Strategic Trading with Wealth Effects*

Sergei Glebkin, Semyon Malamud, and Alberto Teguia

November 20, 2023

Abstract

We analyze asset prices and liquidity in an economy with large investors and many risky assets. The model allows for general investors' preferences and distributions of asset payoffs. We propose a constructive solution approach: solving for equilibrium reduces to solving nonlinear first-order ODE. We show that the equilibrium is unique under mild restrictions on payoffs and preferences. Liquidity risk is priced in equilibrium, leading to deviations from the consumption-CAPM. In stark contrast to a constant absolute risk aversion (CARA) benchmark, in a model with wealth effects, we obtain (1) illiquidity of risk-free assets (such as, e.g., Treasuries); (2) illiquidity contagion (a sell-off in one asset may have a price impact on assets with unrelated fundamentals) and asymmetry in cross-asset price impacts; (3) market liquidity may decrease in the number of traders and their wealth; and (4) in the presence of liquidity shortage, price impact may become negative giving rise to an illiquidity premium in asset prices; (5) safe assets are more illiquid because they have a larger price impact. In the presence of wealth heterogeneity, large traders trade more but also reduce their demands more. As a group, they account for a smaller fraction of orders compared to small investors. Fatter-tailed wealth distribution makes markets less liquid.

JEL Classification: D21, G31, G32, G35, L11

Keywords: Market Liquidity, Funding Liquidity, Price Impact, Strategic Trading, Wealth Effects

*Sergei Glebkin (glebkin@insead.edu) is at INSEAD, Semyon Malamud (semyon.malamud@epfl.ch) is at EPFL, and Alberto Teguia (alberto.mokakteguia@sauder.ubc.ca) is at UBC Sauder. For valuable feedback, we thank Viral Acharya, Frederico Belo, Bernard Dumas, Asaf Manela, Joel Peress, Haoxiang Zhu, and seminar participants at INSEAD.

Large, institutional investors dominate modern markets. After the 2007-2008 financial crisis, many of these investors have been classified as systemically important financial institutions (SiFi): Institutions whose collapse would pose a serious risk to the global economy. One of the key channels through which SiFis may impact financial markets is through their portfolio liquidation decisions: When hit by a shock, large institutions may need to simultaneously adjust their portfolio holdings, which may lead to large adverse movements in market prices due to the SiFi's (market) price impact and/or their inability to provide (enough) liquidity. While formerly viewed as an artifact of risky, informationally sensitive securities, recent turmoils on the government bond markets show that illiquidity is a major consideration even for extremely liquid, money-like securities.[1] In equilibrium, this illiquidity should be priced: The most illiquid securities should trade at a discount, while liquid securities should trade at a so-called "flight-to-liquidity" premium. In order to understand these effects, we need a theoretical model of strategic behavior in a market populated by large (in terms of the wealth or assets under management) investors who internalize their price impact. The goal of this paper is to develop such a model.

We consider a multi-asset economy populated by a finite number of large, strategic, risk-averse investors endowed with arbitrary preferences. Essentially, no restrictions are imposed on the distribution of asset payoffs. We study symmetric equilibria in a game in which all strategic investors submit identical, multi-dimensional demand schedules indicating the portfolio they are willing to buy/sell for a given vector of prices. We show that, in equilibrium, assets are priced according to the standard consumption Euler equation plus a correction term accounting for market liquidity (price impact), linked to an endogenous measure of systemic risk that puts a large weight on low consumption states. We find that, in stark contrast to models with constant absolute risk aversion (CARA), wealth effects imply that price impact is typically

[1] For example, in October 2022, funding liquidity frictions of British Pension funds (large, strategic investors in British government bonds) triggered a serious turmoil in the bond market, forcing the central bank to intervene. See, e.g., Pinter (2023). Even the market for US Treasury bills, commonly viewed as "money-like" and extremely liquid, is impacted by large money market funds whose strategic behavior affects T-bill rates. See, e.g., Doerr, Eren, and Malamud (2023).

asymmetric. This asymmetry implies the existence of endogenous *systemic assets:* That is, assets whose sell-off triggers large moves in all security prices. These moves can be both positive or negative, implying that we may have both "positively systemic" and "negatively systemic" assets. These results have important implications for liquidity injection programs such as TARP and QE: Understanding which assets are systemic may help policy-makers better shape their asset purchase decisions.

In the model, wealth effects also lead to a natural notion of funding liquidity as the amount of liquid cash reserves available to strategic investors. As a result, market liquidity (price impact) is endogenously linked to funding liquidity in a fully micro-founded way, without the need to introduce margin requirements and/or counter-party risk. Importantly, indefinitely increasing funding liquidity does not necessarily make markets infinitely liquid: What matters is whether funding liquidity (cash) is sufficiently uniformly distributed across the investor population. If cash is only concentrated in the hands of a few SiFis, markets stay highly illiquid, making these SiFis vulnerable to shocks.

In standard, CARA-normal models, price impact is always proportional to the fundamental covariance matrix of the traded assets (Glebkin, Malamud, and Teguia, 2023). In particular, risk-free assets (i.e., assets with zero variance) are always fully liquid in such models. We show that this is not true when preferences exhibit wealth effects: In this case, risk-free bonds are also illiquid, and price impact may happen to be negative, in which case bond prices may contain a flight-to-liquidity premium.

1 Literature Review

Our paper belongs to the large literature on strategic trading and price impact. In our model, information is symmetric, and price impact arises due to traders' limited risk-bearing capacity. We model trade using the classic double auction protocol (also known as the uniform price (divisible) auction) in which traders submit price-contingent demand schedules. See, for example, Kyle (1989), Vayanos and Vila (1999), Vives (2011), Rostek and Weretka (2012), Kyle, Obizhaeva, and Wang (2017), Ausubel, Cramton, Pycia, Rostek, and Weretka (2014), Bergemann, Heumann, and Morris (2015), and Du and Zhu (2017) for the single asset case, as well as Rostek and Weretka (2015b) and Malamud and Rostek (2017) for the multi-asset case.[2] All these papers use the standard CARA-Normal assumption to derive linear equilibria, whereby the slopes of the demand schedules are independent of the price level, and the equilibrium price impact (given by the inverse of the slope of the residual supply) is also constant, independent of the trade size. Linearity of equilibrium depends crucially on the CARA-Normal assumption: It is this assumption that ensures that the marginal value of asset holdings is constant and hence guarantees the existence of linear equilibria.[3]

The double auction model considered in our paper allows us to study strategic liquidity provision in the presence of wealth effects. In particular, for the first time in the literature, we solve a fully micro-founded model linking market liquidity (price impact) and funding liquidity (the capital of strategic traders). Our model offers a new perspective on the classical results Brunnermeier and Pedersen (2009), showing how the two types of liquidity are interlinked in subtle and unexpected ways. In particular, we show how funding liquidity can impact the market liquidity of risk-free assets, a phenomenon that played a key role in the recent bond market events (Pinter (2023)), and that cannot occur in the CARA setting of Brunnermeier and Pedersen (2009).

[2]Sannikov and Skrzypacz (2016) develop an alternative trading protocol that they name "conditional double auction", where traders can condition their demand schedules on trading rates of other players.

[3]The only exception is the case two agents: Then, linear equilibria fail to exist but, as Du and Zhu (2017) show, there often exist non-linear equilibria. There is also a large literature on *competitive* noisy rational expectations equilibria (REE) beyond the CARA-Normal setup with a continuum of non-strategic traders. For example, several papers relax the assumption of normal payoff distributions but maintain the CARA assumption or assume risk neutrality. See, Gennotte and Leland (1990), Ausubel (1990a), Ausubel (1990b), Barlevy and Veronesi (2003), Bagnoli, Viswanathan, and Holden (2001), Yuan (2005), Breon-Drish (2015), Pálvölgyi and Venter (2015), Chabakauri, Yuan, and Zachariadis (2017). These papers assume CARA utilities and do not feature wealth effects. Glebkin, Gondhi, and Kuong (2021) feature wealth effects in a setup with CARA preferences: they study a setup with margin constraints, the tightness of which depends on the wealth level. Peress (2003), Malamud (2015), Avdis and Glebkin (2023) study competitive models with asymmetric information and non-CARA preferences.

With the exception of Rostek and Weretka (2015a), Malamud and Rostek (2017), and Glebkin et al. (2023),[4] all of the above-mentioned papers consider the case of a single risky asset. The closest to our paper is the recent work Glebkin et al. (2023), which shows that, in a generic multi-asset setting with CARA preferences, the price impact matrix is (1) proportional to the risk-neutral covariance matrix of asset payoffs and (2) decreases with the number of agents. In particular, with CARA preferences, the price impact matrix is always symmetric, and positive-semi-definite. The symmetry is a direct consequence of CARA demand schedules: For any pair of assets, an increase in the price of asset 1 changes the demand for asset 2 by the same amount as an increase in the price of asset 2 changes the demand for asset 1. Indeed, for CARA preferences, what matters is the overall level of risk, and diversification benefits are the same for each wealth level. By contrast, both symmetry and positive semi-definiteness break down in the presence of wealth effects. Effectively, this is a consequence of the Slutsky equation, whereby the substitution effects are symmetric across different assets (goods), while the income effect is asymmetric. As a result, wealth effects imply the existence of systemic assets, that is, assets whose price change has an asymmetrically large impact on the demand for other assets in the economy.[5]

There is a growing literature emphasizing the importance of institutional investors in modern financial markets. Allen (2001) points out that financial crises are associated with liquidity shortage and argues that the effect of liquidity on asset prices should be endogenous. Basak and Pavlova (2013) study the effects of institutional investors' trades on asset prices when these investors' performance is measured relative to an index. These effects include excess correlation among index stocks, excess index stocks' volatility, and excess aggregate volatility. Brunnermeier and Pedersen (2009) show that institutional investors' aggregate (funding liquidity) capital can drive risk premiums. See, also, Adrian, Etula, and Muir (2014) and He,

[4] Rostek and Weretka (2015a) consider dynamics trading in a centralized exchange and link price impact to the frequency of trade and the timing of information arrival. Malamud and Rostek (2017) consider general fragmented market structures and allow for heterogeneity in risk aversions.

[5] Following the standard logic of the Slutsky income effect, an asset that plays an important role in the agent's portfolio may have a large income effect, and hence a large systemic importance.

Kelly, and Manela (2017). Micro-level evidence on individual trades of institutional investors (Çötelioğlu, Franzoni, and Plazzi (2021), Ben-David, Franzoni, Moussawi, and Sedunov (2021)) suggests the aggregated picture misses important aspects of real markets, and that the granular nature of investors and their individual strategic (i.e., internalizing price impact) behavior are key to understanding the links between market and funding liquidity. We believe that our model offers a tractable framework for analyzing this link and gaining a deeper understanding of the precise role of institutional investors' granularity for asset prices.

Some papers about institutional investors attempt to identify systemically important institutions through their contribution/exposure to systemic risk (see, for example,Acharya, Pedersen, Philippon, and Richardson (2017), Tobias and Brunnermeier (2016)). Our model leads to a complementary notion of *systemic assets.* Trading of these assets results in large price movements for other assets, even assets with uncorrelated payoffs, due to the cross-asset price impact. We believe that investigating and identifying systemic assets empirically might shed novel light on shock propagation in financial markets.

Our paper is part of the broad literature on the effects of illiquidity in financial markets. Many papers in this literature take market frictions as exogenous, such as constant or random trading cost, portfolio constraints, and/or assets that cannot be traded (see, Constantinides (1986), Longstaff (2009), Amihud and Mendelson (1986), Acharya and Pedersen (2005), Duffie, Gârleanu, and Pedersen (2005). In our model, the only friction is the fact that there is a finite number of large traders who behave strategically. A trader is large simply because he owns a non-negligible fraction of the aggregate wealth. Wealth effects endogenously generate (1) portfolio constraints (due to nonnegativity of wealth), (2) illiquidity due to endogenous price impact, and (3) systemic liquidity that is priced in the cross-section of asset returns.[6]

Price impact from institutional trades has been documented empirically in many papers. See, for example, Chan and Lakonishok (1995), Griffin, Harris, and Topaloglu (2003),

[6]Acharya and Bisin (2014) endogenize default risk using counterparty risk when positions are opaque. While there is no counterparty risk in our model, an agent's ability to borrow from other agents is effectively limited by the amount of liquid wealth that he can post as collateral.

Chiyachantana, Jain, Jiang, and Wood (2004), Almgren, Thum, Hauptmann, and Li (2005), Coval and Stafford (2007), and Ben-David et al. (2021). In particular, Chung and Huh (2016) find that price impact is priced and has a larger effect on returns relative to adverse selection.[7] With the exception of Gabaix, Gopikrishnan, Plerou, and Stanley (2003) and Glebkin et al. (2023), most of these papers use CARA-Normal settings, implying a linear price impact, and none of these papers can link the empirically observed non-linearity of price impact to funding liquidity. Our model provides a general framework for investigating such non-linearities.

Finally, we mention a related strand of the literature that considers strategic liquidity provision. See, for example, Biais, Martimort, and Rochet (2000), Roşu (2009), Back and Baruch (2013). This literature uses a discriminatory price auction as opposed to the uniform price auction we use. However, we encounter similar technical difficulties. The multi-dimensional nature of our problem makes the analysis even more involved. In particular, it is challenging to prove the global optimality of candidate equilibria with multiple assets. While we are not able to establish general equilibrium existence, we believe that the techniques developed in our paper may be useful in other models facing similar issues.

2 The model

A number $L > 2$ of strategic *liquidity providers* (LPs) trade assets with *liquidity demanders* (LDs). There are $N+1$ assets, $i \in \{0, 1, 2, ..N\}$, k-th asset is claim to a terminal dividend δ_k. Asset 0 is a risk-free asset with $\delta_0 = 1$. We assume that LDs' aggregate trade is characterized by the aggregate supply shock $s \in \mathbb{R}^N$ that has full support[8] and is independent of δ. As in Klemperer and Meyer, our assumptions imply that equilibrium quantities will depend on the

[7]For papers concerned with liquidity in the presence of asymmetric information, with price determined solely by the degree of adverse selection, see Kyle (1985), Back (1992), Foster and Viswanathan (1996), Back, Cao, and Willard (2000), Back and Crotty (2015).

[8]Supply uncertainty is needed to rule out the extreme multiplicity of equilibria (cf. Klemperer and Meyer 1989; Vayanos 1999).

realization of s but not its distribution.[9] Given initial endowment w_0 of consumption good, initial holdings q_0 of assets, and after trading a portfolio q, LP's utility depends on their post-trade holdings of assets $q_0 + q$ and the post-trade amount of consumption good $w_0 - p^\top q$. It is given by $U(q_0 + q, w_0 - p^\top q)$. This specification incorporates the following cases.

- *General utility of terminal consumption.* Agents trade at time 0, consume at time 1, and risk-free asset is a numeraire. The vectors q and p are $N \times 1$ vectors of quantities and prices of risky assets. Thus, $U(q_0 + q, w_0 - p^\top q) = E[u(\delta^\top(q_0 + q) + w_0 - p^\top q)]$. As a particular case, it includes quasilinear setting: $U(q, p^\top q) = w_0 - p^\top q + u_1(q)$, which, in turn, incorporates the general CARA case studied in Glebkin et al. (2023). For example, in the CARA-Normal case, $u_1(q) = \mu^\top q - \frac{\gamma}{2} q^\top \Sigma q$,
while $u_1(q) = -\frac{1}{\gamma} \ln E[-\exp(-\gamma \delta^\top q)]$ corresponds to generic distributions of asset payoffs.

- *General utility, two consumption dates.* Here p and q are $(N+1) \times 1$ vectors of prices and quantities of all assets, including the risk-free. Agents trade at time 0 but can consume both at time 0 and time 1.

$$U(q + q_0, w_0 - p^\top q) = \underbrace{u_0(w_0 - p^\top q)}_{\text{utility from } t=0 \text{ consumption}} + \underbrace{E[u_1(\delta^\top(q + q_0)]}_{\text{utility from } t=1 \text{ consumption}}$$

  – e.g., CRRA: $u_0(x) = u_1(x) = \frac{x^{1-\gamma}-1}{1-\gamma}$

The utility function $U = U(q, x)$ is assumed to be strictly concave in q and strictly monotone increasing in x, and $U \in C^4(\mathcal{C})$ for some open set $\mathcal{C} \subset \mathbb{R}^{N+1}$. We set U to be $-\infty$ for $(q, p \cdot q) \notin \mathcal{C}$. We denote $\mathcal{Y} \equiv \mathcal{C} \cap supp\{sL\}$.

Each agent is strategic and rationally anticipates the impact that his demand schedule has on equilibrium prices. As is well-known (see, e.g., Klemperer and Meyer, 1989; and Kyle, 1989), equilibria in such demand- or supply-functions competition games can be found using

[9] While we assume that asset supply is price inelastic, it is possible to extend our results to the case of price elastic supply of the form $s(p) = \tilde{s} f(p)$.

the so-called ex-post optimization. Namely, consider the optimization problem of trader i, and suppose that other traders $j \neq i$ submit schedules $D_j(p)$. Denote by

$$D_{-i}(p) \equiv \sum_{j \neq i} D_j(p)$$

the total residual demand of all other large investors, and suppose that $D_{-i}(p)$ is a bijection: $D_{-i}: \hat{\mathcal{A}} \to \mathcal{D}$ onto some open set $\mathcal{D} \subset \mathbb{R}^N$. Let $\Pi: \mathcal{D} \to \hat{\mathcal{A}}$ denote the inverse of this map: $\Pi(D_{-i}(p)) = p$. Suppose first that there is no uncertainty about the asset supply for the agent i, and he observes the realization of ε. If agent i submits a demand schedule $D_i(p)$, market clearing implies that equilibrium prices satisfy

$$D_i(p) = \varepsilon - D_{-i}(p).$$

Therefore, whatever demand schedule $D_i(p)$ the agent submits to the auctioneer, the realized price-quantity pair $(q, p) = (D_i(p), p)$ always satisfy $q = \varepsilon - D_{-i}(p)$, which is equivalent to $p = \Pi(\varepsilon - q)$. Thus, given a known realization of ε, the agent's optimization problem becomes

$$\max_{q \in \mathcal{Y}:\ \varepsilon - q \in \mathcal{D}} U(q, q^\top \Pi(\varepsilon - q)), \tag{1}$$

assuming that the agent is always constrained to submit demands for which there exists a market clearing price.[10] If there exists a $q = q^*(\varepsilon)$ solving the optimization problem (1), then the corresponding market clearing price is given by $p^*(\varepsilon) = \Pi(\varepsilon - q^*(\varepsilon))$. If the map $F: \mathcal{X} \to \mathcal{D}$ defined by $F(\varepsilon) = \varepsilon - q^*(\varepsilon)$ is bijective, then the optimal demand schedule is given by $D_i(p) = q^*(F^{-1}(D_{-i}(p)))$. Importantly, this demand schedule is ex-post efficient: Even conditional on knowing the uncertain supply realization, the investor would like to trade the same amount.

[10]In models with one asset, this issue is usually dealt with by setting the price to plus or minus infinity if the excess demand is positive or negative. That is, submitting demands for which markets do not clear is effectively prohibited. With many assets, the sign of excess demand is not the right object to deal with, and directly prohibiting such trades is a natural alternative.

In this paper, we only consider *regular symmetric equilibria*, in which (1) all traders submit identical demand schedules $D_i(p) = D(p)$; (2) $D(p)$ is two times continuously differentiable and corresponds to an interior optimum in (1); (3) $D(p)$ is bijective and has an invertible Jacobian $\partial D(p)$ for all $p \in \hat{\mathcal{A}}$. In a symmetric equilibrium, we necessarily have $q(\varepsilon) = \varepsilon/L$, and hence the map $F(\varepsilon) = (1 - L^{-1})\varepsilon$ is obviously bijective, and $D_{-i}(p) = (L-1)D(p)$ is bijective if and only if so is $D(p): \ \hat{\mathcal{A}} \to D(\hat{\mathcal{A}})$. In this case, market clearing implies $\varepsilon = LD(p)$, and $\mathcal{D} = (L-1)D(\hat{\mathcal{A}})$.

We formalize these observations in the following proposition.

Proposition 1. *A pair $(D(p), \mathcal{Y})$ with a demand schedule $D(p) : \hat{\mathcal{A}} \to \mathcal{Y}$ is a regular symmetric equilibrium if and only if*

- *$D \in C^2(\hat{\mathcal{A}})$ is bijective and $\partial D(p)$ is invertible for each $p \in \hat{\mathcal{A}}$;*
- *For any $\varepsilon \in \mathcal{X}$, we have that*

$$\varepsilon/L \ = \ \arg\max_{q \in \mathcal{Y}:\ \varepsilon - q \in \mathcal{D}} U(q\,,\, q^\top \Pi(\varepsilon - q)) \tag{2}$$

is an interior optimum with $\Pi(q) = D^{-1}((L-1)^{-1}q)$ and $\mathcal{D} = (L-1)D(\hat{\mathcal{A}})$.

Everywhere in the sequel, we will refer to a symmetric equilibrium satisfying the properties of Proposition 1 as to simply an "equilibrium."

3 Equilibirum characterization

Based on Proposition 1, we will solve for equilibria of this game in the following three steps: (i) derive equations for candidate equilibria using first order conditions for the maximization problem (2), and then solve these equations for some important specifications of the model; (ii) prove that equilibrium demand is indeed bijective; (iii) prove that the solution to the first order condition is indeed the global maximizer in (2). As we will see below, step (iii) is the most involved and often requires imposing additional technical asumptions.[11]

Everywhere in the sequel, we use ∇_q to denote the gradient with respect to q. Assuming that the maximum in (2) is attained at an interior point and differentiating (2) with respect to q, we get

$$\nabla_q U(q, q\cdot\Pi(\varepsilon-q)) \; + \; U_x(q, q\cdot\Pi(\varepsilon-q))\,(\Pi(\varepsilon-q) - \partial\Pi(\varepsilon-q)\,q) \; = \; 0\,,$$

where $\partial\Pi$ denotes the Jacobian of the map Π. In equilibrium, we have $\varepsilon/L \; = \; D(p) = q$ and $\Pi(\varepsilon - q) = p$. Hence, we get

$$\nabla_q U(D(p), D(p)\cdot p) \; + \; U_x(D(p), D(p)\cdot p)\,(p - (\partial\Pi|_{(L-1)D(p)})^T\, D(p)) \; = \; 0\,,$$

where A^T denotes the transpose of a matrix A. This first order condition can be rewritten as

$$p \; = \; -[U_x(D(p), D(p)\cdot p)]^{-1}\nabla_q U(D(p), D(p)\cdot p) \; + \; \partial\Pi|_{(L-1)D(p)})^T\, D(p)$$

By market clearing, in a symmetric equilibrium, we always have $D(p) = \varepsilon/L$, and we arrive at the following result.

