A Fully-Rational Liquidity-Based Theory of IPO Underpricing and Underperformance

Matthew Pritsker

A Fully-Rational Liquidity-Based Theory of IPO Underpricing and Underperformance

Matthew Pritsker[*]

First version: September 9, 2004

This version: February 21, 2006

Abstract

I present a fully-rational symmetric-information model of an IPO, and a dynamic imperfectly competitive model of trading in the IPO aftermarket. The model helps to explain IPO underpricing, underperformance, and why share allocations favor large institutional investors. In the model, underwriters need to sell a fixed number of shares at the IPO or in the aftermarket. To maximize revenue and avoid selling into the aftermarket where they can be exploited by large investors, underwriters distort share allocations towards investors with market power, and set the IPO offer price below the aftermarket trading price. Large investors who receive IPO share allocations sell them slowly afterwards to reduce their trade's price-impact. This curtails the shares that are available to small price-taking investors, causing them to bid up prices and bid down returns. In some simulations, the distorted share allocations and slow unwinding behavior generate post-IPO return underperformance that persists for several years.

---

[*]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 at (202) 452-3534, or Fax: (202) 452-3819, or by email at mpritsker@frb.gov.

1 Introduction

Two of the important functions of a financial system are to facilitate risk sharing among investors and capital formation by firms. The initial public offering (IPO) process performs both functions by allowing the initial owners of a firm to raise capital by transferring and sharing some of the firm's risk with the wider investing public. If the IPO process was fully efficient, an IPO should maximize the issuer's proceeds, the investors who most value the shares should receive them, and in the absence of news or private information, there should be little trade after the shares are allocated. Additionally, the fact that a stock is a new issue should not influence its risk-adjusted expected returns in aftermarket trading.

Relative to this benchmark, U.S. IPOs appear to be highly inefficient: post-IPO share trading is initially very heavy[1], and the allocation price of U.S. IPOs is on average nearly 19 percent below the closing price on the first day of trade [Ritter and Welch (2002)]. This underpricing is an apparent loss to issuers who would prefer to have sold at the higher price in the IPO aftermarket. The IPO process has other inefficiencies: allocations tends to favor institutional investors[2] and, after the first trading day, the returns of new issues underperform on a market and characteristic adjusted basis for a period of time as long as three years [Loughran and Ritter (1991), Ritter and Welch (2002)].[3]

This paper presents a fully rational, symmetric information, theoretical model of IPO share allocation and price-setting, and of trading in the IPO aftermarket. The paper is built around the idea that trading conditions in the aftermarket may simultaneously explain underpricing, underperformance, and why allocations favor institutional investors. The model of the aftermarket is imperfectly competitive in the sense that there are some "large" investors who have market power, that is their trades move prices and they account for this when trading. The IPO is modeled as a bargaining game between the underwriter and the aftermarket investors: The underwriter must sell a fixed number of shares at the IPO or shortly afterwards in aftermarket trading. To do so, he sets a uniform IPO offer price and offers take it or leave it share allocations to the investors. Any shares that go unallocated are sold by the underwriter in aftermarket trading that follows the IPO. Large investors' market

---

[1] Ellis, Michaely, and O'Hara's (2000) study of NASDAQ IPO's, found that turnover on the first trading day is equal to 1/3rd of the turnover that a NASDAQ stock experiences over a year.

[2] Institutional investors receive more favorable allocations than retail investors in the most underpriced issues.

[3] There is an ongoing debate within the empirical literature concerning whether IPO underperformance exists and whether it is statistically significant. For a discussion of these issues see Viswanathan and Wei (2004), de Jong and Dahlquist (2003), Schultz (2003), Loughran and Ritter (2000), and Brav, Geczy, and Gompers (2000).

power in the aftermarket gives them bargaining power at the IPO because they can turn down their share allocation and force the underwriter to sell into the aftermarket, where large investors can influence (and lower) the price. To avoid this outcome, the underwriter distorts IPO asset allocations towards investors with market power, and gives them a favorable IPO offer price.[4]

From the initial asset allocations at the IPO, investors trade towards efficient asset allocations along an equilibrium transition path. Because large investors' trades move prices, the market is not perfectly liquid from their perspective, and this illiquidity influences returns and asset prices. In particular, when initial asset allocations favor large investors, along the equilibrium transition path large investors sell slowly through time to minimize the price impact of their trades. This restricts the supply of shares that is available to price-taking small investors. As a result small investors bid up the new issue's price, and because they don't expect to acquire many assets in the short-term, they bid down its expected returns, causing return underperformance. The magnitude and duration of the underperformance depends on the severity of illiquidity in the aftermarket, and on how the assets are allocated at the IPO.

There is a voluminous related literature on IPO underpricing, underperformance, and share allocations at the IPO.[5] Most of the theoretical models in this literature explain only one or two of these phenomena. Only behavioral models can explain all three.[6] Relative to this literature, this paper makes three contributions. First, to the best of my knowledge, it presents the only fully-rational theoretical model to date that can explain all three phenonomena. Second, the model shows that these phenomena are related to the competitiveness of the aftermarket, which in turn can be related to the distribution of the size of the investors that participate in the IPO or trade in the aftermarket (as measured by their wealth, or assets under management). Therefore, summary statistics of the distribution of investor size constitute a new set of state variables that can be used to test theories of IPOs, and to distinguish the predictions of this model from the related literature. Finally, this

---

[4]In related work, Hoberg (2004) presents a model in which underpricing results from imperfect competition in securities underwriting.

[5] In the dominant strand of the underpricing literature, underpricing occurs because information-asymmetries cause adverse selection in allocating shares [Rock (1986)], or, because underpricing is needed to entice informed investors to reveal information to the underwriters, as in the bookbuilding models that starts with Benveniste and Spindt (1989). Other theories of underpricing are based on agency problems in which the underwriters objective function departs from that of the issuer (Boehmer and Fishe, 2000), (Bias et. al. 2002), (Loughran and Ritter, 2004). For extensive reviews of this literature, see Ritter and Welch (2002), and Ljungvist (2004).

[6]Ritter and Welch (2002) claim there are no rational theoretical models that explain underperformance. Ljunqvist, Nanda, and Singh (2003) present a behavioral model that generates all three phenomena.

paper helps to fill a gap in the small literature on IPOs and liquidity. Within that literature, Ellul and Pagano (2003) present a theoretical model in which markets are illiquid, but competitive. Within their framework, they show that underpricing is required to compensate IPO participants for aftermarket illiquidity; they also find empirical support for the theory because aftermarket illiquidity has a large and positive correlation with more underpricing.[7] Extending Ellul and Pagano's logic suggests IPOs should also earn a positive liquidity premium in aftermarket trading, but this is belied by the fact that IPO's underperform, and suggests that the Ellul and Pagano model cannot on its own explain both underpricing and underperformance. By contrast, this paper shows that a liquidity-based model can explain underpricing and underperformance provided one drops the perfect competition assumption in Ellul and Pagano and replaces it with the imperfect competition assumption that is used here.[8]

The rest of the paper proceeds in six parts. Section 2 provides a model overview; section 3 provides details on aftermarket trading; section 4 describes the process for share allocation and price setting at the IPO. Section 5 studies underpricing and underperformance using simulations; section 6 discusses the empirical implications of the model and provides a brief review of the most closely related empirical literature; a final section concludes.

2 Model Overview

The basic model is a stylized IPO in which a firm that wishes to raise capital by selling X^{IPO} shares of stock enlists a single underwriting firm to market the issue. To abstract from agency issues, the underwriter is assumed to act on behalf of the issuer. The underwriter sells the issue to an investor base that consists of M risk-averse investors who participate in the IPO and trade in the aftermarket. Investor 1 represents a continuum of small investors who each take prices as given. Investors 2 through M are large investors whose desired aftermarket trades are large enough to move asset prices. Because of differences in their size, the small investors can be viewed as representing the demands of retail investors, while

---

[7]In the theoretical model of Booth and Chua (1996), underpricing is used to increase the base of investors that are interested in the new issue and increase aftermarket liquidity. In Westerfield (2003) underpricing is also used to influence the composition of the investor base. If Booth and Chua are correct, then underpricing should be negatively correlated with illiquidity, but this runs counter to the empirical results in Ellul and Pagano, and hence casts doubt on the empirical relevance of Booth and Chua's model.

[8]Other differences with Ellul and Pagano are that their model has information asymmetry while the present model does not, and in their model, the aftermarket only lasts for one period whereas in the present the model the aftermarket trading is fully dynamic and modelled over thousands of periods.

the large investors represent the demands of institutional investors. The process for setting the IPO offer price and share allocations is modeled as a two-stage game. In the first stage, the underwriter assesses the demand for the new issue by learning about the characteristics of the investor base, and about aftermarket trading conditions. Based on his information, the underwriter sets a uniform IPO offer price and offers take-it or leave-it share allocations to the investors.[9] In the second stage, investors decide whether to accept their allocations. Any shares that are turned down by investors at the IPO are sold by the underwriter in the aftermarket.

In much of the theoretical IPO literature, the primary explanation for underpricing is that it represents equilibrium compensation for various types of information asymmetries in the IPO process.[10] To establish that the channels for underpricing and underperformance in the present model do not rely on informational differences, asymmetric information is ruled out by assumption. More specifically, I assume that information on investors' risk preferences, asset holdings, all knowledge of asset values, and the entire model of aftermarket trading is publicly available at all points of time and is common knowledge. The next section formally models the IPO aftermarket; and the following section models the share allocation and price-setting process at the IPO.

3   The IPO aftermarket

The framework for aftermarket trading is a partial equilibrium extension of Pritsker's (2004) model of imperfect competition in asset markets.[11] Investors in the economy hold diversified portfolios, but also specialize in trading the assets that belong to a particular market-segment or industry-group. Although there are many market segments, most of the analysis focuses on the assets within a particular segment. For informational or other reasons, the M investors in the model are the only investors in the economy that trade and hold the assets within this segment.[12] Hereafter, these assets are referred to as segment-assets. All investors in the economy trade a riskless asset with gross return r per period that is in perfectly elastic

---

[9]The first stage share allocation process resembles IPO bookbuilding: in both processes the underwriter collects information about market demand, and then allocates shares and sets an offer price based on the information that he collects.

[10]See footnote 5.

[11]Closely related models of imperfect competition in asset markets include Urosevic (2002a & b), DeMarzo and Urosevic (2000), and Vayanos (2001).

[12]The results would be very similar if some investors hold but do not trade the segment assets.

supply, and a broadly-held index that is in zero net supply.[13] The index proxies for systematic risk and can be thought of as a futures market. Because the index is broadly held, the M investors collective trades do not affect the price of the index.

Risky Assets

The prices of the segment and index assets are $P^1(t)$ and $P^2(t)$; and have stacked price vector $P(t)$. The segment-assets are in fixed supply X^1. The assets pay i.i.d. dividends $D^1(t)$, $D^2(t)$, represented by stacked vector $D(t)$, that has distribution:

$$D(t) \sim \text{i.i.d.} \quad \mathcal{N}(\bar{D}, \Omega), \text{ where } \quad \Omega = \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}. \tag{1}$$

Because prices can be negative, the assets' excess returns are best expressed on a per share basis, and are denoted:

$$Z(t) = P(t) + D(t) - rP(t-1) \tag{2}$$

For simplicity $P^2(.)$ is fixed, which implies $Z^2(t) \sim \text{i.i.d.} \ \mathcal{N}(\bar{Z}^2, \Omega_{22})$. To solve investors' portfolio problem, it is useful to decompose the segment assets return into a component that is perfectly correlated with the index and into a residual return $e(t)$:

$$Z^1(t) = \beta_{12}Z^2(t) + e(t) \tag{3}$$

where $\beta_{12} = \frac{\text{Cov}[Z^1(t), Z^2(t)]}{\text{Var}[Z^2(t)]} = \Omega_{21}\Omega_{22}^{-1}$. Investors' portfolio problem can be represented as choosing their exposure to the index risk, and to the residual risk of the segment assets. Because a limited number of investors hold the residual risk, it is not diversifiable. Therefore, the expected return for holding the residual risk will not necessarily be equal to 0. The variance of $e(t)$ is denoted Ω_e; in equilibrium it turns out to be constant through time and is given by:

$$\Omega_e = \Omega_{11} - \Omega_{21}\Omega_{22}^{-1}\Omega_{12}.$$

---

[13]These assumptions are loosely based on Merton (1987).

Investors

Each investor m chooses risky asset holdings $Q_m(t)$ and consumption $C_m(t)$ to maximize his discounted expected Constant Absolute Risk Aversion (CARA) utility of consumption:

$$U_m(C_m(1), ...C_m(\infty)) = \sum_{t=1}^{\infty} -\delta^t e^{-A_m C_m(t)}, \tag{4}$$

with discount factor δ and absolute risk aversion A_m, subject to the standard intertemporal budget constraints:

$$W_m(t) = Q_m(t)'Z(t) + r[W_m(t-1) - C_m(t-1)] \quad t = 1, \ldots T. \tag{5}$$

An investor's liquid wealth, $W_{ml}(t)$, is that part of her wealth that can be liquidated at current prices. Because large investors' trades have price impact, their wealth and liquid wealth differ; and it is their liquid wealth that appears as an argument in their value functions. Therefore it is useful to express their budget constraints in terms of their beginning of period liquid wealth:

$$\begin{aligned}
W_{ml}(t) = {} & Q_m^1(t)'D^1(t) + Q_m^2(t)'Z^2(t) \\
& + r\left[W_{ml}(t-1) - \Delta Q_m^1(t-1)'P^1(t-1) - C_m(t-1)\right] \quad t = 1, \ldots T.
\end{aligned} \tag{6}$$

Below, I show that investors who differ in their risk-aversion have different trading styles; this affects other investors ability to acquire illiquid assets. As a result, which investors hold the segment assets is a critical argument of investors value functions. It is summarized by $Q^1(t)$ $(= \text{vech}[Q_1^1(t)', \ldots, Q_M^1(t)']')$, the vector of all investors segment-asset holdings.

Trading Dynamics

Segment assets are traded over a total of T time periods. The sequence of actions is as follows: At the beginning of each period $t \leq T$, investors enter the period with risky-asset holdings $Q_m(t), m = 1, \ldots M$; they then receive dividends and choose their risky asset trades $\Delta Q_m(t)$. These trades determine risky asset prices $P(t)$. Investors then choose consumption $C_m(t)$; the period ends; and the same sequence is repeated through time T. After period T, investors continue to consume and trade other assets, but trading of segment assets ends.

The final period of trade in the segment-assets facilitates solution of the model via backwards induction.

The process of trade for the segment-assets is modeled as a dynamic Cournot-Stackelberg game of full information. In each period $t \leq T$, each small investor optimally computes his demand for the segment assets conditional on the state-variable $(Q^1(t), t)$. Inverting the aggregated demands of the small investors defines a linear schedule of prices at which they are willing to absorb all possible quantities of the large investors trades for the segment assets:

$$P^1(.,t) = \frac{1}{r} \left(\beta_0(t) - \beta_{Q^1}(t)Q^1(t) - \sum_{m=2}^{M} \beta_m(t)\Delta Q_m^1(t) \right). \tag{7}$$

The matrix $\beta_m(t)$ measures the impact of large investor m's segment asset trades on the price of the segment assets.[14] Large investors account for this price impact when choosing their trades. Given the price schedule in equation (7), large investors choose their time t risky asset trades and consumption choices to solve the maximization problem:

$$\max_{\substack{C_m(t), \\ \Delta Q_m(t)}} \quad -e^{-A_m C_m(t)} + \delta \, \mathrm{E}_t \, V_m(W_{ml}(t+1), Q^1(t) + \Delta Q^1(t), t+1) \tag{8}$$

subject to the budget constraint:

$$W_{ml}(t+1) = Q_m^1(t+1)'D^1(t+1) + Q_m^2(t+1)'Z^2(t+1) + r \left[W_{ml}(t) - \Delta Q_m^1(t)'P^1(.,t) - C_m(t) \right]. \tag{9}$$

Large investors' equilibrium trades are found by solving for trades that are best-responses to each other under the Cournot-Stackelberg assumption that each large investor chooses his own trades while taking the price schedule and the trades of the other large investors as given. The resulting trades within a period are a Cournot-Nash equilibrium. The model of trade is solved by backwards induction from period T, and the resulting equilibrium is subgame perfect. Additional details on model solution and investors' value functions are contained in the appendix.

Intuition for the main results on asset prices and trades comes from examining the first

---

[14]It is a matrix because his trades in one asset may affect the prices of other segment-assets.

order condition for large investor m's optimal trade vector for the segment assets:

$$\frac{\partial \, \mathrm{E} \, V_m}{\partial \Delta Q_m^1} = \left[\mathrm{E} \, V_{m,W}(.)(D_1(t) - rP^1) + \mathrm{E} \, V_{m,Q_m^1}(.) \right] - \left[\mathrm{E} \, V_{m,W}(.)r\Delta Q_m^{1\,\prime} P_m^1(.) \right] = 0, \quad (10)$$

where $P_m^1(.) = \partial P^1(.)/\partial \Delta Q_m$.

The first term in braces on the right hand side measures the expected net benefit from borrowing money to buy a bit more of a segment asset when trades have no price impact. This is the only term of the first order condition that should be present in a competitive setting. Therefore, in a competitive setting all investors would immediately trade to a point where the first term is equal to 0. When there is imperfect competition, it will never be optimal for a large investor to trade immediately to such a point because the second term in his first order condition would be non-zero, implying his trades had too much price impact. Instead, to reduce their price impact, large investors trade in a way that allows their positions to converge towards a point at which further trade is no longer optimal. Therefore, if a large investor believes an asset is overvalued, or provides too low a return, he will not liquidate its holdings immediately, but will instead sell the asset slowly over time. Because he sells slowly, prices will also adjust slowly, making it possible for low returns to persist in equilibrium, as will be discussed further below.

