directly into the liquidity measure $\mathcal{L}^k$, as one would expect. It is constant, and in particular independent of arbitrage trades. This is a very convenient feature of our model, for it allows us to show the endogenous nature of liquidity, even though depth is constant.

While depth is constant, the supply of an asset on exchange k has a differential impact on the prices of other assets on k depending on the payoff structure d^k. From (5),

$$\frac{\partial q_j^k(y^k)}{\partial y_{j'}^k} = -\beta^k E[d_j^k d_{j'}^k]. \tag{18}$$

The price impact of one unit of trade in asset j' on exchange k affects those assets on exchange k most which are closer substitutes in the sense of having a higher noncentral comovement with j'. For normalized payoffs d, with $E[d^2]$, β^k measures the own-price effect.

Since arbitrageur supply is scaled by depth, there is a natural connection between depth and volume of trade. We will return to this in Section 7 where we discuss the relationship between volume and liquidity.

Turning now to spreads, we define $S^k := E[(\hat{p}_{M^k}^k - p_{M^k}^A)^2]$ as the generalized "bid-ask spread" on exchange k. It is the spread between the valuation on k and the average valuation in the whole economy as measured by p^A (which is also the Walrasian valuation p^{RC}). As the number of arbitrageurs grows without bound, $\hat{p}^k$ converges to p^A, so that the spread S^k converges to zero. There is a close relationship between spreads and liquidity as we have defined them:

Proposition 5.1 (Spreads and liquidity) *In a clientele economy,*

$$\mathcal{L}^k = \frac{1}{\beta^k} N^2 S^k. \tag{19}$$

If the spanning condition **S** *holds, then*

$$S^k = E\left[\left(\sum_{\ell\in K}\lambda^\ell(\hat{p}^k_{M^k} - \hat{p}^\ell_{M^k})\right)^2\right] \tag{20}$$

$$= \frac{1}{1+N}\cdot E\left[\left(\sum_{\ell\in K}\lambda^\ell(p^k_{M^k} - p^\ell_{M^k})\right)^2\right]. \tag{21}$$

Under **S** (for example if the same assets are traded on all exchanges), the spread S^k is the squared pricing error between k and the rest of the economy, with the pricing error relative to another exchange being weighted by its relative depth. This equilibrium spread is in fact the same as the autarky spread scaled by the number of intermediaries.

6 Individual Asset Liquidity

We have defined liquidity as the overall ease with which gains from trade can be exploited. In this section we deduce asset-by-asset liquidity measures from the aggregate measure. The main reason for doing so is to be able to contrast our theoretical results with the existing empirical literature.

Intuitively, the empirical findings of Chordia et al. (2000) that liquidity can be correlated between certain assets is not surprising from a theoretical point of view. The assets held or supplied by arbitrageurs all share the characteristic of being valuable to investors, and those assets will all see higher volumes and liquidity than the remaining assets. Assets that do not contribute towards the satisfaction of gains from trade will not see active trading. In other words, from an economic point of view, the commonality in liquidity is the contribution to a portfolio mimicking the gains from trade.

Recall that $\hat{q}^k = E[d^k\hat{p}^k]$ is the equilibrium asset price vector on exchange k. We denote the autarky asset price vector on k by $q^k := E[d^k p^k]$. We can then formally disaggregate liquidity $\mathcal{L}^k$ into the diverse contributions of the J^k assets on exchange k as follows:

Lemma 6.1 *In a clientele economy,*

$$\mathcal{L}^k = \frac{1}{\beta^k} b^k \cdot (q^k - \hat{q}^k),$$

where $b^k := \{b^k_j\}_{j\in J^k}$ *is the regression coefficient of the multiple regression of* $p^k - \hat{p}^k$ *on* d^k.

The coefficient b^k_j is the portion of the variation of $p^k - \hat{p}^k$ in the mimicking portfolio that is explained by asset j on exchange k. Accordingly we define the liquidity of this asset as

$$\mathcal{L}^k_j := \frac{1}{\beta^k} b^k_j (q^k_j - \hat{q}^k_j),$$

so that indeed

$$\mathcal{L}^k = \sum_{j=1}^{J^k} \mathcal{L}^k_j.$$

Thus the local liquidity of asset j on exchange k equals its depth times the usefulness of asset j in generating overall gains from trade on k, b^k_j, times the gains from trade directly reaped from trading j, $q^k_j - \hat{q}^k_j$, regardless of the indirect gains from trade reflected in the other assets. The term $\frac{1}{\beta^k} b^k_j$ is in fact equal to θ^k_j, the equilibrium holding of asset j by clientele k (see Rahi and Zigrand (2007b)). Liquidity of asset j on exchange k can then be interpreted as follows:

Proposition 6.1 (Individual asset liquidity) *In a clientele economy, liquidity of asset j on exchange k equals the amount of resources saved due to the more favorable equilibrium asset prices induced by intermediation:*

$$\mathcal{L}^k_j = \theta^k_j (q^k_j - \hat{q}^k_j).$$

Note that $\mathcal{L}^k_j$ is positive. The equilibrium holding θ^k_j is equal to the arbitrageur supply y^k_j. From (18) we can see that the own-price effect of arbitrageur supply is negative. For example, if $y^k_j > 0$, then $\hat{q}^k_j < q^k_j$.

Liquidity therefore has a purely pecuniary interpretation as the additional amount of time zero consumption investors can enjoy due to more efficient pricing. This benefit is larger the greater the degree of competition among intermediaries. The proposition suggests a tight relationship between volume and liquidity, which is the subject of the next section.

7 Liquidity and Volume

We define the *inter-exchange volume* originating from exchange k as

$$\tilde{\mathcal{V}}^k := E[(d^k \cdot y^k)^2].$$

Using (6),

$$\tilde{\mathcal{V}}^k = \left[\frac{N}{(1+N)\beta^k}\right]^2 E[(p^k_{M^k} - p^A_{M^k})^2]. \tag{22}$$

The following proposition justifies previous empirical work that uses the induced net volume of intermediated contingent consumption as a proxy for liquidity:

Proposition 7.1 (Volume and liquidity) *In a clientele economy, liquidity equals volume per unit of depth: $\mathcal{L}^k = \beta^k \tilde{\mathcal{V}}^k$.*

The result follows (15), (16) and (22). This relationship between volume and liquidity is what intuition would have suggested. For a given volume, more gains from trade are realized the closer state prices move towards Walrasian ones. State prices do not move very much in deep markets. Therefore volume needs to be large relative to depth to exploit and exhaust gains from trade, which are measured by liquidity. Of course, volume is itself increasing in depth, and the net effect of depth on liquidity is positive, indicating that the volume effect of depth dominates the direct depth effect.

One of the interesting consequences of Proposition 7.1 is that, under the null hypothesis that the model presented in this paper is an accurate description of actual markets, one can avoid measuring the multifaceted concept of liquidity directly and infer its value from observed volume and market depth. In this respect the overall equilibrium volume of state-contingent consumption implied by the volume of asset trade does have a fundamental meaning over and above its direct meaning as a yardstick of how busy markets are. Implicit in the number of transactions are the motivations that gave rise to those transactions as well as the microstructure considerations of asset spans and degree of competition in the intermediation sector. Depth is an important determinant of liquidity, not only because it influences the equilibrium volume of trade, but also because it scales volume to reflect liquidity.

8 Welfare

Equilibrium welfare of investors is given by (10). In a clientele economy we measure the welfare of clientele k by $U^k := \sum_{i \in I^k} U^{k,i}$ and economy-wide welfare by $U := \sum_{k \in K} U^k$. Using (9) and (10),

$$U^k = \bar{U}^k + \frac{1}{2}\mathcal{L}^k$$

and

$$U = \bar{U} + \frac{1}{2}\mathcal{L},$$

where $\bar{U}^k := \sum_{i \in I^k} \bar{U}^{k,i}$ and $\bar{U} := \sum_{k \in K} \bar{U}^k$. Similarly, from (8), (15) and (16), total arbitrageur profits originating from exchange k are

$$\begin{aligned}\Phi^k &= \frac{N}{(1+N)^2}\overline{\mathcal{L}}^k \\ &= \frac{1}{N}\mathcal{L}^k, \end{aligned} \tag{23}$$

so that aggregate economy-wide profits are

$$\sum_{k \in K} \Phi^k = \frac{1}{N}\mathcal{L}. \tag{24}$$

This leads us to the following result on the relationship between the various concepts:

Proposition 8.1 (Liquidity, welfare and volume) *In a clientele economy, the following concepts, local as well as global, are isomorphic: liquidity, investor welfare, profits, social welfare and volume.*

As we have argued in the introduction, we feel that any measure or metric of market liquidity would have to be tightly related to welfare and profits in order to be economically meaningful. If the proposed model of intermediation is correct, one can estimate welfare indirectly from the various depth parameters and from the volume of trade.

9 Transmission of Liquidity Shocks

We now turn to the study of how liquidity shocks are transmitted across the economy. Starting from an initial equilibrium, we perturb fundamentals on one of the exchanges and analyze the economy-wide repurcussions of this local shock. For simplicity, this is not a temporal shock that could have been anticipated. In this regard we follow most of the literature on contagion.