Proposition 2. *Provided that an equilibrium exists, inverse demands $I(q)$ satisfy*

$$I(q) = -\,\frac{\nabla U\left(q, q^\top I(q)\right)}{U_x\left(q, q^\top I(q)\right)} - \Lambda\left(q\right) q, \tag{3}$$

for every $q \in \mathcal{Y}$. The equilibrium prices are given by $p = I(\epsilon/L)$ and the equilibrium price impact matrix $\Lambda(q)$ satisfy

$$\Lambda(q) \; = \; -\frac{1}{L-1}\nabla\, I(q). \tag{4}$$

[11]It is known (see, e.g., Back and Baruch, 2013; Baruch and Glosten, 2016) that first order conditions are often insufficient for a local extremum to be the global optimum in games of strategic liquidity provision.

We emphasize several important implications of Proposition 2. First, (3) and (4) give a system of nonlinear Partial Differential Equations (PDEs) for the vector of inverse demands $I(q)$. No known general techniques exist for solving such complex systems. In the single asset case, (3) and (4) boils down to a single Ordinary Differential Equation (ODE)

$$I(q) = -\frac{U_q\left(q, q^\top I(q)\right)}{U_x\left(q, q^\top I(q)\right)} + \frac{qI(q)'}{L-1} \tag{5}$$

As we now show, it is possible to reduce the general multi-asset problem to solving the single-asset problem at a portfolio level. To this end, we define effective inverse demand for a given portfolio q. Namely, we define $\iota(t;q)$ to be the amount a trader is willing to pay for *per unit* when acquiring t units of a portfolio q. That is, $\iota(t;q) = q^\top I(tq)$.

To gain some intuition behind our approach, consider an economy in which all investors (including LDs) can trade only a single index (portfolio) $q \in \mathbb{R}^N$. We refer to this as a *restricted* economy and to our baseline case as an *unrestricted* economy. Then, $\iota(t)$ represents the inverse demand that LPs submit for $t > 0$ units of the index. Since the restricted economy is a one-asset economy, $\iota(t)$ must satisfy the ODE (5). Now consider the unrestricted economy. In the symmetric equilibrium, for supply shock realizations $s = tq$ ($t \in \mathbb{R}^+$), it should be optimal for LPs to absorb $1/L$ fraction of supply shock s—that is, to trade t/L units of portfolio q. Hence, the price LPs bid for t/L units of portfolio q in the unrestricted economy, or $q^\top I(t/Lq)$, should be an optimal bid in the restricted economy. Therefore, $q^\top I(t/Lq) = \iota(t/L)$ should satisfy ODE (5). Note that $\iota(1) = q^\top I(q)$ is the expenditure $E(q)$ for portfolio q (i.e., the dollar amount spent on buying the portfolio q). Once $E(q)$ is known, we can complete the computation by plugging $\iota(1) = q^\top I(q)$ into (3). We now summarize this discussion in the following proposition.

Proposition 3. *The inverse demand $I(q)$ satisfies (3) and (4) if and only if*

$$I(q) = 1/L \cdot \nabla E(q) - (1 - 1/L)\,\frac{\nabla U(q, E(q))}{U_x(q, E(q))}, \tag{6}$$

where the function

$$E(q) \;=\; q^{\top} I(q) \tag{7}$$

can be computed as follows. First, let $\iota(t) \;=\; \iota(t;q;\alpha)$ *be the unique solution to the following ODE*

$$\iota(t) = -\frac{q^{\top}\nabla U\left(tq, t\iota(t)\right)}{U_x\left(tq, t\iota(t)\right)} + \frac{t\iota(t)'}{L-1}, \tag{8}$$

with the boundary condition $f(1;q,\alpha) = \alpha$. *Then,* $E(q) = f\left(|q|; q/|q|; \rho(q/|q|)\right)$, *where* $\rho(x)$ *is an arbitrary smooth function on the unit sphere* $\{q \in \mathbb{R}^N : |q|= 1\}$ *with* $|q|\equiv (q\cdot q)^{1/2}$.

The equilibrium characterization of Proposition 3 leaves open an important question: Since there are multiple candidate equilibria (each corresponding to a different boundary condition $\rho(q)$, how do we know which one of them corresponds to an equilibrium? Can there be multiple equilibria? And, if there is a unique equilibrium, how can we compute it? As we show below, it is indeed possible to find a unique equilibrium if we impose a natural economic monotonicity condition.

3.1 Equilibrium uniqueness

In order to understand the structure of the unique equilibrium we characterize, we first introduce the competitive inverse demand, $I^c(q)$, defined as a solution to:

$$I^c(q) = -\,\frac{\nabla U\left(q, q^{\top} I^c(q)\right)}{U_x\left(q, q^{\top} I^c(q)\right)}. \tag{9}$$

Intuitively, we expect equilibrium inverse demands to converge to the competitive one when $L \to \infty$. Hence, $I^c(q)$ is the natural primary object in our model, and we formulate our restrictions on asset payoffs, δ, and the utility, $U(\cdot,\cdot)$ directly in terms of conditions on $I^c(q)$. These conditions can be easily verified for each particular example we are going to consider later on.

Denote by $\tilde{\mathcal{Y}}$ the set of all $q \in \mathcal{Y}$ for which $\{t > 0 | \; tq \in \mathcal{Y}\}$ is unbounded. Let also $\iota^c(t) = q^\top I^c(qt)$ be the competitive bid for a portfolio q as. By (9), it satisfies

$$-\frac{q^\top \nabla U(tq, t\iota^c(t))}{U_x\,(tq, t\iota^c(t))} = \iota^c(t).$$

We will need the following technical assumptions.

Assumption 1. *For every q the function $\iota^c(t)$ is strictly decreasing in t for t large enough.*

Intuitively, Assumption 1 simply says that the standard "law of demand" holds.[12] The next condition imposes a form of regularity on the competitive inverse demand, allowing us to exclude unbounded equilibria with "explosive" behavior.

Assumption 2. *For every q, the function $\iota^c(t)$ is bounded for t large enough and, for every q, the function $-\frac{q^\top \nabla U(tq,x)}{U_x(tq,x)}$ is decreasing in x for t large enough.*

The following lemma presents a sufficient condition for the validity of Assumption 2.

Lemma 1. *Assumption 2 holds for $U(q,x) = E\left[u(w_0 - x + \delta(e + tq))\right]$ if δ is bounded and the absolute risk aversion $R(x) = -u''(x)/u'(x)$ is decreasing in x.*

Everywhere in the sequel, we focus on equilibria satisfying the law of demand: The fact that $\iota(t)$ is monotone decreasing and bounded for large t. It turns out that the equilibrium is unique under these natural economic restrictions.

Proposition 4. *Under Assumptions 1 and 2, there exists at most one equilibrium in which equilibrium demand $\iota(t)$ is strictly decreasing and bounded for every $q \in \tilde{\mathcal{Y}}$. It can be constructed as follows. For every $M > 0$ define $\iota^M(t)$ a solution to (8) with a boundary condition $\iota^M(M) = \iota^c(M)$. The equilibrium demand is $\iota(t) = \lim_{M\to\infty} \iota^M(t)$.*

[12]There is a large literature that formulates conditions on the utility function, ensuring that the law of demand holds. (E.g., Quah (2003)) It is beyond the scope of our paper to look for conditions on utility beyond those established in the existing literature.

Proposition 4 is crucial for the subsequent analysis. The technical conditions requiring that the effective demand for a portfolio q is bounded and strictly decreasing in the competitive case are natural and intuitive. If δ has bounded support, then the potential benefit one can derive from holding portfolio q should be bounded. We verify these restrictions in all examples studied below. Given the above discussion, it is also natural to focus on equilibria with strategic demands being downward-sloping and bounded as well. Proposition 4 ensures that there is at most one such equilibrium.

3.2 Comparative statics

The explicit characterization of the unique equilibrium in Proposition 4 allows us to apply methods from the theory of ordinary differential equations (comparison theorems) to derive comparative statics results for the equilibrium behavior. Let us rewrite the ODE (8) as

$$\frac{t\iota(t)'}{L-1} = \psi(t, \iota(t); k), \tag{10}$$

where k indicates a generic parameter (for example, initial endowment, w_0), and

$$\psi(t, \iota; k) \equiv \iota(t) + \frac{q^\top \nabla U\left(tq, t\iota; k\right)}{U_x\left(tq, t\iota; k\right)}\,. \tag{11}$$

Note that, by (9), competitive demand $\iota^c(t)$ satisfies

$$\psi(t, \iota^c(t); k) = 0.$$

We make the following assumption.

Assumption 3. $\psi_t(\cdot) > 0$ *and* $\psi_\iota(\cdot) > 0$.

This assumption implies that the competitive demand is downward sloping because, according to the Implicit Function Theorem, $\iota^c(t)' = -\psi_t(\cdot)/\psi_I(\cdot)$.

Proposition 5. *Suppose that* $U(tq, x) = u_0(w_0 - x) + E\left[u_1(w_1 + tq^\top \delta)\right]$*. Moreover, assume that Inada conditions hold for the function* u_0*. Then for equilibrium expenditures,* $E(q; w_0) = q^\top I(q; w_0)$*, we have* $\frac{\partial}{\partial w_0} E(q; w_0) \geq 0$ *and* $\frac{\partial}{\partial w_1} E(q; w_1) \leq 0$*.*

Proposition 5 implies an intuitive relationship between optimal expenditures, $E(q)$, and funding liquidity, as captured by the endowment, (w_0, w_1). A larger initial endowment w_0 naturally increases the agent's willingness to spend, leading to higher prices and, as a result, a lower illiquidity discount. By contrast, an increase in the time 1 endowment has the opposite effect on the asset demand. Indeed, the demand for assets is pinned down by the marginal rate of intertemporal substitution. If the agent already has enough endowment at time 1, there is no need to reallocate wealth from time zero.

The next proposition considers the comparative statics of our first measure of illiquidity, bid shading, which is a difference between the competitive and non-competitive demands:

$$BS(t) \equiv I_q^c(t) - \iota(t).$$

Proposition 6. *[Bid Shading and Funding Liquidity] Suppose that* $U(tq, x) = u_0(w_0 - x) + E\left[u_1(w_1 + tq^\top \delta)\right]$*. Let* $RRA_i(c) = -cu_i''(c)/u_i'(c),\ i = 0, 1$*. Suppose that* $RRA_0(c_0) = RRA_1(c_1) = \gamma > 1$ *. Then* $\frac{\partial}{\partial w_0} BS(t; w_0) > 0$*.*

The technical condition of RRA above one implies a small elasticity of inter-temporal substitution (EIS) between consumption at different time periods. However, at the same time, a larger risk aversion makes LPs' demand functions less elastic. Strikingly, the second effect always dominates: Larger endowment makes the agents more strategic, increases price impact and, as a result, leads to a drop in market liquidity, as manifested by an increase in bid shading.

4 Consumption-CAPM with Strategic Trading with Wealth Effects

In this section, we consider a general class of utilities $U(q,x) = u_0(w_0 - x) + u_1(q)$ that is separable in expenditures, x, and portfolio, q. This is a standard AP setting used in consumption-based asset pricing models. Although this setting might seem different from the CARA-normal setting with a utility from terminal wealth commonly used on market microstructure and Rational Expectations Equilibrium models,[13] it actually does contain the generic CARA setting described in Glebkin et al. (2023) because it reduces to a quasi-linear utility specification $w_0 - q^\top p + u_1(q)$, where $u_1(q)$ is the cumulant-generating function of δ. We start this section with a statement of the equilibrium characterization for the quasi-linear case.

Proposition 7. *[Glebkin et al. (2023)] Suppose that v is uniformly bounded and concave. Then, there exists a unique equilibrium and the equilibrium inverse demand and price impact are given by*

$$I(q) = \frac{L-1}{L} \int_1^\infty \frac{z^{-L}}{L} \nabla v \left(-\frac{L-1}{L} zq \right) dz, \tag{12}$$

$$\Lambda(q) = \frac{(L-1)^2}{L} \int_1^\infty \frac{z^{1-L}}{L-1} \nabla^2 v \left(-\frac{L-1}{L} zq \right) dz, \tag{13}$$

The equilibrium price is $p = I(\epsilon/L)$. Moreover, the equilibrium price impact matrix has the following properties:

- *Price impact is not affected by investors' wealth: $\frac{\partial \Lambda}{\partial w_0} = 0$.*
- *Price impact matrix is symmetric, i.e. $\Lambda_{ij} = \Lambda_{ji}$ for all j and i.*
- *Price impact matrix is positive-definite, in particular, $\Lambda_{ii} > 0$ for all i.*

Proposition 7 characterizes key theoretical properties of the price impact matrix in the

[13] The common assumption is these models is that the numeraire good is storable, so that $w_0 - q^\top p$ can be brought to $t = 1$ and consumed at that moment in time.

absence of wealth effects: Price impact behaves like a covariance matrix, consistent with standard CARA-Normal models. Our next theoretical result derives the implications of strategic trading for the cross-section of stock returns in the form of liquidity-adjusted Consumption CAPM. The following is true.

Proposition 8. *Provided that the equilibrium exists, the following liquidity-adjusted consumption CAPM holds: The expected return on any asset is given by*

$$E\left[R_i\right]-R_f=\beta_i\left(E\left[R_q\right]-R_f\right)-R_f\frac{\Lambda_{qq}^{\%}}{1+\Lambda_{r_fq}^{\%}}\left(\beta_i-\beta_i^{\lambda}+\beta_{r_f}^{\lambda}\left(1-\beta_i\right)\right), \qquad (14)$$

where, given the equilibrium allocation $q=\epsilon/L$, *we have* $R_i=E\left[\frac{\delta_i}{p_i}\right]$, $R_q=E\left[\frac{q^{\top}\delta_i}{q^{\top}p}\right]$, $\beta_i=\frac{cov\left(u_1'(\delta q),R_i\right)}{cov\left(u_1'(\delta q),R_q\right)}$, $\beta_i^{\lambda}=\frac{\Lambda_{iq}^{\%}}{\Lambda_{qq}^{\%}}$, $\Lambda_{iq}^{\%}=1_i^{\top}\Lambda(q)q/p_i$, $\Lambda_{qq}^{\%}=q^{\top}\Lambda(q)q/(q^{\top}p)$ *for* $i=\{r_f,1,..,N\}$. *If traders do not take the price impact into account, then we have the following C-CAPM:*

$$E\left[R_i\right]-R_f=\beta_i\left(E\left[R_q\right]-R_f\right). \qquad (15)$$

The formula (14) shows that, in equilibrium, asset excess returns behave as if they had a multi-factor structure, as in Acharya and Pedersen (2005), with fundamental consumption betas and liquidity betas. However, in contrast to the reduced form approach in Acharya and Pedersen (2005), in our model, this "liquidity risk" arises endogenously, due to the strategic behavior of investors. To understand the nature of two two-factor cross-sectional asset pricing model, we rewrite (14) as

$$E\left[R_i\right]-R_f\;=\;\alpha\;+\;\beta_i\,\lambda_1\;+\;\beta_i^{\lambda}\,\lambda_2 \qquad (16)$$

where

$$\begin{aligned}
\underbrace{\alpha}_{aggregate\ illiquidity\ discount} &= -R_f \frac{\Lambda_{qq}^{\%}}{1+\Lambda_{r_f q}^{\%}} \beta_{r_f}^{\lambda} \\
\underbrace{\lambda_1}_{fundamental\ risk\ premium} &= \underbrace{E\left[R_q\right]-R_f}_{market\ risk\ premium} - \underbrace{R_f \frac{\Lambda_{qq}^{\%}}{1+\Lambda_{r_f q}^{\%}}(1-\beta_{r_f}^{\lambda})}_{illiquidity\ correction} \qquad (17) \\
\underbrace{\lambda_2}_{co-illiquidity\ premium} &= R_f \frac{\Lambda_{qq}^{\%}}{1+\Lambda_{r_f q}^{\%}}
\end{aligned}$$

The first surprising implication of our findings is the emergence of α, reflecting the *illiquidity of the risk-free asset,* a novel effect that cannot emerge in the CARA setting of Glebkin et al. (2023) (in this case, Λ is a covariance matrix and, hence, $\beta_{r_f}^{\lambda} = 0$). It implies that the risk premium should be evaluated against the illiquidity-adjusted risk-free rate, $R_f + \alpha$. One could interpret α as a form of (in)convenience yield (Krishnamurthy and Vissing-Jorgensen, 2012) associated with its illiquidity (Doerr et al., 2023). Second, the fundamental risk premium should itself be corrected for illiquidity. This has important implications for statistical tests of CAPM: Cross-sectional asset pricing tests will generate a risk premium that is different from the market excess return. Finally, the third novel implication of our asset pricing model is the emergence of the co-illiquidity premium. Namely, asset returns get discounts depending on their systemic importance in the price impact matrix Λ. To understand the role of co-illiquidity, we note the following algebraic identities:

$$\begin{aligned}
\beta_i^{MKT} &= \frac{\mathrm{Cov}(\delta_i, \delta^\top, q)}{\mathrm{Var}(\delta^\top q)} = \frac{1_i^\top \Sigma q / p_i}{q^\top \Sigma q / (q^\top p)} \\
\beta_i^{\lambda} &= \frac{1_i^\top \Lambda(q) q / p_i}{q^\top \Lambda(q) q / (q^\top p)}\,.
\end{aligned} \qquad (18)$$

We observe a clear similarity between β_i^{MKT} (β with respect to the market) and β_i^{λ}, defined through the co-illiquidity (instead of covariance) with the market portfolio q, whereby we replace the fundamental covariance matrix Σ with the price impact matrix $\Lambda(q)$. The co-illiquidity defines how much the price of the market portfolio moves when we trade a unit of the asset i. In the CARA case of Glebkin et al. (2023), Λ is proportional to Σ and hence is symmetric: The cross-impact of asset i on asset j is the same as the impact of asset j on asset i. By contrast, in our model, it is not, and it can even be asymmetric, implying that some assets are systemic and trading them can have a large impact on other assets in the economy. The next proposition characterizes precisely when this asymmetry occurs in equilibrium.

Proposition 9. *The following is true for the asymmetry in price impact, defined as* $\Lambda(q) - \Lambda^\top(q)$:

$$\Lambda(q) - \Lambda^\top(q) \propto \frac{u_0''(q)}{u_0'(q)} \Big(\nabla u_1(q) I(q)^\top - I(q) \left(\nabla u_1(q)\right)^\top\Big) . \tag{19}$$

Consequently, $\Lambda(q)$ *is symmetric if and only if one of the following two conditions hold: (1)* $u_0(x)$ *is linear in* x*; (2)* $I(q) \propto \nabla u_1(q)$.

Proposition 9 shows how wealth effects are crucial for the emergence of price impact asymmetry. Case (1) in Proposition 9 occurs in the quasi-linear case of Glebkin et al. (2023) where initial funding liquidity, w_0, is irrelevant for investors. Case (2) is more subtle and requires a precise understanding of the utility function u_1. It is straightforward to show that, for generic utility functions, $I(q)$ is not proportional to $\nabla u_1(q)$ and, hence, wealth effects typically make price impact asymmetric. In the sections below, we derive closed-form solutions for equilibrium for some special choices of the utility functions.

5 Unit-Elastic Preferences

Throughout this section, we make the following assumption.

Assumption 4. *The elasticity of inter-temporal substitution (EIS) equals one so that* $u_0(x) = \log x$. *Furthermore, the function* $u_1(q)$ *satisfies* $\hat{u}_1(q) > -(1-\varepsilon)$ *for some* $\varepsilon > 0$, *and is defined on a domain* $\mathcal{V} \subset \mathbb{R}^N$ *that is conic: if* $q \in \mathcal{V}$ *then* $\rho q \in \mathcal{V}$ *for all* $\rho \geq 1$.

We also let

$$c \;=\; \frac{L-1}{L}\,. \tag{20}$$

The following is true.

Lemma 2. *Under Assumption 4, the unique equilibrium is given by*

$$E(q) \;=\; w_0 \int_1^\infty e^{-\int_1^s \frac{c\hat{u}_1(\rho^{1/L} q)+1}{\rho} d\rho} \frac{ds}{s} \tag{21}$$

We can now study the behavior of the equilibrium price impact.

Lemma 3. *When L is large enough, the equilibrium price impact $\Lambda^*(q)$ converges to*

$$\lim_{L\to\infty} \Lambda(q) \;=\; \frac{1}{1+\hat{u}_1(q)} \left(-\nabla^2 u_1(q) \;+\; \frac{1}{1+\hat{u}_1(q)} \left(\nabla u_1(q) + \nabla^2 u_1(q)\, q\right) \nabla u_1(q)^T \right)$$

uniformly on compact subsets of $\mathcal{V}$.

Suppose now that the agents have Epstein-Zin (1989) preferences with risk aversion γ so that

$$u_1(q) \;=\; \frac{1}{1-\gamma} \log E[(w + \delta \cdot q))^{1-\gamma}]\,,$$

where we have absorbed the payoff of the potential initial holdings into the random endowment w.[14] Then, $u_1(q)$ satisfies the conditions of Assumption 4 if $\mathcal{V}$ is such that $\delta \cdot q > 0$ for all $q \in \mathcal{V}$ (for example, if $\delta > 0$ and $q \in \mathbb{R}_+^N$), and the following is true.

Proposition 10. *If $w \le 0$, then $q \cdot \Lambda(q) q < 0$.*

The result of Proposition 10 stands in perfect agreement with the conventional wisdom: When $w \ge 0$, traders submit demand schedules that are downward sloping. As a result, each trader faces positive trading costs and shades the bid accordingly, implying that markets clear at prices below their frictionless level. However, quite strikingly, this result is not anymore true when the endowment can take negative values. In this case, the agent desperately needs to

[14]Since markets are incomplete due to price impact, the random endowment is effectively unspanned, even if w is a linear combination of tradable asset payoffs.

buy assets that cover the liabilities at time $t = 1$, leading to a negative realized price impact $q \cdot \Lambda^*(q)q$.