3.1   Asset Pricing

The intuition in the previous section and the results below both show that when markets are illiquid, asset returns will satisfy one equilibrium relationship when asset holdings are at their long-run equilibrium, and satisfy a different relationship during the transition to the long run. To illustrate the exact pricing relationship, additional notation is required. More specifically, let Q_m^{1W} represent investor m's long-run holdings of the segment-assets; these holdings are also the same as if all investors were price-takers, and are also associated with the efficient sharing of risks. When risk sharing is efficient, each investor holds assets in proportion to his risk-tolerance as a fraction of the sum total of all investors' risk tolerances:

$$Q_m^{1W} = \frac{(1/A_m)X^1}{\sum_{m=1}^{M}(1/A_m)} \quad m = 1, \ldots M \quad (11)$$

The following proposition shows that when markets are illiquid, the deviation from investors' efficient asset holdings behaves like a priced factor when computing one-period re-

turns:

Proposition 1 *The segment-assets equilibrium excess expected returns satisfy a linear factor model that provides a reward for assets covariances with the liquid index, for their covariances with segment-specific risk, and for covariances with deviations of large investors asset holdings from those associated with efficient sharing of the residual risk:*

$$\bar{Z}^1(t) = \beta_{12}\bar{Z}^2(t) + \lambda_{[X^1]}\Omega_e X^1 + \sum_{m=2}^{M} \lambda(m,t)\Omega_e(Q_m^1(t) - Q_m^{1W}) \tag{12}$$

where

$$\lambda_{[X^1]} = \frac{1 - (1/r)}{\sum_{m=1}^{M} 1/A_m} \tag{13}$$

Proof: See section D of the appendix.

The proposition is intuitive. Because investors can perfectly hedge the segment risk that is correlated with the index, the reward for bearing that risk is exactly the same as is provided by trading the index. Investors are left to share the segment assets portfolio of residual risk, $X^{1\prime}e_t$. Segment-assets covariance with that residual risk is $\Omega_e X^1$. I refer to this covariance as segment risk; segment risk is rewarded because the investors in the segment cannot diversify it away.[15] In section D.1 of the appendix I show that in a competitive setting, index-risk and segment-risk are the only priced risks in the model.

When some investors trades move prices, imperfect risk sharing among investors introduces additional transient priced factors that vanish only when investors asset holdings converge to those associated with perfect risk sharing. Because of illiquidity, the convergence process takes time; therefore, imperfect risk sharing at period t affects one period risk-premia at period $t + \tau$, as shown in the following corollary:

Corollary 1 *The segment assets τ period ahead 1-period excess returns follow a factor model in which deviations from perfect risksharing at period t affect excess returns at period $t + \tau$:*

$$E_t[Z^1(t + \tau + 1)] = \beta_{12}\bar{Z}^2 + \lambda_{[X^1]}\Omega_e X^1 + \sum_{m=2}^{M} \lambda_m(t,\tau)\Omega_e(Q_m^1(t) - Q_m^{1W}) \tag{14}$$

---

[15]The presence of this term suggests there is a CAPM-like relationship in the model for sharing segment-risk. This is similar to Stapleton and Subrahmanyam (1978), who derive circumstances in which the CAPM holds dynamically through time when investors have CARA utility and trade risky assets whose dividend payments are normally distributed.

9

Proof: See section D of the appendix.

In both the corollary and the proposition, the prices of risk of the transient factors (the λ_m's) are negative because when large investors hold more than their efficient share of risky assets, the marginal investors, in this case the small investors, hold less and hence require a smaller premium for holding the residual risk.

Potential Explanations for Underperformance

There are two notions of return underperformance after an IPO. The first is that an asset's expected returns in the short-run are lower than its expected returns in the long-run. The second are that its expected returns underperform after adjusting for some benchmark measure of risk. This section deals with the first type of underperformance; the second type of underperformance is discussed in section 5.3.

Corollary 1 illustrates how underperformance of the first type can occur. More specifically, because investors trade to efficient asset holdings, the third term in equation (14) is transient, while the first two terms are not. Therefore, short-run underperformance will occur whenever the third-term is negative. Because the market prices of risk $\lambda_m(t, \tau)$ are negative, the third term generates return underperformance when each large investor's IPO share allocation is greater than is consistent with optimal risk sharing; this is consistent with allocations being tilted towards institutional investors and away from retail investors at the IPO.

The magnitude of underperformance depends on whether allocations are distorted towards large investors, and it depends on the risk preferences of the large investors that receive the allocations. In a model with illiquidity, differences in investors' risk tolerances correspond to differences in investors' willingness to sell assets quickly and pay a high liquidity cost in order to share risk. The more risk tolerant large investors are less willing to pay a high a liquidity cost to share risk; therefore, risk sharing progresses more slowly when asset allocations are initially tilted towards them. This reduces the amount of risk that must be borne by the marginal (small) investors and reduces the assets required rate of return. Consistent with this reasoning, in simulations the $\lambda_m(t, \tau)$ functions (which measure the impact of allocation distortions towards investor m at time t on excess returns at time $t + \tau$) show that when allocations are distorted towards large investors with more risk tolerance, the effect on excess returns is longer-lived and larger in magnitude.

The amount and persistence of underperformance also depends on the competitiveness

of the aftermarket. When the aftermarket is highly competitive, the $\lambda_m(t, \tau)$ functions rapidly asymptote toward 0 as τ increases, generating little underperformance. When the aftermarket is less competitive, the functions asymptote slowly, generating longer-lived underperformance.

Because IPO allocations affect post-IPO excess returns, they should also affect the initial trading price of the segment assets in the aftermarket. This is established in the appendix[16], where I show that segment-asset prices at the beginning of each time period have the form:

$$P^1(t) = \frac{1}{r}(\alpha - \sum_{m=1}^{M} \gamma_m(t)\Omega_e Q_m^1),$$ (15)

where the scalars $\gamma_m(t)$ are positive numbers that are smaller for investors who are more risk tolerant. It follows that aftermarket prices will initially be higher than their long-run competitive values if the asset allocations are distorted away from efficient asset holdings towards relatively risk tolerant large investors. The initial "overpricing" in the aftermarket due to allocation distortions may help to explain part of the IPO underpricing puzzle. A full explanation depends on how shares are allocated at the IPO, and on how the IPO offer price is set. I turn to that topic below.

4 IPO Share Allocation and Price Setting

The motivation for the analysis in this section is based on Pritsker (2004). Pritsker studies a situation in which a distressed investor must rapidly sell a given number of shares into an imperfectly competitive aftermarket. Because the distressed investor is essentially selling to an oligopoly, the price he receives is discounted relative to the competitive price. The size of the discount is determined by the intensity of competition, which in turn depends on differences in large investors' risk tolerances. If one large investor is far more risk tolerant than the others, he has significant market power because by purchasing fewer shares, he can force the distressed sales to be absorbed by relatively risk averse investors who require a larger risk premium for taking the shares. Conversely, if investors risk tolerances are more similar, each investors market power is reduced, and the aftermarket becomes more competitive.[17]

---

[16]See propositions 2,3, and 5

[17]In Pritsker (2004), large investors can be interpreted as institutions that trade on behalf of a base of small investors. In this interpretation, large investors absolute risk tolerances are just the integrated risk

The distressed investor analysis is applicable to the IPO setting. The issuing firm is like a distressed investor that needs to sell X^{IPO} shares in a segment that is imperfectly competitive. The IPO is a mechanism that helps avoid selling shares into the imperfectly competitive aftermarket by instead allocating them beforehand at a fixed IPO price. I assume the underwriter acts on the issuer's behalf by lining up investors to buy the issue in order to maximize IPO proceeds. The IPO process resembles bookbuilding as practiced in the United States. The underwriter gathers demand information on the issue. In the model this information consists of knowledge about the other risky assets in the new issue's market segment, as well as knowledge about investors holdings of the segment-assets, the investors risk preferences, and whether there are investors who have market power in aftermarket trading. Based on this information, the underwriter sets an IPO offer price P^{IPO} and makes take it or leave offers of share allocations to the large and small investors. The large investors allocations' are denoted by $X_m^{IPO}, m = 2, \ldots M$. The total number of shares offered to small investors is denoted X_1^{IPO}. I assume that the small investors that are offered share allocations are offered identical amounts of shares. It will turn out that if shares are offered to some small investors, it will be optimal to offer them to all small investors. Therefore, I assume that when small investors are offered shares, they are all offered the same amount.

If a large or small investor turns down the share allocation that he is offered, then the underwriter is stuck with the shares. I assume that the underwriter sells any unallocated shares immediately in the aftermarket following the IPO. The possibility that an investor can force distressed sales in the aftermarket serves as a threat that constrains how the issuer allocates shares and chooses the IPO offer price.[18] In particular, if an investor receives shares, the allocations and offer price must be set so that it cannot be in the interest of the investor to refuse their allocation and instead force the shares to be sold by the underwriter in the aftermarket. Of course, it is possible in theory that the underwriter might find it optimal to sell some shares in the aftermarket; denote these shares as X_U^{IPO} and the aftermarket price on the first day of trading as P_A^{IPO}.[19] This suggests that the underwriter chooses the vector

---

tolerances of their small investor base. If all small investors have identical risk tolerances, then large investors risk tolerances will differ substantially depending on the size of their investor base.

[18]There are many other possible ways to model the threats that available to the large investors and the threats that are available to the underwriter.

[19]The aftermarket price on the first day of trading was solved for in section 3; it is given by equation (B3). Because a distressed sale in period 1 is tantamount to sales by investors that don't account for the price impact of their trades, for the purposes of price determination if X_U is not equal to 0, then X_U is added to the initial allocation of investor 1 in order to determine the aftermarket price.

of share allocations and aftermarket sales X and IPO offer price P^{IPO} to maximize

$$\Pi(P^{IPO}, X) = P^{IPO} \times (\sum_{m=1}^{M} X_m^{IPO}) + P_A^{IPO} X_U^{IPO}, \tag{16}$$

where $\Pi(.)$ represents the underwriters profit, the first term on the right hand side measures revenues raised at the IPO, and the second represents revenues raised by distressed sales in the IPO aftermarket.

The maximization takes place subject to the constraints that the total issue is allocated:

$$\sum_{m=1}^{M} X_m^{IPO} + X_U^{IPO} = X^{IPO}, \tag{17}$$

that there are no short-sales by investors 1 through M,

$$X_m^{IPO} \geq 0, m = 1, \ldots M, \tag{18}$$

and subject to incentive compatibility (IC) constraints that those who receive allocations in the IPO will accept the allocations. The IPO allocations are evaluated based on investors ex-ante expected value of entering the first period of aftermarket trading when the asset allocation is X.[20] Each small investor's IC constraint is denoted ψ_s, and takes the form:

$$\psi_s = V_s[Q_s^{IPO}; Q^{IPO}, t^{IPO} + 1] - V_s[Q_s; Q^{IPO}, t^{IPO} + 1] \geq 0 \tag{19}$$

The segment-asset holdings of each small investor that chooses to and chooses not to participate in the IPO are denoted Q_s^{IPO} and Q_s respectively; and the vector of all investors post-IPO segment asset holdings is denoted Q^{IPO}. Because each small investor is infinitesimal, his individual participation decision has no influence on the state vector Q^{IPO}.

Large investors IC constraints are denoted ψ_m and take the form:

$$\psi_m = V_m[Q^{IPO}, t^{IPO} + 1] - V_m[Q_{-m}^{IPO}, t^{IPO} + 1] \geq 0 \tag{20}$$

where large investor m's share allocation is Q_m^{IPO} if he participates in the IPO, and Q_{-m}^{IPO} if he does not.

While other investors are not allowed to take a short position at the IPO, the underwriter

---

[20]The value functions are derived from proposition 2, but they contain an additional argument to account for the underwriters net sales in the aftermarket X_U.

can do so. If he does, then he is obligated to buy shares in the first period of aftermarket trade in order to cover his short position.[21]

The assumption that any shares not sold at the IPO are sold immediately afterward by the underwriter in the form of distressed sales is very strong. A more realistic assumption is that any shares that the underwriter fails to sell at the IPO will instead be sold over a longer amount of time, τ_S trading periods, following the IPO. This modeling assumption is consistent with empirical evidence, reported in Ellis et. al. (2000), that IPO underwriters engage in price support activities in the IPO aftermarket, and with evidence reported by Ellis et. al. (2002) which shows that underwriters are active participants in the IPO aftermarket for long periods of time.[22].

I assume that when the underwriter sells shares over τ_s time periods he will sell them optimally. By optimality I mean that the underwriter buys shares at the IPO offer price, and then participates in the IPO aftermarket, trading as a large investor over the following τ_S time periods in order to maximize his own utility subject to the constraint that by time τ_S the underwriter holds no shares of the new issue. It is assumed that the certainty equivalent value of the underwriters utility from buying and trading the shares is turned over to the issuing firm at the time of the IPO. For tractability I assume that the underwriter has CARA utility like the other large investors. Let $CE_U(Q^{IPO}, \tau_s)$ denote the underwriters certainty equivalent. Then, under the less restrictive assumption, the underwriter maximizes:

$$P^{IPO} \times (X_1^{IPO} + \sum_{m=2}^{M} X_m^{IPO}) + CE_U(Q^{IPO}, \tau_s), \qquad (21)$$

subject to the constraints that the total issue is allocated [equation (17)], that there are no short sales [equation (18)], and subject to a new set of participation constraints that account for the underwriter's trading activity in the aftermarket.[23] My conjecture is that the underwriter's ability to spread out sales of unsold shares through time will reduce large investors bargaining power at the IPO and result in a higher IPO offer price, and higher

---

[21]The "Green Shoe" option that is often given to underwriters allows the underwriter to cover his short position following the offering by acquiring up to a certain amount of additional shares from the issuer at the IPO offer price. This option is valuable when there is uncertainty about demand for the new issue. In the present model, there is no demand uncertainty. Therefore, underwriters are not allowed to short-cover by using the "Green Shoe" option.

[22]In Ellis et al.'s (2002) sample of 313 NASDAQ IPOs, the lead underwriter participated in an average of more than 90 percent of post IPO NASDAQ trades during the first day of the IPO; this amount tapers down over the next 140 days, but remained above 40 percent on average on the 140th day

[23]The solutions for investors value functions and optimal trades when a distressed investor optimally sells a position over τ time periods is similar to Pritsker (2004), and is not presented here in order to save space.

profits for the issuer. That said, most of my recent simulations using the objective function in equation (21) show spreading out the underwriters sales has little effect on profit. Therefore, I do not report any results for this case in the current version; but hope to study it more in future work.

In closing this section I should emphasize that the optimal IPO allocation problem is difficult because it involves maximizing a nonlinear objective function subject to nonlinear inequality constraints. Details on how I solved the problem are provided in appendix E.

5 Simulation Analysis

The model's properties are studied for a single case in which an underwriter sells 40 shares to a continuum of small investors (investor 1) and five large investors (investors 2-6) under a variety of liquidity conditions. For simplicity, the only segment asset is the new issue, and investors risk tolerances are normalized so that they sum to 1.

Aftermarket liquidity depends on the concentration of risk bearing capacity among investors, and on the number of periods of aftermarket trade. An investors share of a segment's risk bearing capacity is his risk tolerance as a percentage of the sum total of all investors risk tolerances. Concentration of risk bearing capacity provides investors with market power in the aftermarket; and this makes the market more illiquid. Results on concentration of risk bearing capacity are reported using the Herfindahl index, which ranges from 10,000 when risk bearing capacity is concentrated with one investor, to 0 when all investors are small and there is no concentration.[24]

The number of post-IPO trading periods consider ranges from 200 trading days (a bit less than a year) up to 2000 trading days (8 years).[25] To explain the relationship between the number of post-IPO trading periods and liquidity, note that the more shares an investor trades within a period, the more he moves the price. Therefore, he would prefer to break up his trades through time to minimize their price impact. When fewer periods of trade remain, opportunities to break up trades are limited; therefore risk sharing becomes more costly. Consequently, shares acquire a greater liquidity premium when fewer periods of trade remain. Conversely, as the number of tradings periods grow, the liquidity premium

---

[24]The Herfindahl index is equal to sum of each investors squared percentage share of risk bearing capacity. Because each small investor is infinitesimal, their contribution to the index is 0.

[25]The time scalefor the trading periods is determined by the annualized per-period interest rate. There are 250 trading periods per year, and the annualized per-period interest rate is fixed at 2 percent.

vanishes and the market becomes perfectly competitive.[26] Because my focus is on imperfect competition, I assume the number of trading periods is finite.

To illustrate the role of imperfect competition, the main results focus on two benchmark cases that differ only in whether the investors behave competitively. In both cases, investor 2 has more than a fifty percent share of the segment's risk bearing capacity (Table 1, Panel A); and the other investors differ in their risk tolerances.

When investors behave competitively, and the underwriter must sell all 40 shares, the IPO is no different than if the underwriter sold the shares directly in the aftermarket. Therefore, he raises $420 at the competitive price of $10.50 per share; additionally investors share allocations are proportional to their share of risk bearing capacity, which implies investor 2 should receive 21.82 shares.

When the aftermarket is imperfectly competitive, investor 2's large share of risk bearing capacity gives him enormous bargaining power at the IPO, because if he turns down his share allocation then the underwriter will have to sell shares into an aftermarket where investor 2 has substantial influence over price. Alternatively, if the underwriter instead chooses to allocate shares to other more risk averse investors, then they will be exploited by investor 2 in the aftermarket—and this will depress the price that other investors are willing to pay for shares. To illustrate these points, I studied a sequence of IPOs in which investor 2 can trade in the aftermarket but is restricted in the amounts that he can acquire at the IPO. More specifically investor 2's aquistion was progressively restricted to be no larger than $0, 2, 4, \ldots X^{IPO}$ shares.[27] If the aftermarket is perfectly competitive, the allocation restrictions are of little consequence because other investors can aquire the shares at the IPO and then quickly sell them to investor 2 in the aftermarket at the competitive price; therefore the IPO offer price is barely discounted from the competitive price. By contrast, if the aftermarket is imperfectly competitive, then the IPO offer price is severely discounted relative to the competitive price; moreover, the feasible offer price (Figure 1) and proceeds (not shown) are monotone increasing in the amount that can be offered to investor 2 up to the total outstanding supply of the issue. Investor 2's enormous bargaining power has implications for allocations and trading volume, as well as for underpricing, and return

---

[26]Alternative intuition for the relationship between illiquidity and the number of post-IPO trading periods is based on Coasian analysis of a durable goods monopolist. Coase shows that a durable goods monopolist has more market power if he can commit to selling over a single time period instead of allowing for several periods of retrade. Kihlstrom (2001) argues that stocks are durable goods and shows that the Coasian analysis applies to a monopolist in stocks. The results here are similar to Kihlstrom, but in an oligopoly setting.

[27]The analysis used approximate optimization over a discrete grid of allocations.

underperformance.

5.1 Allocation Distortions and Trading Volume

Because prices and proceeds are increasing in the amount that investor 2 can purchase, all of the shares are allocated to him at the IPO. This shows that allocations are distorted towards investors who have market power and away from retail investors (who each have none) (Table 1,Panel C). Additionally, among investors with market power, in the example the allocations are distorted towards the large investor with the most power. This illustrates that differences in investors market power provide a noninformation based explanation for why some large investors receive share allocations while others do not.