In order to simplify the analysis, we shall assume that the spanning condition **S** holds. Then we can choose $p^A = p^* = \sum_k \lambda^k p^k$ by Proposition 3.3. We shall also continue to restrict ourselves to a clientele economy.

We consider a local shock on exchange ℓ. There are a number of ways to model this shock. The following turns out to be analytically tractable. Consider a shock to the number of investors I^ℓ, while preserving the relative distribution of preferences and endowments on ℓ, $\{\beta^{\ell,i}, \omega^{\ell,i}\}_{i\in I^\ell}$. A negative population shock on exchange ℓ lowers its depth $1/\beta^\ell$ while keeping its autarky state-price deflator, $p^\ell = 1 - \beta^\ell \omega^\ell$, constant. Consequently p^ℓ plays a less prominent role in p^A, but without making the economy more risk averse as would have happened had we simply lowered the depth of exchange ℓ.

Let

$$\vartheta^{k\ell} := \frac{E[(p^k_{M^k} - p^A_{M^k})(p^\ell_{M^k} - p^A_{M^k})]}{E[(p^k_{M^k} - p^A_{M^k})^2]}.$$

Thus $\vartheta^{k\ell}$ is the regression coefficient of the (projected) mispricing on exchange ℓ, $p^\ell_{M^k} - p^A_{M^k}$, on the mispricing on exchange k, $p^k_{M^k} - p^A_{M^k}$; this measure of covariation is a noncentral "beta" in the language of the CAPM. Ignoring integer constraints on N, we have the following result:

Proposition 9.1 (Contagion) *Consider a clientele economy satisfying the spanning condition* **S**. *Then the effect on exchange k of a population shock on exchange ℓ is given by*

$$\frac{d\log\mathcal{L}^k}{d\log I^\ell} = \underbrace{\mathbf{1}_{k=\ell} - 2\lambda^\ell\vartheta^{k\ell}}_{\frac{d\log\mathcal{L}^k}{d\log I^\ell}\big|_N} + \frac{\mathcal{L}^\ell}{N\mathcal{L}}.$$

Effects can be split into two categories: direct effects for a given N, captured by the term $(\mathbf{1}_{k=\ell} - 2\lambda^\ell\vartheta^{k\ell})$, and indirect effects via entry or exit which are represented by the term $\mathcal{L}^\ell/(N\mathcal{L})$.

Consider first an exchange $k \neq \ell$, and suppose N is fixed. The effect on liquidity on k is $-2\lambda^\ell \vartheta^{k\ell}$. If $\vartheta^{k\ell} < 0$, exchanges k and ℓ are complements in the sense that arbitrageurs tend to buy on one when they are selling to the other, i.e. there is intermediated trade between the two exchanges. If exchange ℓ experiences a reduction in its investor base, and a consequent deterioration of its depth, these intermediated trades become less valuable and less plentiful in equilibrium, thus reducing liquidity on k.

With endogenous N, this effect is exacerbated: fewer investors and lower depth on ℓ lead to less trade and to lower liquidity, which in turn leads to lower profits and thereby to fewer intermediaries, which in turn affects liquidities adversely and so forth. It is this cascade of deteriorating liquidities that has received significant attention in the contagion literature. The net effect of this feedback loop is represented by the term $\mathcal{L}^\ell/(N\mathcal{L})$. The effect is more pronounced the larger is the role of exchange ℓ in generating trades, as measured by its relative size $\mathcal{L}^\ell/\mathcal{L}$, and the smaller the initial N. A smaller initial N means that the feedback loop of liquidity on N and again of N on liquidity etc. is stronger as each arbitrageur is more powerful and holds a larger portfolio the less competition there is.

So far we have assumed k and ℓ to be complements. On the other hand, if $\vartheta^{k\ell} > 0$, then valuations on exchanges k and ℓ are similar in the sense of being on average on the same side as the economy-wide valuation p^*. The two exchanges therefore compete for trades, and can be said to be substitutes. In this case, a shallower ℓ induces intermediaries to migrate to k, thereby increasing liquidity on k, for given N. The contagion effect operating through a lower N is however the same as in the case of complementary exchanges.