In the absence of endowment at time $t = 1$ and with logarithmic preferences,

$$u_1(q) \;=\; E[\log(\delta \cdot q)]\,,$$

a direct calculation implies that the price impact Λ^* is

$$\Lambda(q) \;=\; \frac{c}{1+c} w_0\, E\left[\frac{1}{(\delta \cdot q)^2}\delta\delta^T\right]$$

In this case, the price impact is positive definite, and symmetric. Both price level and price impact are proportional to w_0, so that the percentage price impact is independent of wealth. One surprising consequence of this formula is the fact that price impact is non-zero even for the risk-free asset: If $\delta_0 = 1$, we get

$$\Lambda_{r_f}(q) \;=\; \frac{c}{1+c} w_0\, E\left[\frac{1}{(\delta \cdot q)^2}\right] \tag{22}$$

Furthermore, trading the risk-free asset (e.g., through quantitative easing or tightening) affects prices of all other assets proportionally to their expected payoffs:

$$\Lambda_{r_f,i}(q) \;=\; \frac{c}{1+c} w_0\, E\left[\frac{1}{(\delta \cdot q)^2}\delta_i\right]\,. \tag{23}$$

6 Liquidity Shortage

In this section, we study the case when investors are highly constrained in their funding liquidity and need to sell assets in order to replenish their liquid reserves. Formally, we assume that $w_0 = 0$. Furthermore, throughout this section, we assume that inter-temporal preferences exhibit a constant elasticity of inter-temporal substitution (CEIS) given by $1/\gamma$, so that

$$u_0(x) \;=\; \frac{x^{1-\gamma}}{1-\gamma}\,.$$

In addition, we will make the following technical assumption.

Assumption 5. *The function $u_1(q)$ is defined on a set $\mathcal{V}$ such that $t\mathcal{V} \subset \mathcal{V}$ for all $t \in (0,1]$.*

In this case, the following is true.

Proposition 11. *Suppose that $w_0 = 0$ and $u_0(x) \;=\; \frac{x^{1-\gamma}}{1-\gamma}$ and that $\gamma > 1$. Then, the function $H(q) = E(q)^{1-\gamma}$ satisfies*

$$(1-c)(1-\gamma)^{-1} q \cdot \nabla H(q) \;-\; H(q) \;=\; c\hat{u}_1(q)$$

and hence

$$H(q) \;=\; \frac{c}{\beta}\int_0^1 t^{-1-1/\beta}\hat{u}_1(qt)dt$$

with

$$\beta \;=\; \frac{1-c}{1-\gamma}\,.$$

We can now characterize the equilibrium price impact.

Lemma 4. *We have*

$$\Lambda(q) = \; (H(q))^{\frac{\gamma}{1-\gamma}}\left(\nabla^2 S(q) \;+\; \frac{\gamma}{1-\gamma}\frac{\nabla H(q)}{H(q)}\left(\beta\nabla H(q) \;-\; c\nabla u_1(q)\right)^T\right) \tag{24}$$

and converges to

$$\lim_{L\to\infty}\Lambda(q) \;=\; [-\hat{u}_1(q)]^{\gamma/(1-\gamma)}\left(-\nabla^2 u_1(q) \;+\; \frac{\gamma}{1-\gamma}\frac{1}{(-\hat{u}_1(q))}(\nabla u_1(q) + \nabla^2 u_1(q)q)(\nabla u_1(q))^T\right)$$

when $L \to \infty$.

Suppose now that the agents have standard, time-separable CRRA preferences, so that

$$u_1(q) \;=\; (1-\gamma)^{-1}E[(w+\delta\cdot q)^{1-\gamma}]\;,\;\hat{u}_1 \;=\; E[\delta\cdot q(w+\delta\cdot q)^{-\gamma}]$$

for some positive random endowment $w > 0$. Suppose also that $\delta > 0$. Due to liquidity shortage, the agents will be forced to sell some of the claims on δ against parts of their endowment w.[15] Let $\mathcal{V} = \{q \in \mathbb{R}_-^N : \; w + \delta q > 0\}$ denote the set of admissible portfolios that can be sold while keeping the consumption positive. In this case, Assumption 5 clearly holds, and equilibrium price impact takes the form

$$\lim_{L\to\infty}\Lambda(q) \;=\; \gamma E[\delta\delta^T(w+\delta\cdot q)^{-\gamma-1}] + \frac{\gamma}{\gamma-1}\frac{E[\delta(w+\delta\cdot q)^{-\gamma}]}{E[(\delta\cdot q)(w+\delta\cdot q)^{-\gamma}]}E[\delta^T\frac{w+(1-\gamma)\delta\cdot q}{(w+\delta\cdot q)^{\gamma+1}}] \quad (25)$$

In particular, the cost of selling the portfolio $-q$ is given by

$$q\cdot\Lambda(q)q \;=\; -\frac{1}{\gamma-1}\left(E[(\delta\cdot q)^2\,(w+\delta\cdot q)^{-\gamma-1}] + \gamma E[(-\delta\cdot q)\,(w+\delta\cdot q)^{-\gamma}]\right) \;<\; 0\,.$$

As a result, the price of the portfolio $(-\varepsilon)/L$ that all agents sell in equilibrium is above the frictionless level due to an illiquidity premium. This result stands in stark contrast to the CARA case studied in Glebkin et al. (2023): When price impact is positive, prices always contain an illiquidity discount. The mechanism underlying this surprising result is based on a simple wealth effect. Namely, in the presence of liquidity shortage, agents' demand curves are upward sloping because, when prices drop, agents sell even more of the assets in order to increase their time-zero consumption from zero to an optimal level. As a result, markets clear even before the prices have a chance to reach their competitive levels. At the same time, it is possible to show that exactly $N-1$ eigenvalues of price impact (25) are positive, and they are non-monotonic in w. In particular, an increase in the funding liquidity (as measured by w) can lead to an increase in the price impact and hence to a drop in the market liquidity.

[15]For example, w can be just a portfolio x of tradable assets, $w = x\cdot\delta$, with some $x \in \mathbb{R}_+^N$.

This shows that the conventional monotonic relationship between two forms of liquidity (as in Brunnermeier and Pedersen (2009)) may fail to hold in the presence of wealth effects. We also note that the price impact matrix (25) exhibits very strong asymmetries due to wealth effects.

In the rest of the section, we investigate the behavior of our closed-form solution (11) for finite L. The following is true.

Lemma 5. *We have*

$$H(q) \;=\; -\frac{(\gamma-1)(L-1)}{(\gamma-1)L+1} E\left[(\delta\cdot q)\, w^{-\gamma}\, {}_2F_1\left(\gamma, L(\gamma-1)+1; L(\gamma-1)+2; -\frac{\delta\cdot q}{w}\right)\right],$$

where ${}_2F_1(a,b;c,z)$ is the Hypergeometric function.

Substituting this closed-form expression into (24), we can study the behavior of the equilibrium price impact and its dependence on various model parameters. The two explicit examples we consider are devoted to the two key novel properties arising due to wealth effects: Breakdown of positive definiteness and asymmetry.

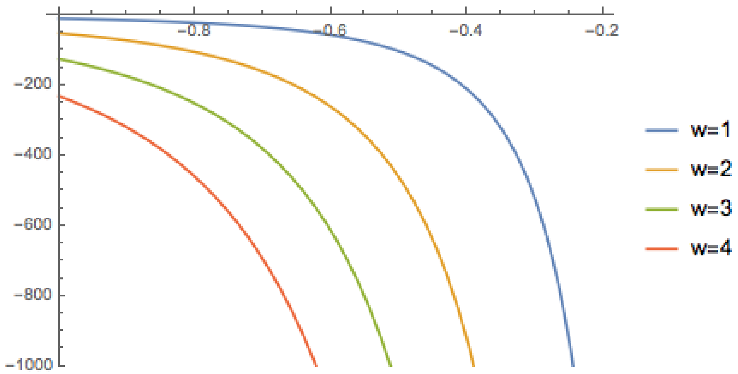
Suppose first that there is only a single risk-less asset, and the endowment w is non-random. Then,

$$H(q) = \;-\frac{(\gamma-1)(L-1)}{(\gamma-1)L+1} q\, w^{-\eta}\, {}_2F_1\left(\eta, L(\gamma-1)+1; L(\gamma-1)+2; -\frac{q}{w}\right).$$

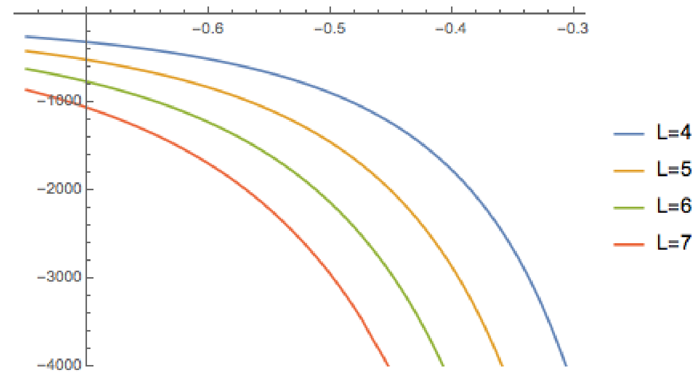
Figures 1 and 2 show the price impact Λ that occurs in equilibrium:

$$\Lambda\left((L-1)\frac{\varepsilon}{L}\right).$$

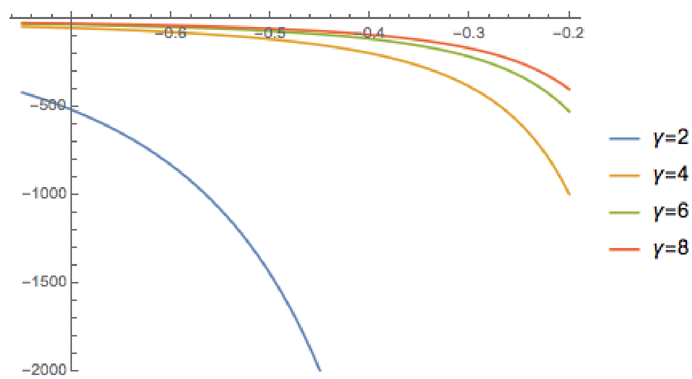
We can see that the price impact is negative, which represents a departure from the CARA case, as we previously discussed. We observe that the price impact decreases (becomes more negative) in the endowment level w: The illiquidity premium is higher when agents can pledge a larger amount of wealth because the desire to smooth consumption over time increases the demand for liquidity at time $t = 0$. It also increases the coefficient of risk aversion: When γ is large, the marginal value of time $t = 1$ consumption is higher, and the agent is less interested in allocating wealth towards $t = 0$.



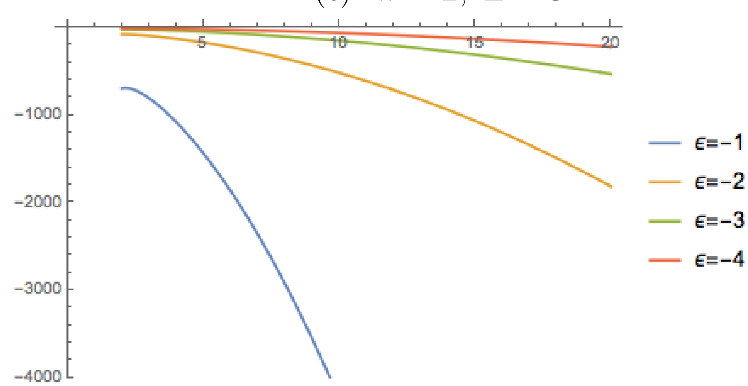
(a) **L = 3**, **ﬂ = 2**



(b) **w = 2**; **ﬂ = 2**



(c) **w = 2**; **L = 3**



(d) **w = 2**; **L = 3**

Figure 1: **Price Impact** We plot Λ for different values of q. Parameters: $\gamma = \eta = 2$

We now proceed to investigate the asymmetry of price impact. We assume that there are N states of the world $s_1, \cdots, s_N$ and that the N securities traded are Arrow-Debreu securities. Then,

$$H(q) = \sum_{i=1}^{N} \bar{p}_i h(q_i, w_i; \gamma, \eta, L),$$

where $\bar{p}_i$ is the probability that state s_i will occur at time $t = 1$ and

$$h(x, y; \gamma, \eta, L) \equiv -\frac{(\gamma-1)(L-1)}{(\gamma-1)L+1} x\, y^{-\eta}\, {}_2F_1\left(\eta, L(\gamma-1)+1; L(\gamma-1)+2; -\frac{x}{y}\right),$$

To illustrate the asymmetry, we consider $N = 2$. Figure 3 shows the ratio

$$\frac{\Lambda_{12}}{\Lambda_{21}}.$$

We see that the ratio is different from one even when both states of the world are equally likely, and there is no uncertainty associated with time $t = 1$ wealth (Figure 3 (a)). The asymmetry is not generated by having a state of the world more likely to occur than the other one (Figure 3 (b)). Wealth uncertainty is an important factor determining this asymmetry (Figure 3 (c) and (d)): The asset that pays in low-wealth states is systematic: That asset has a higher impact on the liquidity (and hence the price) of other assets. Such assets are commonly viewed as "safe," and a large amount of literature investigates the demand for safety. See, for example, Gorton and Ordonez (2022). The striking implication of our results is that contrary to conventional wisdom, safe assets are more illiquid because they have a larger price impact.

7 Wealth Inequality

In the previous sections, we study symmetric equilibria in economies populated by a finite number of identical agents. In general, investigating asymmetric equilibrium with heterogeneous, strategic agents is an extremely complex problem that is reduced to solving systems of partial differential equations. However, it is still possible to solve for asymmetric equilibria analytically when the nature of heterogeneity is sufficiently simple. In this section, we investigate asymmetric equilibria for the case when agents only differ in their initial wealth and have log preferences.[16] Namely, throughout this section, we make the following assumption.

Assumption 6. *The economy is populated by L agents, with agent i maximizing a constant elasticity of substitution utility*

$$U_i(q,x) \;=\; u(\alpha_i w_0 \;-\; x) \;+\; u_1(q)$$

with $u(c) = \log(c)$ and $u_1(q) = e^{-\rho}E[\log(\delta \cdot q)]$. The weights $\alpha_i > 0$ add up to one, $\sum_i \alpha_i = 1$. Thus, α_i is the fraction of the aggregate endowment that the agent i owns.

Assumption 6 implies agents only differ with respect to their initial wealth. In the competitive case, equilibrium demand functions satisfy the first-order conditions

$$\nabla V(D_i(p)) \;=\; u'(\alpha_i w_0 - p \cdot D_i(p))\, p\,,$$

and the assumed homogeneity immediately implies that $D_i(p) = \alpha_i D(p)$ for some function $D(p)$ satisfying $\nabla V(D(p)) \;=\; pu'(w_0 - p \cdot D(p))$. Therefore, intuitively, we expect that in the non-competitive case, agents' demand functions will only differ up to scaling in the sense of the following definition.

[16] While one might conjecture that our analysis should extend to the more general case of CRRA preferences, this is actually not the case.

Definition 1. *An equilibrium tuple* $(D_i(p))_{i=1}^L$ *of demand schedules is scale-symmetric if there exists a schedule* $D(p)$ *and a tuple of constants* $\beta_i,\ i=1,\cdots,L$ *such that* $D_i(p)=\beta_i D(p),\ i=1,\cdots,L$. *A regular scale-symmetric equilibrium is such that the demand schedule* $D(p)$ *is two times continuously differentiable, bijective, has an invertible Jacobian* $\partial D(p)$ *for all* $p\in\hat{\mathcal{A}}$, *and corresponds to an interior optimum in*

$$\max_{q:\ \varepsilon-q\in\mathcal{D}} U_i(q\,,\,q\cdot\Pi_i(\varepsilon-q))\,, \tag{26}$$

where we have defined

$$D_{-i}(p)\ \equiv\ \sum_{j\neq i} D_j(p)$$

and $\Pi_i:\ \mathcal{D}\to\hat{\mathcal{A}}$ *denotes the inverse of this map:* $\Pi_i(D_{-i}(p))=p$.

The following proposition is a straightforward extension of Proposition 1.

Proposition 12. *A set of demand schedules* $\{D_i=\beta_i\,D(p),\ i=1,\cdots,L\}$ *with* $D(p):\hat{\mathcal{A}}\to\mathbb{R}^N$ *is a regular scale-symmetric equilibrium if and only if*

- $D\in C^2(\hat{\mathcal{A}})$ *is bijective and* $\partial D(p)$ *is invertible for each* $p\in\hat{\mathcal{A}}$;
- *For any* $\varepsilon\in\mathcal{X}$, *we have that*

$$\eta_i\,\varepsilon\ =\ \arg\max_{q:\ \varepsilon-q\in\mathcal{D}} U_i(q\,,\,q\cdot\Pi_i(\varepsilon-q))\qquad\text{where}\qquad \eta_i\ =\ \frac{\beta_i}{\sum_j\beta_j}, \tag{27}$$

is an interior optimum with $\Pi_i(q)=D^{-1}((\beta-\beta_i)^{-1}q)$.

Let

$$\beta\ \equiv\ \sum_i \beta_i\,. \tag{28}$$

The following is true.

Proposition 13. *There exists a unique scale-symmetric equilibrium satisfying* $D_i(p) = \beta_i D(p)$ *with*

$$D^{-1}(x) = w_0 e^{-\tau} \mathrm{E}[(\delta \cdot x)^{-1} \delta] .$$

In equilibrium, prices are given by

$$p = \beta w_0 \mathrm{E}\left[(\varepsilon \cdot \delta)^{-1} \delta\right] . \tag{29}$$

while individual price impacts satisfy

$$\Lambda_i(q) = (\beta - \beta_i) w_0 E\left[(\delta \cdot q)^{-2} \delta \delta^T\right] .$$

The scaling constants β_i *are given by*

$$\beta_i = \frac{2\alpha_i e^{-\tau} \beta}{\alpha_i e^{-\tau} + (1 + e^{-\tau})\beta + \sqrt{[\alpha_i e^{-\tau} - \beta(1 - e^{-\tau})]^2 + 4\beta^2 e^{-\tau}}} ,$$

where β *is the unique solution to*

$$\sum_{j=1}^{L} \frac{2\alpha_j e^{-\tau}}{\alpha_j e^{-\tau} + (1 + e^{-\tau})\beta + \sqrt{[\alpha_j e^{-\tau} - \beta(1 - e^{-\tau})]^2 + 4\beta^2 e^{-\tau}}} = 1 . \tag{30}$$

Finally, the coefficients β_i *are monotone increasing in* α_i*.*

We would now like to use the explicit solution of Proposition 13 to study how the economy's oligopolistic structure influences equilibrium allocations and efficiency. After the 2007-2009 financial crisis, some economists and policymakers suggested that breaking up the largest banks and financial institutions would make them less systemically important and lead to more efficient markets. At the same time, the dynamics of the banking industry have been going in the opposite direction over the past decades, with mergers and acquisitions making the banking industry increasingly concentrated.[17] Furthermore, Ben-David et al. (2021) argue that the nature of institutional investors is granular, and they cannot be simply split. We investigate two different scenarios in our model. The first one is that of a "merger," in which two large investors merge into one investor with the total endowment equal to the sum of individual endowments. Another one is a breakup scenario whereby a large investor is split into two smaller investors. We would like to understand the impact of such changes on efficiency and welfare; if a breakup is optimal, we can study how it needs to be implemented (that is, which fractions of the original endowment should the split parts inherit).

Proposition 14. *Let $\boldsymbol{\alpha} = \{\alpha_i\}_{i=1}^{L}$ be an initial distribution of wealth with $L \geq 3$ and consider a break-up $\bar{\boldsymbol{\alpha}} = \{\bar{\alpha}_i\}_{i=1}^{\bar{L}}$ of $\boldsymbol{\alpha}$ with*[18]

$$\bar{L} \;=\; L+1; \qquad \bar{\alpha}_i \;=\; \alpha_i, \quad i<L; \qquad \text{and} \qquad \bar{\alpha}_L \;=\; y; \quad \bar{\alpha}_{L+1} \;=\; \alpha_L - y; \quad y \;\in\; (0,\, \alpha_L)\,.$$

Then,

$$p \;<\; \bar{p}$$

where p and $\bar{p}$ are the equilibrium prices under wealth distributions $\boldsymbol{\alpha}$ and $\bar{\boldsymbol{\alpha}}$, respectively. In addition, the highest $\bar{p}$ is achieved when the trader has been broken up is split into two equal traders; that is when

$$y \;=\; \frac{1}{2}\alpha_L\,.$$

Suppose that $i<L$ (that is, trader i was left unchanged). Then,

$$U^*_{i,\boldsymbol{\alpha}} \;>\; U^*_{i,\bar{\boldsymbol{\alpha}}}\,.$$

[17]For example, according to Wheelock and Wilson (2012): "...in 2001, the five largest commercial banks held 30 percent of total U.S. banking system assets, topped by Bank of America, which had \$552 billion of assets. By contrast, in 2011, the five largest banks held 48 percent of total system assets. Four banks had total assets in excess of \$1 trillion, and the largest commercial bank, JPMorgan Chase Bank, had \$1.8 trillion of assets, equal to 14 percent of the total assets of all U.S. commercial banks."

[18]Thus, we are breaking the Lth trader into two, with respective sizes $\bar{\alpha}_{L,1}$ and $\bar{\alpha}_{L,2}$.

where $U^*_{i,\boldsymbol{\alpha}}$ and $U^*_{i,\bar{\boldsymbol{\alpha}}}$ are the equilibrium utilities for Trader i under wealth distributions $\boldsymbol{\alpha}$ and $\bar{\boldsymbol{\alpha}}$, respectively.

Proposition (14) shows that the break-up (respectively, merging) of two large traders results in lower (resp. higher) welfare for the remaining traders. The reason is that prices are higher post-break up, and this income effect dominates and results in lower wealth.

Proposition 15. *Suppose that $(\hat{\alpha}_i) \prec (\alpha_i)$ in the sense of second-order stochastic dominance (SOSD). Then, $p((\alpha_i)) > p((\hat{\alpha}_i))$.*

Proof. The proof follows because $(\hat{\alpha}_i)$ obtains from (α_i) through a sequence of mean-preserving spreads applied to pairs of agents (see, Rothschild and Stiglitz (1978)). ■

Thus, the fire sale discount is larger when wealth is more unevenly dispersed.

8 Conclusion

We develop a novel analytical framework for studying strategic liquidity provision in the presence of wealth effects. We show how finding the unique equilibrium reduces to solving a single ordinary differential equation that is sufficiently tractable, allowing us to derive general comparative statics results. We show that the setting with wealth effects (non-quasi-linear utilities) leads to several surprising phenomena that have no analogs in the simpler CARA models (see, e.g., Glebkin et al. (2023)) studied in the literature so far.

While our setting is stylized, it allows us to address several important and realistic aspects of the functioning of modern markets that are dominated by large financial institutions. In particular, consistent with recent empirical evidence Doerr et al. (2023), we find that (a) safe assets can be more illiquid; (b) riskless assets can be illiquid and trade with an illiquidity premium; (c) wealth heterogeneity matters. Our model also allows us to investigate the role of the granularity of institutional investors and the impact of changes in the oligopolistic structure originating from mergers and acquisitions. Our results open a new perspective on the classic link between market and funding liquidity Brunnermeier and Pedersen (2009) and have important policy implications. In markets dominated by strategic traders, open market interventions such as quantitative easing (QE) Benford, Berry, Nikolov, Young, and Robson (2009) may have unexpected consequences due to the subtle, non-linear, and non-monotonic impact of QE on market liquidity. Furthermore, the presence of systemic assets and the fact that safe assets are systemic have concrete implications for the nature of assets a central bank should use during interventions. We leave the analysis of these important questions for future research.