Because allocations at the IPO are not pareto optimal, the distortions create a basis for trade in the aftermarket. The resulting trading volume is heaviest on the first day of trading, constituting 1.2 percent of shares issued; daily volume drops off to about 0.1 percent of shares issued after the first week of trading (Figure 2). The example's pattern of heavy trading volume that rapidly drops off is qualitatively consistent with the empirical literature on post-IPO trading (Ellis et al. 2000), but the model fails to match the empirical magnitude of the first day's trading, which averaged 33.3 percent of shares issued in the 1980's, and 148.7 percent of shares issued in 1999-2000 (Ritter, 2005).

5.2 IPO Underpricing

The example shows the model is capable of generating substantial IPO underpricing. Underpricing ranges from a low of 40 percent when many trading periods remain to a high of 159 percent when relatively few periods remain (Table 1, Panel B). The underpricing has three pronounced features. First, the offer price is low relative to the asset's competitive price. Second, the aftermarket price is inflated relative to the asset's competitive price. If the competitive price is interpreted as the asset's "fundamental value" then it appears as if the IPO is associated with irrational price-overshooting even though the model is fully rational. The third feature is that the first two features are most pronounced when there are fewer periods of trade following the IPO.

Intuition for the first feature is related to investor 2's bargaining power; additional intuition for the underpricing comes from interpreting illiquidity as a tax on risk sharing.

Because the tax makes risk sharing more costly, it pushes down the IPO offer price; in addition because of the tax less risk sharing takes place, which means investor 2 sells less in the aftermarket—and this constraint on supply pushes up the aftermarket price. When fewer trading periods remain, illiquidity is more severe, making the first two features more pronounced.

5.3   Underperformance

Recall that there are two types of return underperformance—underperformance relative to an assets long-run return—and underperformance relative to a risk adjusted benchmark. Section 3.1 shows that the former is associated with IPOs; to address the latter, post-IPO expected excess returns were computed for up to 2000 business days and then adjusted for market-risk by subtracting the asset's beta times the expected excess return on the market portfolio. To compute the excess return on the market portfolio, I aggregated up the excess returns across all market segments. Recall from equation (12) (reproduced below) that in each segment j, assets excess expected returns can be decomposed into a component that is correlated with the index, a second that is associated with market segmentation (due to imperfect sharing of idiosyncratic risk), and a third that is associated with allocation distortions and illiquidity:

$$\bar{Z}^j(t) = \beta_{j2}\bar{Z}^2(t) + \lambda_{[X^j]}\Omega_{e_j}X^j + \sum_{m=2}^{M_j}\lambda(m,t)\Omega_{e_j}(Q_m^j(t) - Q_m^{jW}).$$

Because investors trade towards efficient asset holdings, the third component is transitory. and is assumed to be small or zero for segments that have not had recent IPOs; but it is negative for those that have had them.

Under these circumstances, if investors trade the liquid index and specialize in trading the assets within different market-segments, then in the context of a one-period example (appendix F), the assets in each segment and the market portfolio earn a segmentation premium. Therefore, after adjustment for market-risk, the assets average segmentation premia are 0, and do not depend on whether the assets are new issues. Using similar reasoning, because relatively few firms are new issues, these firms will only have a minor effect on the market return. Therefore, after market adjustment, IPO firms will retain a large negative liquidity premium; that is they will underperform on a risk-adjusted basis.

Under simplifying circumstances (see appendix F), if ϕ is the fraction of firms that are

new issues, then the average market adjusted underperformance at time t for segments with IPO's should be approximately $(1 - \phi)$ times the allocation distortion / liquidity component of excess returns:

$$(1 - \phi) \sum_{m=2}^{M} \lambda(m, t) \Omega_e (Q_m^1(t) - Q_m^{1W}).$$

In the example, if ϕ is 0, and the market adjusted segmentation premium is approximately 0 for simplicity, then the CAR grows to -32 percent over 2000 periods. In the simulations which corresponds to about 8 years of underperformance (Figure 3, Panel A). More realistically, if ϕ is about 10 percent, then the expected market-adjusted CAR should be about -29 percent over 8 years. This confirms that the model can in theory generate both IPO underpricing and long-lived market adjusted return underperformance.

Adjusting for Characteristics

A stylized fact of measured underformance is that it is strongest when returns are market-adjusted, and weaken after adjustment for additional characteristics such as market-to-book (Ritter and Welch, 2002). Under some interpretations, these stylized facts are consistent with the liquidity-based explanation in the model; in particular, if assets' long-run competitive prices are interpreted as book-value, then the model predicts that because of illiquidity and allocation distortions at the IPO, then just after the IPO, market to book is high and consistent with the market-to-book effect, IPO returns underperform. It is important to bear in mind that in the empirical asset pricing literatature, market-to-book's ability to explain asset returns is an empirical regularity, not a theory of asset pricing. The contribution of the present model is that it provides a theoretical reason why market-to-book appears to "explain" post IPO return underperformance. This result has empirical implications; in particular, it shows that controlling for market-to-book can be problematic when testing the present model because a finding of no underperformance after adjustment for market-to-book is consistent with the model, and not evidence against it. This topic is discussed further in footnote 33 of section 6.1.

5.4   Money Left on the Table

IPO underpricing makes it appears as if the IPO should have raised more revenue. The loss in revenue is referred to as "money left on the table", hereafter MLOTT, and is usually measured as the differential between the IPO-offer price and closing price on the first day of

trade times the number of shares issued. MLOTT often has the interpretation of a measure of the issuer's losses due to imperfections in the IPO process.[28] For the present model, the imperfection is illiquidity; and it should be clear that the usual MLOTT calculation overstates issuer losses because it incorrectly assumes that all of the shares could be sold at the artificially high price that prevails in an illiquid aftermarket. An alternative measure of MLOTT instead compares the revenues that were raised at the IPO against the revenues that could have been raised if the aftermarket was perfectly competitive, which in this model means perfectly liquid. Computing losses by this alternative metric shows that the usual calculation can very significantly overstate issuer losses; in the case of the example it overstates them by a factor of 5 to 10 (Table 1, Panel B).

5.5 Alternative market configurations

To study how differences in investors risk tolerances affect the results, I solved the model under two alternative market configurations. In the first, some of investor 2's risk tolerance is spread evenly among the other large investors (Table 2). This change makes the aftermarket more competitive, and has three further effects on the aftermarket. First, when the number of aftermarket trading periods is 2000, or 1800, the market is sufficiently competitive that differences in who receives shares at the IPO have a very small effect on prices (no more than 6 cents per share). In this circumstance, the optimal allocation problem is ill-posed. I simply assume that the competitive allocation results in these cases. Second, the magnitude and persistence of underperformance diminish somewhat, with the liquidity component of the CAR reduced to -16 percent over 5 years. Third, the prices charged by the underwriter at the IPO are actually above the competitive price. This occurs because when the aftermarket is a little bit more competitive, investor 2 has less market power in the aftermarket, and hence has less bargaining power at the IPO. In this circumstance, the underwriter can extract some of investor 2's surplus from acquiring the new issue. Because the aftermarket is still not perfectly liquid, the aftermarket price is inflated above the IPO offer price, but now the underwriter is actually generating more revenue for the issuer than if the aftermarket were competitive. Compared with the earlier results, this shows that the underwriter and issuer actually benefit if the aftermarket is a bit less than perfectly competitive, but is hurt if the aftermarket becomes too imperfectly competitive. Because a little bit of imperfect competition can sometimes help the underwriter, this result may help to explain underwriter

---

[28]For example, in informational theories of bookbuilding, the money left on the table is equal to informational rents that investors receive for sharing their information with the underwriter.

practices that restrain trade in the aftermarket, such as restrictions on investors ability to flip shares.

The second alternative market configuration contained two dominant large investors that each have 30% of the risk bearing capacity.[29] Because this configuration was highly competitive, unless a small number of trading periods followed the IPO, results are only discussed for when there are 200 post-IPO trading periods. Unlike the previous examples in which the optimum involved allocating all of the shares to one investor, in this example the optimal share allocations involved splitting the shares evenly between investors 2 and 3 while distorting the share allocations away from all other investors. As in the previous example, the resulting IPO offer price was above the price that would prevail if the aftermarket was perfectly competitive. Nevertheless, the aftermarket price following the IPO was even higher, resulting in IPO underpricing of about 12%. Additionally, cumulative abnormal returns indicate short-lived underperformance; the liquidity component of the underperformance was 4.5% over 200 trading days.

6 Empirical Implications

The purpose of this section is to outline a rough strategy for testing the theory in this paper. The main implication of the theory is that IPO illiquidity and imperfect competition in the aftermarket following an IPO lead to allocation distortions towards large investors and IPO underpricing and underperformance. These implications lead to the following testable predictions:

1. Allocation distortions are associated with underperformance in the aftermarket.

2. Allocation distortions are associated with IPO underpricing.

3. The above two effects are associated with illiquidity and imperfect competition in aftermarket trading.

Prediction 1 imposes restrictions on the coefficients of empirical versions of equation (14):

$$Z_i(t+\tau+1) = \beta_{i,2}\pi_{IND} + [\lambda_{[X^1]}\Omega_e X^1]\pi_{SEG} + [\sum_{m=2}^{M} \lambda_m(t,\tau)\Omega_e(Q_m^1(t) - Q_m^{1W})]\pi_{LIQ} + \epsilon_i(t+\tau) \tag{22}$$

---

[29]The other 4 investors each had 10% of the risk bearing capacity.

The empirical version provided above differs from equation (14) because it is parameterized with additional π coefficients, and because expected excess returns have been replaced by their realization on the left hand-side, which introduces an expectational error $\epsilon_i(t+\tau)$ on the right-hand side. The main coefficient of interest is π_{LIQ}. Under the null hypothesis that allocation distortions at the IPO are unrelated to underperformance, π_{LIQ} should be equal to 0; while under the alternative π_{LIQ} should be positive and the coefficients $\lambda_{m,t}$ should be negative. Estimation of π_{LIQ} in cross-section requires information on Ω_e and investors holdings of all assets within each segment. However, if Ω_e is nearly diagonal, or if allocation distortions are only significantly different from zero for new issues, then it is sufficient to create an allocation distortion measure for each new issue that only depends on its own idiosyncratic risk and own asset holdings. This simplified approach is outlined below.

The allocation distortion measure that I propose for firm i has the form:

$$\text{distort}_i = \Omega_{ii} \sum_{m=2}^{M} g(RBC_m) \left[\frac{x_m}{\sum_{m=1}^{M} x_m} - RBC_m \right], \tag{23}$$

where x_m is the share allocation of large investor m at the IPO; Ω_{ii} is the variance of the idiosyncratic component of firm i's return; RBC_m is large investor $m's$ risk bearing capacity, which is his risk tolerance as a fraction of the sum total of all M investors risk tolerances; and $g()$ is a negative and monotone decreasing function, which serves as a proxy for λ in equation (22). The distortion measure is 0 if allocations are efficient; positive if allocations are distorted towards small investors, and negative if distorted towards large investors. In addition, consistent with the results that show underperformance and the aftermarket price are maximized when allocations are distorted towards the most risk tolerant large investor, the distortion measure is maximized when all assets are allocated to the most risk tolerant large investor.

Proxies for RBC_m are needed to operationalize the distortion measure. Because the model assumes that investors have CARA utility primarily for tractability, a sensible way to proceed is to use a more realistic assumption about investors utility. For example, if all investors have power utility with coefficient of relative risk aversion γ, then investors absolute risk tolerance is W_m/γ, and each investors risk bearing capacity is equal to their own wealth as a fraction of the sum total of the wealth of other investors in the segment $(RBC_m = \frac{W_m}{\sum_{m=1}^{M} W_m})$.[30] Since wealth is a measure of investors's size, the allocation distortion

---

[30]In the case of institutional investors, W_m represents wealth under management, and the measure of risk bearing capacity implies that more wealthy investors have a greater capacity to bear risk.

measure them becomes a measure of how assets are allocated relative to the size of the investors involved. Armed with this measure of allocation distortions it should be possible to test prediction 1 by estimating variants of equation (22) in cross-section and then test whether π_{LIQ} is positive.[31]. A full analysis of how to estimate the equation is well beyond the scope of the present paper. To test prediction 2, the same measures of allocation distortions that are used to estimate equation (22) can be used to attempt to explain the cross-section of IPO underpricing.

To test the third prediction, measures of aftermarket illiquidity should be interacted with the allocation distortion variable. The theory predicts that the allocation distortions should only have an effect when there is illiquidity in the aftermarket. Hence, the interaction with the illiquidity variables should provide a sharper test of theory. The coefficients on the interaction terms are expected to have the same sign as the coefficients on the allocation distortions, and including these terms should cause the coefficients on the allocation distortions to weaken.

Allocation distortions are predicted by other theories, such as for example bookbuilding and adverse selection, in which the allocations are based on investors information. However, neither of those theories predicts a relationship between allocation distortions and IPO return underperformance. Therefore, if allocation distortions explain both underpricing and underperformance, it should be interpreted as evidence that an explanation based on illiquidity and imperfect competition in the aftermarket helps contribute to our understanding of IPOs.

To close this section, I briefly review the most closely related empirical literature on underpricing and underperformance.

6.1   Related literature

The empirical literature that is most closely related to this paper studies the relationship between after-market liquidity and underpricing or underperformance. The relationship between IPO underpricing and illiquidity has been empirically studied by Booth and Chua (1996), Hahn and Ligon (2004), and Ellul and Pagano (2003). In closely related work, Butler et. al. (2005) study the relationship between a stock's liquidity, and the underwriting fees that are paid during a seasoned equity offering.

---

[31]Getting data on IPO allocations is difficult, but it is sometimes available on a proprietary basis as in Cornelli and Goldreich (2003), or can be constructed as in Reuter (2005).

Although the Booth and Chua model makes predictions about the relationship between underpricing and aftermarket liquidity, they don't test this implication of their model; instead their tests focus on underpricing as compensation for costs of information gathering. Because such costs could generate underpricing irrespective of illiquidity, the implications of their tests for the relationship between underpricing and aftermarket liquidity are unclear. Hahn and Ligon attempt to directly test the Booth and Chua hypothesis that underpricing is used to increase liquidity by running OLS regressions of market microstructure measures of aftermarket liquidity on IPO underpricing. In regressions that account for other determinants of illiquidity, their results are mixed; with coefficients on underpricing sometimes statistically significant and positive, sometimes statistically significant and negative, and sometimes not statistically significant at all. A potential difficulty with the Hahn and Ligon methodology is that causality may run from underpricing to illiquidity (as in Booth and Chua) as well as from illiquidity to underpricing (as in Ellul and Pagano). The possibility that causality runs in both directions suggests that an instrumental variable approach is needed. In Ellul and Pagano, they regress underpricing on a set of determinants for underpricing, including measures of aftermarket liquidity. Additionally, they recognize the potential for simultaneity bias and instrument for it in some of their regressions.[32] In all of Ellul and Pagano's regressions they find that more aftermarket illiquidity increases the amount of IPO underpricing. Butler et. al. find qualitatively similar relationship between illiquidity and underwriter fees in SEO's, but quantitatively the effects of illiquidity are much smaller than in Ellul and Pagano. This suggests the Ellul and Pagano results, while favorable for liquidity based theories, should be interpreted with caution. An additional reason for caution is if underpricing is a risk premium for aftermarket illiquidity, then the logical extension of Ellul and Pagano's theory would suggest that in the aftermarket, IPO's should earn a positive and significant risk premium for aftermarket illiquidity. If we believe the empirical evidence that IPO returns underperform in the aftermarket, this suggests that the mechanism driving aftermarket returns is more complicated than the theory of illiquidity considered by Ellul and Pagano. Eckbo and Norli (2002) take this argument one step further; they claim that newly issued stocks are more liquid than other stocks with similar risk characteristics; and thus their returns should underperform. To establish this point empirically, Eckbo and Norli compare the returns of a rolling portfolio of newly issued stocks that are held for up to five years against the returns a portfolio of more seasoned issues that are matched on size and book to market. They find that after adjusting for these factors, and controlling for differences in liquidity, new issues do not underperform.

---

[32]They do not report any results on tests for the strength of the instruments, nor do they report any results of tests for instrument validity.

The Eckbo and Norli and Ellul and Pagano findings, taken together are puzzling because the latter suggests that IPOs are very illiquid, while the former suggests the opposite. Both papers can only be correct if liquidity conditions change rapidly after the IPO, and if participants in IPOs are very concerned about a short-term need to liquidate. This paper points towards a different resolution in which illiquidity, when combined with imperfect competition, generates both underpricing and underperformance. To make a strong case that this is the correct resolution requires careful future research on how to measure underperformance and which measures of liquidity are relevant for large investors.[33,34]

7   Conclusions

In this paper I have presented a fully-rational, symmetric-information model to simultaneously explain IPO underpricing, underperformance, and a tilt in IPO share allocations towards institutional investors, and away from retail investors. The key model features that generate these results are illiquidity and imperfect competition in aftermarket trading. The model also generates a new set of testable predictions that tie the market structure of after-

---

[33] A difficult issue in assessing return underperformance is whether to adjust returns for market-to-book effects. The standard basis for adjustment is to determine whether an empirical pricing anomaly is new, or is a manifestation of a known anomaly such as market-to-book. Because market-to-book is an empirical regularity, but not a theory, risk adjustment for it may lead to spurious inference about IPO underperformance when testing the present theory. To see what can go wrong, suppose that the forces that drive market-to-book effects for other stocks don't apply to IPO's, but that market-to-book is negatively correlated with IPO returns for other theoretical reasons, such as those given in this paper. In this circumstance, adjustment for market-to-book effects will spuriously bias the test against detecting underperformance when it is present. A better method to test the present theory is to forgo a problematic market-to-book adjustment, and instead test the predicted relationships between allocation distortions and the returns on new issues.

