Equilibrium Theory of Financial Markets: Recent Developments*

Marzena Rostek and Ji Hee Yoon

October 12, 2020

Abstract

This article reviews the theory of imperfectly competitive financial markets with special attention to recent contributions and rapidly growing new areas of research. We survey the literature on advances that have led to a common analytic framework for static, dynamic, centralized, and decentralized markets. This allows us to highlight the results that arise when traders have price impact (but not in competitive markets) and, separately, those that arise with market fragmentation.

JEL Classification: D47, D53, G11, G12

Keywords: Imperfect competition, Price impact, Decentralized trading, Efficiency, Financial innovation, Market design

1 Motivation

Among the assumptions that power the classical equilibrium theory of financial markets, two have become particularly salient; namely, that markets are *competitive* — i.e., no individual trader has impact on prices — and that trading is *centralized* — essentially, there is a single (actual or virtual) exchange for all traders and all assets. Today's financial markets, however, are *imperfectly competitive* and highly fragmented.

That financial markets are not competitive has been well understood at least since the late 1980s, when trade-level data first became available. Even before market fragmentation occurred, on a typical day in the largest world exchanges, like the NYSE, trade would be dominated by a relatively small number of large institutional investors whose orders move prices. Institutional investors routinely estimate their price impact, often multiple times per

*Rostek gratefully acknowledges the financial support from the NSF Grants SES-1357758 and SES-1851857. Additional acknowledgments to be added. Rostek: University of Wisconsin-Madison, Department of Economics; mrostek@ssc.wisc.edu. Yoon: University College London, Department of Economics and Centre for Finance; jihee.yoon@ucl.ac.uk.

day. Software for price impact estimation has long been available not only to institutional investors but also individual investors.[1] Trading costs associated with price impact are the first-order costs and exceed the explicit trading costs, such as commission, brokerage, and order-processing fees.

More recently, the financial crisis of 2008 highlighted the crucial role that market fragmentation plays within the financial system as well as the extent to which the financial market is fragmented and interconnected. Essentially all financial assets (and goods) are traded in multiple coexisting trading venues. This is the case not only for assets with heterogeneous units, such as real estate, but also for homogeneous assets. Most bonds (government, municipal, and corporate) are traded over the counter, as are interbank loans, currencies, asset-backed securities, derivatives, and many stocks. To take the top US equity markets as an example, the NYSE currently creates less than 25% of the trading volume of its listed assets. The remaining trade occurs in 12 other public exchanges (i.e., those in which all traders can participate), over 30 private exchanges (i.e., those with participation restricted mostly to large institutional investors), and over 200 dealer networks.[2]

Even setting aside the challenges involved in modeling market fragmentation, the scope of the existing theory of imperfectly competitive equilibrium (and the games in demand and supply functions on which it relies) has yet to reach that of the competitive theory. Academic papers that gave the impetus for the analysis of imperfectly competitive financial markets[3] were published three to four decades after the foundational papers for the competitive equilibrium. The reasons for such inertia in moving beyond the competitive analysis are clear: the unrivaled generality and elegance of the competitive model and the way it facilitates analysis of many questions regarding aggregate variables in the economy.[4] However, the proliferation of new trading venues, market-clearing arrangements, trading speeds, and financial instruments, along with the availability of new data, have increased the salience of imperfect competition and market fragmentation. New questions concerning market structure, design, and regulation that were raised by the financial crisis, trends in concentration, and by advances in technology have all created new demand for understanding the effects of imperfect competition.

The literature has made great strides in understanding the equilibrium phenomena due to imperfect — but not perfect — competition and, separately, those that arise with market fragmentation. This article's goal is twofold: first, to survey the theory of imperfectly competitive and/or decentralized financial markets with special attention to recent contributions and new research areas; and second, to synthesize the new ideas and methods for

[1] E.g., software by Citigroup, EQ International, ITG, MCI Barra, and OptiMark.

[2] `https://www.nyse.com/market-data`

[3] Notably, Wilson (1979), Kyle (1989), Klemperer and Meyer (1989), and Vayanos (1999).

[4] In fact, markets for financial assets were still given as the primary example of a perfectly competitive market during the 2000s when one of the authors of this article was in graduate school.

equilibrium analysis that have emerged over the past decade in the class of models based on demand submission games and, more generally, in methods for analyzing imperfectly competitive markets. The article highlights which classic results from the competitive (general equilibrium) theory carry over when trading is imperfectly competitive or decentralized, and points to the areas in which there is scope for — indeed, need for — new theory. Accounting for the fact that markets are imperfectly competitive and fragmented sheds light on several features of the data that theory has grappled with: Why is there so much trade? Why is there so much financial innovation? Why are there active markets for many types of financial products and derivatives that would be redundant in the standard theory? Which markets should and which should not be active?