Finally, consider the effect of a population shock on exchange ℓ on its own liquidity. For fixed N, this effect is given by $(1 - 2\lambda^\ell)$. If λ^ℓ is small, this has the straightforward interpretation of the direct loss of liquidity due to the flight of investors. This is compounded by the consequent flight of intermediaries in the same way as for the rest of the economy. If λ^ℓ is non-negligible, however, there is a countervailing effect. Indeed, if $\lambda^\ell > 1/2$, $\mathcal{L}^\ell$ actually increases when the population on ℓ falls, for given N. This might at first appear odd, but the effect stems from the endogenous nature of Walrasian prices. Fewer investors on exchange ℓ lower the depth of exchange ℓ, and everything else constant, liquidity is lower. But the smaller size of this clientele also means that it will now play a less prominent role in the determination of the economy-wide valuation p^*. The valuation p^* will become more dissimilar from p^ℓ, thereby increasing the potential gains from trade between ℓ and the rest of the economy, stimulating intermediated trades and increasing liquidity on ℓ. If $\lambda^\ell > 1/2$, this effect is strong enough to compensate for the loss of depth, before accounting for the knock-on effect on the number of intermediaries.

Evidently, in an economy with many exchanges, loss of liquidity is more likely to go hand in hand with a decline of active investors. But there might be situations where a dominant exchange optimally limits or rations participants. There may be situations in which a lower number of investors can sustain a higher level of liquidity, or conversely where the arrival of more (identical) investors can hurt local liquidity. The converse implication is that liquidity can suffer on an exchange that experiences a rise in its investor population while substitute exchanges at the same time benefit from higher liquidity. These various examples illustrate that there is a clear liquidity externality in our economy that can go in either direction.

As an illustration of contagion, one such episode has been documented by Peek and Rosengren (1997). They study the liquidity shock emanating from Japan at the end of the 1980s and beginning of the 1990s. While Japan was a major financial power, it is nevertheless safe to assume that it did not constitute more than half the world's financial depth. Given that the flow of capital was from Japan to the US, Japan and the US were complements and on average assets were cheaper abroad than in Japan. The adverse shock to Japanese liquidity caused by the drop in I^{Japan} depresses (and effectively depressed at the time) stock prices in Japan,

$$\partial_{I^{\text{Japan}}} q^{\text{Japan}} = \frac{N}{1+N} \frac{\lambda^{\text{Japan}}}{I^{\text{Japan}}} (q^{\text{Japan}} - q^{\lambda,\text{Japan}})$$

The authors documented that the result of this Japanese liquidity shock was a sharp decline in Japanese investment in the US which in turn adversely affected liquidity in the US, an instance of contagion along the lines suggested by our model.

10 Security Design

In this section we allow the intermediaries to innovate and add assets to the ones already available for trade on the exchanges. We will see that the optimally innovated assets not only augment intermediary profits, but also allow a better exploitation of gains from trade, leading to higher liquidity, volume and welfare.

One might guess that any innovation would lead to more liquid markets. The reasoning might be as follows: since intermediaries can always choose not to trade the new assets, volumes, and therefore liquidity, cannot be lower than in the absence of innovation. The reality is more complicated though, since liquidity as defined here captures the extent to which markets allow the economy to move closer to the ideal Walrasian equilibrium for the given asset structure. Since an asset innovation perturbs the Walrasian equilibrium also (in particular the deflator p^A), it is not necessarily true that pricing at the new equilibrium is closer to the new Walrasian equilibrium than the old pricing was to the old Walrasian equilibrium. It turns out, however, that the aforementioned logic is correct if the innovations are optimal for arbitrageurs.

We have already seen, in Section 3, that there is a unique CWE for any given asset structure $\{d^k\}_{k\in K}$. We now allow each arbitrageur to add assets to each exchange before any trading takes place. This determines a new asset structure $\{d^k_{innov}\}_{k\in K}$. The payoffs of the arbitrageurs in this security design game are the profits they earn in the ensuing CWE.[10] Which asset(s) would arbitrageurs introduce at a Nash equilibrium of this game? Rahi and Zigrand (2007b) show that there is a unique asset added to each exchange (if not already present):

Proposition 10.1 (Optimal innovation: Rahi and Zigrand (2007b)**)**
For a given $\{d^k\}_{k \in K}$, the asset structure

$$\begin{array}{ccc} [d^k \quad (p^k - p^*)] & \text{if} & p^k - p^* \notin M^k; \\ d^k & \text{if} & p^k - p^* \in M^k; \end{array}$$

is

1. *a minimal optimal asset structure for arbitrageurs; and*

2. *a minimal Nash equilibrium of the security design game.*

The reader is referred to Rahi and Zigrand (2007b) for a proof and a detailed discussion of this result. The term "minimal" refers to the fact that there are other optimal (or equilibrium) configurations, but involving more assets—all of these configurations spans an asset structure that is minimal. If there is an innovation cost, howsoever small, the chosen structure would unambiguously be a minimal one.