[34] A difficult issue in testing liquidity-based theories of IPO underpricing and underperformance is determining an appropriate measure of liquidity. Eckbo and Norli (2002) measure it based on trading volume; but it is not clear that high volume after an IPO has the same liquidity implications as high volume for other stocks. For example, if IPO's extraordinarily high turnover (see footnote 1) implied they were very liquid for large investors, then those investors should be selling out en masse on the first day following an IPO because of the high first day return. In fact, Aggarwal's (2003) study of sales by initial share holders, which is also known as share flipping, shows that on median only 7.34% of initial shares are sold during the first few days following an IPO. This suggests the extraordinarily high trading volume associated with IPOs does not necessarily translate into a highly liquid environment in which large investors can sell their shares. Post IPO volume may mismeasure liquidity because, as noted by Aggarwal (2003) it is associated with the same small number of shares being traded very frequently within the day, and does not measure the price impact of a large number of shares being sold into the market. This is consistent with Corwin et. al. (2004), who find that NYSE-listed IPOs typically have low depth to trading volume during the first several days of trading. Additonally, for investors who risk punishment from underwriters for flipping too many shares, the high volume does not measure their ability to sell off their holdings. Finally, although volume may be problematic as a measure of liquidity, other measures based on bid-ask spreads and price impacts may be problematic because it may not measure liquidity on the scale that large shareholders wish to transact.

market trading and allocation distortions at the IPO, to the cross-sectional pattern of return underperformance following an IPO. Hopefully the results in this paper will stimulate new empirical research that studies the relationship between the structure of the IPO aftermarket trading environment and IPO underpricing and underperformace.

Appendix

A Notation

There are M investors and $N = N_1 + N_2$ risky assets. The first N_1 assets are illiquid. The next N_2 assets are perfectly liquid. The risky asset holdings of investor m at time t are denoted by

$$Q_m(t) = \begin{pmatrix} Q_m^1(t) \\ Q_m^2(t) \end{pmatrix}$$

where $Q_m^1(t)$ and $Q_m^2(t)$ are investor $m's$ holdings of illiquid and liquid risky assets respectively. $Q^1(t)$ denote the $N_1 M \times 1$ vector of all investors illiquid asset holdings at time t where

$$Q^1(t) = \begin{pmatrix} Q_1^1(t) \\ \vdots \\ Q_M^1(t) \end{pmatrix}.$$

$Q_1^1(t)$ represents the net asset holdings of a continuum of infinitesimal small investors indexed by s:

$$Q_1^1(t) = \int_0^1 Q_s^1(t)\mu(s)ds.$$

The small investors are often collectively referred to as the competitive fringe. $Q_2^1(t)$ through $Q_M^1(t)$ denotes the net illiquid risky asset holdings of large investors, and is denoted by the $N_1 \times (M-1)$ vector $Q_B^1(t)$. The change in investors illiquid risky asset holdings from the beginning of time period t to the beginning of time period $t+1$ is denoted by the $N_1 M \times 1$ vector $\Delta Q^1(t)$. Similarly, $\Delta Q_1^1(t)$ and $\Delta Q_B^1(t)$ denote the change in the competitive fringe's illiquid asset holdings, and the change in the illiquid asset holdings of the large investors.

The algebra which follows requires many matrix summations and the use of selection matrices. Rather than write summations explicitly, I use the matrix $S = \iota_M' \otimes I_N$ to perform summations where ι_M is an M by 1 vector of ones, and I_N is the $N \times N$ identity matrix.[35] In some cases, the matrix S may have different dimensions to conform to the vector whose elements are being added. In all such cases, S will always have N, or N_1 rows. The matrix

---

[35]For example, $SQ(t) = \sum_{m=1}^M Q_m(t)$

S_i is used for selecting submatrices of a larger matrix. S_i has form

$$S_i = \iota'_{i,M} \otimes I_N,$$

where $\iota_{i,M}$ is an M vector has a 1 in its i'th element, and has zeros elsewhere.[36] As above S_i will sometimes have different dimensions to conform with the matrices being summed, but it will always have N or N_1 rows.

In the rest of the exposition, I will occasionally suppress time subscripts to save space.

B   Proof of Proposition 2

Proposition 2 : *Small investors value functions for entering period t with liquid wealth W_s, when investors' state vector of illiquid asset holdings is given by Q^1 is given by:*

$$
\begin{aligned}
V_s(W_s, Q^1, t) &= -K_1(t)\ F(Q^1, t)\ e^{-A_s(t)W_s}, \\
\text{where } F(Q^1, t) &= e^{-Q^1(t)'\bar{v}_s(t) - Q^1(t)'\theta_s(t)Q^1(t)}.
\end{aligned}
\tag{B1}
$$

Large investor m's value function for entering period t when the state vector of illiquid asset holdings is Q and his holdings of liquid wealth is W_m is given by:

$$V_m(W_m, Q^1, t) = -K_m(t)e^{-A_m(t)W_m - A_m(t)Q^{1\prime}\Lambda_m(t) + .5A_m(t)^2 Q^{1\prime}\Xi_m(t)Q^1} \quad m = 2, \ldots M, \tag{B2}$$

and the price function for illiquid assets has the functional form:

$$P^1(t) = \frac{1}{r}(\alpha(t) - \Gamma(t)Q^1) \tag{B3}$$

Proof: The proof is by induction. Part I of the proof establishes that if the value function has this form at time t, then it has the same form at time $t-1$. Part II of the proof establishes the result for time T, the first period in which trade cannot occur.

---

[36]To illustrate the use of the selection matrix, $Q_m(t) = S_m Q(t)$.

B.1 Part I:

Suppose the form of the value function is correct for time t. Then, to establish the form of the value function at time $t - 1$, I first solve for the competitive fringe's demand curve for absorbing the net order flow of the large investors. I then solve the large investors and competitive fringe's equilibrium portfolio and consumption choices, and then solve for the value function at time $t - 1$.

The competitive fringe's demand curve

The competitive fringe represents a continuum of infinitesimal investors that are distributed uniformly on the unit interval with total measure 1, i.e. $\mu(s) = 1$ for $s \in [0, 1]$. At time $t - 1$, each participant s of the competitive fringe solves:

$$\max_{\substack{C_s(t-1), \\ Q_s, \\ q_s}} \quad -e^{-A_s C_s(t-1)} - \delta \, \mathrm{E}[K_s(t) F(Q^1, t) e^{-A_s(t) W_s(t)}] \tag{B4}$$

where, Q_s is the stacked vector of small investor s's holdings of illiquid (Q_s^1) and perfectly liquid (Q_s^2) risky assets:

$$Q_s = \begin{pmatrix} Q_s^1 \\ Q_s^2 \end{pmatrix};$$

$Z(t)$ is the stacked vector of excess returns for the illiquid and liquid assets:

$$Z(t) = \begin{pmatrix} Z^1(t) \\ Z^2(t) \end{pmatrix} = \begin{pmatrix} P^1(t) + D^1(t) - rP^1(t) \\ P^2(t) + D^2(t) - rP^2(t) \end{pmatrix}; \tag{B5}$$

and small investors liquid wealth is given by

$$W_s(t) = Q_s' Z(t) + r[W_s(t - 1) - C_s(t - 1)].$$

Note: Although I refer to the first set of assets as illiquid, they are only illiquid for large investors whose trades have price impact. Because each small investor is infinitesimal, their

trades do not have price impact and hence both assets are perfectly liquid from their per-spective.

In equation (B5),

$$E Z(t) \equiv \bar{Z}(t) \equiv \begin{pmatrix} \bar{Z}_1(t) \\ \bar{Z}_2(t) \end{pmatrix},$$

and

$$\text{Var } Z(t) \equiv \Omega \equiv \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix}.$$

Substituting the expression for W_s in (B4) and taking expectations shows that small investors maximization becomes:

$$\max_{\substack{C_s(t-1), \\ Q_s}} -e^{-A_s C_s(t-1)} - \delta F(Q^1, t)e^{-A_s(t)r[W_s(t-1)-C_s(t-1)]-A_s(t)Q_s'\bar{Z}(t)+.5A_s(t)^2 Q_s'\Omega Q_s} \quad \text{(B6)}$$

In solving the model, it is useful to break small investors maximization into pieces by first solving for optimal Q_s^2 as a function of Q_s^1, and then solving for optimal Q_s^1. For given Q_s^1, the first order condition for optimal Q_s^2 shows that optimal Q_s^2 is given by

$$Q_s^2 = \frac{1}{A_s(t)}\Omega_{22}^{-1}\bar{Z}_2(t) - \beta_{12}'Q_s^1, \quad \text{(B7)}$$

where $\beta_{12} = \Omega_{12}\Omega_{22}^{-1}$.

Plugging the solution for Q_s^2 into the small investors value function and simplifying then shows that the small investors maximization problem reduces to:

$$
\begin{aligned}
\max_{\substack{C_s(t-1),\\ Q_s^1}} \quad & -e^{-A_s C_s(t-1)} - \delta F(Q^1,t) K_s(t) \operatorname{Exp}\left\{-.5\bar{Z}_2'\Omega_{22}^{-1}\bar{Z}_2 - A_s(t)r[W_s(t-1) - C_s(t-1)]\right\}
\end{aligned}
$$

$$
\times \operatorname{Exp}\left\{-A_s(t)Q_s^{1\prime}[\bar{Z}_1(t) - \beta_{12}\bar{Z}_2(t)] + .5A_s(t)^2 Q_s^{1\prime}\Omega_e Q_s^1\right\}
$$

(B8)

where Ω_e is given by

$$
\Omega_e = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}.
$$

To gain intuition for the above expression, note that the excess return on each illiquid asset can be decomposed into a component that is correlated with the liquid assets and into a second idiosyncratic component.

$$
Z_1(t) = \beta_{12}Z_2(t) + \epsilon_1(t)
$$

$\bar{Z}_1 - \beta_{12}\bar{Z}_2(t)$ is the vector of expected returns on the idiosyncratic components at time t and Ω_e is the variance covariance matrix of the idiosyncratic returns. The expression shows that small investors portfolio maximization problem can equivalently be written in terms of choosing an exposure to the returns of the liquid assets, and to the idiosyncratic component of returns of the illiquid assets.

Solving for optimal $Q_s^1(t)$ then shows

$$
Q_s^1(t) = \frac{1}{A_s(t)}\Omega_e^{-1}[\bar{Z}_1(t) - \beta_{12}\bar{Z}_2(t)]
$$

(B9)

The aggregate demand for Q^1 at time t by all small investors can be found by integrating both sides of equation (B9) with respect to μ_s, the density of small investors, yielding:

$$
\begin{aligned}
Q_1^1(t) &= \int_0^1 Q_s^1(t)\mu_s ds \\
&= \left[\int_0^1 \frac{1}{A_s(t)}\mu_s ds\right]\Omega_e^{-1}[\bar{Z}_1(t) - \beta_{12}\bar{Z}_2(t)] \\
&= \frac{1}{A_1(t)}\Omega_e^{-1}[\bar{Z}_1(t) - \beta_{12}\bar{Z}_2(t)]
\end{aligned}
$$

(B10)

The Price Schedule Faced by Large Investors

The price schedule faced by large investors at time $t-1$ maps large investors desired orderflow of the illiquid assets into the time $t-1$ prices at which the competitive fringe is willing to absorb the net orderflow. To solve for the price schedule, I solve for prices $P(.,t-1)$ in equation (B10) such that when the large investors choose trade $\Delta Q_B^1(t-1)$ at time t-1, then the competitive fringe chooses trade $-S\Delta Q_B^1(t-1)$.

Rearranging, equation (B10) while making the substitutions

$$
\begin{aligned}
Q^1(t) &= Q^1(t-1) + \Delta Q^1(t-1), \\
Q_1^1(t) &= S_1[Q^1(t-1) + \Delta Q^1(t-1)], \\
\Delta Q^1(t-1) &= \begin{pmatrix} -S\Delta Q_B^1(t-1) \\ I\Delta Q_B^1(t-1) \end{pmatrix}
\end{aligned}
$$

and

$$
\begin{aligned}
\bar{Z}^1(t) &= P^1(t) + \bar{D}^1 - rP^1(t-1,.) \\
P^1(t) &= \frac{1}{r}\left(\alpha(t) - \Gamma(t)[Q^1(t-1) + \Delta Q^1(t-1)]\right)
\end{aligned}
$$

then produces the price schedule faced by large investors at time $t-1$:

$$
P^1(.,t-1) = \frac{1}{r}\left(\beta_0(t-1) - \beta_{Q^1}(t-1)Q^1(t-1) - \beta_{Q_B^1}(t-1)\Delta Q_B^1(t-1)\right), \qquad \text{(B11)}
$$

where,

$$
\begin{aligned}
\beta_0(t-1) &= \bar{D}^1 + (1/r)\alpha(t) - \beta_{12}\bar{Z}^2 && \text{(B12)} \\
\beta_{Q^1}(t-1) &= (1/r)(\Gamma(t) + rA_1(t)\Omega_e S_1) && \text{(B13)} \\
\beta_{Q_B^1}(t-1) &= (1/r)\Gamma(t)\begin{pmatrix} -S \\ I \end{pmatrix} - A_1(t)\Omega_e S && \text{(B14)}
\end{aligned}
$$

Given the price schedule in equation (B11), large investors at time $t-1$ solve the maximization problem:

Large Investors Maximization Problem

$$\max_{\substack{C_m(t-1),\\ Q_m}} \quad -e^{-A_m C_m(t-1)} - \mathrm{E}\left\{\delta K_m(t) \operatorname{Exp}\left(-A_m(t)W_m - A_m(t)Q^{1\prime}\Lambda_m(t) + .5A_m(t)^2 Q^{1\prime}\Xi_m(t)Q^1\right)\right\}$$

(B15)

where, substituting in the budget constraint, liquid wealth at the beginning of time t is given by

$$\begin{aligned}
W_m(t) =& Q_m^1(t)' D^1(t) + Q_m^2(t)' Z^2(t) \\
& + r(W_m(t-1) - \Delta Q_m^1(t-1)' P^1(t-1,.) - C_m(t-1))
\end{aligned}$$

(B16)

Note: Because dividends are paid in cash, the dividend payments received for holdings of illiquid asset are counted as part of liquid wealth even though the illiquid assets themselves are not counted.

Note that in equation (B15), $\Lambda_m(t)$ and $\Xi_m(t)$ are deterministic functions of time that are parameters of the value function. Keeping this in mind, large investors holdings of the liquid assets are solved in the same way as for small investors. Taking expectations in equation (B15), solving for optimal Q_m^2 given Q^1, and substituting the optimal choice back into the large investor's value function, transforms the large investors maximization problem so that it has the following form:

$$\begin{aligned}
\max_{\substack{C_m(t-1),\\ Q_m^1}} \quad & - e^{-A_m C_m(t-1)} \\
& - \delta K_m(t)\left\{\operatorname{Exp}(-.5\bar{Z}^{2\prime}\Omega_{22}^{-1}\bar{Z}_2 - A_m(t)r[W_m(t-1) - \Delta Q_m(t-1)'P^1(t-1,.) - C_m(t-1)]\right. \\
& \left. \times \operatorname{Exp}(-A_m(t)Q^{1\prime}\bar{v}_m(t) + .5A_m(t)^2 Q^{1\prime}\theta_m(t)Q^1)\right\}
\end{aligned}$$

(B17)

where,

$$\bar{v}_m(t) = S_m'(\bar{D}_1 - \beta_{12}\bar{Z}_2) + \Lambda_m(t) \tag{B18}$$

$$\theta_m(t) = S_m'\Omega_e S_m + \Xi_m(t) \tag{B19}$$

The large investors play a Cournot game in which each choose his time $t-1$ trade

$\Delta Q_m(t-1)$ in the illiquid assets to solve the maximization problem in (B17) while taking the trades of the other large investors as given, but while taking into account the effect that his own trades have on the prices of the illiquid assets. Recall the price impact function for the illiquid assets at time $t-1$ is given by equation (B11).

The first order condition for large investors illiquid asset choices is given by:

$$0 = -A_m(t)[(-S_1+S_m)\bar{v}_m(t)] + A_m(t)^2(-S_1+S_m)[(\theta_m(t)+\theta_m(t)')/2](Q^1 + \Delta Q^1)$$
$$+ A_m(t)\left[rP^1(.,t-1) - S_m\beta_{Q^1_B}(t-1)'S_m\Delta Q^1_B\right], \tag{B20}$$

After substituting for $P^1(.,t-1)$ from equation (B11), writing $Q^1 + \Delta Q^1$ as $Q^1 + \begin{pmatrix} -S\Delta Q^1_B \\ \Delta Q^1_B \end{pmatrix}$ and simplifying, this produces the following reaction function for large investor m:

$$\pi_m(t-1)\Delta Q^1_B = \chi_m(t-1) + \xi_m(t-1)Q^1, \tag{B21}$$

where,

$$\pi_m(t-1) = A_m(t)(-S_1 + S_m)[(\theta_m(t)+\theta_m(t)')/2]\begin{pmatrix} -S \\ I \end{pmatrix}$$
$$- \beta_{Q^1_B}(t-1) - S_m\beta_{Q^1_B}(t-1)'S_m \tag{B22}$$

$$\chi_m(t-1) = (-S_1 + S_m)\bar{v}_m(t) - \beta_0(t-1) \tag{B23}$$
$$\xi_m(t-1) = \beta_{Q^1}(t-1) - A_m(t)(-S_1 + S_m)[(\theta_m(t)+\theta_m(t)')/2] \tag{B24}$$

Stacking the (M-1) reaction functions produces a system of $(M-1)N$ linear equations in $(M-1)N$ unknowns:

$$\Pi(t-1)\Delta Q^1_B(t-1) = \chi(t-1) + \xi(t-1)Q^1(t-1) \tag{B25}$$

Assume that $\Pi(t-1)$ is invertible. Then the solution for $\Delta Q^1_B(t-1)$ is unique, and given by

$$\Delta Q^1_B(t-1) = \Pi(t-1)^{-1}\chi(t-1) + \Pi(t-1)^{-1}\xi(t-1)Q^1(t-1) \tag{B26}$$

Equilibrium Asset Holdings

The solution for $\Delta Q_1^1(t-1)$ is $-S\Delta Q_B^1(t-1)$. Therefore, the solution for $\Delta Q^1(t-1) = (\Delta Q_1^1(t-1)', \Delta Q_B^1(t-1)')'$ can be written as:

$$\Delta Q^1(t-1) = H_0(t-1) + H_1(t-1)Q^1(t-1). \tag{B27}$$

where,

$$H_0(t-1) = \begin{pmatrix} -S\Pi(t-1)^{-1}\chi(t-1) \\ \Pi(t-1)^{-1}\chi(t-1) \end{pmatrix}, \quad \text{and} \quad H_1(t-1) = \begin{pmatrix} -S\Pi(t-1)^{-1}\xi(t-1) \\ \Pi(t-1)^{-1}\xi(t-1) \end{pmatrix}. \tag{B28}$$

With the above notation, the equilibrium purchases by large participant m in period $t-1$ are given by

$$\Delta Q_m^1(t-1) = S_m[H_0(t-1) + H_1(t-1)Q^1(t-1)] \tag{B29}$$

Additionally, the equilibrium transition dynamics for beginning of period illiquid risky asset holdings are given by:

$$Q^1(t) = G_0(t-1) + G_1(t-1)Q^1(t-1) \tag{B30}$$

where $G_0(t-1) = H_0(t-1)$ and $G_1(t-1) = H_1(t-1) + I$.