Appendices

A Proofs

Proof of Lemma 1. We have

$$\iota^c(t) = -\frac{q^\top \nabla U(tq, t\iota^c(t))}{U_x\,(tq, t\iota^c(t))} = \frac{E\left[u'(w_0 - t\iota^c(t) + \delta^\top (e+tq))q^\top \delta\right]}{E\left[u'(w_0 - t\iota^c(t) + \delta^\top (e+tq))\right]}.$$

If δ is bounded, then there exist $\underline{q^\top \delta}$ and $\overline{q^\top \delta}$ such that $\underline{q^\top \delta} \le q^\top \delta \le \overline{q^\top \delta}$. Then

$$\underline{q^\top \delta} \le \frac{E\left[u'(w_0 - t\iota^c(t) + \delta^\top (e+tq))q^\top \delta\right]}{E\left[u'(w_0 - t\iota^c(t) + \delta^\top (e+tq))\right]} \le \overline{q^\top \delta}$$

The second claim follows from Lemma 6 below. ■

Lemma 6. *Let* $e = 0$. *The function*

$$v(q,x) \;=\; \frac{E\left[u'(w_0 - x + \delta^\top q)q^\top \delta\right]}{E\left[u'(w_0 - x + \delta^\top q)\right]}$$

is monotone increasing (decreasing) in x *if so is absolute risk aversion* $R(x) = -u''(x)/u'(x)$.

Proof of Lemma 6. We have

$$v_x(q,x) \;=\; \frac{-E[u''(w_0-x+\delta^\top q)q^\top\delta]E\left[u'(w_0-x+\delta^\top q)\right]+E\left[u''(w_0-x+\delta^\top q)\right]E\left[u'(w_0-x+\delta^\top q)\delta^\top q\right]}{E\left[u'(w_0-x+\delta^\top q)\right]^2}$$

We introduce a new measure with the density $u'(w_0-x+\delta^\top q)/E[u'(w_0-x+\delta^\top q)]\eta(\delta)$. Denote by E^* expectations under this new measure. Then, the required positivity is equivalent to

$$-E^*[\delta^\top q]E^*[R(\delta^\top q-x+w_0)]+\;E^*[R(\delta^\top q-x+w_0)\delta^\top q]\;>\;0\,,$$

where

$$R(x)\;=\;-u''(x)/u'(x)\,.$$

The result follows from the standard correlation inequality. ■

A.1 Reduction to Integral Equation

Proposition 16. *$\iota(t)$ solves (8) if and only if it solves*

$$\iota(t)\;=\;k\,t^{L-1}\;+\;(L-1)\int_1^\infty \xi^{-L}f(t\xi)d\xi\,, \tag{31}$$

where

$$f(t)\;=\;-\frac{q^\top\nabla U\left(tq,t\iota(t)\right)}{U_x\left(tq,t\iota(t)\right)} \tag{32}$$

for some $k\in\mathbb{R}$.

Proof. We have

$$\iota'(t)\;=\;(L-1)t^{-1}\iota(t)\;-\;t^{-1}(L-1)f(t)\,, \tag{33}$$

and the standard variation of constants formula implies any solution to this linear ODE is given by

$$\iota(t) \ = \ t^{L-1}k + t^{L-1}(L{-}1)\int_t^{\infty} y^{-(L-1)}y^{-1}f(y)dy \ = \ \{y = t\xi\} \ = \ t^{L-1}k + t^{L-1}(L{-}1)\int_1^{\infty}(t\xi)^{-(L-1)}(t\xi)^{-1}f(t$$

and the claim follows. ■

Lemma 7. *Suppose that* $U(q,x) = E\left[u(w_0 - x + \delta^\top(e+tq))q^\top\delta\right]$. *Then,*

$$f(t) \ = \ \frac{E\left[u'(w_0 - t\iota(t) + \delta^\top(e+tq))q^\top\delta\right]}{E\left[u'(w_0 - t\iota(t) + \delta^\top(e+tq))\right]}$$

If the support of δ *is bounded,* $\|\delta\| \leq K$, *then* $|f(t)| \leq K\|q\|$.

Suppose now that $L > 2$. *Then, any bounded solution* $i(t)$ *to* (8) *satisfies the integral equation*

$$\iota(t) \ = \ (L-1)\int_1^{\infty}\xi^{-L}f(t\xi)d\xi\,, \tag{34}$$

Proposition 17. *Let*

$$g(t,x) \ = \ -\frac{q^\top\nabla U\,(tq,tx)}{U_x\,(tq,tx)}\,. \tag{35}$$

Suppose that $g_x < 0$ *and that for each* $t > 0$ *there exists a constant* $K(t) > 0$ *such that* $g_t(t\xi, x) \geq K(t)$ *for all* $x < \iota^c(t\xi)$. *Then, we have*

$$\iota(t) \ = \ \iota^c(t) \ + \ (L-1)^{-1}\phi(t) \ + \ O((L-1)^{-2}) \tag{36}$$

with

$$\phi(t) \ = \ \frac{t(\iota^c(t))'}{1 - \iota^c(t)\left(\frac{q^\top\nabla U_x(tq,t\iota^c(t))t}{q^\top\nabla U(tq,t\iota^c(t)))} - \frac{U_{xx}(tq,t\iota^c(t))t}{U_x(tq,t\iota^c(t))}\right)} \tag{37}$$

Proof. We will need

Lemma 8. *Let* $f(x)$ *be a strictly decreasing function with* $f'(x) < 0$ *for all* x *and* x_* *the unique solution to* $x_* = f(x_*)$. *Then, there exists a constant* $K > 0$ *such that* $|x - f(x)| < \varepsilon$ *implies*

$|x - x_*| < K\varepsilon$.

Proof. It suffices to consider the case when $x < x_*$. In this case, when ε is small enough, strict positivity of f' we have $f(x_*) + \delta_2(x_* - x) > f(x) > f(x_*) + \delta_1(x_* - x)$ for some $\delta_1, \delta_2 > 0$ and x close to x_*, and the claim follows when we consider that $x - f(x) = x - f(x) - (x_* - f(x_*))$. ■ First, define

$$\phi(t; L) \; = \; (L-1)(\iota(t) - \iota^c(t)) \, . \tag{38}$$

Let us show that ϕ stays bounded as $L \to \infty$. Indeed, the integral equation

$$\iota(t) \; = \; (L-1) \int_1^\infty g(t\xi, \iota(t\xi)) \xi^{-(L-1)} d\xi \tag{39}$$

can be rewritten

$$\iota(t) \; = \; \int_0^1 g(ty^{-1/(L-1)}, \iota(ty^{-1/(L-1)})) dy \, , \tag{40}$$

where we have used the change of variables $\xi^{-(L-1)} = y$. Differentiating, we get

$$\begin{aligned} 0 \ge \iota'(t) \; &= \; \int_0^1 y^{-1/(L-1)} g_t(ty^{-1/(L-1)}, \iota(ty^{-1/(L-1)})) dy \\ &+ \; \int_0^1 \iota'(ty^{-1/(L-1)}) y^{-1/(L-1)} g_x(ty^{-1/(L-1)} \iota(ty^{-1/(L-1)})) dy \\ &\ge \; \int_0^1 y^{-1/(L-1)} g_t(ty^{-1/(L-1)}, \iota(ty^{-1/(L-1)})) dy \; \ge \; K \int_0^1 y^{-1/(L-1)} dy \; = \; K \frac{1}{1 - 1/(L-1)} \end{aligned} \tag{41}$$

and thus $\iota'(t)$ is bounded as $L \to \infty$. Thus, (8) and Lemma 8 imply that $\iota(t) - \iota^c(t) = O(1/(L-1))$ as $L \to \infty$. Passing to a subsequence, we may assume that a limit of $\phi(t) = \phi(t; L), \; L \to \infty$ exists.

Now, we can compute $\phi(t)$. Substituting $\iota(t) = \iota^c(t) + (L-1)^{-1}\phi(t) + O((L-1)^{-2})$ into the ODE (8), we get

$$
\begin{aligned}
&\iota^c(t) + (L-1)^{-1}\phi(t) + O((L-1)^{-2}) \;=\; \iota(t) \\
&= \; -\frac{q^\top \nabla U\left(tq, t(\iota^c(t) + (L-1)^{-1}\phi(t) + O((L-1)^{-2}))\right)}{U_x\left(tq, t(\iota^c(t) + (L-1)^{-1}\phi(t) + O((L-1)^{-2}))\right)} + \frac{t(\iota^c(t) + (L-1)^{-1}\phi(t) + O((L-1)^{-2}))'}{L-1} \\
&= \; -\frac{q^\top \nabla U\left(tq, t\iota^c(t))\right) + q^\top \nabla U_x(tq, t\iota^c(t))t\phi(t)(L-1)^{-1}}{U_x\left(tq, t(\iota^c(t))\right) + U_{xx}(tq, t\iota^c(t))t\phi(t)(L-1)^{-1}} + \frac{t(\iota^c(t))'}{L-1} \;+\; O((L-1)^{-2}) \\
&= \; \iota^c(t)\left(1 + \frac{q^\top \nabla U_x(tq, t\iota^c(t))t}{q^\top \nabla U\left(tq, t\iota^c(t))\right)}\phi(t)(L-1)^{-1} - \frac{U_{xx}(tq, t\iota^c(t))t}{U_x(tq, t\iota^c(t))}\phi(t)(L-1)^{-1}\right) \\
&+ \; \frac{t(\iota^c(t))'}{L-1} \;+\; O((L-1)^{-2})\,.
\end{aligned}
\tag{42}
$$

Thus,

$$
\phi(t) \;=\; \frac{t(\iota^c(t))'}{1 - \iota^c(t)\left(\frac{q^\top \nabla U_x(tq, t\iota^c(t))t}{q^\top \nabla U(tq, t\iota^c(t)))} - \frac{U_{xx}(tq, t\iota^c(t))t}{U_x(tq, t\iota^c(t))}\right)} \tag{43}
$$

■

A.2 Proof of Proposition 4

Proof of Proposition 4. Let us rewrite the (5) as follows:

$$
t\frac{1}{L-1}\iota'(t) = \iota(t) - v(t, t\iota(t)),
$$

where

$$
v(t, t\iota(t)) = -\frac{q^\top \nabla U(tq, t\iota(t))}{U_x\left(tq, t\iota(t)\right)}.
$$

Let $\bar{t} > 0$ be such that $-\frac{q^\top \nabla U(tq,x)}{U_x(tq,x)}$ is decreasing in x for $t > \bar{t}$.

In our proofs, we will be frequently using the following classic comparison theorem for ODEs.

Lemma 9. *Let $f(t)$ be the unique solution to the ODE $f'(t) = F(t, f(t))$, $f(t_0) = f_0$ with a Lipshits-continuous F. Let also $g(t)$ satisfy $g(t_0) \le f_0$ and $g'(t) \le F(t, g(t))$. Then, $g(t) \le f(t)$*

for all $t > t_0$.

Lemma 10. *Suppose that* $\iota(t) > \iota^c(t)$ *for some* $t > \bar{t}$. *Then,* $\iota'(t) > 0$.

Proof. We have

$$t\frac{1}{L-1}\iota'(t) \;=\; \iota(t) - v(t, t\iota(t)) \underbrace{>}_{t>\bar{t}} \iota(t) - v(t, t\iota^c(t)) \;= \iota(t) - \iota^c(t) > 0\,.$$

■

Lemma 11. *Let* $\underline{I} = \min_{t>0} \iota^c(t)$. *Then, any bounded solution* $\iota(t)$ *to* (5) *satisfies* $\iota(t) \geq \underline{I}$ *for all* $t > \bar{t}$.

Proof. Suppose on the contrary that $\iota(t) < \underline{I}$ for some $t_0 > \bar{t}$. Since $\iota(t)$ is decreasing, this inequality also holds for all $t > t_0$. Therefore, for all $t > t_0$, we have

$$v(t, t\iota(t)) > v(t, t\underline{I}) \;=\; \underline{I}\,.$$

Therefore, for all $t > t_0$, we have

$$t\frac{1}{L-1}\iota'(t) \;=\; \iota(t) - v(t, t\iota(t)) \;<\; \iota(t) - \underline{I}.$$

Therefore, by Lemma 9, $\iota(t) \leq \xi(t)$ where $\xi(t)$ solves $t\frac{1}{L-1}\xi'(t) = \xi(t) - \underline{I}$ with $\xi(t_0) = \iota(t_0)$. We have

$$\xi(t) \;=\; (\iota(t_0) - \underline{I})t_0^{1-L}t^{L-1} + \underline{I}$$

Since $\iota(t_0) - \underline{I} < 0$, we get that $\xi(t)$ converges to $-\infty$ as $t \to \infty$, and hence $\iota(t)$ cannot be bounded. ■

Lemma 12. *All decreasing and bounded solutions* $\iota(t)$ *to* (8) *satisfy* $\underline{I} \;\leq\; \iota(t) \;<\; \iota^c(t)$ *and, hence,* $\lim_{t\to+\infty} \iota(t) = \underline{I}$ *for all such solutions.*

Proof. The first claim follows from Lemmas 11 and 10. The second claim follows because $\lim_{t\to+\infty} \iota^c(t) = \underline{I}$. ■

It remains to prove that a decreasing, bounded solution is unique. Suppose the contrary. Let $\iota_1(t)$ and $\iota_2(t)$ be two such solutions. Pick a $t_0 > \bar{t}$ and assume without loss of generality that $\iota_1(t_0) > \iota_2(t_0)$. By the uniqueness theorem for ODEs, these two solutions cannot intersect and, hence, $\iota_1(t) > \iota_2(t)$ for all t. Therefore, using the assumption that $v(t,x)$ is monotone decreasing in x, we get

$$t\frac{1}{L-1}\left(\iota_1(t)-\iota_2(t)\right)' = \underbrace{\iota_1(t)-\iota_2(t)}_{>0} - \underbrace{\left(v(t,t\iota_1(t))-v(t,t\iota_2(t))\right)}_{<0} > 0$$

That is, the difference between $\iota_1(t)$ and $\iota_2(t)$ is increasing as t grows. Thus, by Lemma 12,

$$0 < \iota_1(t) - \iota_2(t) \;\le\; \lim_{t\to+\infty}(\iota_1(t)-\iota_2(t)) = 0,$$

which is a contradiction. Thus, if a descreasing and bounded solution exists, it is unique.

Let us now prove existence. For each $t_0 > \bar{t}$, let $\iota_{t_0}(t)$ be the unique solution to (8) satisfying $\iota_{t_0}(t) = \iota^c(t_0)$. Then, the same comparison argument as in Lemma 10 implies that $\iota_{t_0}(t)$ satisfies $\iota_{t_0}(t) < \iota^c(t)$ for all $t < t_0$ and, hence, $\iota_{t_0}(t)$ is decreasing for $t_0 < t$ (same argument implies that it is increasing for $t > t_0$. Sending t_0 to ∞, we get that we can select a subsequence of $\iota_{t_0}(t)$ functions that converges on compact subsets of $\mathbb{R}^+$. Clearly, its limit is globally decreasing and is bounded from below by $\underline{I}$.

The proof of Proposition 4 is complete. ■

A.3 Proof of Proposition 5

We start with the following Lemma.

Lemma 13. *Suppose Assumption 3 holds. Then, if* $\psi_k(\cdot) < 0$ *(*> 0*) then* $\frac{\partial}{\partial k} I_q^M(t;k) > 0$ *(*< 0*)*

for every $0 < t < M$*, where* $I_q^M(t;k)$ *is defined in the Proposition 4.*

Proof of Lemma 13. We prove the Lemma for the case $\psi_k(\cdot) < 0$. The second case is symmetrical.

Differentiate the ODE (44) with respect to k. We get

$$\frac{t}{L-1}\frac{\partial^2 I_q^M(t;k)}{\partial k \partial t} = \psi_w(t, I_q^M(t;k);k) + \psi_I(t, I_q^M(t;w);k)\frac{\partial I_q^M(t;k)}{\partial k}. \tag{44}$$

Denote $\frac{\partial I_q^M(t;k)}{\partial k} = \theta(t;k)$, then the above ODE can be rewritten as

$$\frac{t}{L-1}\frac{\partial \theta}{\partial t} = \psi_k(\cdot) + \psi_I(\cdot)\theta. \tag{45}$$

Since for each k we have that $I_q^M(M;k) = I_q^c(M,k)$ and $I_q^c(M,k)$ shifts up when k increases (indeed, by the Implicit Function Theorem, $\frac{\partial}{\partial k} I_q^c(t;k) = -\psi_k(\cdot)/\psi_I(\cdot) > 0$), we have that

$$\theta(M;k) > 0.$$

It suffices to prove that θ stays positive for all $t < M$. Suppose, on the contrary, that θ becomes zero at some point N (for the first time as we move to the left of M). It must be $\frac{\partial \theta(N;w)}{\partial t} > 0$ (at that point $\theta(t)$ crosses zero from above) and $\theta(N;w) = 0$. At this point (45) becomes

$$\frac{t}{L-1}\underbrace{\frac{\partial \theta}{\partial t}}_{>0} = \underbrace{\psi_k(\cdot)}_{<0}.$$

A contradiciton. ■

Since the equilibrium demand is a limit of $I_q^c(M,w)$ we have the following proposition.

Proof of Proposition 5. First, compute

$$\psi_I = 1 - \frac{E\left[\left(q^\top \delta\right) u_1'(w_1 + tq^\top \delta)\right]}{u_0'(w_0 - t\iota(t))^2} u_0''(w_0 - t\iota(t))t > 0.$$

The inequality is true because if Inada conditions hold for $u_1(\cdot)$, we must have $q^\top\delta > 0$. Otherwise, consumption can become negative, for which the utility function $u_1(\cdot)$ is not defined.

Similarly,

$$\psi_t = -\frac{E\left[\left(q^\top\delta\right)^2 u_1''(w_1 + tq^\top\delta)\right] u_0'(w_0 - t\iota(t)) + \iota(t)u_0''(w_0 - t\iota(t))E\left[\left(q^\top\delta\right) u_1'(w_1 + tq^\top\delta)\right]}{u_0'(w_0 - t\iota(t))^2}$$
$$= -\frac{E\left[\left(q^\top\delta\right)^2 u_1''(c_1)\right] - \iota(t)ARA(c_0)E\left[\left(q^\top\delta\right) u_1'(c_1)\right]}{u_0'(c_0)} > 0$$

The inequality is true because if Inada conditions hold for $u_1(\cdot)$, we must have $q^\top\delta > 0$. But then the equilibrium demand $\iota(t)$ will be positive as well. Therefore, Assumption 1 holds.

Now compute ψ_{w_0} and ψ_{w_1}

$$\psi_{w_0} = \frac{E\left[\left(q^\top\delta\right) u_1'(w_1 + tq^\top\delta)\right]}{u_0'(w_0 - t\iota(t)))^2} u_0''(w_0 - t\iota(t)))$$

Note that if Inada conditions hold for $u_1(\cdot)$, we must have $q^\top\delta > 0$ (otherwise, consumption can become negative). Thus $\psi_{w_0} < 0$.

By a similar argument,

$$\psi_{w_1} = -\frac{E\left[\left(q^\top\delta\right) u_1''(w_1 + tq^\top\delta)\right]}{u_0'(w_0 - t\iota(t)))} > 0.$$

The rest follows from Lemma 13. ■

A.4 Proof of Proposition 6

The Poof relies on several Lemmas.

Lemma 14. *sign* $\left[(t\iota^c(t))'\right] =$ *sign* $\left[1 - E\left[RRA_1(tq^\top\delta)\right]\right]$

Proof. Denote $Y(t) = t\iota^c(t)$. It solves

$$Y(t)u_0'(w_0 - Y(t)) = E\left[t\left(q^\top\delta\right)u_1'(tq^\top\delta)\right]$$

Differentiate wrt to t

$$\begin{aligned} Y(t)'\left(u_0'(c_0) - Y(t)u_0''(c_0)\right) &= E\left[\left(q^\top\delta\right)u_1'(tq^\top\delta)\right] + E\left[t^2\left(q^\top\delta\right)^2 u_1''(tq^\top\delta)\right] \\ &= E\left[\left(q^\top\delta\right)u_1'(tq^\top\delta)\right]\left(1 - E\left[RRA_1(tq^\top\delta)\right]\right) \end{aligned}$$

■

Lemma 15. *Suppose that $RRA_0(c_0) = RRA_1(c_1) = \gamma \geq 1$. Then $\frac{\partial}{\partial w_0} I_q^c(t; w_0)$ is decreasing in t.*

Proof. By the Implicit Function Theorem,

$$\frac{\partial}{\partial w_0} I_q^c(t; w_0) = -\frac{\psi_{w_0}(t, \iota^c(t); w_0)}{\psi_I(t, \iota^c(t); w_0)}.$$

Compute

$$\begin{aligned} \psi_{w_0} &= \frac{E\left[\left(q^\top\delta\right)u_1'(tq^\top\delta)\right]}{u_0'(w_0 - t\iota^c(t))^2} u_0''(w_0 - t\iota^c(t)) \\ &= \frac{E\left[\left(q^\top\delta\right)u_1'(tq^\top\delta)\right]}{u_0'(w_0 - t\iota^c(t))} \frac{u_0''(w_0 - t\iota^c(t))}{u_0'(w_0 - t\iota^c(t))} \\ &= -\frac{\gamma\iota^c(t)}{w_0 - t\iota^c(t)} \end{aligned}$$

Note that by Lemma 14 $-\psi_{w_0} > 0$ is decreasing in t.

Similarly,

$$\psi_I = 1 - \frac{E\left[\left(q^\top\delta\right)u_1'(tq^\top\delta)\right]}{u_0'(w_0 - t\iota^c(t))}\frac{u_0''(w_0 - t\iota^c(t))}{u_0'(w_0 - t\iota^c(t))}$$
$$= 1 - \iota^c(t)\frac{u_0''(w_0 - t\iota^c(t))}{u_0'(w_0 - t\iota(t))}t$$
$$= 1 - \psi_{w_0}t$$

Therefore,

$$-\frac{\psi_I}{\psi_{w_0}} = -1/\psi_{w_0} + t,$$

which is increasing in t. Thus, $\frac{\partial}{\partial w_0}I_q^c(t;w_0) = -\frac{\psi_{w_0}}{\psi_I}$ is decreasing in t. ■

Lemma 16. *Suppose that* $RRA_0(c_0) = \gamma > 1$. *Then* $\frac{\partial}{\partial w_0}I_q^c(t;w_0) > -\frac{\psi_{w_0}(t,\iota(t);w_0)}{\psi_I(t,\iota(t);w_0)}$.