Scope and Methods. To survey the theory of imperfectly competitive financial markets, the article focuses on environments in which traders have continuous multiunit demands and supplies for homogenous assets, which match the primitives of the standard competitive model.[5] A canonical model in the literature on markets that are dominated by large institutional investors (i.e., most financial markets) is based on the *uniform-price* double auction, in which traders submit demand and supply schedules specifying quantities for all prices. In practice, schedules are submitted using combinations of market and limit orders. The uniform-price demand submission game is the finite-market counterpart of the competitive model, which is based on the uniform price.

The framework based on demand submission encompasses the competitive model of centralized trading in the quasilinear-quadratic setting (e.g., CAPM). This permits a direct comparison with classical methods and results. The equilibrium framework based on the demand submission game allows for the continuation of the "general equilibrium" tradition in the sense that it accommodates the core feature of general equilibrium — i.e., interdependence vs. partial equilibrium — but with strategic traders, thus enabling the development of imperfectly competitive models using game-theoretic tools while allowing perfect competition to appear as a result rather than an assumption.[6]

To survey the literature, the article focuses on the uniform-price mechanism in the quadratic-Gaussian setting,[7] which retains the key substantive general insights while allowing for their exposition at a relatively nontechnical level. Most of what is understood about strategic trading today was established in a variant of the quadratic-Gaussian setting, which is a continuation of the finance tradition. Even within this setting, advances such as incorporating trader heterogeneity, flexible information structures, or rich market structures have occurred

[5]Networked models of *indivisible* goods are surveyed in several chapters in the Handbook of Networks (2016) coedited by Bramoullé, Galeotti, and Rogers; networked models of divisible goods with *quantities* rather than schedules as strategies are surveyed by Jackson and Zenou (2014).

[6]Section 2.2 indicates which of the characteristics typically seen as defining features of a general equilibrium model have yet to be incorporated into demand submission models.

[7]E.g., utilities are CARA and random variables are Normally distributed or, often equivalently for trading behavior, utilities are quadratic in the quantity traded.

in the past 10 years.

The article provides a unified treatment for static, dynamic, centralized, and decentralized markets — competitive and imperfectly competitive. This indicates a common mathematical structure across trading environments and draws attention to new properties brought by imperfect competition, dynamics, and market decentralization. The Appendix includes a more detailed presentation and background results. The approach to achieve a modeling synthesis in one analytic framework is an (equivalent) representation of equilibrium as a fixed point in price impacts.

Throughout the article, we relate the demand submission game to alternative models of imperfectly competitive markets: the Cournot model and the (single-agent) transaction costs approach.

We identify the competitive model with the uniform-price mechanism and a single exchange — the assumptions of the Arrow-Debreu model, which have provided the foundation for results in equilibrium and asset pricing theory. A large body of work provides noncooperative foundations for the competitive equilibrium.[8] In light of these results, which are established in different models of markets, it is not essential to identify the competitive model with a centralized exchange or a single aggregation; nonetheless, we make this identification to put some discipline on our analysis.

Outline. The article begins with a historical perspective on the development of the relevant ideas (Section 2). We next lay out the equilibrium implications due to imperfect competition in static (Section 4) and dynamic (Section 5) centralized markets. Section 6 then examines additional features implied by market decentralization, particularly the *joint* implications of fragmentation and price impact. Locally, in the trading relationships in which they participate, traders are not negligible and have price impact (see Section 4.1.2).

The Appendix leverages the unified analytic framework to systematize the steps in equilibrium characterization and properties of the fixed point.

2 Imperfect Competition in Financial Markets

2.1 Financial Markets in Numbers

Most trading in financial markets is done by a relatively small number of large institutions, such as banks and institutional investors. Institutional investors include hedge funds, mutual and pension funds, endowment funds, commercial banks, and insurance companies. These traders continuously monitor prices and are ready to respond to price differentials at any time,

[8] These include papers on large double auctions (e.g., Wilson (1977), Milgrom (1981), Pesendorfer and Swinkels (2000), and Reny and Perry (2006)) and models based on noncooperative bilateral bargaining in large dynamic markets (Rubinstein and Wolinsky (1985, 1990), Gale (1986a,b), and McLennan and Sonnenschein (1991)). See also Gale (2000).

as they have the technology and liquidity to do so. Thus, while the number of traders who occasionally participate in financial markets is quite large, it is the behavior of institutional traders that, to a large extent, determines the equilibrium properties of prices and trades.[9]

These realities of modern financial markets stand in stark contrast to the price-taking assumption, which has often been justified by the large number of participants in financial markets, each of whom is negligible in the market and thus has no price impact. The earliest academic studies reporting the significance of institutional investors and price impact include Holthausen, Leftwich, and Mayers (1987), Chan and Lakonishok (1993, 1995), and Keim and Madhavan (1995, 1996, 1998). Shleifer (1986) initiated the estimation of the slope of aggregate market demand, showing that the demand curve for US equities is downward sloping and thus there is a price impact when large orders are executed.