Equilibrium Price Function

Recall that the equilibrium price function in each time period maps investors beginning of period holdings of risky assets to an equilibrium price after trade. The equilibrium price function for period $t-1$ is found by plugging the solution for large investors equilibrium trades from equation (B26) into the price schedule faced by large investors (equation (B11)). The resulting price function for illiquid asset in period $t-1$ has form:

$$P^1(t-1, Q^1) = \frac{1}{r}\left(\alpha(t-1) - \Gamma(t-1)Q^1\right) \tag{B31}$$

where,

$$\alpha(t-1) = \beta_0(t-1) - \beta_{Q_B^1}(t-1)\pi(t-1)^{-1}\chi(t-1) \tag{B32}$$

$$\Gamma(t-1) = \beta_Q(t-1) + \beta_{Q_B^1}(t-1)\pi(t-1)^{-1}\xi(t-1) \tag{B33}$$

Large Investors Consumption

Large investors optimal time $t-1$ consumption depends on optimal time $t-1$ trades. After plugging the expressions for equilibrium prices, and equilibrium trades [equations (B30), (B31), and (B27)] into equation (B17), large investors consumption choice problem has form:

$$\max_{C_m(t-1)} -e^{-A_m C_m(t-1)} - \delta k_m(t)e^{rA_m(t)C_m(t-1)} \times \psi_m(Q^1(t-1), W_m(t-1), D(t-1), t-1), \tag{B34}$$

where

$$\begin{aligned}
\psi_m(Q^1, W_m(t-1), t-1) =& e^{-.5\bar{Z}^{2\prime}\Omega_{22}^{-1}\bar{Z}^2 - A_m(t)rW_m(t-1)} \\
&\times e^{+A_m(t)r[S_m(H_0(t-1)+H_1(t-1)Q^1(t-1)]'(\alpha(t-1)-\Gamma(t-1)Q^1(t-1))/r} \\
&\times e^{-A_m(t)(G_0(t-1)+G_1(t-1)Q^1(t-1))'\bar{v}_m(t)} \\
&\times e^{.5A_m(t)^2[G_0(t-1)+G_1(t-1)Q^1(t-1)]'\theta_m(t)[G_0(t-1)+G_1(t-1)Q^1(t-1)]}
\end{aligned} \tag{B35}$$

The first order condition for choice of consumption implies that optimal consumption is given by:

$$C_m(t-1) = \frac{-1}{A_m(t)r + A_m}\ln\left(\frac{\delta k_m(t)A_m(t)r\psi_m(Q^1(t-1), W_m(t-1), t-1)}{A_m}\right) \tag{B36}$$

Large investors value function at time $t-1$

Define $V_m(t-1, Q^1, W_m(t-1))$ as the value function to large investor m from entering period $t-1$ when the vector of illiquid risky asset holdings is Q^1, and his liquid asset holdings are $W_m(t-1)$. After substituting the optimal consumption choice in (B36) into equation (B34),

this value function is given by:

$$V_m(W_m(t-1), Q^1, t-1) = -\left[\frac{1+r_m^*(t)}{r_m^*(t)}\right]\left[\delta k_m(t)r_m^*(t)\psi_m(Q^1, W_m(t-1), t-1)\right]^{\frac{1}{1+r_m^*(t)}}$$

(B37)

where,

$$r_m^*(t) = A_m(t)r/A_m$$

(B38)

Tedious algebra then shows that large investor m's value function at time $t-1$ has form:

$$V_m(t-1, Q^1, W_m(t-1)) = -k_m(t-1) \times e^{-A_m(t-1)W_m(t-1)-A_m(t-1)Q^{1\prime}\Lambda_m(t-1)+.5A_m(t-1)^2Q^{1\prime}\Xi_m(t-1)Q^1}$$

(B39)

where the parameters of the value function at time $t-1$ are given by the following Riccati difference equations.

$$A_m(t-1) = A_m(t)r/(1+r_m^*(t))$$

(B40)

$$k_m(t-1) = \left[\frac{r_m^*(t)+1}{r_m^*(t)}\right]\left[\delta k_m(t)r_m^*(t)\right]^{\frac{1}{1+r_m^*(t)}}$$
$$\times e^{\frac{-.5\bar{Z}^{2\prime}\Omega_{22}^{-1}\bar{Z}^2}{1+r_m^*(t)}}$$
$$\times e^{A_m(t-1)H_0(t-1)'S_m'\alpha(t-1)/r-A_m(t-1)G_0(t-1)'\bar{v}_m(t)/r+.5A_m(t-1)^2((1+r_m^*(t))/r^2)(G_0(t-1)'\theta_m(t)G_0(t-1))}$$

(B41)

$$\Lambda_m(t-1) = -H_1(t-1)'S_m'\alpha(t-1)/r + \Gamma(t-1)'S_mH_0(t-1)/r + G_1(t-1)'\bar{v}_m(t)/r$$
$$- A_m(t-1)(1+r_m^*(t))G_1(t-1)'\left(\frac{\theta_m(t)+\theta_m(t)'}{2}\right)G_0(t-1)/r^2$$

(B42)

$$\Xi_m(t-1) = \frac{-2H_1(t-1)'S_m'\Gamma(t-1)}{rA_m(t-1)} + (1+r_m^*(t))G_1(t-1)'\theta_m(t)G_1(t-1)/r^2$$

(B43)

Small investors optimal consumption

The solution for each small investors consumption depends on small investors optimal trades. To solve for optimal consumptions, I first use equation (B9) to substitute out for Q_s^1 in equation (B8). I then substitute out for $\bar{Z}^1(t) - \beta_{12}\bar{Z}^2(t)$ with the expression:

$$\bar{Z}^1(t) - \beta_{12}\bar{Z}^2(t) = a_0(t-1) + a_1(t-1)Q^1(t-1), \tag{B44}$$

where,

$$a_0(t-1) = \frac{\alpha(t)}{r} - \alpha(t-1) + \bar{D}^1 - \beta_{12}\bar{Z}^2(t) - \frac{\Gamma(t)G_0(t-1)}{r} \tag{B45}$$

$$a_1(t-1) = \Gamma(t-1) - \frac{\Gamma(t)G_1(t-1)}{r}. \tag{B46}$$

Finally I substitute out $Q^1(t)$ with $[G_0(t-1) + G_1(t-1)Q(t-1)]$. With these substitutions, small investors choice of optimal consumptions simplifies to:

$$\max_{C_s(t-1)} -e^{-A_s C_s(t-1)} - \delta k_s(t)e^{rA_s(t)C_s(t-1)} \times \psi_s(Q^1(t-1), W_s(t-1), t-1), \tag{B47}$$

where,

$$\begin{aligned}
\psi_s(Q^1(t-1), W_s(t-1), t-1) &= e^{-A_s(t)rW_s(t-1) - .5\bar{Z}_2'\Omega_{22}^{-1}\bar{Z}_2} \\
&\times e^{-.5[a_0(t-1)+a_1(t-1)Q^1(t-1)]'\Omega_e^{-1}[a_0(t-1)+a_1(t-1)Q^1(t-1)]} \\
&\times e^{-[G_0(t-1)+G_1(t-1)Q^1(t-1)]'\bar{v}_s(t)} \\
&\times e^{-[G_0(t-1)+G_1(t-1)Q^1(t-1)]'\theta_s(t)[G_0(t-1)+G_1(t-1)Q^1(t-1)]}
\end{aligned} \tag{B48}$$

The first order condition for choice of optimal consumption implies that optimal consumption is given by:

$$C_s(t-1) = \frac{-1}{A_s(t)r + A_s}\ln\left(\frac{\delta k_s(t)A_s(t)r\psi_s(Q^1(t-1), W_s(t-1), t-1)}{A_s}\right) \tag{B49}$$

Small investors value function at time $t-1$

Define $V_s(W_s(t-1), Q^1(t-1), t-1)$ as the value function to small investor s from entering period $t-1$ when the vector of illiquid risky asset holdings is $Q^1(t-1)$, and his liquid wealth is $W_s(t-1)$. After substituting the optimal consumption choice in (B49) into equation (B47),

this value function is given by:

$$V_s(W_s(t-1), Q^1(t-1), t-1) =$$

$$- \left[\frac{1 + r_s^*(t)}{r_s^*(t)} \right] \left[\delta k_s(t) r_s^*(t) \psi_s(Q^1(t-1), W_s(t-1), t-1) \right]^{\frac{1}{1+r_s^*(t)}}$$

(B50)

where,

$$r_s^*(t) = A_s(t) r / A_s$$

(B51)

Simplification then shows that the value function has form:

$$V_s(W_s(t-1), Q^1(t-1), t-1) = -K_s(t-1) \, F(Q^1, t-1) \, e^{-A_s(t-1) W_s(t-1)},$$

$$\text{where } F(Q^1(t-1), t-1) = e^{-Q^1(t-1)' \bar{v}_s(t-1) - Q^1(t-1)' \theta_s(t-1) Q^1(t-1)}$$

(B52)

The parameters in the small investors value functions at time $t-1$ are a function of time t parameters as expressed in the following Riccati equations:

$$A_s(t-1) = \frac{r A_s(t)}{1 + r_s^*(t)}$$

(B53)

$$k_s(t-1) = \left[\frac{r_s^*(t) + 1}{r_s^*(t)} \right] \left[\delta k_s(t-1) r_s^*(t) e^{-.5 \bar{Z}^{2'} \Omega_{22}^{-1} \bar{Z}^2} \right]^{\frac{1}{1+r_s^*(t)}}$$

$$\times \text{Exp} \left\{ \frac{-a_0(t-1)' \Omega_e^{-1} a_0(t-1) - G_0(t-1)' \bar{v}_s(t) - G_0(t-1)' \theta_s(t) G_0(t-1)}{1 + r_s^*(t)} \right\},$$

(B54)

$$\bar{v}_s(t-1) = \frac{a_1(t-1)' \Omega_e^{-1} a_0(t-1) + G_1(t-1)' \bar{v}_s(t) + G_1(t-1)' (\theta_s(t) + \theta_s(t)') G_0(t-1)}{1 + r_s^*(t)},$$

(B55)

$$\theta_s(t-1) = \frac{.5 a_1(t-1)' \Omega_e^{-1} a_1(t-1) + G_1(t-1)' \theta_s(t) G_1(t-1)}{1 + r_s^*(t)}$$

(B56)

39

This completes part I of the proof because equations (B39) and (B52) verify that the value functions at time $t-1$ have the same form as at time t.

B.2 Part II

To establish part II of the proof, I need to show that investors value functions for entering entering period T, the last period of trade, has the same functional form as given in the proposition. To establish this result, I first need to solve for investors value function at time $T+1$, the first period when investors cannot trade the illiquid assets (recall they can continue to trade the riskless asset and the liquid assets indefinitely). Then, given this value function, I use backwards induction to solve for investors value function at time T.

Investors Value Functions at Time T+1

Recall that investors are infinitely lived but that from time T onwards they cannot alter their holdings of illiquid assets, but they can continue to alter their consumption, and their holdings of liquid and riskless assets. Because investors cannot trade in period $T+1$ and after, the distinction between small and large investors after this period is irrelevant. Hence, the index m used below could be for either a large or small investor. Using the Bellman principle, the value function $V_m(.)$ of entering period $t+1$ ($t \geq T$) with illiquid asset holdings Q_m^1 and liquid wealth W_m satisfies the functional equation:

$$V_m(W_m(t+1), Q_m^1) = \max_{\substack{C_m(t+1) \\ Q_m^2(t+2)}} -\exp^{-A_m C_m(t+1)} + \delta \, \mathrm{E}\{V_m(W_m(t+2), Q_m^1)\}, \quad t \geq T,$$

(B57)

where,

$$W_m(t+2) = Q_m^{1\prime} D^1(t+2) + Q_m^{2\prime} Z^2(t+2) + r[W_m(t+1) - C_m(t+1)],$$

(B58)

and,

$$Z^2(t+2) = P^2(t+2) + D^2(t+2) - rP^2(t+1).$$

40

Inspection shows that the function

$$V_m(W_m, Q_m^1) = -K_m \exp^{-A_m[1-(1/r)]W_m - A_m[1-(1/r)]Q_m^{1\prime}\frac{(1/r)[\bar{D}^1 - \beta_{12}\bar{Z}^2]}{1-(1/r)} + \frac{1}{2}A_m^2[1-(1/r)]^2 Q_m^{1\prime}\frac{(1/r)\Omega_e}{1-(1/r)}Q_m^1}$$

(B59)

with

$$K_m = \frac{r}{r-1} \times (\delta r)^{\frac{1}{r-1}} \times \exp^{-.5\frac{\bar{Z}^{2\prime}\Omega_{22}^{-1}\bar{Z}^2}{r-1}},$$

satisfies the Bellman equation (B57) for all time periods $\geq T + 1$.

Given the value function at time $T+1$, to solve for investors value functions at time T, I follow the same steps as in equations (B4) through equation (B56). Therefore, substituting in from equation (B59), small investors maximization problem at time T has form:

$$\max_{\substack{C_s(T), \\ Q_s}} -e^{-A_s C_s(T)} - \delta \, \mathrm{E}\left\{K_s(T+1)e^{-A_s(T+1)W_s(T+1) - A_s(T+1)Q_s^{1\prime}\Lambda_s(T+1) + \frac{1}{2}A_s(T+1)^2 Q_s^{1\prime}\Xi_{se}(T+1)Q_s^1}\right\}$$

(B60)

such that,

$$W_s(T+1) = Q_s^{1\prime}Z^1(T+1) + Q_s^{2\prime}Z^2(T+1) + r[W_s(T) - C_s(T)], \tag{B61}$$

where,

$$
\begin{align}
K_s(T+1) &= \frac{r}{r-1} \times (\delta r)^{\frac{1}{r-1}} \times \exp^{-.5\frac{\bar{Z}^{2\prime}\Omega_{22}^{-1}\bar{Z}^2}{r-1}}, \tag{B62}\\
A_s(T+1) &= A_s[1-(1/r)], \tag{B63}\\
\Lambda_s(T+1) &= \frac{(1/r)[\bar{D}^1 - \beta_{12}\bar{Z}^2]}{1-(1/r)}, \tag{B64}\\
\Xi_{se}(T+1) &= \frac{(1/r)\Omega_e}{1-(1/r)}, \tag{B65}\\
Z^1(T+1) &= D^1(T+1) - rP^1(T), \tag{B66}\\
Z^2(T+1) &= P^2(T+1) + D^2(T+1) - rP^2(T). \tag{B67}
\end{align}
$$

Substituting the expression for $W_s(T+1)$ into the value function, taking expectations, and then solving for optimal Q_s^2 given Q_s^1, and substituting that into the value function shows that small investors optimal choice of Q_s^1 and $C_s(T)$ problem has form:

$$\max_{\substack{C_s(T), \\ Q_s^1}} \quad -e^{-A_s C_s(T)} - \delta K_s(T+1) \operatorname{Exp}\left\{-.5\bar{Z}_2'\Omega_{22}^{-1}\bar{Z}_2 - A_s(T)r[W_s(T) - C_s(T)]\right\}$$

$$\times \operatorname{Exp}\left\{-A_s(T+1)Q_s^{1\prime}[\bar{v}_s(T+1) - rP^1(T)] + .5A_s(T+1)^2 Q_s^{1\prime}\Omega_e(T+1)Q_s^1\right\}$$
$$(\text{B68})$$

where

$$\bar{v}_s(T+1) = \left[\frac{\bar{D}^1(T+1) - \beta_{12}\bar{Z}_2(T+1)}{1 - (1/r)}\right] \tag{B69}$$

$$\Omega_e(T+1) = \left[\frac{\Omega_e}{1 - (1/r)}\right] \tag{B70}$$

Integrating the solution for optimal Q_s^1 over the set of small investors then reveals that the net demand for the illiquid assets by the competitive fringe is:

$$Q_1^1(T+1) = \frac{1}{A_1(T+1)}[\Omega_e(t+1)]^{-1}[\bar{v}_s(T+1) - rP(t)] \tag{B71}$$

Following the approach that was used earlier to solve for the price schedule faced by large investors in equation (B11), inverting the small investors demand schedule for the illiquid assets reveals that the price schedule faced by large investors has the form:

$$P^1(.,T) = \frac{1}{r}\left(\beta_0(T) - \beta_{Q^1}(T)Q^1(T) - \beta_{Q_B^1}(T)\Delta Q_B^1(T)\right), \tag{B72}$$

$$\beta_0(T) = \bar{v}_s(T+1) \tag{B73}$$

$$\beta_{Q^1}(T) = A_1(T+1)\Omega_e(T+1)S_1) \tag{B74}$$

$$\beta_{Q_B^1}(T) = -A_1(T+1)\Omega_e(T+1)S \tag{B75}$$

Given the price schedule at time T, and the value function in equation (B59), large investors maximization problem at time T can be written in the form:

$$\max_{\substack{C_m(T), \\ Q_m}} \quad -e^{-A_m C_m(T)}$$

$$-\,\mathrm{E}\left\{\delta K_m(T+1)e^{-A_m(T+1)W_m(T+1)-A_m(T+1)Q^{1\prime}\Lambda_m(T+1)+.5A_m(t)^2 Q^{1\prime}\Xi_m(T+1)Q^1}\right\}$$

$$\text{(B76)}$$

where,

$$A_m(T+1) \;=\; A_m[1-(1/r)] \tag{B77}$$

$$\Lambda_m(T+1) \;=\; S_m'\left[\frac{(1/r)[\bar{D}^1-\beta_{12}\bar{Z}^2]}{1-(1/r)}\right], \tag{B78}$$

$$\Xi_m(T+1) \;=\; S_m'\left(\frac{(1/r)\Omega_e}{1-(1/r)}\right)S_m. \tag{B79}$$

$$K_m(T+1) \;=\; \frac{r}{r-1}\times(\delta r)^{\frac{1}{r-1}}\times\exp^{-.5\frac{\bar{Z}^{2\prime}\Omega_{22}^{-1}\bar{Z}^2}{r-1}} \tag{B80}$$

Substituting in the budget constraint, liquid wealth at the beginning of time $T+1$ is given by

$$\begin{aligned}
W_m(T+1) =\,& Q_m^1(T+1)'D^1(T+1) + Q_m^2(T+1)'Z^2(T+1) \\
& + r(W_m(T) - \Delta Q_m^1(T)'P^1(T,.) - C_m(T))
\end{aligned} \tag{B81}$$

Large investors maximization problem at time T has exactly the same form as given in equation (B15). Therefore, the optimal trades and consumption of large investors follow precisely the same equations as given in Part I of the proof. Large investors value function at time T also has the same functional form as in part I. The equilibrium price function at time T also has the same functional form as in part I. Therefore, to complete the proof, it suffices to solve for small investors consumption and then value function and verify that the value function has the appropriate functional form.