Proof. Compute

$$\psi_{w_0} = \frac{E\left[\left(q^\top\delta\right)u_1'(tq^\top\delta)\right]}{u_0'(w_0 - t\iota(t))^2}u_0''(w_0 - t\iota(t))$$

and

$$\psi_I = 1 - \frac{E\left[\left(q^\top\delta\right)u_1'(tq^\top\delta)\right]}{u_0'(w_0 - t\iota(t))}\frac{u_0''(w_0 - t\iota(t))}{u_0'(w_0 - t\iota(t))}$$
$$= 1 - \psi_{w_0}t.$$

We have that

$$-\frac{\psi_{w_0}}{\psi_I} = \left(-1/\psi_{w_0} + t\right)^{-1}.$$

At the same time (Lemma 15)

$$\frac{\partial}{\partial w_0} I_q^c(t; w_0) = \left(-1/\psi_{w_0}^c + t\right)^{-1},$$

where

$$\psi_{w_0}^c = \frac{E\left[\left(q^\top \delta\right) u_1'(tq^\top \delta)\right]}{u_0'(w_0 - t\iota^c(t))^2} u_0''(w_0 - t\iota^c(t)).$$

Since for CRRA utility $\frac{u_0''(x)}{u_0'(x)^2} = -\gamma x^{\gamma-1}$ is increasing in x and $\iota(t) < \iota^c(t)$ we get

$$-1/\psi_{w_0}^c < -1/\psi_{w_0}$$

Therefire

$$-\frac{\psi_{w_0}}{\psi_I} = (-1/\psi_{w_0} + t)^{-1} < \left(-1/\psi_{w_0}^c + t\right)^{-1} = \frac{\partial}{\partial w_0} I_q^c(t; w_0)$$

■

Proof of Proposition 6. Denote

$$\theta^c(t; w_0) = \frac{\partial I_q^c(t; w_0)}{\partial w_0}$$

and

$$\theta^0(t; w_0) = -\frac{\psi_{w_0}(t, \iota(t); w_0)}{\psi_I(t, \iota(t); w_0)}$$

which is a locus where $\frac{\partial \theta}{\partial t} = 0$. From Lemma 16 we know that $\theta^0(t)$ lies below $\theta^c(t)$. Moreover, at point M $I_q^c(t; w_0) = I_q(t; w_0)$ and we have that $\theta^0(M) = \theta^c(M) > 0$. Suppose $\theta^M(t)$ crosses $\theta^c(t)$ at some point N (for the first time as we move left of the point M). At this point $\theta^M(N)' < 0$ (the crossing has to be from below, since at point M $\theta^M(t)$ crosses $\theta^c(t)$ horiszontally). But his means that $\theta^M(N)$ is below $\theta^0(t)$ which contradicts the fact that $\theta^0(t)$ lies below $\theta^c(t)$. ■

A.5 Additional proofs

Proposition 18. *Pick a pair of compact sets $\mathcal{K}$, $\mathcal{Y} \subset \mathbb{R}^N$ such that $\mathcal{Y}$ contains an open neighborhood of zero. Then, there exists a threshold $L^* = L^*(\mathcal{K}, \mathcal{Y})$ such that, for all $\varepsilon \in \mathcal{K}$ and $L > L^*$ the function $\Phi(q) = v(q) - q \cdot \nabla Z^*((q-\varepsilon)/(c(L-1)))$ attains its maximum over $q \in \mathcal{Y}$ at $q = \varepsilon/L$ and hence $(D(p), \mathcal{Y})$ with $D(p) = -c^{-1}\nabla Z(p)$ is the equilibrium. In this equilibrium, the price impact matrix Λ satisfies the following properties:*

- *Λ is symmetric: $\Lambda_{i,j} = \Lambda_{j,i}$ for all i, j.*
- *Λ is positive-definite. In particular, $\Lambda_{i,i} > 0$ for all i.*
- *if δ_i and δ_j are independent, then $\Lambda_{i,j} = 0$.*

Proof of Proposition 18. The following lemma follows by standard arguments.

Assumption 7. *The probability distribution of δ has a a compact support.*

Lemma 17. *Under the Assumption 7, there exist constants $c_1 > c_2 > 0$ such that the eigenvalues of $-\nabla^2 v(q)$ and $-\nabla^2 Z^*(q)$ belong to $[c_2, c_1]$. Furthermore, $\nabla^3 Z^*$ is uniformly bounded on $\mathbb{R}^N$.*

Now, pick a compact set $\mathcal{K}$. Our goal is to show that the function $\Phi(q) = v(q) - q \cdot \nabla Z^*((q-\varepsilon)/(c(L-1)))$ attains its global maximum at $q = s/L$ over all $q \in \mathcal{Y}$ when L is sufficiently large. Define $\Psi(q, \varepsilon) \equiv -q \cdot \nabla Z^*((q-\varepsilon)/(c(L-1)))$. Then,

$$\nabla_q \Psi \;=\; -\nabla Z^*((q-\varepsilon)/(c(L-1))) - q\frac{1}{c(L-1)}\nabla^2 Z^*((q-\varepsilon)/(c(L-1)))$$

and

$$\nabla_q^2 \Psi \;=\; -\frac{1}{c(L-1)}\left(2\nabla^2 Z^*((q-\varepsilon)/(c(L-1))) + q\frac{1}{c(L-1)}\nabla^3 Z^*((q-\varepsilon)/(c(L-1)))\right).$$

Thus, Φ is concave for $(\varepsilon, q) \in \mathcal{K} \times \mathcal{Y}$ when L is sufficiently large, and the claim follows. ■

Everywhere in the sequel, we always make the following assumption.

Assumption 8. *The function $U(q, x)$ is strictly concave in x.*

The following theorem shows that finding equilibria reduces to solving just one ODE.

Theorem 1. *Let $D(p)$ be an equilibrium. Then, the function $F(q) \equiv q \cdot D^{-1}(q)$ solves the PDE*

$$(1-c)q \cdot \nabla F(q) \;-\; F(q) \;=\; c\frac{q \cdot \nabla_q U(q, F(q))}{U_x(q, F(q))}\,, \tag{46}$$

and the inverse of the map D is given by

$$D^{-1}(q) \;=\; (1-c)\nabla F(q) \;-\; c\frac{\nabla_q U(q, F(q))}{U_x(q, F(q))}\,. \tag{47}$$

The general solution to (46) can be constructed as follows. Consider the ODE

$$(1-c)tf'(t) \;-\; f(t) \;=\; c\frac{tq \cdot \nabla_q U(tq, f(t))}{U_x(tq, f(t))}\,, \tag{48}$$

and let $f(t; q; \alpha)$ be the solution to this ODE satisfying $f(1; q, \alpha) = \alpha$. Then, the general solution to (46) is given by $F(q; \rho) = f(|q|; q/|q|; \rho(q/|q|))$, where $\rho(x)$ is an arbitrary smooth function on the unit sphere $\{q \in \mathbb{R}^N : |q| = 1\}$ with $|q| \equiv (q \cdot q)^{1/2}$.

Proposition 19. *Equilibrium asset prices are given by*

$$p \;=\; \frac{\nabla_q U(\varepsilon/L\,,\, p \cdot \varepsilon/L)}{U_x(\varepsilon/L\,,\, p \cdot \varepsilon/L)} - \Lambda\left(\frac{L-1}{L}\varepsilon\right)\frac{\varepsilon}{L} \tag{49}$$

where the equilibrium price impact matrix Λ is given by

$$\Lambda(q) \;=\; -\partial \Pi(q)^T\,, \tag{50}$$

whereby the cost of acquiring an infinitesimally small portfolio x is given by $x \cdot \Lambda(q)x$. In particular, trading costs are always positive if and only if $\Lambda(q)$ is positive semi-definite.

Proposition 19 shows that equilibrium prices are given by the standard Euler equation plus a liquidity correction term accounting for price impact Λ . Thus, understanding the impact of the agents' strategic behavior on asset prices is equivalent to understanding the behavior of the price impact matrix Λ. This is precisely the goal of this paper.

By definition, $\Pi(q)$ is the inverse of the map $(L-1)D$, so that $\Pi((L-1)D(p)) = p$. Differentiating, we get

$$\partial\Pi|_{(L-1)D(p)} = (L-1)^{-1}(\partial D(p))^{-1} ,$$

and hence

$$\nabla_q U(D(p), D(p)\cdot p) + U_x(D(p), D(p)\cdot p)\,(p - (L-1)^{-1}(\partial D(p)^T)^{-1}\,D(p)) = 0 . \qquad (51)$$

As we can see, (51) is a highly non-linear system of partial differential equations for the N components of the map D, that may be difficult to tackle analytically when the number N of securities is large. However, as we show below, quite surprisingly, (51) can be reduced to solving just one ordinary differential equation (ODE).

Multiplying (51) by the matrix $\partial D(p)^T$, and adding $U_x D(p)$ to both sides of the equation, we get

$$\partial D(p)^T \nabla_q U + U_x \partial D(p)^T p + U_x\, D(p) = \frac{L}{L-1} U_x D(p) . \qquad (52)$$

Let

$$Z(p) \equiv U(D(p), D(p)\cdot p) \qquad (53)$$

be the agent's equilibrium indirect utility for a given price vector p. Let also

$$Y(p) \equiv U_x(D(p), D(p)\cdot p) , \qquad (54)$$

and

$$c \;\equiv\; \frac{L-1}{L}\,. \tag{55}$$

Then, we have

$$\nabla Z(p) \;=\; \partial D(p)^T \nabla_q U(D(p), D(p)\cdot p) + U_x(D(p), D(p)\cdot p)(\partial D(p)^T p + D(p))\,, \tag{56}$$

and hence we can rewrite (52) as

$$\nabla Z(p) \;=\; \frac{1}{c}\,Y(p)\,D(p)\,. \tag{57}$$

Substituting this identity into the definition of $Z(p)$, we arrive at the following system of PDEs for $(Z(p), Y(p))$:

$$\begin{aligned} Y(p) \;&=\; U_x(c\nabla Z(p)/Y(p), cp\cdot\nabla Z(p)/Y(p)) \\ Z(p) \;&\equiv\; U(c\nabla Z(p)/Y(p), cp\cdot\nabla Z(p)/Y(p))\,. \end{aligned} \tag{58}$$

Proof of Theorem 1.

Consider now an equilibrium demand schedule

$$D(p) = -c\nabla Z(p)/Y(p)\,,$$

where $Z(p)$ and $Y(p)$ are the functions from Proposition 2. Let $q = -c\nabla Z(p)/Y(p)$ and let $\pi(q) = D^{-1}(q)$. Define $\tilde{Y}(q) = Y(\pi(q))$. Then, we have $\nabla Z(p(q)) = -\tilde{Y}(q)q/c$, and the system (58) takes the form

$$\begin{aligned} \tilde{Y}(q) \;&=\; -U_x(q, \pi(q)\cdot q) \\ Z(\pi(q)) \;&=\; U(q, \pi(q)\cdot q)\,. \end{aligned} \tag{59}$$

Define $J(q, x)$ to be the unique solution to

$$x \;=\; -U_x(q, J)\,.$$

Then, the first equation in (59) implies that

$$\pi(q)\cdot q = \; J(q,\tilde{Y}(q))\,. \tag{60}$$

Differentiating the second equation in (59) with respect to q, we get

$$-\tilde{Y}(q)c^{-1}(\partial\pi(q))^Tq \;\equiv\; \nabla_qU(q,J(q,\tilde{Y}(q))\rangle) + U_x(q,J(q,\tilde{Y}(q)))[J_q(q,\tilde{Y}(q)) + J_x\nabla\tilde{Y}(q)]\,. \tag{61}$$

Differentiating the identity (60) we get

$$(\partial\pi(q))^Tq \;+\; \pi(q) \;=\; J_q(q,\tilde{Y}(q)) \;+\; J_x\nabla\tilde{Y}(q)$$

and hence we can rewrite equation (61) as

$$\begin{aligned} &-\tilde{Y}(q)c^{-1}[J_q(q,\tilde{Y}(q)) + J_x\nabla\tilde{Y}(q) - \pi(q)] \\ &=\; \nabla_qU(q,J(q,\tilde{Y}(q))) + U_x(q,J(q,\tilde{Y}(q)))[J_q(q,\tilde{Y}(q)) + J_x\nabla\tilde{Y}(q)]\,. \end{aligned} \tag{62}$$

Using the fact that, by definition, $U_x(q,J(q,\tilde{Y}(q))) = -\tilde{Y}(q)$, and multiplying equation (62) by q, we get

$$\begin{aligned} &\tilde{Y}(q)c^{-1}[q\cdot J_q(q,\tilde{Y}(q)) + J_xq\cdot\nabla\tilde{Y}(q) - J(q,\tilde{Y}(q))] \\ &=\; -q\cdot\nabla_qU(q,J(q,\tilde{Y}(q))\rangle) + \tilde{Y}(q)[q\cdot J_q(q,\tilde{Y}(q)) + J_xq\cdot\nabla\tilde{Y}(q)]\,. \end{aligned} \tag{63}$$

Now, define

$$F(q) \;\equiv\; J(q,\,\tilde{Y}(q))\,.$$

Then, a direct calculation implies that we can rewrite the PDE as

$$((c^{-1}-1)q\cdot\nabla F(q) \;-\; c^{-1}F(q)) \;=\; \frac{q\cdot\nabla_qU(q,F(q))}{U_x(q,F(q))}\,, \tag{64}$$

which yields both Equations (46) and (47). Furthermore, equation (62) implies that

$$\begin{aligned} \pi(q) &= (1-c)\nabla F(q) - c\frac{\nabla_q U(q,F(q))}{U_x(q,F(q))} = \nabla F(q) - \frac{c}{U_x(q,F(q))}\nabla \hat{U}(q) \\ &= -\frac{\nabla_q U(q,F(q))}{U_x(q,F(q))} + \frac{1-c}{U_x(q,F(q))}\nabla \hat{U}(q) \end{aligned} \tag{65}$$

with $\hat{U}(q) = U(q, F(q))$.

Let us now make a change of variables to spherical coordinates

$$q_1 = t\cos\phi_1,\ q_2 = t\sin\phi_1\cos\phi_2,\ \cdots,\ q_N = t\sin\phi_1\sin\phi_2\cdots\sin\phi_{N-1}\,.$$

Then, let $\Psi(t,\phi) = F(q(t,\phi))$. Then, we have

$$\frac{\partial \Psi}{\partial t} = \frac{\partial q}{\partial t}\cdot\nabla F = \frac{1}{t}q\cdot\nabla F(q)\,,$$

and thus (64) takes the form

$$(1-c)t\frac{\partial \Psi}{\partial t} - \Psi(t,\phi) = c\,\frac{t\phi\cdot\nabla_q U(t\phi,\Psi(t,\phi))}{U_x(t\phi,\Psi(t,\phi))}$$

and we arrive at (48).

■

Let $G(q) = w_0 - F(q)$ be the agent's equilibrium time-zero consumption when he acquires portfolio q, and let $\hat{V}(q) = q\cdot\nabla V(q)$. Then, a direct calculation yields the following result.

Corollary 1. *The function $G(q)$ satisfies the PDE*

$$u'(G)\left[(1-c)q\cdot\nabla G(q) - G(q) + w_0\right] = c\hat{V}(q)\,, \tag{66}$$

and the inverse $D^{-1}(q) = \pi(q)$ *of equilibrium demand is given by*

$$\pi(q) \;=\; \frac{\nabla S(q)}{-u'(G(q))} \tag{67}$$

with

$$S(q) \;\equiv\; (1-c)u(G(q)) - cV(q)\,.$$

Moreover,

$$q \cdot \pi(q) \;=\; w_0 \;-\; G(q) \qquad \text{and} \qquad q^T(\nabla \pi(q))^T q \;=\; (L-1)\left[w_0 \;-\; G(q) \;-\; \frac{\hat{V}(q)}{u'(G(q))}\right].$$

Proof of Corollary 1. The assumption

$$U(q,x) \;=\; u(w_0 - x) \;+\; V(q)$$

implies that

$$\nabla_q U(q,x) \;=\; \nabla V(q) \qquad \text{and} \qquad U_x(q,x) \;=\; -u'(w_0 - x)\,.$$

The result then follows directly from the PDE (46) and the definitions of G and $\hat{u}_1$.

We know that G satisfies the following differential equation:

$$u'(G(q))\,[(1-c)q \cdot \nabla G(q) \;-\; G(q) \;+\; w_0] = \; c\hat{u}_1(q)$$
$$\Longrightarrow \quad (1-c)q \cdot \nabla G(q) = \; \frac{c}{u'(G(q))}\hat{u}_1(q) \;+\; G(q) \;-\; w_0\,.$$

Furthermore, by (47), since $G(q) = w_0 - F(q)$, we have

$$\pi(q) \;=\; (1-c)\nabla F(q) \;-\; c\frac{\nabla_q U(q,F(q))}{U_x(q,F(q))} \;=\; -(1-c)\nabla G(q) \;+\; c\frac{\nabla_q V(q)}{u'(G(q))} \;=\; -\frac{\nabla S(q)}{u'(G(q))} \tag{68}$$

because

$$\nabla S(q) \;=\; (1-c)u'(G(q))\nabla G(q) \;-\; c\nabla V(q)\,.$$

Thus,

$$\begin{aligned} q\cdot\pi(q) = \; -\frac{q\cdot\nabla S(q)}{u'(G(q))} \;&=\; -(1-c)q\cdot\nabla G(q) \;+\; \frac{c}{u'(G(q))}q\cdot\nabla V(q) \\ =\; w_0 \;-\; G(q)\,. \end{aligned}$$

This equality implies that

$$-\nabla G(q) \;=\; \nabla(q\cdot\pi(q)) \;=\; \pi(q) \;+\; (\nabla\pi(q))^T q \qquad \Longrightarrow \qquad (\nabla\pi(q))^T q \;=\; -\pi(q) \;-\; \nabla G(q)\,.$$

Multiplying both sides of the last equality by q, we obtain

$$\begin{aligned} q^T(\nabla\pi(q))^T q =&\; -q\cdot\pi(q) \;-\; q\cdot\nabla G(q) \\ =&\; -q\cdot\pi(q) \;-\; \frac{1}{1-c}\left[\frac{c}{u'(G(q))}q\cdot\nabla V(q) \;-\; q\cdot\pi(q)\right] \\ =&\; (L-1)\left[q\cdot\pi(q) \;-\; \frac{1}{u'(G(q))}\hat{u}_1(q)\right]\,. \end{aligned}$$

■

Lemma 18. *Suppose that S is convex with a strictly positive definite Hessian. Then, $D(p)$ is bijective and ∂D is always invertible.*

Proof of Lemma 18. Our goal is to show that $\pi(q) = \frac{\nabla S(q)}{-u'(G(q))}$ is bijective. That is, equation

$$p \;=\; \frac{\nabla S(q)}{-u'(G(q))}$$

has unique solution $q = \pi^{-1}(p)$. Let $q \in \pi^{-1}(p)$ and let $a(q) = -u'(G(q))$. Let also $\Theta(x) = \nabla S^{-1}(x)$. By assumption, S is strictly convex, and hence Θ is well-defined, and $\partial\Theta = (\nabla^2 S)^{-1}$

is positive definite. We have

$$q \;=\; \Theta(a(q)p).$$

Furthermore, by Corollary 1, we have $G(q) = w_0 - q \cdot \pi(q) \;=\; w_0 - q \cdot p \;=\; w_0 - p \cdot \Theta(a(q)p)$. Thus, $a(q)$ satisfies

$$a(q) \;=\; -u'(G(q)) \;=\; -u'(w_0 - p \cdot \Theta(a(q)p)).$$

We have $\frac{\partial}{\partial a}(p \cdot \Theta(ap)) \;=\; p^T \partial\Theta(ap)p \geq 0$ because $\partial\Theta$ is positive definite. Thus, $-u'(w_0 - p \cdot \Theta(a(q)p))$ is monotone decreasing in a and hence $a \;=\; -u'(w_0 - p \cdot \Theta(p))$ has a unique solution $a = a(p)$, and then $q = \Theta(a(p)p)$ is also uniquely defined.

■

Proposition 20. *We have* $\Lambda(q) = \frac{1}{L-1}\Lambda^*(q/(L-1))$ *with*

$$\Lambda^*(q) \;=\; \frac{1}{u'(G(q))}\left(\nabla^2 S(q) - \frac{u''(G(q))}{u'(G(q))}\nabla G(q)((1-c)u'(G(q))\nabla G(q) - c\nabla V(q))^T\right)$$

Thus, Λ *is symmetric if and only if* $\nabla G(q)$ *is proportional to* $\nabla V(q)$*. Furthermore,*

$$u'(G)q \cdot \Lambda(q)q = (L-1)\left(\hat{V}(q) + (G(q) - w_0)u'(G)\right) \tag{69}$$

Proof of Proposition 20. We have

$$S(q) \;=\; (1-c)u(G(q)) - cV(q)\,,$$

and hence

$$\begin{aligned} \nabla S(q) \;&=\; (1-c)u'(G(q))\nabla G(q) - c\nabla V(q) \\ \nabla^2 S(q) \;&=\; (1-c)u'(G(q))\nabla^2 G(q) + (1-c)u''(G(q))\nabla G(q)(\nabla G(q))^T - c\nabla^2 V(q)\,. \end{aligned} \tag{70}$$

Differentiating (66), we get

$$u''(G)\,[(1-c)q\cdot\nabla G(q)\ -\ G(q)+w_0]\nabla G(q)\ +\ u'(G)[-c\nabla G(q)+(1-c)\nabla^2 G(q)q]\ =\ c\nabla\hat{u}_1(q) \tag{71}$$

and hence, defining $\chi\equiv q\cdot\nabla G(q)$, we get

$$u'(G)(1-c)q\cdot\nabla^2 G(q)q\ =\ cq\cdot\nabla\hat{u}_1(q)-u''(G)\,[(1-c)\chi\ -\ G(q)+w_0]\chi+u'(G)c\chi$$

and

$$\begin{aligned} q\cdot\nabla S(q)\ &=\ (1-c)u'(G(q))\chi-c\hat{u}_1(q)\\ q\cdot\nabla^2 S(q)q\ &=\ [cq\cdot\nabla\hat{u}_1(q)-u''(G)\,[(1-c)\chi\ -\ G(q)+w_0]\chi+u'(G)c\chi]\\ &+(1-c)u''(G(q))\chi^2-cq\cdot\nabla^2 V(q)q\\ &=q\cdot\nabla\hat{u}_1(q)-u''(G)\,[-G(q)+w_0]\chi+u'(G)c\chi-cq\cdot\nabla^2 V(q)q\,. \end{aligned} \tag{72}$$

Hence, we have

$$\begin{aligned} u'(G)q\cdot\Lambda(q)q\ &=\ -u'(G)q\cdot\partial\pi(q)q\ =\ q\cdot\nabla^2 S(q)q\ -\frac{u''(G)}{(u'(G))}(q\cdot\nabla S(q))\,\chi\\ &=\ q\cdot\nabla\hat{u}_1(q)-u''(G)\,[-G(q)+w_0]\chi+u'(G)c\chi-cq\cdot\nabla^2 V(q)q\\ &-\frac{u''(G)}{(u'(G))}\chi[(1-c)u'(G(q))\chi-c\hat{u}_1(q)]\\ &=\ cq\cdot\nabla\hat{u}_1(q)-u''(G)\,[-G(q)+w_0]\chi+u'(G)c\chi-(1-c)u''(G(q))\chi^2\\ &-cq\cdot\nabla^2 V(q)q+\frac{u''(G)}{(u'(G))}\chi[c\hat{u}_1(q)]\\ &=\ c(q\cdot\nabla\hat{u}_1(q)-q\cdot\nabla^2 V(q)q+u'(G)\chi)\\ &=\ c\left(q\cdot\nabla\hat{u}_1(q)-q\cdot\nabla^2 V(q)q+(1-c)^{-1}(c\hat{u}_1(q)+(G(q)-w_0)u'(G))\right)\\ &=\ c(1-c)^{-1}\left(\hat{u}_1(q)+(G(q)-w_0)u'(G)\right)\,, \end{aligned} \tag{73}$$

where we have used that

$$(1-c)\chi\ =\ (c\hat{u}_1/u'(G)+G-w_0)$$

and

$$q \cdot \nabla \hat{u}_1 \;=\; \hat{u}_1 \;+\; q \cdot \nabla^2 V q\,.$$

In the competitive limit, as $c \to 1$, G satisfies

$$w_0 - G(q) \;=\; \frac{\hat{u}_1(q)}{u'(G(q))}$$

which implies

$$-u'(G)\chi \;=\; q \cdot \nabla \hat{u}_1(q) - \frac{u''(G)}{u'(G)}\hat{u}_1(q)\chi$$

so that

$$\chi \;=\; \frac{-q \cdot \nabla \hat{u}_1(q)}{u'(G) - \frac{u''(G)}{u'(G)}\hat{u}_1(q)}$$

■

The claim of Proposition **??** follows from the following lemma.