Since arbitrageur profits are higher in the post-innovation economy (condition 1 of Proposition 10.1), so is liquidity due to the isomorphism between profits and liquidity (Proposition 8.1)):

Proposition 10.2 (Innovation and liquidity) *In a clientele economy, liquidity $\mathcal{L}$ increases when intermediaries can innovate assets.*

A clear distinction needs to be made between local and global liquidity. While liquidity overall improves with optimal innovation, even though the intermediaries act strategically, it is also shown in Rahi and Zigrand (2007b) that profits on any particular exchange may fall. Invoking the isomorphism between *local* profits and liquidity (Proposition 8.1), this means that innovation may hurt liquidity on some exchanges. The intuition goes as follows. If due to the innovation one of the exchanges sees decreased volume due to decreased usefulness of trade, then liquidity falls on that exchange. This occurs for instance if the exchange in question had an initial asset structure that permitted intermediaries to execute some crucial trades, say to borrow some state-contingent resources. When intermediaries can innovate optimally, they build such trades into the assets they innovate, thereby reducing the need to execute the trades on the exchange in question.

[10]Note that all arbitrageurs are able to trade the assets introduced by any one arbitrageur. Also, due to the symmetry of the CWE (Proposition 3.1), all arbitrageurs have the same equilibrium payoff.

11 Heterogeneous Investors

In this section we allow a diversity of investors within each exchange. Liquidity is then defined by (11) and (12). Most of our results continue to hold in this more general setup, but we only mention those extensions which add to our understanding of liquidity, in particular the relationship between liquidity on the one hand and volume and welfare on the other.

The volume of trade originating from agent (k, i) is

$$\mathcal{V}^{k,i} := E[(d^k \cdot \theta^{k,i})^2].$$

The *total volume* of trade on exchange k is $\mathcal{V}^k := \sum_{i \in I^k} \mathcal{V}^{k,i}$. This contrasts with the *inter-exchange volume* originating from exchange k defined above as $\tilde{\mathcal{V}}^k := E[(d^k \cdot y^k)^2]$. Evidently, $\tilde{\mathcal{V}}^k \leq \mathcal{V}^k$ since part of the trading volume on k arises from direct trades among the local investors.

Plugging the expression for $d^k \cdot \theta^{k,i}$, given by (9), into the definition of volume, we get

$$\mathcal{L}^k = \sum_{i \in I^k} \beta^{k,i} \mathcal{V}^{k,i}.$$

The intuition is the same as above. Since liquidity also considers the gains from trade realized from intra-exchange trade, liquidity is related to total volume, not just to the inter-exchange portion of volume as was the case in a clientele economy.

We now turn to the welfare properties of liquidity. If we give each investor (k, i) the same welfare weight we see that the results derived above in the clientele case again hold exactly:

$$U^k = \bar{U}^k + \frac{1}{2}\mathcal{L}^k$$

and

$$U = \bar{U} + \frac{1}{2}\mathcal{L}.$$

We now summarize these results:

Proposition 11.1 (Liquidity, welfare and volume: general case) *The following concepts are isomorphic:*

- *liquidity, investor welfare and total volume, local as well as global,*
- *local profits and local inter-exchange volume.*

Since arbitrageur profits on exchange k only depend on inter-exchange trade, the expression for Φ^k given by (23) is still valid here, i.e. $\Phi^k = \mathcal{L}^k / N$, where $\mathcal{L}^k$ is the liquidity measure for a clientele economy. Thus there is a wedge between the objectives of arbitrageurs and investors. Investors consider trades among themselves as valuable, whereas arbitrageurs do not. This will have obvious repercussions if intermediaries design securities. As shown in Rahi and Zigrand (2007b), while the designed securities are socially optimal in a clientele economy, this is no longer true in a more general economy.

12 Conclusion

In this paper we have attempted to provide an equilibrium notion of liquidity that encompasses the disparate attributes commonly attached to the word "liquidity": depth, ease of trade, volume, low transaction costs and so forth. The definition provided here is the extent to which gains from trade are realized in equilibrium. The more gains from trade are realized, the more liquid we call markets. Liquidity originates both from local batch auctions as well as from intermediation.