To do so, note that from equation (B68), it is straightforward to show that the optimal choice of $Q_s^1(T+1)$ is

$$Q_s^1(T+1) = \frac{1}{A_s(T+1)}[\Omega_e(T+1)]^{-1}\times[\bar{v}_s(T+1)-rP^1(T)],$$

and that after substituting this expression back in the value function, and making the sub-

stitution $P^1(T) = \frac{1}{r}(\alpha(t) - \Gamma(t)Q^1(t))$, then the maximization in equation (B68) simplifies to have the form:

$$\max_{C_s(T)} -e^{-A_s C_s(T)} - \delta K_s(T+1) \operatorname{Exp}\{A_s(T)rC_s(T)\} \times \Psi_s(T, Q^1) \qquad \text{(B82)}$$

where,

$$
\begin{aligned}
\Psi_s(T, Q^1) = {} & \operatorname{Exp}\left\{-.5\bar{Z}_2'\Omega_{22}^{-1}\bar{Z}_2 - A_s(T)rW_s(T)\right\} \\
& \times \operatorname{Exp}\left\{-.5[\bar{v}_s(T+1) - \alpha(T)]'[\Omega_e(T+1)]^{-1}[\bar{v}_s(T+1) - \alpha(T)]\right\} \\
& \times \operatorname{Exp}\left\{-Q^1(T)'\Gamma(T)'[\Omega_e(T+1)]^{-1}[\bar{v}_s(T+1) - \alpha(T)]\right\} \\
& \times \operatorname{Exp}\left\{-.5Q^1(T)'\Gamma(T)'[\Omega_e(T+1)]^{-1}\Gamma(T)Q^1(T)\right\}
\end{aligned}
\qquad \text{(B83)}
$$

Using the same approach that was used to solve for large investors optimal consumption and then value function in part I of the proof, tedious algebra shows that small investors value function at time T has form

$$-F(Q^1, T)K_s(T) \operatorname{Exp}(-A_s(T)W_s(T))$$

where, $F(Q^1, T) = e^{-Q^1(T)'\bar{v}_s(T) - Q^1(T)'\theta_s(T)Q^1(T)}$,

$$
\begin{aligned}
r_s^*(T+1) &= A_s(T+1)r/A_s, & \text{(B84)} \\
A_s(T) &= A_s(T+1)r/(1 + r_s^*(T+1)), & \text{(B85)}
\end{aligned}
$$

$$
\begin{aligned}
K_s(T) = {} & \left[\frac{r_s^*(T+1)+1}{r_s^*(T+1)}\right] [\delta K_s(T+1)r_s^*(T+1)]^{\frac{1}{1+r_s^*(T+1)}} \\
& \times \operatorname{Exp}\left(\frac{-.5\bar{Z}^{2\prime}\Omega_{22}^{-1}\bar{Z}^2 - .5[\bar{v}_s(T+1) - \alpha(T)]'[\Omega_e(T+1)]^{-1}[\bar{v}_s(T+1) - \alpha(T)]}{1 + r_s^*(T+1)}\right),
\end{aligned}
$$
$$\text{(B86)}$$

$$\bar{v}_s(T) = \frac{\Gamma(T)'[\Omega_e(T+1)]^{-1}[\bar{v}_s(T+1) - \alpha(T)]}{1 + r_s^*(T+1)}, \qquad \text{(B87)}$$

$$\theta_s(T) = \frac{\Gamma(T)'[\Omega_e(T+1)]^{-1}\Gamma(T)}{1 + r_s^*(T+1)}. \qquad \text{(B88)}$$

This completes the proof by establishing that large and small investors value functions take the hypothesized form in all periods that involve trade. □

C   Solutions for Value Function Parameters

Proposition 3 *For all time periods $t = 1, \ldots, T$, and for large investors $m = 2, \ldots M$:*

$$\bar{v}_m(t) = \frac{S'_m(\bar{D}^1 - \beta_{12}\bar{Z}^2)}{1 - (1/r)} \tag{C1}$$

$$\alpha(t) = (\bar{D}^1 - \beta_{12}\bar{Z}^2) \tag{C2}$$

$$A_m(t) = A_m[1 - (1/r)] \tag{C3}$$

$$r^*(t) = r - 1 \tag{C4}$$

$$k_m(t) = \left(\frac{r}{r-1}\right) \times (\delta r)^{\frac{1}{r-1}} e^{-.5 \frac{\bar{Z}^{2\prime}\Omega_{22}^{-1}\bar{Z}^2}{r-1}} \tag{C5}$$

Proof:

For $\bar{v}_m(t)$ and $\alpha(t)$:

The proof is by induction. First, suppose that the results for $\bar{v}_m(t)$ and $\alpha(t)$ are true at time t. Then, from equation (B12), $\beta_0(t-1) = \alpha(t)$. This implies that from equation (B23) that $(-S_1 + S_m)\bar{v}_m(t) - \beta_0(t-1) = 0$. As a result $\chi(t-1) = 0$, which implies from equation (B32) that $\alpha(t-1) = \beta_0(t-1)$ and from equations (B28) and (B30) that $H_0(t-1) = G_0(t-1) = 0$. Substituting for $H_0(t-1)$ and $G_0(t-1)$ in equation (B42) and simplifying then shows:

$$\Lambda_m(t-1) = S'_m\alpha(t)/r. \tag{C6}$$

Finally, substituting this result in equation (B18) proves the result for $\bar{v}_m(t-1)$. To complete the induction, I use equations (B78) and (B18) to solve for $\bar{v}_m(T+1)$; I then substitute the resulting expression as well as the one for $\beta_0(T)$ (equation (B73)) in equation (B23) and use it to show that $\chi(T) = 0$, which implies $G_0(T) = H_0(T) = 0$. Substituting into equation (B32), then shows that $\alpha(T) = \beta_0(T) = S'_m(\bar{D}^1 - \beta_{12}\bar{Z}^2)/[1 - (1/r)]$, which confirms the result for $\alpha(T)$. Finally, given the solutions for $\alpha(T)$ and $\bar{v}_m(T+1)$, substitution in equations (B78) and (B18) confirms the result for $\bar{v}_m(T)$ and completes the induction.

For $A_m(t)$ and $r^*(t)$:

45

The proof is by backwards induction. We know $A_m(T+1) = A_m[1-(1/r)]$ from equation (B77). Using this expression, and iterating on equations (B40) and (B38) proves the result for all times $t = 1, \ldots T$.

For $k_m(t)$:

The proof is by backwards induction. Equation (B80) establishes that it is true at time $T+1$. Plugging the solution for $K_m(T+1)$ into equation (B41) while using the solutions for $r_m^*(t)$ and the result $H_0(t-1) = G_0(t-1) = 0$ confirms the result for periods $1, \ldots T$. $\square$

The next proposition provides information on the value functions of the small investors:

Proposition 4 *For all time periods $t = 1, \ldots, T$, and for each small investor s*

$$a_0(t) = 0 \tag{C7}$$

$$\bar{v}_s(t) = 0 \tag{C8}$$

$$A_s(t) = A_s[1-(1/r)] \tag{C9}$$

$$r_s^*(t) = r - 1 \tag{C10}$$

$$k_s(t) = \left(\frac{r}{r-1}\right) \times (\delta r)^{\frac{1}{r-1}} e^{-.5 \frac{\bar{Z}^{2\prime} \Omega_{22}^{-1} \bar{Z}^2}{r-1}} \tag{C11}$$

Proof:

For $a_0(t)$ and $\bar{v}_s(t)$: Plugging the solutions for $\alpha(t)$ and $G_0(t-1)$ from proposition 3 into equation (B45) shows that $a_0(t) = 0$ for all times t. Since $G_0(t-1) = 0$ for all times t, it then follows from equation (B55) that if $\bar{v}_s(t) = 0$, then so does $\bar{v}_s(t-1)$. To complete the induction, note that substituting the solutions for $\bar{v}_s(T+1)$ (equation (B69)) and $\alpha(T)$ (proposition 3) into equation (B87) confirms the result.

For $A_s(t)$, $r_s^*(t)$, and $k_s(t)$:

The form of the proof is identical to that given in proposition 3. $\square$

Proposition 5 *Assume that for $t \leq T$, conditional on state variable $Q^1(t)$ the Nash Equilibrium trades of the large investors exists and is unique. Then for all $m = 2, \ldots, M$ and $t = 1, \ldots, T$, $\theta_m(t)$ has form:*

$$\vartheta_m(t) \otimes \Omega_e, \tag{C12}$$

where, $\vartheta_m(t)$ is $M \times M$; and

$$\Gamma(t) = \gamma(t) \otimes \Omega_e, \tag{C13}$$

where, $\gamma(t)$ is $1 \times M$.

Proof: The proof is by induction. First, assume that the theorem is true at time t. Then, from equations (B14) and (B13) $\beta_{Q_B}(t-1) = B_{Q_B}(t-1) \otimes \Omega_e$, and $\beta_Q(t-1) = B_Q(t-1) \otimes \Omega_e$, where $B_{Q_B}(t-1)$ is $1 \times M - 1$ and $\beta_Q(t-1)$ is $1 \times M$. Applying these substitutions in large investors reaction functions and then stacking the results reveals that in equation (B25), $\pi(t-1) = \mathcal{P}(t-1) \otimes \Omega_e$ and $\xi(t-1) = Z(t-1) \otimes \Omega_e$. The assumption that the Nash Equilibrium trades in each period are unique implies that $\mathcal{P}(t-1)$ is invertible. Solving for $H_0(t-1)$ and $H_1(t-1)$ then shows that $H_0(t-1) = 0$ and

$$H_1(t-1) = \begin{pmatrix} -S[P(t-1)^{-1}Z(t-1)] \otimes I_{N_1} \\ (P(t-1)^{-1}Z(t-1)) \otimes I_{N_1} \end{pmatrix} \tag{C14}$$

$$= \begin{pmatrix} [-\iota'_M \mathcal{P}(t-1)^{-1}Z(t-1)] \otimes I_{N_1} \\ (\mathcal{P}(t-1)^{-1}Z(t-1)) \otimes I_{N_1} \end{pmatrix} \tag{C15}$$

$$= \mathcal{H}_1(t-1) \otimes I_{N_1} \tag{C16}$$

where ι_M is a $1 \times M$ vector of ones, and $\mathcal{H}_1(t-1)$ is $M \times M$. Since $G_1(t-1) = H_1(t-1) + I_{N_1 M}$, it follows that $G_1(t-1) = \mathcal{G}_1(t-1) \otimes I_{N_1}$ for $\mathcal{G}_1(t-1) = \mathcal{H}_1(t-1) + I_M$. From here, substitution in equation (B33) shows that $\Gamma(t-1) = \gamma(t-1) \otimes \Omega$ and substitution in equation (B43) and (B19) shows that $\theta_m(t-1) = \vartheta_m(t-1) \otimes \Omega$. To complete the induction, I substitute the expression for $\xi_m(T+1)$ (equation (B79)) into equation (B19) and show that the result is true for $\theta_m(T+1)$. Then, following steps similar to those in the first part of the induction, it is straightforward to show that the result holds for $\Gamma(T)$ and $\theta_m(T)$, which completes the induction. $\square$.

Corollary 2 *For each small investors, and for each time period $t = 1, \ldots T$,*

$$\theta_s(t) = \vartheta_s \otimes \Omega_e,$$

where ϑ_s is $M \times M$.

Proof: Straightforward induction involving application of the results from proposition 5.

D Proofs of Asset Pricing Propositions

Proposition 6 *When asset markets are imperfectly competitive as specified in section 2 of the text, then if market participants initial asset holdings are Q^{1W}, then investors will hold Q^{1W} forever, and asset prices and expected returns will be the same as when there is perfect competition.*

Proof: When investors risky asset holdings are Q^{1W}, then investors asset holdings are identical to those associated with a competitive equilibrium and complete markets in which trading is restricted to the set of market participants that has been modeled. Hence, when trade in the first set of assets is restricted to be among the market participants, asset holdings are pareto optimal in all time periods; and investors asset holdings will remain at Q^{1W} because investors have no basis to trade away from asset holdings that are associated with perfect risk sharing. Because Q^{1W} is the vector of asset holdings from a competitive equilibrium, the resulting prices and expected returns which support Q^W are the same as in the competitive equilibrium. $\square$

Corollary 3 *For all $t \geq T$,*

$$[\Gamma(t) - \frac{1}{r}\Gamma(t+1)G_1(t)]Q^{1W} = \lambda_{[X^1]}\Omega_e X^1.$$

Proof: Algebra shows that when asset holdings of asset 1 at time t are Q^{1W}, then excess returns of asset 1 are equal to:

$$P^1(t+1) + \bar{D}^1 - rP^1(t) = \beta_{12}\bar{Z}^2 + [\Gamma(t) - \frac{1}{r}\Gamma(t+1)G_1(t)]Q^{1W}.$$

Proposition 6 shows that when asset holdings are Q^{1W} then the excess returns of asset 1 are $\beta_{12}\bar{Z}^2 + \lambda_{[X^1]}\Omega_e X^1$. Equating the two expressions confirms the claim in the corollary. $\square$.

Proposition 1: *When investors asset holdings of the first asset are not Q^{1W}, then equilibrium expected asset returns satisfy a linear factor model in which one factor is the returns on asset 2, another factor corresponds to perfect risk-sharing, but imperfect diversification of the idiosyncratic risk of asset 1, and the remaining factors correspond to the deviation of large investors asset holdings from those associated with the large investors perfectly sharing the idiosyncratic risk of asset 1.*

Proof: Let Q^{1W} denote the vector of asset holdings of asset 1 that is associated with perfect risk sharing among the investors that trade in asset 1. Manipulation of the equation for equilibrium prices given in proposition 2, and substitution of $G_0(t) + G(t)Q(t)$ for $Q(t+1)$ shows:

$$P^1(t+1)+\bar{D}^1-rP^1(t) = [\frac{1}{r}\alpha(t+1)+\bar{D}^1-\alpha(t)]-[\frac{1}{r}\Gamma(t+1)G_0(t)]+[\Gamma(t)-\frac{1}{r}\Gamma(t+1)G_1(t)]Q^1(t)$$

Plugging in the solution for $\alpha(t) = \alpha(t-1) = [\bar{D}^1 - \beta_{12}\bar{Z}^2]/[1-(1/r)]$ shows the first term in braces on the right hand side of the equation is equal to $\beta_{12}\bar{Z}^2$. The second term in braces is zero since proposition 3 shows that $G_0(t) = 0$. Adding and subtracting Q^{1W} to $Q^1(t)$, the above equation can be rewritten as:

$$P^1(t+1)+\bar{D}^1-rP^1(t) = \beta_{12}\bar{Z}^2+[\Gamma(t)-\frac{1}{r}\Gamma(t+1)G_1(t)](Q(t)-Q^W)+[\Gamma(t)-\frac{1}{r}\Gamma(t+1)G_1(t)]Q^{1W}$$
(D1)

Using the fact that $Q_1^1 = X^1 - SQ_B^1$, the vector $Q^1(t) - Q^{1W}$ can be expressed in terms of the deviations of large investors asset holdings from pareto optimal asset holdings:

$$
\begin{aligned}
Q^1(t) - Q^{1W} &= \begin{bmatrix} (X^1 - SQ_B^1) - (X^1 - SQ_B^{1W}) \\ Q_B^1 - Q_B^{1W} \end{bmatrix} \\
&= \begin{bmatrix} -S \\ I \end{bmatrix}(Q_B^1 - Q_B^{1W})
\end{aligned}
$$

Applying the substitution for $Q^1(t) - Q^{1W}$, and the result of corollary 3 in equation (D1) shows

$$P^1(t+1)+\bar{D}^1-rP^1(t) = \beta_{12}\bar{Z}^2+\lambda_{[X^1]}\Omega_e X^1+[\Gamma(t)-\frac{1}{r}\Gamma(t+1)G_1(t)]\begin{pmatrix} -S \\ I \end{pmatrix}(Q_B^1(t)-Q_B^{1W})$$

Finally, applying the algebra used in the derivation of proposition 5 shows

$$[\Gamma(t)-\frac{1}{r}\Gamma(t+1)G_1(t)]\begin{pmatrix} -S \\ I \end{pmatrix} = \lambda(t)\otimes\Omega_e$$
(D2)

where $\lambda(t)$ is $1 \times M - 1$. Making this substitution then shows:

$$
\begin{aligned}
P^1(t+1) + \bar{D}^1 - rP^1(t) &= \beta_{12}\bar{Z}^2 + \lambda_{[X^1]}\Omega_e X^1 + [\lambda(t) \otimes \Omega_e](Q_B^1(t) - Q_B^{1W}) \\
&= \beta_{12}\bar{Z}^2 + \lambda_{[X^1]}\Omega_e X^1 + \sum_{m=2}^{M} \lambda(m,t)\Omega_e(Q_m^1(t) - Q_m^{1W})
\end{aligned}
$$
(D3)

where $\lambda(m,t) = \lambda(t)s'_{m-1}$. $\square$.