Lemma 19. *For any compact subset $\hat{\mathcal{V}} \subset \mathcal{V}$ and any compact subset $\hat{\mathcal{X}} \subset \mathcal{X}$ there exists an $L^* > 1$ such that*

$$\arg\max_{q \in \hat{\mathcal{V}}} \{u(w_0 - q \cdot \pi((\varepsilon - q)/(L-1))) \;+\; V(q)\} \;=\; \varepsilon/L$$

for any $\varepsilon \in \hat{\mathcal{X}} \cap (L\hat{\mathcal{V}})$ for any $L > L^*$.

Proof of Lemma 19. Recall that

$$\Pi(q) = \pi(q/(L-1))\,,$$

and hence we need to verify that

$$B(q) \;=\; u(w_0 - q \cdot \pi((\varepsilon - q)/(L-1))) \;+\; V(q)$$

attains its maximum at $q = \varepsilon/L$ for any ε. Differentiating, we get

$$\begin{aligned}
\nabla B(q) \;&=\; \nabla V(q) \\
&+\; u'(w_0 - q\cdot\pi((\varepsilon - q)/(L-1))) \\
&\times [-\pi((\varepsilon - q)/(L-1)) + (L-1)^{-1}\partial\pi^T((\varepsilon - q)/(L-1))q]
\end{aligned} \tag{74}$$

and

$$\begin{aligned}
\nabla^2 B(q) \;&=\; u''(w_0 - q\cdot\pi((\varepsilon - q)/(L-1)))K(q)K(q)^T \;+\; \nabla^2 V(q) \\
&+\; u'(w_0 - q\cdot\pi((\varepsilon - q)/(L-1))) \\
&\times [2\frac{1}{L-1}\partial\pi((\varepsilon - q)/(L-1)) - (L-1)^{-2}\partial^2\pi^T((\varepsilon - q)/(L-1))q]\,.
\end{aligned} \tag{75}$$

where we have defined

$$K(q) \;\equiv\; -\pi((\varepsilon - q)/(L-1)) + (L-1)^{-1}\partial\pi^T((\varepsilon - q)/(L-1))q\,.$$

The admissible set is such that $q\cdot\pi((\varepsilon - q)/(L-1))$ is uniformly bounded away from w_0 from above. ■

Proof of Lemma 2.

When the elasticity of inter-temporal substitution (EIS) equals one, that is $u(x) = \log(x)$, the PDE (66) characterizing G becomes

$$(1-c)q\cdot\nabla G(q) \;-\; (1 + c\hat{u}_1(q))G(q) \;=\; -w_0\,, \tag{76}$$

and the claim follows by direct calculation from Theorem 1 because the corresponding ODE is linear and hence can be solved explicitly. The result about the growth rate of ϕ follows from the lower bound on $\hat{u}_1$.

■

Proof of Lemma 3. We will need the following lemma.

Lemma 20. *Suppose that the function $A(q,t)$ is convex in q for any t. Then, for any random variable T taking values in the domain of definition of A, the function*

$$F(q) = \log E[e^{A(q,T)}]$$

is convex in q.

Proof. We have

$$\nabla F(q) \;=\; \frac{E[\nabla_q A(q,T)e^{A(q,T)}]}{E[e^{A(q,T)}]}$$

and

$$\begin{aligned} & E[e^{A(q,T)}]^2\, \nabla^2 F(q) \\ & = \; E[e^{A(q,T)}]E[(\nabla_q^2 A(q,T) + (\nabla_q A)(\nabla_q A)^T)e^{A(q,T)}] \\ & - E[\nabla_q A(q,T)e^{A(q,T)}]E[(\nabla_q A(q,T))^T e^{A(q,T)}] \end{aligned} \tag{77}$$

Pick a vector x. Then, the sign of the quadratic form $x^T\nabla^2 F(q)x$ coincides with that of

$$\begin{aligned} & E[e^{A(q,T)}]E[x^T(\nabla_q^2 A(q,T) + (\nabla_q A)(\nabla_q A)^T)xe^{A(q,T)}] \\ & - E[x^T\nabla_q A(q,T)e^{A(q,T)}]E[(\nabla_q A(q,T))^T xe^{A(q,T)}] \end{aligned} \tag{78}$$

Since, by assumption, A is convex, we have $E[x^T\nabla_q^2 A(q,T)x \geq 0$, and hence it suffices to show that

$$E[e^{A(q,T)}]E[x^T((\nabla_q A)(\nabla_q A)^T)xe^{A(q,T)}] \;\geq\; E[x^T\nabla_q A(q,T)e^{A(q,T)}]E[(\nabla_q A(q,T))^T xe^{A(q,T)}]\,. \tag{79}$$

Define a random variable $Y = x^T\nabla_q A(q,T)$, and a new measure with the density $d\tilde{P}/dP = e^{A(q,T)}(E[e^{A(q,T)}])^{-1}$ with respect to the original measure dP. Let also $\tilde{E}$ denote the expectation under this measure change. Then, the desired inequality (79) is equivalent to

$$\tilde{E}[Y^2] \;\geq\; \tilde{E}[Y]^2 \,,$$

and the claim follows. ■ Suppose now that $\hat{u}_1$ is concave. Define $A(q,t) \equiv -\int_1^t \frac{c\hat{u}_1\left(\rho^{1/L}q\right)}{\rho}d\rho$. Then, A is convex q and hence, by Lemma 20, so is

$$\log G(q) \;=\; \log \int_1^\infty e^{A(s)}\frac{ds}{s^2}\,.$$

Thus, $S = (1-c)\log G(q) - V(q)$ is also convex if V is concave.

By direct calculation,

$$\nabla^2 S(q) \;=\; \frac{(1-c)}{G^2(q)}\left[G(q)\nabla^2 G(q) \;-\; \nabla G(q)(\nabla G(q))^T\right] \;-\; c\nabla^2 V\,.$$

Passing to the limit as $L \to \infty$ in $G(q) = \int_1^\infty e^{-\int_1^s \frac{c\hat{u}_1(\rho^{1/L}q)+1}{\rho}d\rho}\,\frac{ds}{s}$ and using the Lebesgue dominated convergence theorem, we get

$$G(q;\infty) \;\equiv\; \lim_{L\to\infty} G(q) \;=\; \frac{w_0}{1+\hat{u}_1(q)}$$

and, similarly,

$$\nabla G(q;\infty) \;=\; -\frac{G(q;\infty)}{1+\hat{u}_1(q)}[\nabla V + (\nabla^2 V)^T\, q]\,.$$

Substituting these expressions into the formula for price impact (Proposition 20)

$$\Lambda^*(q) \;=\; \frac{1}{u'(G(q))}\left(\nabla^2 S(q) - \frac{u''(G(q))}{u'(G(q))}\nabla G(q)((1-c)u'(G(q))\nabla G(q) - c\nabla V(q))^T\right)\,,$$

we get the required result. ■

Proof of Proposition 10. By (69), we need to show that

$$\hat{u}_1(q) + (G(q) - w_0)u'(G) < 0,$$

that is

$$G(q)/w_0 < \frac{1}{1+\hat{u}_1(q)}.$$

Let $\ell \equiv 1/L$ and denote $G(q,\ell) = G(q)$. By (21),

$$G(q,\ell) \;=\; w_0 \int_1^\infty e^{-\int_1^s \frac{(1-\ell)\hat{u}_1(\rho^\ell q)+1}{\rho} d\rho} \frac{ds}{s}$$

and the competitive limit corresponds to $G(q,0) = \frac{1}{1+\hat{u}_1(q)}$. Thus, to prove the claim, it suffices to show that $\partial G/\partial \ell \le 0$.

Then, we have

$$\begin{aligned}
\hat{u}_1(q) \;&=\; \frac{E[\delta\cdot q(w+\delta\cdot q)^{-\gamma}]}{E[(w+\delta\cdot q)^{1-\gamma}]},\\
\nabla^2 V(q) \;&=\; -\gamma\frac{E[\delta\delta^T(w+\delta\cdot q)^{-\gamma-1}]}{E[(w+\delta\cdot q)^{1-\gamma}]}\\
&-(1-\gamma)\frac{1}{(E[(w+\delta\cdot q)^{1-\gamma}])^2}E[\delta(w+\delta\cdot q)^{-\gamma}]E[\delta^T(w+\delta\cdot q)^{-\gamma}]
\end{aligned} \tag{80}$$

and hence

$$\begin{aligned}
\nabla\hat{u}_1(q) \;&=\; \nabla V(q) \;+\; \nabla^2 V(q) q \;=\; \frac{E[\delta(w+\delta\cdot q)^{-\gamma}]}{E[(w+\delta\cdot q)^{1-\gamma}]}\\
&-\gamma\frac{E[\delta(\delta\cdot q)(w+\delta\cdot q)^{-\gamma-1}]}{E[(w+\delta\cdot q)^{1-\gamma}]}\\
&-(1-\gamma)\frac{\hat{u}_1(q)}{E[(w+\delta\cdot q)^{1-\gamma}]}E[\delta(w+\delta\cdot q)^{-\gamma}]\\
&=\;((\gamma-1)\hat{u}_1(q)+1)\,\nabla V(q) \;-\; \gamma\frac{E[\delta(\delta\cdot q)(w+\delta\cdot q)^{-\gamma-1}]}{E[(w+\delta\cdot q)^{1-\gamma}]}.
\end{aligned} \tag{81}$$

Thus,

$$
\begin{aligned}
\nabla^2 \hat{u}_1(q) \;&=\; (\gamma-1)\,\nabla V(q)\nabla \hat{u}_1(q)^T \;+((\gamma-1)\hat{u}_1(q)+1)\,\nabla^2 V(q)\\
&-\;\gamma\frac{E[\delta\delta^T(w+\delta\cdot q)^{-\gamma-1}]-(\gamma+1)E[\delta\delta^T(\delta\cdot q)(w+\delta\cdot q)^{-\gamma-2}]}{E[(w+\delta\cdot q)^{1-\gamma}]}\\
&+\;\gamma(1-\gamma)\frac{E[\delta(\delta\cdot q)(w+\delta\cdot q)^{-\gamma-1}]E[\delta^T(w+\delta\cdot q)^{-\gamma}]}{(E[(w+\delta\cdot q)^{1-\gamma}])^2}\\
&=\;(\gamma-1)\,\nabla V(q)\left(((\gamma-1)\hat{u}_1(q)+1)\,\nabla V(q)\;-\;\gamma\frac{E[\delta(\delta\cdot q)(w+\delta\cdot q)^{-\gamma-1}]}{E[(w+\delta\cdot q)^{1-\gamma}]}\right)^T\\
&+((\gamma-1)\hat{u}_1(q)+1)\,\nabla^2 V(q)\\
&-\;\gamma\frac{E[\delta\delta^T(w+\delta\cdot q)^{-\gamma-1}]-(\gamma+1)E[\delta\delta^T(\delta\cdot q)(w+\delta\cdot q)^{-\gamma-2}]}{E[(w+\delta\cdot q)^{1-\gamma}]}\\
&+\;\gamma(1-\gamma)\frac{E[\delta(\delta\cdot q)(w+\delta\cdot q)^{-\gamma-1}]E[\delta^T(w+\delta\cdot q)^{-\gamma}]}{(E[(w+\delta\cdot q)^{1-\gamma}])^2}
\end{aligned}
\tag{82}
$$

Pick a vector $x \in \mathbb{R}^N$ and denote $X = (x\cdot\delta)/(w+q\cdot\delta)$ and $Y = (\delta\cdot q)/(w+\delta\cdot q)$. Then, under the measure $(w+q\cdot\delta)^{1-\gamma}/E[(w+q\cdot\delta)^{1-\gamma}]$, we have

$$
x\cdot\nabla^2 V(q)x \;=\; -\gamma E[X^2]-(1-\gamma)E[X]^2 \;=\; -\gamma \mathrm{Var}[X]-E[X]^2 \tag{83}
$$

and therefore

$$
\begin{aligned}
x\cdot\nabla^2\hat{u}_1(q)x \;&=\; (\gamma-1)\,E[X]\Big(((\gamma-1)E[Y]+1)\,E[X]\;-\;\gamma E[XY]\Big)\\
&+((\gamma-1)E[Y]+1)\,(-\gamma E[X^2]+(\gamma-1)E[X]^2)\\
&-\;\gamma(E[X^2]-(\gamma+1)E[X^2Y])\;+\;\gamma(1-\gamma)E[XY]E[X]\\
&=\;2(\gamma-1)\,E[X]^2((\gamma-1)E[Y]+1)+2\gamma(1-\gamma)E[XY]E[X]\\
&-\gamma(1+(\gamma-1)E[Y]+1)E[X^2]+\gamma(\gamma+1)E[X^2Y]
\end{aligned}
\tag{84}
$$

Thus,

$$
\begin{aligned}
(1+\hat{u}_1(q))\Lambda^*(q;\infty) \;&=\; \left(-\nabla^2 V(q) \;+\; \frac{1}{1+\hat{u}_1(q)}\left(\nabla V(q)+\nabla^2 V(q)\,q\right)\nabla V(q)^T\right)\\
&= \;\gamma\frac{E[\delta\delta^T(w+\delta\cdot q)^{-\gamma-1}]}{E[(w+\delta\cdot q)^{1-\gamma}]}\\
&+(1-\gamma)\frac{1}{(E[(w+\delta\cdot q)^{1-\gamma}])^2}E[\delta(w+\delta\cdot q)^{-\gamma}]E[\delta^T(w+\delta\cdot q)^{-\gamma}]\\
&+\;\frac{1}{1+\hat{u}_1(q)}\left(\frac{E[\delta(w+q\cdot\delta)^{-\gamma}]}{E[(w+q\cdot\delta)^{1-\gamma}]}\right.\\
&-\gamma\frac{E[(\delta\cdot q)\delta(w+\delta\cdot q)^{-\gamma-1}]}{E[(w+\delta\cdot q)^{1-\gamma}]}\\
&\left.-(1-\gamma)\frac{1}{(E[(w+\delta\cdot q)^{1-\gamma}])^2}E[\delta(w+\delta\cdot q)^{-\gamma}]E[(\delta\cdot q)(w+\delta\cdot q)^{-\gamma}]\right)\frac{E[\delta^T(w+q\cdot\delta)^{-\gamma}]}{E[(w+q\cdot\delta)^{1-\gamma}]}
\end{aligned}
\tag{85}
$$

Denoting $Y=\delta\cdot q/(w+\delta\cdot q)$, we get that, under the measure $(w+q\cdot\delta)^{1-\gamma}/E[(w+q\cdot\delta)^{1-\gamma}]$,

$$
(1+\hat{u}_1(q))q^T\Lambda^*(q;\infty)q \;=\; \frac{\gamma(E[Y^2]-E[Y]^2)+2E[Y]^2}{1+E[Y]} > 0
$$

However,

$$
(1+\hat{u}_1(q))x^T\Lambda^*(q;\infty)x \;=\; \gamma\frac{E[X^2]-E[X]^2}{1+E[Y]}+\gamma\frac{E[X^2]E[Y]-E[X]E[XY]}{1+E[Y]}+\frac{2E[X]^2}{1+E[Y]}\,.
$$

■

Proposition 21. *Suppose that*

$$
u(x) \;=\; \log(x) \qquad \text{and} \qquad V(q) \;=\; E[\log(\delta\cdot q)]\,.
$$

Then, the function S is convex and the function

$$
U(q,q\cdot\Pi(\varepsilon-q)) = \;\log\left(w_0-q\cdot\Pi(\varepsilon-q)\right) \;+\; E[\log(\delta\cdot q)]
$$

is concave over the set of feasible q. Therefore, the conjectured equilibrium is a true equilibrium.

Proof of Proposition 21.

Feasible portfolios $q \in \mathcal{A}$ satisfy

$$\delta \cdot q \ < \ |\delta \cdot \varepsilon|$$

almost surely. This inequality implies that[19]

$$\frac{\delta\varepsilon}{\delta \cdot (\varepsilon - q)} \ > \ 0 \qquad \text{a.s.}\,.$$

It follows from the functional form of V that

$$\hat{u}_1(q) \ = \ 1\,.$$

Thus,

$$\begin{aligned} G(q) = \ & w_0 \int_1^\infty e^{-\int_1^s \frac{c\hat{u}_1(\rho^{1/L}q)+1}{\rho} d\rho} \frac{ds}{s} \\ = \ & w_0 \int_1^\infty e^{-\int_1^s \frac{c+1}{\rho} d\rho} \frac{ds}{s} \\ = \ & w_0 \int_1^\infty \frac{1}{s^{2+c}}\, ds \\ = \ & \frac{w_0}{1+c}\,. \end{aligned}$$

We know that

$$F(q) = \ qD^{-1}(q); \qquad G(q) = \ w_0 - F(q); \qquad \hat{u}_1(q) = \ q\nabla V(q)$$

$$S(q) = \ (1-c)u(G(q)) \ - \ cV(q); \qquad \text{and} \qquad \pi(q) = \ -\frac{\nabla S(q)}{u'(G(q))}\,.$$

[19] The denominator and the numerator have the same sign.

It follows that

$$qD^{-1}(q) = \frac{c}{1+c}w_0; \qquad\qquad F(q) = \frac{c}{1+c}w_0;$$

$$S(q) = (1-c)\log\left(\frac{c}{1+c}w_0\right) - cE[\log(\delta\cdot q)] \qquad\qquad \pi(q) = \frac{c}{1+c}w_0\,E\left[\delta\frac{1}{\delta\cdot q}\right].$$

The convexity of S is the a direct consequence of the concavity of the logarithm function.

By definition,

$$\Pi(q) = \pi\left(\frac{q}{L-1}\right) = (L-1)\frac{c}{1+c}w_0\,E\left[\delta\frac{1}{\delta\cdot q}\right].$$

Therefore,

$$\begin{aligned} U(q, q\cdot\Pi(\varepsilon-q)) &= \log\left(w_0 - q\cdot\Pi(\varepsilon-q)\right) + E[\log(\delta\cdot q)] \\ &= \log\left(w_0 - \frac{c(L-1)}{1+c}w_0\,E\left[\frac{\delta\cdot q}{\delta\cdot(\varepsilon-q)}\right]\right) + E[\log(\delta\cdot q)]\,. \end{aligned}$$

WLOG, we set

$$w_0 = 1\,.$$

To show that U is concave, it is enough to show that

$$E\left[\frac{\delta\cdot q}{\delta\cdot\varepsilon-\delta\cdot q}\right]$$

is convex since log is an increasing function. We have[20]

$$\begin{aligned}
\nabla E\left[\frac{\delta\cdot q}{\delta\cdot\varepsilon-\delta\cdot q}\right] &= E\left[\frac{(\delta\cdot\varepsilon)\delta}{(\delta\cdot\varepsilon-\delta\cdot q)^2}\right]\\
\nabla^2 E\left[\frac{\delta\cdot q}{\delta\cdot\varepsilon-\delta\cdot q}\right] &= E\left[\nabla\frac{(\delta\cdot\varepsilon)\delta}{(\delta\cdot\varepsilon-\delta\cdot q)^2}\right]\\
&= E\left[(\delta\cdot\varepsilon)\delta\left(\nabla\left(\frac{1}{(\delta\cdot\varepsilon-\delta\cdot q)^2}\right)\right)^T\right]\\
&= E\left[\frac{2\delta\cdot\varepsilon}{(\delta\cdot\varepsilon-\delta\cdot q)^3}\delta\;(\delta)^T\right].
\end{aligned}$$

Therefore,

$$x^T\left(\nabla^2 E\left[\frac{\delta\cdot q}{\delta\cdot\varepsilon-\delta\cdot q}\right]\right)x = 2E\left[\frac{\delta\cdot\varepsilon}{(\delta\cdot(\varepsilon-q))^3}(\delta\cdot x)^2\right].$$

The expression inside the expectation is almost surely positive. Hence, the expectation is strictly positive (assuming weak regularity of the random variables' distributions)

■

Proof of Proposition 11.

Under the assumptions of proposition, the PDE (66) characterizing G becomes

$$(1-c)G^{-\gamma}(q)\,q\cdot\nabla G(q) \;-\; G^{1-\gamma}(q) \;=\; c\hat{u}_1(q)\,. \tag{86}$$

Thus,[21]

$$H(q) \;=\; G^{1-\gamma}(q)$$

[20]We assume that the density functions are such that we can interchange differentiation and integration.

[21]$G(q)$ is positive under the assumption of positive consumption, which holds in equilibrium for $\gamma > 1$. Thus, we need not worry about branches of complex number when taking powers.

satisfies the following PDE:

$$\beta G^{-\gamma}(q)\, q \cdot \nabla G(q) \;-\; H(q) \;=\; c\hat{u}_1(q)\,.$$

It is direct to verify that the proposed solution satisfies this PDE.