This simple definition has a number of pleasant characteristics. We show that the definition is welfare-grounded: liquidity equals social welfare. But it also provides a number of operational forms which can in principle be tested and estimated, and which coincide with some of the intuitive attributes of liquidity. For instance, liquidity on exchange k is shown to equal volume per unit of depth. Alternatively, overall liquidity is equal to the total cost of intermediation times the number of intermediaries. Liquidity can also be disaggregated into individual asset liquidity, which equals the amount of resources saved due to the more favorable asset prices induced in equilibrium by the intermediaries. Some of the standard liquidity measures, on the other hand, such as individual asset bid-ask spreads, have only a tenuous connection to welfare—they fail to take the bigger picture into account wherein investors can use substitute assets to satisfy their portfolio needs.

Having defined and characterized the relevant concepts, the paper studies the transmission of liquidity shocks originating in one sector of the economy. The signs of the "contagion" effects are shown to depend on the degree of substitutability of exchanges or on the degree of overpricing of the market portfolio, while the intensity is shown to depend on the number of intermediaries as well as the local depths.

Our analysis also does not require us to assume that assets are exogenously given. In fact, we show that if intermediaries innovate assets with the sole aim of augmenting their own private profits, the equilibrium innovation improves liquidity, and therefore welfare, at least in a clientele economy.

A Appendix

In the Appendix we adopt matrix notation that simplifies the proofs considerably. Rather than viewing asset payoffs on exchange k as random variables d^k, we stack them into an $S \times J^k$ matrix R^k. The j'th column of R^k corresponds to the j'th asset, listing its payoffs in each state $s \in S$. In this notation the set of marketable payoffs M^k is the column space of R^k.

Let π_s be the probability of state s, and denote by Π the $S \times S$ diagonal matrix with diagonal elements $(\pi_1, \dots, \pi_S)$. A state-price deflator for (q, R) is a vector $p \in \mathbb{R}^S$ such that $q = R^\top \Pi p$.[11] In other words, it is convenient to think of state-price deflators as vectors instead of random variables. Similarly, the expression $E[xy]$ can be written as $x^\top \Pi y$, where the random variables x and y are viewed as vectors in $\mathbb{R}^S$. In our finite-dimensional setting, the inner product space L^2 is the space $\mathbb{R}^S$ endowed with the inner product $\langle x, y \rangle_2 := x^\top \Pi y$. Then $x_{M^k} = P^k x$, where P^k is the orthogonal projection operator in L^2 onto M^k, given by the idempotent matrix

$$P^k := R^k ({R^k}^\top \Pi R^k)^{-1} {R^k}^\top \Pi. \tag{25}$$

An explicit derivation of P^k can be found in Rahi and Zigrand (2007b). P^k depends on R^k only through the span M^k. The notation $\|\cdot\|_2$ stands for the L^2-norm defined by $\|x\|_2 := (x^\top \Pi x)^{\frac{1}{2}}$, for $x \in \mathbb{R}^S$.

We now write p^k as $\mathbf{1} - \beta^k \omega^k$, where $\mathbf{1}$ is the $S \times 1$ vector each element of which is 1. Liquidity for exchange k, in the general case, is

$$\mathcal{L}^k = \sum_{k \in K} \frac{1}{\beta^{k,i}} \|P^k(p^{k,i} - \hat{p}^k\|_2^2.$$

For a clientele economy,

$$\mathcal{L}^k = \frac{1}{\beta^k} \|P^k(p^k - \hat{p}^k\|_2^2.$$

Proof of Lemma 2.1 Let the asset payoff matrix be R with marketed subspace M and corresponding projection matrix P.

The first statement says that $P(p - p') = 0$ if and only if $R^\top \Pi (p - p') = 0$. If $P(p-p') = 0$, then $R^\top \Pi P(p-p') = 0$. But $R^\top \Pi P = R^\top \Pi$, so that $R^\top \Pi (p-p') = 0$. Conversely, $R^\top \Pi (p - p') = 0$ implies that $(p - p') \in M^\perp$. Hence $P(p - p') = 0$.

As to the second statement, consider a payoff d, not necessarily in M. The mispricing of d using Pp versus Pp' is $m(d) := d^\top \Pi P(p - p')$. Since $\Pi P = P^\top \Pi P$, by the Cauchy-Schwartz inequality we have $m(d) \le \|Pd\|_2 \|P(p - p')\|_2$; equality occurs if and only if Pd and $P(p - p')$ are collinear. It follows that

$$\|P(p - p')\|_2 = \max_{d: \|Pd\|_2 = 1} d^\top \Pi P(p - p'). \tag{26}$$

[11] The symbol $\top$ denotes "transpose." We adopt the convention of taking all vectors to be column vectors by default, unless transposed.