Corollary 1: *When asset holdings at time t are not efficient, then asset returns at time $t + \tau$ follow a factor model in which the market portfolio, the portfolio of segment residual risk, and the deviation of large investors time t asset holdings from efficient asset holdings are factors.*

 Proof: Iterating equation (D1), by τ periods shows:

$$
P^1(t+\tau+1) + \bar{D}^1 - rP^1(t+\tau) = \beta_{12}\bar{Z}^2 + [\Gamma(t+\tau) - \frac{1}{r}\Gamma(t+1+\tau)G_1(t+\tau)](Q^1(t+\tau) - Q^{1W})
$$

$$
+ [\Gamma(t+\tau) - \frac{1}{r}\Gamma(t+\tau+1)G_1(t+\tau)]Q^{1W}.
$$
(D4)

Iterating the equation for equilibrium trades in each period shows

$$
Q^1(t+\tau) = [\prod_{j=0}^{\tau-1} G_1(t+j)]Q^1(t).
$$

Additionally, because the investors will not trade away from efficient asset holdings, it also follows that

$$
[\prod_{j=0}^{\tau-1} G_1(t+j)]Q^{1W} = Q^{1W}.
$$

Making both of these substitutions in equation (D4) shows that:

$$
\begin{aligned}
P^1(t+\tau+1) + \bar{D}^1 - rP^1(t+\tau) &= \beta_{12}\bar{Z}^2 + \lambda_{[X^1]}\Omega_e X^1 + [\lambda(t,\tau) \otimes \Omega_e](Q^1(t) - Q^{1W}) \\
&= \beta_{12}\bar{Z}^2 + \lambda_{[X^1]}\Omega_e X^1 + \sum_{m=2}^{M} \lambda_m(t,\tau)\Omega_e(Q_m^1(t) - Q_m^{1W})
\end{aligned}
$$

where,

$$
\lambda(t,\tau) \otimes \Omega_e = [\Gamma(t+\tau) - \frac{1}{r}\Gamma(t+1+\tau)G_1(t+\tau)]\prod_{j=0}^{\tau-1} G_1(t+j),
$$

and $\lambda_m(t,\tau) = \lambda(t,\tau)S'_{m-1}$ $\square$.

D.1 Competitive Benchmark Model

It is useful to contrast the behavior in the multi-market model with large investors with the behavior of asset prices and trades in the same model when all investors are price takers and can trade forever.

In this infinite period set-up with competitive markets, the equilibrium risk-premium should be time invariant. Denote this risk premium by ρ, where,

$$\rho = \begin{pmatrix} \rho^1 \\ \rho^2 \end{pmatrix} = \begin{pmatrix} \bar{Z}^1 \\ \bar{Z}^2 \end{pmatrix} = \begin{pmatrix} P^1(t+1) + \bar{D}^1 - rP^1(t) \\ P^2(t+1) + \bar{D}^2 - rP^2(t) \end{pmatrix} \tag{D5}$$

Note that $\bar{Z}^2$ is taken as exogenous. The goal is to solve for $\bar{Z}^1$ and P^1 that makes the prices of the first group of assets (the ones that are illiquid in the imperfect competition model) consistent with equilibrium in all time periods.

Solving the equation for $P^1(t)$ forward while imposing the transversality condition $\lim_{t\to\infty} r^{-t}P^1(t) = 0$, shows that

$$P^1(t) = \frac{\bar{D}^1 - \rho^1}{r - 1}$$

for all time periods t.

Given the hypothesized behavior of prices, it remains to solve for ρ^1 and then to show that the hypothesized behavior of prices is consistent with equilibrium.

The function,

$$V_m(W, t) = -\frac{r}{r-1}(r\ \delta)^{\frac{-1}{r-1}} \exp^{-A_m(1-(1/r))W - \frac{.5\bar{Z}^{2\prime}\Omega_{22}^{-1}\bar{Z}^2}{r-1} - \frac{.5\rho^{1\prime}\Omega_e^{-1}\rho^1}{r-1}}$$

and the risk premium solution

$$\rho^1 = \bar{Z}^1 = \beta_{12}\bar{Z}^2 + \lambda_{[X^1]}\Omega_e X^1, \tag{D6}$$

where,

$$\lambda_{[X^1]} = \frac{(1 - (1/r))}{\sum_{m=1}^M (1/A_m)} \tag{D7}$$

satisfies the Bellman equation,

$$V_m(W, t) = \max_{\substack{C_m(t), \\ Q_m^1(t), \\ Q_m^2(t)}} -e^{-A_m C_m(t)} + E_t\{\delta V_m(W(t+1), t+1)\},$$

such that,

$$W(t+1) = Q_m^1(t)'Z^1(t) + Q_m^2(t)'Z^2(t) + r[W(t) - C_m(t)].$$

In addition, in the competitive equilibrium, investors optimal choices of Q_m^1 are constant through time, and are market clearing for the hypothesized ρ^1. Investor $m's$ competitive equilibrium holdings of Q_m^1 is denoted by Q_m^{1W} and is equal to

$$Q_m^{1W} = \frac{(1/A_m)X^1}{\sum_{m=1}^M (1/A_m)}, \quad m = 1, \ldots M. \tag{D8}$$

Substituting the hypothesized ρ^1 into the expression for equilibrium P^1, it follows that in a competitive equilibrium, the equilibrium price is given by

$$P^1(t) = \frac{\bar{D}^1 - \beta_{12}\bar{Z}^2}{r-1} - \frac{\Omega_e X^1}{r \sum_{m=1}^M \frac{1}{A_m}}, \quad t = 1, \ldots \infty \tag{D9}$$

E   Solving the IPO Allocation and Price-Setting Problem

The underwriter's problem in equations (16) - (18) problem requires that he maximize a nonlinear objective function subject to the equality constraint that the total issue is allocated at the IPO or sold in the aftermarket, and subject to a set of nonlinear participation constraints. To solve the maximation, the participation constraints were expressed in terms of investors certainty equivalent wealth. The transformed participation constraints are quadratic in the state variables. Additionally, the equality constraint was used to express X_u in terms of the size of the issue, and the allocations to other investors:

$$X_u = X^{IPO} - \sum_{m=1}^M X_m^{IPO}$$

To economize on notation, I will drop the "IPO" superscript below in what follows. The Lagrangian for the transformed maximization problem is:

$$\mathcal{L} = \max_{P,X} \Pi(P, X) + \sum_{m=1}^{M} \lambda_m \; \psi_m \tag{E1}$$

The necessary conditions for an optimum are given by the Kuhn-Tucker conditions:

$$\mathcal{L}_P = \Pi_P + \sum_{m=1}^{M} \lambda_m \; \psi_{m,P} \leq 0 \text{ with c.s.} \tag{E2}$$

and for $j = 1, \ldots M$:

$$\mathcal{L}_{X_j} = \Pi_{X_j} + \sum_{m=1}^{M} \lambda_m \; \psi_{m,X_j} \leq 0 \text{ with c.s.} \tag{E3}$$

If an investor purchases any shares at the IPO, then his utility is decreasing in the IPO offer price. Therefore, if shares are issued at the IPO, then the incentive constraints must be binding for some investors, because if they were not the underwriter could profit by raising the offer price until the incentive constraints do bind.

Although it is clear the incentive constraints must bind for at least one investor, it is not clear for how many other investors the constraints will bind. I solved the model in the special case when there are 6 investors ($M = 6$) and the only asset in the segment is the new issue. To solve the model, I assumed that the underwriters short position is limited, i.e. $X_u \geq -LB$. With this added constraint, the feasible choices of X_m lie on a simplex. I discretized the simplex so that each investor's post IPO asset holdings could take one of 21 values. Given a vector of asset holdings, I then solved for the highest IPO offer price that satisfies all investors incentive compatibility constraints,[37], and then evaluated the objective function. The point on the discretized simplex for which the objective function was highest was treated as being in a neighborhood of the global optimum, and was used as a starting point for a second stage optimization. Additionally, the investor or investors for whom the incentive compatibility constraints were binding, were deemed as the investors for whom the constraints will be binding at the optimum.

The second stage optimization minimized the squared norm of the Kuhn-Tucker conditions while checking for changes in the set of investors for whom the constraints are binding.

---

[37]This step is simple because the price that makes each investors incentive constraint bind has a closed form expression.

A quadratic hillclimbing algorithm with safeguards was used in the second stage optimization. When the investors all differ in their risk aversion, as they do in some of the examples in the paper, then the incentive compatability constraints tended to bind for a single investor. Under this circumstance, the Kuhn Tucker condition for price produce an analytical expression for λ_m, which simplifies the second-stage optimization. A complicating factor in the second stage is that the Hessians of the first and second stage objective functions are singular. I believe (but cannot show), that Gauss' nonlinear constrained optimization crashed because of this property. To compensate for this problem in the second stage, at each iteration I projected the direction vector onto the range and null spaces of the hessian matrix, and then solved for the projection coefficients that guarantee an increasing step in the objective function. Finally, I found the book <u>Practical Optimization</u>, by Gill, Murray, and Wright (1981) to be a very useful reference for solving nonlinear optimization problems.

F   Market Segmentation and Market Adjusted Abnormal Returns

The purpose of this section is to study the properties of market adjusted abnormal returns when there is market segmentation. The analysis and results are closely related to Merton (1987). The model contains two time-periods and N blocks of risky assets, and a riskfree asset that is in perfectly elastic supply and has return r. The supply of shares outstanding in block i is denoted by the vector X_i. Note that the number of assets can vary from block to block. Investors trade the assets in period 1 and consume in period 2. The period 2 payoff vector per share in block i is equal to D_i. The payoff of D_i is further decomposed into its mean μ_i, a component that is sensitive to systematic risk factor f, and an idiosyncratic component ϵ_i:

$$D_i = \mu_i + a_i f + \epsilon_i \tag{F1}$$

I assume that $f \sim \mathcal{N}(0, \sigma_f^2)$, and $\epsilon_i \sim \mathcal{N}(0, \Omega_i)$. The ϵ_i are uncorrelated across blocks and are uncorrelated with f.

In addition to the N blocks of assets, there is a single asset in zero net supply whose payoff vector is:

$$D_{N+1} = f. \tag{F2}$$

The N'th market can be viewed as a market for sharing systematic risk.

At time period 1, the price vector of the risky assets is denoted by P_i, $i = 1, \ldots N + 1$.

I assume that for some reason that is not specified here, asset markets are segmented. This means that each block of assets has its own set of investors that only take positions in the segment assets and in asset $N+1$. For simplicity, the investors in each block are modelled as a single representative price-taking investor with CARA utility of period 2 wealth and risk tolerance τ_i.

The market portfolio has payoff $D_M = \sum_{i=1}^N X_i' D_i$, and price $P_M = \sum X_i' P_i$. Its excess return over the risk free rate is denoted by $Z_m = D_M - r P_M$; and the excess return of the assets in block i over the risk free rate is denoted by $Z_i = D_i - r P_i$.

Armed with this notation,

$$\text{Var}(Z_m) = (\sum_{i=1}^N X_i' a_i)^2 \sigma_f^2 + \sum_{i=1}^N X_i' \Omega_i X_i; \tag{F3}$$

and,

$$\text{Cov}(Z_i, Z_M) = a_i (\sum_{i=1}^N X_i' a_i) \sigma_f^2 + \Omega_i X_i \tag{F4}$$

If I use the primitive assets to construct N assets with payoffs $\mu_i + \epsilon_i$, and a single asset with payoff f, then the new assets will span the same space as the old, but they are easier to work with. In particular, working from the new assets, it is straightforward to show that for each block i, under the assumption of market segmentation,

$$\bar{Z}_i = \tau_i^{-1} \Omega_i X_i + a_i [\sum X_i' a_i][\sum \tau_i^{-1}] \sigma_f^2 \tag{F5}$$
$$= \eta_i \Omega_i X_i + \lambda_M \text{Cov}(Z_i, Z_M) \tag{F6}$$

where $\lambda_M = [\sum \tau_i]^{-1}$, and $\eta_i = \tau_i^{-1} - \lambda_M$, and $\bar{Z}_i = \text{E} Z_i$.

Using the above expression to solve for the market's excess expected return then shows:

$$\bar{Z}_M = \sum_{i=1}^N \eta_i X_i' \Omega_i X_i + \lambda_M \text{Var}(Z_M) \tag{F7}$$

The expression for $\bar{Z}_M$ contains two components. The second component is the standard risk premium for an assets covariance with market risk when investors fully diversify their

asset holdings. The first component is an additional premium for imperfect diversification. Examination will show that each term of the imperfect diversification premium is positive, which implies that imperfect diversification increases the return on the market portfolio.

Below, I seek to examine how imperfect diversification affects market adjusted abnormal returns. Letting $\beta_i = \text{Cov}(Z_i, Z_M)/Var(Z_M)$, algebra then shows:

$$\bar{Z}_i - \beta_i \bar{Z}_M = \eta_i \Omega_i X_i - \beta_i \sum \eta_i X_i' \Omega_i X_i \qquad (F8)$$
$$= \alpha_i \qquad (F9)$$

The term on the right hand of equation (F8) is the vector of Jensen's alphas for block i that is due to market segmentation. These alphas capture the risk premium for inefficiently sharing risk across market segments (or blocks). My goal here is to characterize the average behavior of the $alpha_i$. Simple algebra shows $\sum X_i' alpha_i = 0$. Because $X_i >> 0$, it follows that some of the market corrected $alpha_i$ must be positive and some must be negative. Additionally, if the number of shares outstanding of each asset is the same, then the average value of α_i is equal to 0. This suggests that to a first approximation when the market is perfectly competitive, and there is segmentation, then the average market adjusted α is 0.

When there is imperfect competition, I showed that asset returns have two components.[38] The first component is compensation for imperfect risk sharing across blocks (or segments), and is identical to the component above for the competitive case. The second component is for inefficient risk sharing among the investors who trade within a segment. Because investors within each segment eventually share risks optimally, this second component is transitory. For simplicity here, I will write it as π_i. π_i should be close to 0 for segments that have not experienced an IPO recently, but may be nonzero if they did. Note: if the first asset in a segment is an IPO, and the risks of the other assets are shared efficiently within the segment, then the allocation distortions at the IPO will cause π_i to depart from 0 for all assets within the segment, but the size of the departure for each asset is proportional to the first column of Ω_i. If Ω_i is diagonal, then the other assets in the segment are not affected by the IPO; but they will be otherwise. To examine the role that this term plays in market adjusted excess returns, note that π_i should be added to each segments excess returns, and

---

[38]When I speak of imperfect competition here, it is only over a single period of trade, which is a special case of the dynamic model in the paper. I have full confidence that all of the reasoning in this subsection can be written in terms of the general dynamic model in the text, but for now it is easier not to do so formally.

$\sum X'_i \pi_i$ should be added to the excess return on the market. Arithmetic then shows that:

$$\bar{Z}_i - \beta_i \bar{Z}_M = \alpha_i + \left[\pi_i - \beta_i \sum X'_i \pi_i \right] \tag{F10}$$

The second term on the right hand side of the above equation reflects how an IPO affects the returns within a particular market segment. To get a feel for the magnitude of this term, with great loss of generality suppose for a moment that an IPO occurs only in segment 1, that there is only 1 risky asset per segment, and that the characteristics of the assets in each segment are identical (same number of shares, etc). Then π_i is nonzero in segment 1, and 0 in the other $N-1$ segments. Algebra shows that $X'_i \beta_i = 1/N$; solving for β_i and substituting in the right hand side of (F10) then shows that if segment 1 has an IPO then

$$\begin{aligned} \bar{Z}_1 - \beta_1 \bar{Z}_M &= \alpha_1 + \left[\pi_1 - \frac{1}{NX_1} \sum X'_i \pi_i \right] \\ &= \alpha_i + \pi_1 [1 - (1/N)], \end{aligned}$$

where the last line follows because by assumption in this specialized case, $X_i = X_1$, and $\pi_i = \pi_1$ for all i. Generalizing a bit, since α_i is 0 on average, this analysis suggests that averaging across IPOs through time and across segments, the average market adjusted returns on IPOs will be $\pi_1 [1 - \phi]$ where ϕ is the fraction of segments that have an IPO. This implies that IPO's should underperform after adjusting for market risk.

BIBLIOGRAPHY

Aggarwal, R., 2003, "Allocation of initial public offerings and flipping activity," *Journal of Financial Economics*, 68, 111-135.

Benveniste, L.M., and P.A. Spindt, 1989, "How Investment Bankers Determine the Offer Price and Allocation of New Issues," *Journal of Financial Economics* 24, 343-61.

Biais, B., Bossaerts, P., and J.C. Rochet, 2002, "An Optimal IPO Mechanism," *Review of Economic Studies* 69, 117-46.

Boehmer, E., and R.P.H. Fishe, 2000, "Do Underwriters Encourage Stock Flipping? A New Explanation for the Underpricing of IPOs," Mimeo.

Booth, J.R., and L. Chua, 1996, "Ownership dispersion, costly information, and IPO underpricing," *Journal of Financial Economics* 41, 291-310.

Brav, A., Geczy, C., and P.A. Gompers, 2000, "Is the abnormal return following equity issuance anomalous?", *Journal of Financial Economics* 56, 209-249.

Butler, A., Grullon, G., and J. Weston, 2005 "Stock Market Liquidity and the Cost of Issuing Equity," *Journal of Financial and Quantitative Analysis,* 40, 331-348.

Cornelli, F., and D. Goldreich, (2003): "Bookbuilding: How Informative is the Order Book?," *Journal of Finance,* 58, 1415-44.

Corwin, S.A., Harris, J.H., and M.L. Lipson, 2004, "The Development of Secondary Market Liquidity for NYSE-Listed IPOs," *Journal of Finance,* 59, no. 5, 2339-2373.

DeMarzo, P.M., and B. Urosevic, 2000 "Optimal Trading by a Large Shareholder," Working Paper, Haas School of Business, University of California at Berkeley.

de Jong, F. and M. Dahlquist, 2003, "Pseudo-Market Timing, Fact or Fiction," Mimeo, the University of Amsterdam.

Derrien, F., and K.L. Womack, 2003, "Auctions vs. Bookbuilding and the Control of Underpricing in Hot IPO Markets," *Review of Financial Studies* 16, no. 1, 31-61.

Eckbo, B.E., and Ø. Norli, 2002, "Liquidity risk, leverage, and long-run IPO returns," Mimeo, Dartmouth Tuck School of Business.

Ellis, K., Michaely, R., and M. O'Hara, 2000, "When the Underwriter is the Market Maker: An Examination of Trading in the IPO Aftermarket," *Journal of Finance,* 55, no. 3, 1039-1074.

Ellis, K., Michaely, R., and M. O'Hara, 2002, "The Making of a Dealer Market: From Entry to Equilibrium in the Trading of NASDAQ Stocks," *Journal of Finance* 57, no. 5, 2289 - 2316.

Ellul, A., and M. Pagano, 2003, "IPO Underpricing and After-Market Liquidity," CSEF Working Paper no. 99, University of Salerno Centre for Studies in Economics and Finance.

Gill, P.E., Murray, W., and M.H. Wright, 1981, Practical Optimization, Academic Press, London.

Hahn, T., and J.A. Ligon, 2004, "Liquidity and Initial Public Offering Underpricing," Mimeo, College of Business and Economics, The University of Idaho.

Hoberg, G., 2004, "Strategic Underwriting in Initial Public Offers," Yale ICF Working Paper No. 04-07.

Jenkinson, T., and H. Jones, 2004, "Bids and Allocations in European IPO Bookbuilding," *Journal of Finance* , 59, no. 5, 2309-2338.