■

Proof of Lemma 4.

Equation (86) implies that

$$G^{1-\gamma}(q;\infty) \;\equiv\; \lim_{L\to\infty} G^{1-\gamma}(q) \;=\; -\hat{u}_1(q) \quad\Longrightarrow\quad \frac{\nabla G(q;\infty)}{G(q;\infty)} \;=\; \frac{[\nabla V(q) + (\nabla^2 V(q))^T\, q]}{(1-\gamma)\hat{u}_1(q)}\,.$$

We know from Proposition 20 that

$$\begin{aligned}\Lambda^*(q) = \;& \frac{1}{u'(G(q))}\left(\nabla^2 S(q) - \frac{u''(G(q))}{u'(G(q))}\nabla G(q)((1-c)u'(G(q))\nabla G(q) - c\nabla V(q))^T\right)\\ =\;& G^{\gamma}(q)\left(\nabla^2 S(q) + \frac{\gamma}{G(q)}\nabla G(q)\Big[(1-c)G^{-\gamma}(q)\nabla G(q) - c\nabla V(q)\Big]^T\right)\,.\end{aligned}$$

Putting these equalities together yields the desired result.

■

Proposition 22. *Suppose that $N = 1$, that V is concave, increasing, and that ∇V is convex. Moreover, suppose that*

$$\gamma \;>\; 1 \qquad \text{and} \qquad \varepsilon \;<\; 0\,.$$

Then, the function

$$U(q,\, q\cdot\Pi(\varepsilon - q)) = \;\frac{1}{(1-\gamma)}\,(-q\cdot\Pi(\varepsilon-q))^{1-\gamma} \;+\; V(q)$$

is concave on the interval

$$\left(\min\left\{1,\, \frac{\gamma}{2}\right\}\, \varepsilon,\, 0\right)\,.$$

In particular, if

$$\gamma \;\geq\; 2\,,$$

the function $U(q,\, q\cdot\Pi(\varepsilon-q))$ is concave over the set of feasible q, and thus, the proposed equilibrium is the true equilibrium.

Proof of Proposition 22. The function U is

$$\begin{aligned}
U(q, q\cdot\Pi(\varepsilon-q)) &= \frac{1}{1-\gamma}(-q\cdot\Pi(\varepsilon-q))^{1-\gamma} \;+\; V(q)\\
&= \frac{1}{(1-\gamma)(L-1)^{\gamma-1}}\left(-\frac{q}{\varepsilon-q}\frac{\varepsilon-q}{L-1}\cdot\pi\left(\frac{\varepsilon-q}{L-1}\right)\right)^{1-\gamma} \;+\; V(q)\\
&= \frac{1}{(1-\gamma)(L-1)^{\gamma-1}}\left(\frac{q}{\varepsilon-q}G\left(\frac{\varepsilon-q}{L-1}\right)\right)^{1-\gamma} \;+\; V(q)\\
&= \frac{1}{(1-\gamma)(L-1)^{\gamma-1}}H\left(\frac{\varepsilon-q}{L-1}\right)\left(\frac{q}{\varepsilon-q}\right)^{1-\gamma} \;+\; V(q)\,,
\end{aligned}$$

where the third equality follows from Corollary 1.

To prove that U is concave, it is sufficient to show that

$$H\left(\frac{\varepsilon-q}{L-1}\right)\left(\frac{q}{\varepsilon-q}\right)^{1-\gamma}$$

is convex. As we show below, this convexity result holds for $q<0$ if H is decreasing and convex for $q<0$. We start by showing that H satisfies these properties.

We know that

$$\nabla\hat{u}_1(q) \;=\; \nabla V(q) \;+\; q\cdot\nabla^2 V(q) \qquad \text{and} \qquad \nabla^2\hat{u}_1(q) \;=\; 2\nabla^2 V(q) \;+\; q\cdot\nabla^3 V(q)\,.$$

Thus, $\hat{u}_1$ is increasing for $q\leq 0$ since V is increasing and concave. Moreover, $\hat{u}_1$ is concave for

$q \leq 0$ since V is concave and ∇V is convex. It follows from the definition

$$H(q) \;=\; \frac{c}{\beta}\int_0^1 t^{-1-1/\beta}\hat{u}_1(qt)dt$$

that H is decreasing and convex for $q \leq 0$.

The result then follows from the lemma below

■

Lemma 21. *Suppose that*

$$\min\left\{1,\,\frac{\gamma}{2}\right\}\varepsilon \;<\; q \;<\; 0 \qquad \text{and} \qquad \gamma \;>\; 1\,.$$

Then, for an arbitrary function $f(q)$ that is a convex and decreasing over the assumed interval for q, the function

$$h(q) \;\equiv\; f(\alpha\,(\varepsilon - q))\left(\frac{q}{\varepsilon - q}\right)^{1-\gamma}$$

is convex for any positive constant α over the assumed interval for q.

Proof. WLOG, we assume that

$$\alpha \;=\; 1\,.$$

Let

$$h(q) \;=\; f(\varepsilon - q)\left(\frac{q}{\varepsilon - q}\right)^{1-\gamma}\,.$$

Then,

$$h''(q) \;=\; \frac{\left(\frac{q}{\varepsilon-q}\right)^{-\gamma}\Big((\gamma-1)\varepsilon(2q-\gamma\varepsilon)f(\varepsilon-q) \;-\; q(q-\varepsilon)\Big[q(q-\varepsilon)f''(\varepsilon-q) \;-\; 2(\gamma-1)\varepsilon f'(\varepsilon-q)\Big]\Big)}{q(q-\varepsilon)^3}\,.$$

We know that

$$
\begin{aligned}
\frac{q}{\varepsilon - q} &> 0 \\
(\gamma - 1)\varepsilon(2q - \gamma\varepsilon) f(\varepsilon - q) &< 0 \\
- q(q - \varepsilon) &> 0 \\
q(q - \varepsilon) f''(\varepsilon - q) &< 0 \\
- 2(\gamma - 1)\varepsilon f'(\varepsilon - q) &< 0 \,.
\end{aligned}
$$

The result then follows from the assumptions.

■

Proof of Proposition 13.

The function $\hat{u}_1$ simplifies to

$$\hat{u}_1(q) \;=\; e^{-\tau} \,.$$

Thus,

$$
\begin{aligned}
G_i(q) = \;& \alpha_i w_0 \int_1^\infty e^{-\int_1^s \frac{e^{-\tau} c_i + 1}{\rho} d\rho} \, \frac{ds}{s} \\
= \;& \alpha_i w_0 \int_1^\infty e^{\log s^{-\left(1 + e^{-\tau} c_i\right)}} \, \frac{ds}{s} \\
= \;& \alpha_i w_0 \int_1^\infty s^{-\left(2 + e^{-\tau} c_i\right)} \, ds \\
= \;& \alpha_i w_0 \frac{1}{1 + e^{-\tau} c_i} \,.
\end{aligned}
$$

Thus,

$$D^{-1}\left(\frac{1}{\beta_i} q\right) \;=\; D_i^{-1}(q) \;=\; -(1 - c_i)\nabla G_i(q) \;+\; c_i G_i(q)\, \nabla V(q) \,.$$

It follows that

$$\begin{aligned}
D^{-1}\left(\frac{1}{\beta_i}q\right) &= D_i^{-1}(q) \;=\; c_i G_i(q)\,\nabla V(q) \\
&= \alpha_i w_0 \frac{e^{-\tau}c_i}{1+e^{-\tau}c_i}\mathrm{E}[(\delta\cdot q)^{-1}\delta] \\
&= \alpha_i w_0 \frac{1}{\beta_i}\frac{e^{-\tau}c_i}{1+e^{-\tau}c_i}\mathrm{E}[(\delta\cdot \frac{1}{\beta_i}q)^{-1}\delta] \\
&= \alpha_i \frac{1}{\beta_i}\frac{\beta-\beta_i}{\beta+e^{-\tau}(\beta-\beta_i)}\mathrm{E}[(\delta\cdot \frac{1}{\beta_i}q)^{-1}\delta]e^{-\tau}w_0 \\
\Longrightarrow D^{-1}(x) &= \alpha_i \frac{1}{\beta_i}\frac{(\beta-\beta_i)e^{-\tau}}{\beta+e^{-\tau}(\beta-\beta_i)} w_0\mathrm{E}[(\delta\cdot x)^{-1}\delta]\,.
\end{aligned}$$

We now compute the equilibrium price directly. Recall that

$$\Pi_i(q) \;=\; D^{-1}\left((\beta-\beta_i)^{-1}q\right) \qquad \text{and} \qquad \Pi_i(\varepsilon-q) \;=\; p\,.$$

In equilibrium, $q_i = \eta_i\varepsilon$ and

$$\begin{aligned}
p = &\; \Pi_i((1-\eta_i)\varepsilon) \\
= &\; D^{-1}\left(\frac{1}{\beta-\beta_i}(1-\eta_i)\varepsilon\right) \\
= &\; D^{-1}\left(\frac{1}{\beta-\beta_i}\frac{\beta-\beta_i}{\beta}\varepsilon\right) \\
= &\; D^{-1}\left(\frac{1}{\beta}\varepsilon\right) \\
= &\; \kappa\, w_0\mathrm{E}\left[\left(\frac{1}{\beta}\varepsilon\cdot\delta\right)^{-1}\delta\right] \\
= &\; w_0\mathrm{E}\left[\left(\frac{1}{\kappa\beta}\varepsilon\cdot\delta\right)^{-1}\delta\right]\,.
\end{aligned}$$

For a scale-invariant equilibrium, it is necessary that the coefficient β_i satisfies

$$\exists\,\kappa \quad \text{such that} \quad \kappa \;=\; \alpha_i\frac{1}{\beta_i}\frac{(\beta-\beta_i)e^{-\tau}}{\beta+e^{-\tau}(\beta-\beta_i)} \quad \forall\, i\,. \tag{87}$$

We normalize $\kappa = 1$. Then,

$$\beta_i[\beta + e^{-\tau}(\beta - \beta_i)] \;=\; \alpha_i(\beta - \beta_i)e^{-\tau}\,.$$

$$(1 + e^{-\tau})\beta\beta_i - \beta_i^2 \;=\; \alpha_i(\beta - \beta_i)e^{-\tau}\,.$$

$$0 \;=\; \beta_i^2 \;-\; \left[\alpha_i e^{-\tau} + \beta\left(1 + e^{-\tau}\right)\right]\beta_i \;+\; \alpha_i\beta e^{-\tau}\,.$$

It follows that

$$\beta_i = \frac{\left[\alpha_i e^{-\tau} + \beta\left(1 + e^{-\tau}\right)\right] \;\pm\; \sqrt{\left[\alpha_i e^{-\tau} + \beta\left(1 + e^{-\tau}\right)\right]^2 - 4\alpha_i e^{-\tau}\beta}}{2} \tag{88}$$

$$\Longrightarrow \beta_i = \frac{2\alpha_i e^{-\tau}\beta}{\alpha_i e^{-\tau} + \left(1 + e^{-\tau}\right)\beta \;\pm\; \sqrt{\left[\alpha_i e^{-\tau} - \beta\left(1 - e^{-\tau}\right)\right]^2 + 4\beta^2 e^{-\tau}}}\,. \tag{89}$$

We can thus solve for β as the solution of the equation

$$\sum_{j=1}^{L} \frac{2\alpha_j e^{-\tau}}{\alpha_j e^{-\tau} + \left(1 + e^{-\tau}\right)X \;\pm\; \sqrt{\left[\alpha_j e^{-\tau} - \left(1 - e^{-\tau}\right)X\right]^2 + 4X^2 e^{-\tau}}} \;=\; 1\,. \tag{90}$$

We will show that this equation has a no solution when the $\pm$ sign in the denominator is negative and a unique solution when this sign is positive. Moreover, the unique solution is positive.

First, we show that there is no equilibrium with $\beta < 0$. Suppose that $\beta < 0$. If

$$\beta_i = \frac{2\alpha_i e^{-\tau}\beta}{\alpha_i e^{-\tau} + \left(1 + e^{-\tau}\right)\beta \;-\; \sqrt{\left[\alpha_i e^{-\tau} - \beta\left(1 - e^{-\tau}\right)\right]^2 + 4\beta^2 e^{-\tau}}}\,,$$

then,

$$
\begin{aligned}
0 > \; & \left[\alpha_i e^{-\tau} + \left(1 + e^{-\tau}\right)\beta\right]^2 \; - \; \left[\left(\alpha_i e^{-\tau} - \left(1 - e^{-\tau}\right)\beta\right)^2 + 4e^{-\tau}\beta^2\right] \\
\implies \alpha_i e^{-\tau} + \left(1 + e^{-\tau}\right)\beta < \; & \sqrt{\left(\alpha_i e^{-\tau} - \left(1 - e^{-\tau}\right)\beta\right)^2 + 4e^{-\tau}\beta^2} \\
\implies 0 > \; & \alpha_i e^{-\tau} + \left(1 + e^{-\tau}\right)\beta \; - \; \sqrt{\left(\alpha_i e^{-\tau} - \left(1 - e^{-\tau}\right)\beta\right)^2 + 4e^{-\tau}\beta^2} \\
\implies 0 > \; & \frac{2\alpha_i e^{-\tau}}{\alpha_i e^{-\tau} + \left(1 + e^{-\tau}\right)\beta \; - \; \sqrt{\left[\alpha_i e^{-\tau} - \beta\left(1 - e^{-\tau}\right)\right]^2 + 4\beta^2 e^{-\tau}}} \\
\implies \beta_i > \; & 0 \, .
\end{aligned}
$$

If

$$
\beta_i \; = \; \frac{2\alpha_i e^{-\tau}\beta}{\alpha_i e^{-\tau} + \left(1 + e^{-\tau}\right)\beta \; + \; \sqrt{\left[\alpha_i e^{-\tau} - \beta\left(1 - e^{-\tau}\right)\right]^2 + 4\beta^2 e^{-\tau}}} \, ,
$$

we have

$$
\begin{aligned}
0 < \; & \left[\alpha_i e^{-\tau} - \left(1 + e^{-\tau}\right)\beta\right]^2 \; - \; \left[\left(\alpha_i e^{-\tau} - \left(1 - e^{-\tau}\right)\beta\right)^2 + 4e^{-\tau}\beta^2\right] \\
\implies \alpha_i e^{-\tau} - \left(1 + e^{-\tau}\right)\beta > \; & \sqrt{\left(\alpha_i e^{-\tau} - \left(1 - e^{-\tau}\right)\beta\right)^2 + 4e^{-\tau}\beta^2} \\
\implies 2\alpha_i e^{-\tau} > \; & \alpha_i e^{-\tau} + \left(1 + e^{-\tau}\right)\beta \; + \; \sqrt{\left[\alpha_i e^{-\tau} - \beta\left(1 - e^{-\tau}\right)\right]^2 + 4\beta^2 e^{-\tau}} \; > \; 0 \\
\implies 1 < \; & \frac{2\alpha_i e^{-\tau}}{\alpha_i e^{-\tau} + \left(1 + e^{-\tau}\right)\beta \; + \; \sqrt{\left[\alpha_i e^{-\tau} - \beta\left(1 - e^{-\tau}\right)\right]^2 + 4\beta^2 e^{-\tau}}} \\
\implies \beta_i > \; & 1 \, .
\end{aligned}
$$

Thus, under the assumption that $\beta < 0$, we have

$$
0 \; > \; \beta \; = \; \sum_i \beta_i \; > \; 1 \, ,
$$

a contradiction. Thus, Equation (90) has no solution $\beta < 0$.

Next, we consider the case of positive β. Suppose that $\beta > 0$. It follows that

$$\beta_i \; > \; 0 \; \forall i; \qquad \text{and} \qquad \beta \; > \; \beta_i \; \forall i \, .$$

Moreover,

$$\begin{aligned} 0 > & \; \left[\alpha_i e^{-\tau} - \left(1 + e^{-\tau}\right) \beta\right]^2 \; - \; \left[\left(\alpha_i e^{-\tau} - \left(1 - e^{-\tau}\right) \beta\right)^2 + 4 e^{-\tau} \beta^2\right] \\ \Longrightarrow \alpha_i e^{-\tau} - \left(1 + e^{-\tau}\right) \beta > & \; -\sqrt{\left(\alpha_i e^{-\tau} - \left(1 - e^{-\tau}\right) \beta\right)^2 + 4 e^{-\tau} \beta^2} \\ \Longrightarrow 2\alpha_i e^{-\tau} > & \; \alpha_i e^{-\tau} + \left(1 + e^{-\tau}\right) \beta \; - \; \sqrt{\left[\alpha_i e^{-\tau} - \beta\left(1 - e^{-\tau}\right)\right]^2 + 4\beta^2 e^{-\tau}} \; > \; 0 \\ \Longrightarrow 1 < & \; \frac{2\alpha_i e^{-\tau}}{\alpha_i e^{-\tau} + \left(1 + e^{-\tau}\right) \beta \; - \; \sqrt{\left[\alpha_i e^{-\tau} - \beta\left(1 - e^{-\tau}\right)\right]^2 + 4\beta^2 e^{-\tau}}} \, , \end{aligned}$$

from which is follows that[22]

$$\beta_i \neq \; \frac{2\alpha_i e^{-\tau} \beta}{\alpha_i e^{-\tau} + \left(1 + e^{-\tau}\right) \beta \; + \; \sqrt{\left[\alpha_i e^{-\tau} - \beta\left(1 - e^{-\tau}\right)\right]^2 + 4\beta^2 e^{-\tau}}} \, .$$

Thus, we are left with

$$\beta_i = \; \frac{2\alpha_i e^{-\tau} \beta}{\alpha_i e^{-\tau} + \left(1 + e^{-\tau}\right) \beta \; + \; \sqrt{\left[\alpha_i e^{-\tau} - \beta\left(1 - e^{-\tau}\right)\right]^2 + 4\beta^2 e^{-\tau}}} \, ,$$

which satisfies the condition $\beta > \beta_i$. It remains to show that there exists a unique β in this case, which is equivalent to showing that the equation

$$\sum_j f(X, \alpha_j) \; = \; 1$$

[22]Otherwise we would have $\beta_i > \beta$, a contradiction.

has a unique solution X, where

$$f(X,\alpha) \;=\; \frac{2\alpha e^{-\tau}}{\alpha e^{-\tau}+(1+e^{-\tau})\,X \;+\; \sqrt{\left[\alpha e^{-\tau}-(1-e^{-\tau})\,X\right]^2+4X^2e^{-\tau}}}\,.$$

We shall use the following lemma, whose proof we omit:

Lemma 22. *1. The function f is decreasing in X, increasing and concave in α.*

2. Let

$$g(X,\alpha) \equiv \; Xf(X,\alpha)\,.$$

The function g is increasing in X, in α, and concave in α.

The monotonicity of f, combined with the facts that

$$\sum_j f(0,\alpha_j) \;=\; L \;>\; 1 \qquad \text{and} \qquad \lim_{X\to\infty}\sum_j f(X,\alpha_j) \;=\; 0\,,$$

imply that the equation

$$\sum_j f(X,\alpha_j) \;=\; 1$$

has a unique positive solution.

We have thus for any distribution $\{\alpha_i\}_{i=1}^L$ of initial wealth, there exists a unique equilibrium. In equilibrium,

$$\beta \;>\; \beta_i \;>\; 0 \qquad \forall\, i \qquad \text{and} \qquad \alpha_i \;<\; \alpha_j \quad \Longleftrightarrow \quad \beta_i \;<\; \beta_j\,.$$

■

Proof of Proposition 14. We shall use the following lemma:

Lemma 23. *Let $\boldsymbol{\alpha}$ be an initial distribution of wealth with $L \geq 2$ and consider an arbitrary break-up $\bar{\boldsymbol{\alpha}}$ of $\boldsymbol{\alpha}$ with*[23]

$$\bar{L} \;=\; L+1 \qquad \text{and} \qquad \bar{\alpha}_{L,1} \;=\; y; \quad \bar{\alpha}_{L,2} \;=\; \alpha_L - y; \qquad y \;\in\; (0,\, \alpha_L)\,.$$

Moreover, let X^* and $\bar{X}^*$ be the respective solutions of the equations

$$f(X, \boldsymbol{\alpha}) \;=\; 1 \qquad \text{and} \qquad f(X, \bar{\boldsymbol{\alpha}}) \;=\; 1\,.$$

Then,

$$X^* \;<\; \bar{X}^*.$$

Recall that the equilibrium price is

$$p \;=\; \beta\, w_0 \mathrm{E}\left[(\varepsilon \cdot \delta)^{-1}\, \delta\right]\;;$$

the price impact matrix of trader i is

$$\Lambda_i^*(q) \;=\; \beta_i\, w_0 \mathrm{E}\left[(q \cdot \delta)^{-2}\, \delta \delta^T\right]\;;$$

β is the solution to the equation

$$f(X, \boldsymbol{\alpha}) \;=\; 1 \qquad \text{while} \qquad \beta_i \;=\; g(\beta, \alpha_i)\,.$$

It thus follows directly from Lemma 23 that

$$p \;<\; \bar{p}$$

[23] Thus, we are breaking the Lth trader into two, with respective sizes $\bar{\alpha}_{L,1}$ and $\bar{\alpha}_{L,2}$.

where p and $\bar{p}$ are the equilibrium prices under wealth distributions $\boldsymbol{\alpha}$ and $\bar{\boldsymbol{\alpha}}$ respectively. In addition, the highest $\bar{p}$ is achieved when the trader been broken up is split into two equal traders, that is when

$$y \;=\; \frac{1}{2}\alpha_L\,.$$

Moreover, suppose that $i < L$ (that is, trader i was left unchanged). Then,

$$\Lambda_i^*(q) \;<\; \bar{\Lambda}_i^*(q)\,,$$

where the inequality is *component-wise*.