For the last statement, consider a payoff $d \in M$. Then $d = R\theta$ for some $\theta \in \mathbb{R}^N$. Using again the fact that $R^\top \Pi P = R^\top \Pi$, we see that $m(d) = d^\top \Pi(p - p')$. Hence, (26) can be written as

$$\|P(p - p')\|_2 = \max_{d \in M : \|d\|_2 = 1} d^\top \Pi(p - p').$$

■

Proof of Proposition 5.1 The relationship (19) is immediate from (7) which implies that $p^k - \hat{p}^k = N(\hat{p}^k - p^A)$. If **S** holds, then we can choose $p^A = p^* = \sum \lambda^k p^k$ by Proposition 3.3. Using (7) once again, it is easy to check that $p^A = \sum_k \lambda^k \hat{p}^k$. This gives us the first expression for S^k, (20). Another consequence of (7) is $\hat{p}^k - \hat{p}^\ell = \frac{1}{1+N}(p^k - p^\ell)$. This gives us (21). ■

Proof of Lemma 6.1

$$\begin{aligned}
\mathcal{L}^k &= \frac{1}{\beta^k} \|P^k(p^k - \hat{p}^k)\|_2^2 \\
&= \frac{1}{\beta^k} (p^k - \hat{p}^k)^\top {P^k}^\top \Pi P^k (p^k - \hat{p}^k) \\
&= \frac{1}{\beta^k} \underbrace{(p^k - \hat{p}^k)^\top \Pi R^k ({R^k}^\top \Pi R^k)^{-1}}_{{b^k}^\top} \underbrace{{R^k}^\top \Pi (p^k - \hat{p}^k)}_{q^k - \hat{q}^k}.
\end{aligned}$$

■

Proof of Proposition 9.1 In order to calculate the effect of a proportional change in I^ℓ, $d \log I^\ell$, it is convenient to write the population of exchange ℓ as αI^ℓ, with corresponding depth α/β^ℓ, and compute derivatives with respect to α evaluated at $\alpha = 1$. Using (16), we can write $\mathcal{L}^k$ as a function of α, p^A and N:

$$\begin{aligned}
\mathcal{L}^k(\alpha, p^A(\alpha), N(\alpha)) &= \frac{\alpha^k}{\beta^k} \left(\frac{N}{1+N} \right)^2 \|P^k(p^k - p^A)\|_2^2 \\
&= \frac{\alpha^k}{\beta^k} \left(\frac{N}{1+N} \right)^2 (p^k - p^A)^\top \Pi P^k (p^k - p^A),
\end{aligned}$$

where $\alpha^k = \alpha$ for $k = \ell$, and $\alpha^k = 1$ for $k \neq \ell$. The total derivative of $\mathcal{L}^k$ with respect to α is

$$\frac{d\mathcal{L}^k}{d\alpha} = \underbrace{\frac{\partial \mathcal{L}^k}{\partial \alpha} + \frac{\partial \mathcal{L}^k}{\partial p^A} \cdot {p^A}'(\alpha)}_{\left.\frac{d\mathcal{L}^k}{d\alpha}\right|_N} + \frac{\partial \mathcal{L}^k}{\partial N} N'(\alpha). \tag{27}$$

Noting that

$$p^A(\alpha) = \frac{\frac{\alpha}{\beta^\ell} p^\ell + \sum_{k \neq \ell} \frac{1}{\beta^k} p^k}{\frac{\alpha}{\beta^\ell} + \sum_{k \neq \ell} \frac{1}{\beta^k}},$$

we have

$$\frac{\partial \mathcal{L}^k}{\partial \alpha} = \mathcal{L}^\ell \mathbf{1}_{k=\ell}$$
$$\frac{\partial \mathcal{L}^k}{\partial p^A} = -\frac{2}{\beta^k}\left(\frac{N}{1+N}\right)^2 \Pi P^k (p^k - p^A)$$
$$p^{A\prime}(\alpha) = \lambda^\ell (p^\ell - p^A) = \frac{\beta^k \lambda^k}{\beta^\ell}(p^\ell - p^A),$$

where all the derivatives are evaluated at $\alpha = 1$. Hence the effect on $\mathcal{L}^k$ for given N is

$$\left.\frac{d\mathcal{L}^k}{d\alpha}\right|_N = \mathcal{L}^\ell \mathbf{1}_{k=\ell} - \frac{2\lambda^k}{\beta^\ell}\left(\frac{N}{1+N}\right)^2 \varphi^{k\ell}, \tag{28}$$

where

$$\varphi^{k\ell} := (p^k - p^A)^\top \Pi P^k (p^\ell - p^A).$$

We now solve for $N'(\alpha)$. With free entry, $\mathcal{L} = cN^2$ (equation (17)). Therefore, the function $N(\alpha)$ is defined by the identity

$$\sum_k \mathcal{L}^k(\alpha, p^A(\alpha), N(\alpha)) - c[N(\alpha)]^2 \equiv 0.$$