Ljungqvist, A.P., Nanda, V., and R. Singh, 2003, "Hot Markets, Investor Sentiment, and IPO Pricing," Mimeo, Stern School of Business, New York University.

Ljungvist, A., 2004, "IPO Underpricing," Handbooks in Finance: Empirical Corporate Finance, Chapter III.4, edited by B. E. Eckbo, forthcoming.

Kihlstrom, R., 2001, "Monopoly Power in Dynamic Securities Markets,"Working Paper, The Wharton School, University of Pennsylvania.

Loughran, T., and J.R. Ritter, 2000, "Uniformly Least Powerful Tests of Market Efficiency," *Journal of Financial Economics* 55, 361-389.

Loughran, T., and J.R. Ritter, 2004, "Why has IPO Underpricing Changed over time?", *Financial Management*, 33, no. 3, 5-37.

Merton, R.C., 1987, "A Simple Model of Capital Market Equilibrium with Incomplete Information", *The Journal of Finance*, 42 no. 3, 483-510.

Pritsker, M., 2004, "Large Investors: Implications for Equilibrium Asset Returns, Shock Absorption and Liquidity," Mimeo, The Federal Reserve Board.

Reuter, J., 2005, "Are IPO Allocations for Sale? Evidence from Mutual Funds," *The Journal of Finance*, Forthcoming.

Ritter, J.R., 2005, "Some Factoids About the 2004 IPO Market," mimeo, the University of Florida, http://bear.cba.ufl.edu/ritter.

Ritter, J.R., and I. Welch, 2002, "A Review of IPO Activity, Pricing, and Allocations," *Journal of Finance* 57, no. 4, 1795 - 1828.

Rock, K., 1986, "Why New Issues Are Underpriced," *Journal of Financial Economics* 15, 187-212.

Schultz, P., 2003, "Pseudo Market Timing and the Long-Run Underperformance of IPOs," *Journal of Finance* 58, no. 2, 483-517.

Stapleton, R.C., and M.G. Subrahmanyam, 1978, "A Multi-period Equilibrium Asset Pricing Model," *Econometrica,* 46, 1077-1096.

Urosevic, B., 2002a, "Moral Hazard and Dynamics of Insider Ownership Stakes," Working Paper, Universitat Pompeu Fabra.

Urosevic, B., 2002b, "Essays in Optimal Dynamic Risk Sharing in Equity and Debt Markets," Ph.D. Dissertation, University of California at Berkeley, May.

Vayanos, D., 2001, "Strategic Trading in a Dynamic Noisy Market," *Journal of Finance,* 56, 131-171.

Viswanathan, S, and B. Wei, 2004, "Endogenous Events and Long Run Returns," Mimeo, Duke University.

Westerfield, M.M., 2003, "Market Composition and the Participation Externality in Equity Market Formation," Mimeo, MIT Sloan School of Business.

Table 1: IPO Under-Pricing and Money Left on the Table: I.

A. Investors Risk Bearing Capacity

Investor Number	Type	Risk Bearing Capacity
1	Retail	10.00
2	Institutional	54.56
3	Institutional	21.82
4	Institutional	8.73
5	Institutional	3.49
6	Institutional	1.40
Risk Bearing Concentration (Herfindahl Index)		3543.26

B. IPO Under-Pricing and Money Left on the Table (MLOTT)

	Prices				MLOTT	
Periods Liq	P_Offer	P_Open	P_Comp	% Und_Price	Raw	Liq. Adj.
2000	10.13	14.25	10.50	40.66	164.75	14.90
1800	10.06	14.83	10.50	47.38	190.73	17.38
1600	10.00	15.43	10.50	54.26	217.12	19.79
1400	9.94	16.04	10.50	61.33	243.93	22.23
1200	9.88	16.66	10.50	68.61	271.18	24.72
1000	9.82	17.29	10.50	76.10	298.87	27.24
800	9.75	17.93	10.50	83.82	327.01	29.81
600	9.69	18.58	10.50	91.76	355.60	32.41
400	9.62	19.24	10.50	99.94	384.65	35.06
200	7.95	20.62	10.50	159.36	506.89	101.74

C. IPO Allocation Distortions (Percent)

	Investor Number					
Post-IPO Trading Periods	1	2	3	4	5	6
2000	-100.00	83.29	-100.00	-100.00	-100.00	-100.00
1800	-100.00	83.29	-100.00	-100.00	-100.00	-100.00
1600	-100.00	83.29	-100.00	-100.00	-100.00	-100.00
1400	-100.00	83.29	-100.00	-100.00	-100.00	-100.00
1200	-100.00	83.29	-100.00	-100.00	-100.00	-100.00
1000	-100.00	83.29	-100.00	-100.00	-100.00	-100.00
800	-100.00	83.29	-100.00	-100.00	-100.00	-100.00
600	-100.00	83.29	-100.00	-100.00	-100.00	-100.00
400	-100.00	83.29	-100.00	-100.00	-100.00	-100.00
200	-100.00	83.29	-100.00	-100.00	-100.00	-100.00

Table 1 Continued

Notes: When there is a continuum of small retail investors (investor 1), and 5 large investors (investors 2-6), the table reports expected IPO underpricing and money left on the table for 10 IPOs that vary by the number of days of trading that take place in the IPO aftermarket (Periods Liq). Fewer trading days correspond to less aftermarket liquidity. Each investors risk bearing capacity (panel A) is his own risk tolerance as a percent of the sum total of all investors risk tolerances. Percentage underpricing (panel B) is equal to the net expected one-day return associated with purchasing at the offer price (P_Offer) and then selling at the first days closing price (P_Open). Raw money left on the table is the difference in P_Offer and P_Open times the number of shares issued. Liquidity adjusted money left on the table is equal to the shares issued times the one day return associated with purchasing at P_Offer and then selling into a competitive (and perfectly liquid) aftermarket at price P_Comp. In a competive market, each investors allocation at the IPO should be equal to risk capacity times the number of shares issued. Allocation distortions (panel C) measure are the percentage difference in investors allocations at the IPO relative to the allocation he would have received in an efficient market.

Table 2: IPO Under-Pricing and Money Left on the Table: II.

A. Investors Risk Bearing Capacity

Investor Number	Type	Risk Bearing Capacity
1	Retail	10.00
2	Institutional	40.00
3	Institutional	12.50
4	Institutional	12.50
5	Institutional	12.50
6	Institutional	12.50
Risk Bearing Concentration (Herfindahl Index)		2225

B. IPO Under-Pricing and Money Left on the Table (MLOTT)

	Prices				MLOTT	
Periods Liq	P_Offer	P_Open	P_Comp	% Und_Price	Raw	Liq. Adj.
2000*	10.50	10.50	10.50	0.00	0.00	0.00
1800*	10.50	10.50	10.50	0.00	0.00	0.00
1600	10.53	10.71	10.50	1.70	7.14	-1.38
1400	10.69	11.65	10.50	8.97	38.33	-7.62
1200	10.85	12.61	10.50	16.27	70.59	-14.07
1000	11.01	13.60	10.50	23.49	103.48	-20.65
800	11.18	14.61	10.50	30.62	136.94	-27.34
600	11.35	15.63	10.50	37.65	170.96	-34.14
400	11.52	16.66	10.50	44.59	205.55	-41.06
200	11.70	17.72	10.50	51.43	240.70	-48.09

C. IPO Allocation Distortions (Percent)

|	Investor Number					
Post-IPO Trading Periods	1	2	3	4	5	6
2000*	0.00	0.00	0.00	0.00	0.00	0.00
1800*	0.00	0.00	0.00	0.00	0.00	0.00
1600	-100.00	147.50	-100.00	-100.00	-100.00	-100.00
1400	-100.00	149.57	-100.00	-100.00	-100.00	-100.00
1200	-100.00	149.77	-100.00	-100.00	-100.00	-100.00
1000	-100.00	149.84	-100.00	-100.00	-100.00	-100.00
800	-100.00	149.88	-100.00	-100.00	-100.00	-100.00
600	-100.00	149.90	-100.00	-100.00	-100.00	-100.00
400	-100.00	149.92	-100.00	-100.00	-100.00	-100.00
200	-100.00	149.93	-100.00	-100.00	-100.00	-100.00

Table 2 Continued

Notes: This table is similar to Table 1 except that investors risk bearing capacity is less concentrated. When there are 2000, and 1800 post-IPO trading perids (marked with an asterisk), the aftermarket is sufficiently competitive that investors and the underwriter are nearly indifferent over how shares are allocated. I have assigned efficient share holdings in this case at competitive prices. When the aftermarket is slightly more competitive than in Table 1, but not perfectly competitive, the underwriter can allocate shares at a price that is higher than the competitive price (which is also the price with perfect liquidity). In such circumstances, liquidity adjusted money left on the table is negative.

Figure 1: **IPO Offer Prices and Investor 2's Participation**
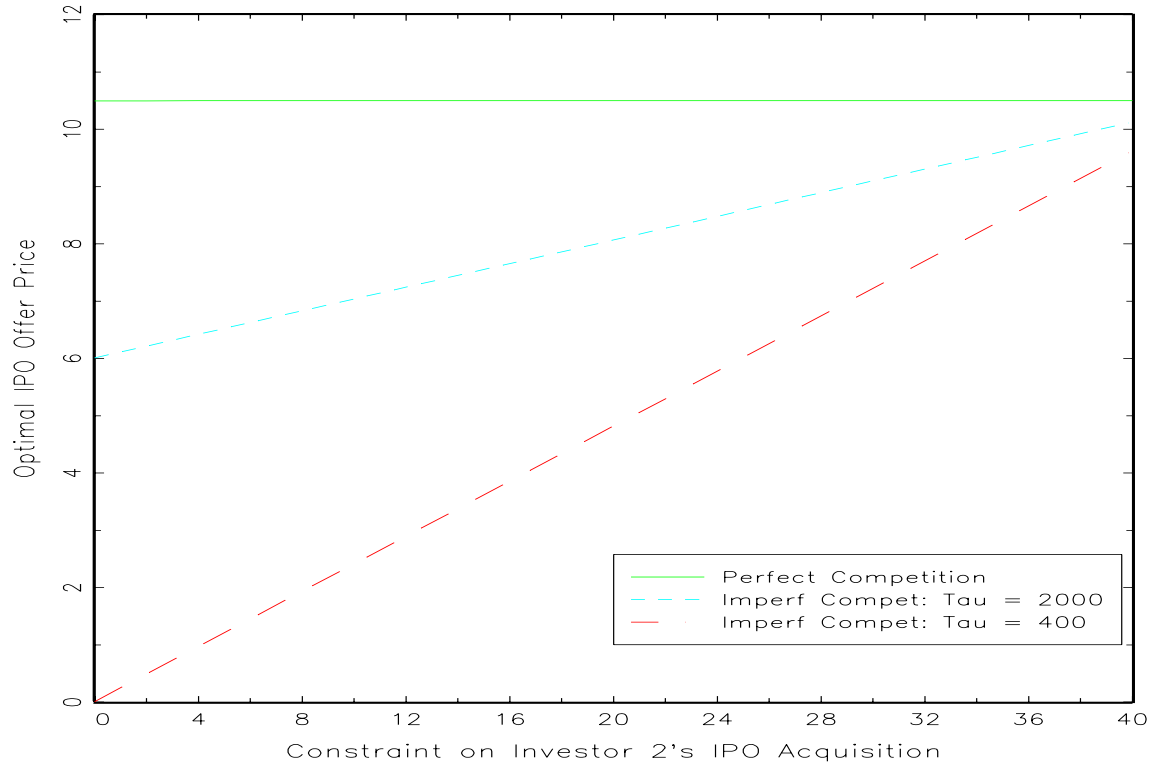
Notes: For the market structure in panel A of Table 1, the figure presents optimal IPO offer prices as a function of constraints on the amount of shares that investor 2 can acquire at the IPO. Results are presented for when the aftermarket is perfectly competitive (nearly flat solid line), and when it is imperfectly competitive and there are 2000 (short dashes) or 400 (long dashes) periods of aftermarket trading following the IPO. The figure shows that when there is imperfect competition, the underwriter should distort share holdings towards investor 2 in order to increase the revenue raised at the IPO.

Figure 2:   **Post-IPO Share Turnover**
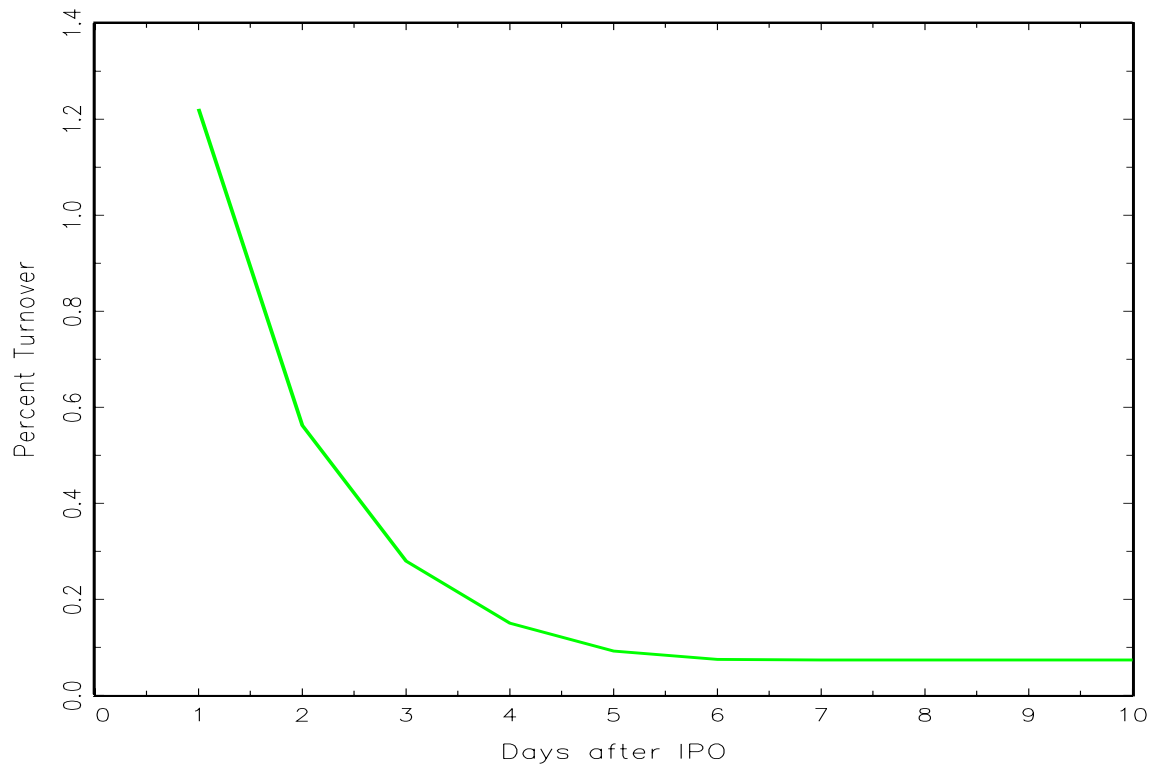
Notes: For the market structure in panel A of Table 1, when there are 2000 periods of trade remaining following the IPO, the figure presents share turnover [(buy volume + sell volume)/2] as a percent of shares outstanding.

Figure 3: **Market-Adjusted Cumulative Abnormal Returns**
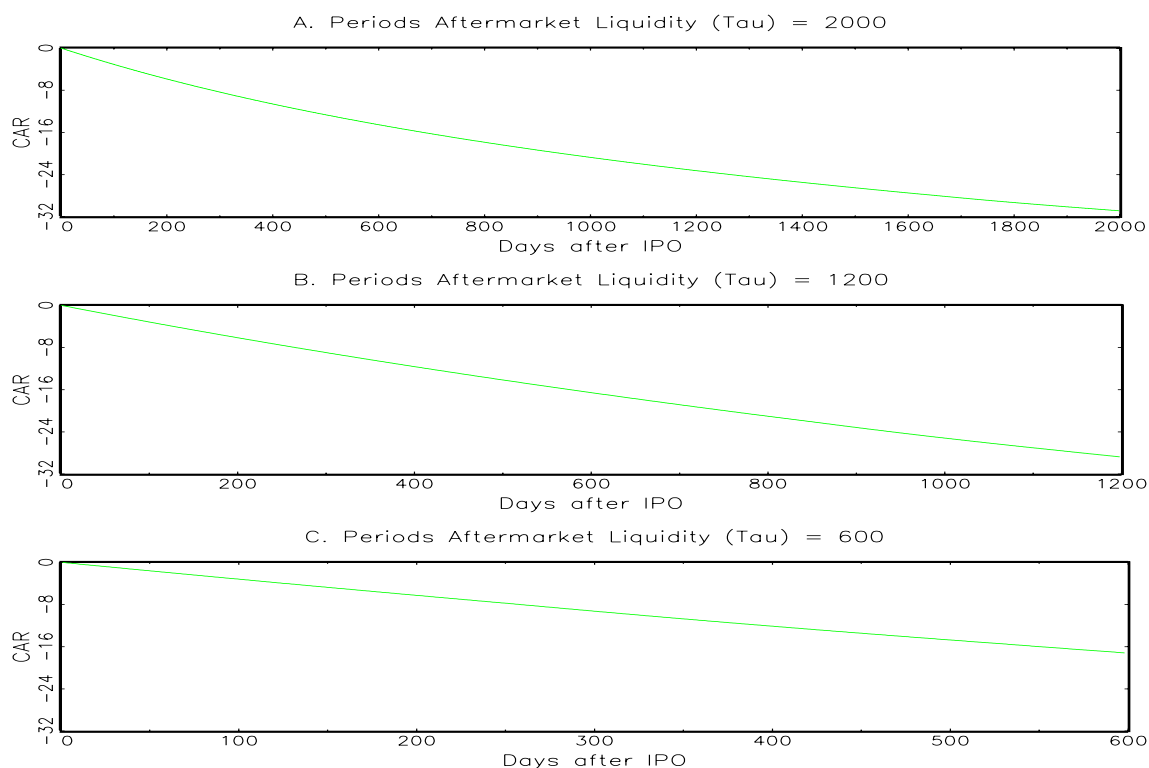
Notes: For the market structure in panel A of Table 1, the figure presents expected cumulative abnormal returns (CARs) relative to the market portfolio for differing numbers of periods of liquid trading in the aftermarket following the IPO. The figure is constructed under the assumption that IPO's are an infinitesimal fraction of the market portfolio. If instead, new issues represent ϕ percent of the market, then the CAR should be adjusted downward in magnitude by approximately ϕ percent. See appendix F for details.