Next, we consider the utility function. The utility function is

$$U_i(q, q\cdot p) \;=\; \log(\alpha_i w_0 \;-\; q\cdot p) \;+\; e^{-\tau}\mathrm{E}[\log(\delta\cdot q)]\,.$$

In equilibrium,

$$q_i^* \;=\; \eta_i\,\varepsilon \qquad \text{where} \qquad \eta_i \;=\; \frac{\beta_i}{\beta}\,.$$

Thus, the equilibrium utility is

$$\begin{aligned}
U_i^* &= \log\left(\alpha_i w_0 \;-\; \frac{\beta_i}{\beta}\,w_0\mathrm{E}\left[\left(\frac{1}{\beta}\varepsilon\cdot\delta\right)^{-1}\varepsilon\cdot\delta\right]\right) \;+\; e^{-\tau}\;\mathrm{E}\left[\log\left(\frac{\beta_i}{\beta}\varepsilon\cdot\delta\right)\right] \\
&= \log w_0 \;+\; e^{-\tau}\mathrm{E}\left[\log(\varepsilon\cdot\delta)\right] \;+\; \log(\alpha_i - \beta_i) \;+\; e^{-\tau}\log\left(\frac{\beta_i}{\beta}\right) \\
&= \log w_0 \;+\; e^{-\tau}\mathrm{E}\left[\log(\varepsilon\cdot\delta)\right] \;+\; \log(\alpha_i - \beta_i) \;+\; e^{-\tau}\log\, f(\kappa_{\boldsymbol{\alpha}}\beta_{\boldsymbol{\alpha}}, \alpha_i)\,.
\end{aligned}$$

Let $\boldsymbol{\alpha}$ and $\bar{\boldsymbol{\alpha}}$ be as in Proposition 14. Suppose that trader i is not the one been broken up (that is, $i < L$). Then,

$$
\begin{aligned}
U^*_{i,\boldsymbol{\alpha}} &= \log w_0 \;+\; e^{-\tau}\mathrm{E}\left[\log(\varepsilon\cdot\delta)\right] \;+\; \log(\alpha_{i,\boldsymbol{\alpha}} - \beta_{i,\boldsymbol{\alpha}}) \;+\; e^{-\tau}\log\, f(\beta_{\boldsymbol{\alpha}}, \alpha_{i,\boldsymbol{\alpha}}) \\
U^*_{i,\bar{\boldsymbol{\alpha}}} &= \log w_0 \;+\; e^{-\tau}\mathrm{E}\left[\log(\varepsilon\cdot\delta)\right] \;+\; \log(\alpha_{i,\bar{\boldsymbol{\alpha}}} - \beta_{i,\bar{\boldsymbol{\alpha}}}) \;+\; e^{-\tau}\log\, f(\beta_{\bar{\boldsymbol{\alpha}}}, \alpha_{i,\bar{\boldsymbol{\alpha}}}) \\
&= \log w_0 \;+\; e^{-\tau}\mathrm{E}\left[\log(\varepsilon\cdot\delta)\right] \;+\; \log(\alpha_{i,\boldsymbol{\alpha}} - \beta_{i,\bar{\boldsymbol{\alpha}}}) \;+\; e^{-\tau}\log\, f(\beta_{\bar{\boldsymbol{\alpha}}}, \alpha_{i,\boldsymbol{\alpha}})
\end{aligned}
$$

since the share of wealth of this trader does not change. Lemmas 22 and 23 imply that

$$
\beta_{\boldsymbol{\alpha}} \;<\; \beta_{\bar{\boldsymbol{\alpha}}} \qquad \text{and} \qquad \beta_{i,\boldsymbol{\alpha}} \;<\; \beta_{i,\bar{\boldsymbol{\alpha}}}\,.
$$

Thus,

$$
\log(\alpha_{i,\boldsymbol{\alpha}} - \beta_{i,\boldsymbol{\alpha}}) \;>\; \log(\alpha_{i,\boldsymbol{\alpha}} - \beta_{i,\bar{\boldsymbol{\alpha}}}) \qquad \text{and} \qquad f(\beta_{i,\boldsymbol{\alpha}},\, \alpha_{i,\boldsymbol{\alpha}}) \;>\; f(\beta_{i,\bar{\boldsymbol{\alpha}}},\, \alpha_{i,\boldsymbol{\alpha}})\,.
$$

Hence,

$$
U^*_{i,\boldsymbol{\alpha}} \;>\; U^*_{i,\bar{\boldsymbol{\alpha}}}\,.
$$

■

References

Viral Acharya and Alberto Bisin. Counterparty risk externality: Centralized versus over-the-counter markets. Journal of Economic Theory, 149:153–182, 2014.

Viral V. Acharya and Lasse Heje Pedersen. Asset pricing with liquidity risk. Journal of Financial Economics, 77(2):375–410, 2005.

Viral V Acharya, Lasse H Pedersen, Thomas Philippon, and Matthew Richardson. Measuring systemic risk. The review of financial studies, 30(1):2–47, 2017.

Tobias Adrian, Erkko Etula, and Tyler Muir. Financial intermediaries and the cross-section of asset returns. The Journal of Finance, 69(6):2557–2596, 2014.

Franklin Allen. Do financial institutions matter? The Journal of Finance, 56(4):1165–1175, 2001.

Robert Almgren, Chee Thum, Emmanuel Hauptmann, and Hong Li. Direct estimation of equity market impact. Risk, 18(7):58–62, 2005.

Yakov Amihud and Haim Mendelson. Asset pricing and the bid-ask spread. Journal of financial Economics, 17(2):223–249, 1986.

Lawrence M Ausubel. Insider trading in a rational expectations economy. The American Economic Review, pages 1022–1041, 1990a.

Lawrence M Ausubel. Partially-revealing rational expectations equilibrium in a competitive economy. Journal of Economic Theory, 50(1):93–126, 1990b.

Lawrence M Ausubel, Peter Cramton, Marek Pycia, Marzena Rostek, and Marek Weretka. Demand reduction and inefficiency in multi-unit auctions. The Review of Economic Studies, 81(4):1366–1400, 2014.

Efstathios Avdis and Sergei Glebkin. Chile. Working Paper, 2023.

Kerry Back. Insider trading in continuous time. The Review of Financial Studies, 5(3):387–409, 1992.

Kerry Back and Shmuel Baruch. Strategic liquidity provision in limit order markets. Econometrica, 81(1):363–392, 2013.

Kerry Back and Kevin Crotty. The informational role of stock and bond volume. The Review of Financial Studies, 28(5):1381–1427, 2015.

Kerry Back, C Henry Cao, and Gregory A Willard. Imperfect competition among informed traders. The journal of finance, 55(5):2117–2155, 2000.

Mark Bagnoli, S Viswanathan, and Craig Holden. On the existence of linear equilibria in models of market making. Mathematical Finance, 11(1):1–31, 2001.

Gadi Barlevy and Pietro Veronesi. Rational panics and stock market crashes. Journal of Economic Theory, 110(2):234–263, 2003.

Suleyman Basak and Anna Pavlova. Asset prices and institutional investors. American Economic Review, 103(5):1728–1758, 2013.

Itzhak Ben-David, Francesco Franzoni, Rabih Moussawi, and John Sedunov. The granular nature of large institutional investors. Management Science, 67(11):6629–6659, 2021.

James Benford, Stuart Berry, Kalin Nikolov, Chris Young, and Mark Robson. Quantitative easing. Bank of England. Quarterly Bulletin, 49(2):90, 2009.

Dirk Bergemann, Tibor Heumann, and Stephen Morris. Information and volatility. Journal of Economic Theory, 158:427–465, 2015.

Bruno Biais, David Martimort, and Jean-Charles Rochet. Competing mechanisms in a common value environment. Econometrica, 68(4):799–837, 2000.

Bradyn Breon-Drish. On existence and uniqueness of equilibrium in a class of noisy rational expectations models. The Review of Economic Studies, 82(3):868–921, 2015.

Markus K Brunnermeier and Lasse Heje Pedersen. Market liquidity and funding liquidity. The review of financial studies, 22(6):2201–2238, 2009.

Georgy Chabakauri, Kathy Yuan, and Konstantinos E Zachariadis. Multi-asset noisy rational expectations equilibrium with contingent claims. LSE working paper, 2017.

Louis KC Chan and Josef Lakonishok. The behavior of stock prices around institutional trades. The Journal of Finance, 50(4):1147–1174, 1995.

Chiraphol N Chiyachantana, Pankaj K Jain, Christine Jiang, and Robert A Wood. International evidence on institutional trading behavior and price impact. The Journal of Finance, 59(2): 869–898, 2004.

Kee H Chung and Sahn-Wook Huh. The noninformation cost of trading and its relative importance in asset pricing. The Review of Asset Pricing Studies, 6(2):261–302, 2016.

George M Constantinides. Capital market equilibrium with transaction costs. Journal of Political Economy, 94(4):842–862, 1986.

Efe Çötelioğlu, Francesco Franzoni, and Alberto Plazzi. What constrains liquidity provision? evidence from institutional trades. Review of Finance, 25(2):485–517, 2021.

Joshua Coval and Erik Stafford. Asset fire sales (and purchases) in equity markets. Journal of Financial Economics, 86(2):479–512, 2007.

Sebastian Doerr, Egemen Eren, and Semyon Malamud. Money market funds and the pricing of near-money assets. Swiss Finance Institute Research Paper, (23-04), 2023.

Songzi Du and Haoxiang Zhu. Bilateral trading in divisible double auctions. Journal of Economic Theory, 167:285–311, 2017.

Darrell Duffie, Nicolae Gârleanu, and Lasse Heje Pedersen. Over-the-counter markets. Econometrica, 73(6):1815–1847, 2005.

F Douglas Foster and S Viswanathan. Strategic trading when agents forecast the forecasts of others. The Journal of Finance, 51(4):1437–1478, 1996.

Xavier Gabaix, Parameswaran Gopikrishnan, Vasiliki Plerou, and H Eugene Stanley. A theory of power-law distributions in financial market fluctuations. Nature, 423(6937):267–270, 2003.

Gerard Gennotte and Hayne Leland. Market liquidity, hedging, and crashes. The American Economic Review, pages 999–1021, 1990.

Sergei Glebkin, Naveen Gondhi, and John Chi-Fong Kuong. Funding constraints and informational efficiency. The Review of Financial Studies, 34(9):4269–4322, 2021.

Sergei Glebkin, Semyon Malamud, and Alberto Teguia. Illiquidity and higher cumulants. The Review of Financial Studies, 36(5):2131–2173, 2023.

Gary Gorton and Guillermo Ordonez. The supply and demand for safe assets. Journal of Monetary Economics, 125:132–147, 2022.

John M Griffin, Jeffrey H Harris, and Selim Topaloglu. The dynamics of institutional and individual trading. The Journal of Finance, 58(6):2285–2320, 2003.

Zhiguo He, Bryan Kelly, and Asaf Manela. Intermediary asset pricing: New evidence from many asset classes. Journal of Financial Economics, 126(1):1–35, 2017.

Paul D Klemperer and Margaret A Meyer. Supply function equilibria in oligopoly under uncertainty. Econometrica: Journal of the Econometric Society, pages 1243–1277, 1989.

Arvind Krishnamurthy and Annette Vissing-Jorgensen. The aggregate demand for treasury debt. Journal of Political Economy, 120(2):233–267, 2012.

Albert S Kyle. Continuous auctions and insider trading. Econometrica: Journal of the Econometric Society, pages 1315–1335, 1985.

Albert S Kyle. Informed speculation with imperfect competition. The Review of Economic Studies, 56(3):317–355, 1989.

Albert S Kyle, Anna A Obizhaeva, and Yajun Wang. Smooth trading with overconfidence and market power. The Review of Economic Studies, 85(1):611–662, 2017.

Francis A Longstaff. Portfolio claustrophobia: Asset pricing in markets with illiquid assets. American Economic Review, 99(4):1119–1144, 2009.

Semyon Malamud. Noisy arrow-debreu equilibria. Available at SSRN 2572881, 2015.

Semyon Malamud and Marzena Rostek. Decentralized exchange. American Economic Review, 107(11):3320–62, 2017.

Dömötör Pálvölgyi and Gyuri Venter. Multiple equilibria in noisy rational expectations economies. Available at SSRN 2524105, 2015.

Joel Peress. Wealth, information acquisition, and portfolio choice. The Review of Financial Studies, 17(3):879–914, 2003.

Gabor Pinter. An anatomy of the 2022 gilt market crisis. Available at SSRN, 2023.

John KH Quah. The law of demand and risk aversion. Econometrica, 71(2):713–721, 2003.

Marzena Rostek and Marek Weretka. Price inference in small markets. Econometrica, 80(2): 687–711, 2012.

Marzena Rostek and Marek Weretka. Dynamic thin markets. The Review of Financial Studies, 28(10):2946–2992, 2015a.

Marzena Rostek and Marek Weretka. Information and strategic behavior. Journal of Economic Theory, 158:536–557, 2015b.

Ioanid Roşu. A dynamic model of the limit order book. The Review of Financial Studies, 22 (11):4601–4641, 2009.

Michael Rothschild and Joseph E Stiglitz. Increasing risk: I. a definition. In Uncertainty in Economics, pages 99–121. Elsevier, 1978.

Yuliy Sannikov and Andrzej Skrzypacz. Dynamic trading: Price inertia and front-running. Preprint, 2016.

Adrian Tobias and Markus K Brunnermeier. Covar. The American Economic Review, 106(7): 1705, 2016.

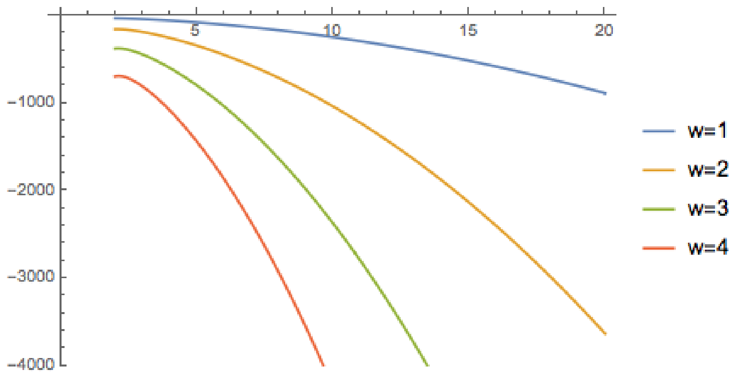
Dimitri Vayanos. Strategic trading and welfare in a dynamic market. The Review of Economic Studies, 66(2):219–254, 1999.

Dimitri Vayanos and Jean-Luc Vila. Equilibrium interest rate and liquidity premium with transaction costs. Economic theory, 13(3):509–539, 1999.

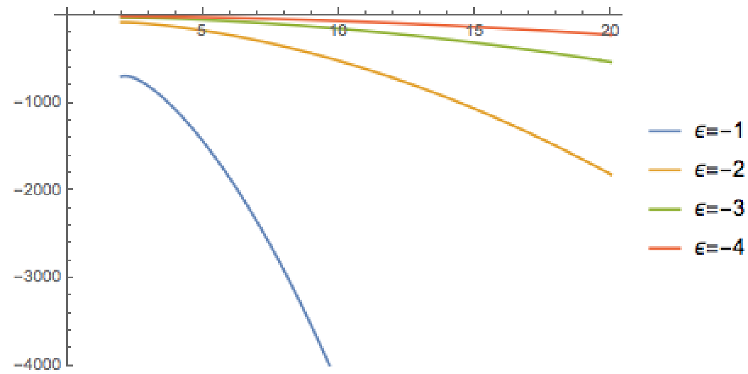
Xavier Vives. Strategic supply function competition with private information. Econometrica, 79(6):1919–1966, 2011.

David C Wheelock and Paul W Wilson. Do large banks have lower costs? new estimates of returns to scale for us banks. Journal of Money, Credit and Banking, 44(1):171–199, 2012.

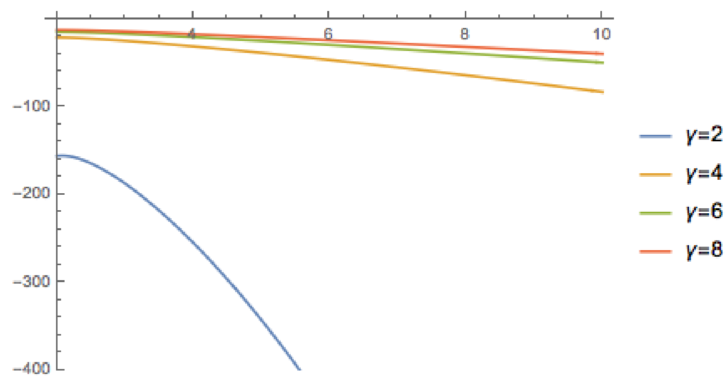
Kathy Yuan. Asymmetric price movements and borrowing constraints: A rational expectations equilibrium model of crises, contagion, and confusion. The Journal of Finance, 60(1):379–411, 2005.



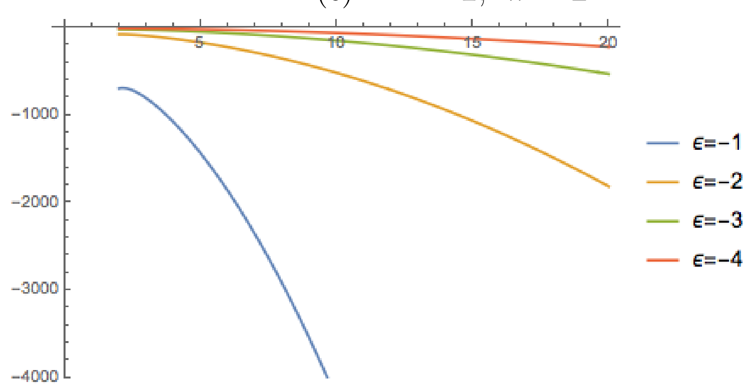
(a) ” $= -\mathbf{1}$; fl $= \mathbf{2}$



(b) $\mathbf{w} = \mathbf{4}$; fl $= \mathbf{2}$

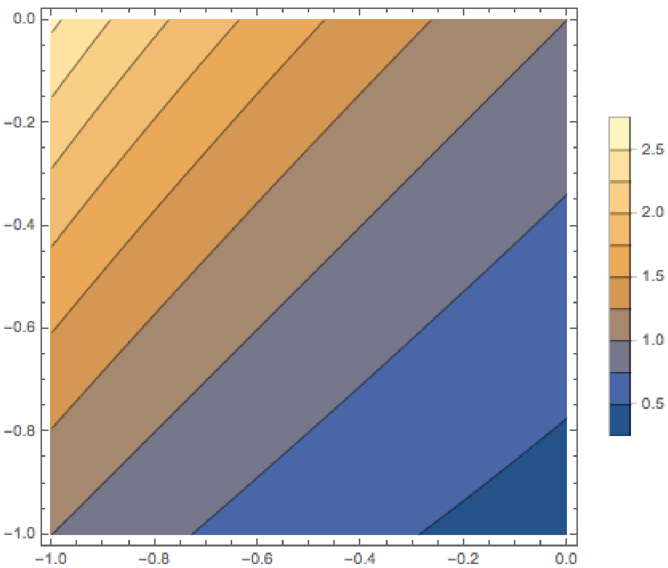


(c) ” $= -\mathbf{1}$; $\mathbf{w} = \mathbf{2}$

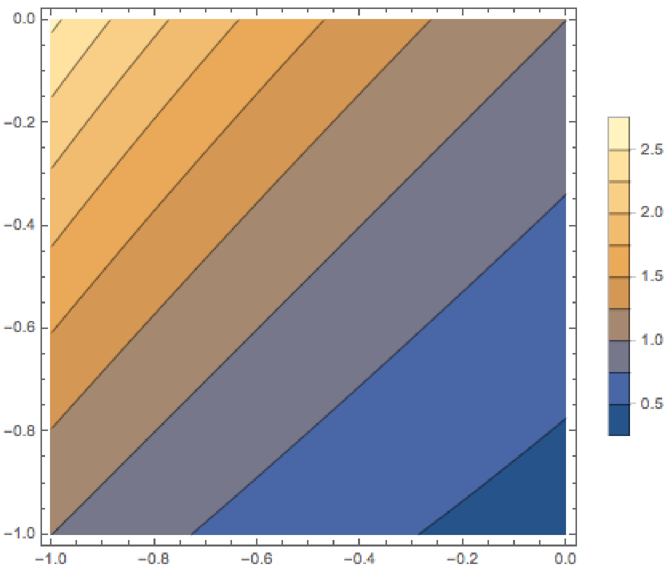


(b) $\mathbf{L} = \mathbf{5}$

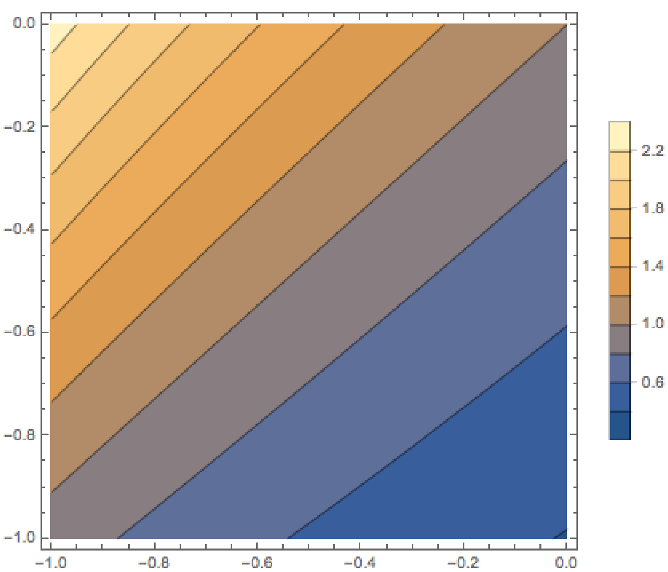
Figure 2: **Equilibrium Price Impact** We plot Λ as a function of L.



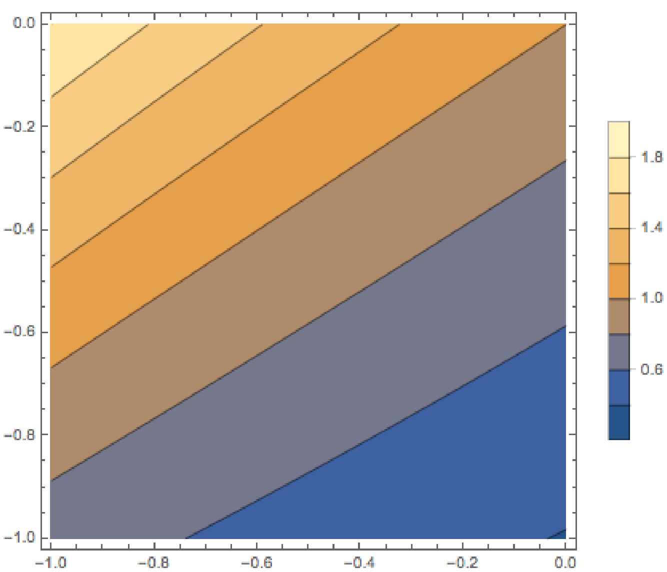
(a) $\mathbf{\bar{p}_1 = \bar{p}_2 = 0.5;\ w_1 = w_2 = 2}$



(b) $\mathbf{\bar{p}_1 = 0.1; \bar{p}_2 = 0.9;\ w_1 = w_2 = 2}$



(c) $\mathbf{\bar{p}_1 = \bar{p}_1 = 0.5;\ w_1 = 2; w_2 = 2.2}$



(d) $\mathbf{\bar{p}_1 = \bar{p}_1 = 0.5;\ w_1 = 2; w_2 = 3}$

Figure 3: **Price Impact Asymmetry** We present contour plots of $\Lambda_{12}/\Lambda_{21}$ as a function of (q_1, q_2).