Implicit differentiation gives us

$$N'(\alpha) = \frac{\sum_k \left.\frac{d\mathcal{L}^k}{d\alpha}\right|_N}{2cN - \sum_k \frac{\partial \mathcal{L}^k}{\partial N}}.$$

Now

$$\frac{\partial \mathcal{L}^k}{\partial N} = \frac{2}{N(1+N)}\mathcal{L}^k, \tag{29}$$

so that

$$\sum_k \frac{\partial \mathcal{L}^k}{\partial N} = \frac{2}{N(1+N)}\mathcal{L} = \frac{2cN}{1+N},$$

where we have once again used the result that $\mathcal{L} = cN^2$.

Under the spanning condition $\mathbf{S}$, either $P^k = P$ or $P^k(p^k - p^A) = p^k - p^A$. In both cases $\sum_k \lambda^k \varphi^{k\ell} = 0$, since $p^A = p^*$. Therefore, from (28),

$$\sum_k \left.\frac{d\mathcal{L}^k}{dI^\ell}\right|_N = \mathcal{L}^\ell.$$

Altogether, this yields

$$N'(\alpha) = \frac{(1+N)\mathcal{L}^\ell}{2cN^2} = \frac{(1+N)\mathcal{L}^\ell}{2\mathcal{L}}. \tag{30}$$

Substituting (28), (29) and (30) into (27) gives us

$$\frac{d\mathcal{L}^k}{d\alpha} = \mathcal{L}^\ell \mathbf{1}_{k=\ell} - \frac{2\lambda^k}{\beta^\ell}\left(\frac{N}{1+N}\right)^2 \varphi^{k\ell} + \frac{\mathcal{L}^k\mathcal{L}^\ell}{N\mathcal{L}}.$$

Dividing through by $\mathcal{L}^k$ we get the desired result (note that $\vartheta^{k\ell} = \frac{\varphi^{k\ell}}{\varphi^{k,k}}$). ■

References

Allen, F. and Gale, D. (2000). Financial contagion. *Journal of Political Economy*, 108(1):1–33.

Chen, Z. and Knez, P. J. (1995). Measurement of market integration and arbitrage. *Review of Financial Studies*, 8(2):287–325.

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

Dewatripont, M. and Tirole, J. (1993). *The Prudential Regulation of Banks*. MIT Press.

Diamond, D. and Verecchia, R. (1981). Information aggregation in a noisy rational expectations economy. *Journal of Financial Economics*, 9:221–235.

Fernando, C. S. (2003). Commonality in liquidity: Transmission of liquidity shocks across investors and securities. *Journal of Financial Intermediation*, 12:233–254.

Freixas, X., Parigi, B., and Rochet, J.-C. (2000). Systemic risk, interbank relations and liquidity provision by the central bank. *Journal of Money Credit and Banking*, 32(3):611–638.

Gabszewicz, J. J. and Vial, J.-P. (1972). Oligopoly 'a la Cournot' in a general equilibrium analysis. *Journal of Economic Theory*, 4(3):381–400.

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

Gromb, D. and Vayanos, D. (2007). Financially constrained arbitrage and cross-market contagion. Mimeo, London School of Economics.

Grossman, S. and Miller, M. (1988). Liquidity and market structure. *Journal of Finance*, 43:617–633.

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

Hasbrouck, J. and Seppi, D. (2001). Common factors in prices, order flows and liquidity. *Journal of Financial Economics*, 59:383–411.

Hodrick, L. S. and Moulton, P. (2003). Liquidity. Mimeo, Columbia University.

Huberman, G. and Halka, D. (2001). Systematic liquidity. *Journal of Financial Research*, 24:161–178.

Kyle, A. (1985). Continuous auctions and insider trading. *Econometrica*, 53:1315–1336.

Marston, R. (1995). *International Financial Integration: A Study of Interest Differentials Between the Major Industrial Countries*. Cambridge University Press.

O'Hara, M. (1995). *Market Microstructure Theory*. Blackwell.

Peek, J. and Rosengren, E. S. (1997). The international transmission of financial shocks: The case of Japan. *The American Economic Review*, 87(4):495–505.

Rahi, R. and Zigrand, J.-P. (2007a). Cournotian foundations for competitive equilibria with restricted consumption or restricted participation. Mimeo, London School of Economics.

Rahi, R. and Zigrand, J.-P. (2007b). Strategic financial innovation in segmented markets. *Review of Financial Studies*. Forthcoming.