In the late 1980s — even before the rise of the alternative trading venues — a typical institutional package in an exchange as deep as the NYSE would represent more than 60% of the average daily trading volume if traded at once (Chan and Lakonishok (1995), Table 1). Institutional investors dominate trading in traditional public exchanges as well as alternative trading venues, such as intradealer markets and dark pools. Additionally, institutional stock ownership more than doubled between 1980 and 2005 (Thompson Reuters). This increase is due mainly to the growing concentration of ownership.

In one of the largest studies to date, Frazzini, Israel, and Moskowitz (2018) analyzed 1.7 trillion of live executed trades from a large money manager over 19 years (1998 to June 2016) for 21 developed equity markets. In the data, there is considerable variation in trade size, ranging from less than 0.1% to 13.1% of daily trading volume. The authors report that the average bid-ask spread at the time of order arrival is 21.33 bps. Nevertheless, they show that traders rarely incur the full — or even half — spread because of the passive limit orders, concluding, "The main cost our trades face is price impact." The trading algorithms have discretion over the duration of the trade. The average *ex ante* expected trade horizon is 2.7 days, with the median trade taking place over 1.7 days and the maximum taking 9.8 days. The study contains a wealth of insight into how trading costs vary globally across trade type, size, and exchange.

[9]Other than their size, institutional investors differ from individual investors in the ways in which they participate in markets. Whereas retail investors trade stocks in round lots of 100 shares or more, institutional investors buy and sell in block trades of at least 10,000 shares. Further, institutional investors have access to investments like swaps and forward markets, while retail investors generally do not. And individual investors typically place their orders, particularly large ones, through a broker at a fee; in turn, institutional investors place orders directly on the exchange or through an intermediary at a negotiated price.

2.2 Theory of Imperfectly Competitive Financial Markets: A Historical Perspective

Despite its high-stakes applications, the theory of imperfectly competitive trading is relatively undeveloped compared to the competitive model. A closer look at the history of modeling imperfectly competitive financial markets sheds light on its central difficulties.

The impact of competitive general equilibrium theory on economic theory, applications, and market practice has been such that equilibrium theory has become synonymous with general equilibrium. Soon after the foundational results for the competitive model were established (Arrow (1951), Debreu (1951), and Arrow and Debreu (1954)), economists began to investigate the possibility of an analogous foundation for an imperfectly competitive equilibrium. The earliest work in the general equilibrium tradition includes Negishi (1961), Arrow and Hahn (1971), Hahn (1977, 1978), and Hart (1979). However, relaxing the assumption of perfect competition in the general equilibrium framework posed problems (surveyed, e.g., by Bonano (1990), Hart (1985), and Salanie (2000)). In essence, these difficulties manifested in either a large multiplicity of equilibria or equilibrium nonexistence, except under certain conditions. At the heart of these difficulties in developing the imperfectly competitive analogue of the competitive (general equilibrium) model lied the ability to endogenize the market's response to a demand change by one trader. A game-theoretic formulation, which requires making the behavior following a unilateral deviation explicit, thus held promise. As a result, the theory of imperfect competition has advanced through partial equilibrium models using noncooperative game theory. Developments in game theory with asymmetric information facilitated further progress.

While the idea of considering demand or supply functions as strategies had been around earlier, the finite-market counterpart of the competitive model was formalized in the late 1970s/early 1980s as a game where players' strategies are demand or supply functions. Grossman (1981) is usually credited with introducing the complete information game with demands (supplies) as strategies in industrial organization. Contemporaneously, Wilson (1979) presented the first strategic analysis of divisible good auctions; that is, the demand function game with asymmetric information. Other early seminal contributions include Klemperer and Meyer (1989) in industrial organization, and Kyle (1989) and Vayanos (1999) in finance. The models of Kyle (1989) — a demand submission game — and Kyle (1985) were cast explicitly in a financial markets context and became workhorse frameworks for modern equilibrium analysis of imperfectly competitive financial markets for divisible assets in which traders have multiunit demands.[10]

[10] Vives (2008) provides a textbook treatment of the early models. For markets with indivisible assets in which traders have demands for one unit, influential contributions followed Glosten and Milgrom (1985). These models, too, capture price impact, induced by asymmetric information (cf. ft. 23). For an overview of models based on the limit-order book, see Foucault (2010).

We should mention that general equilibrium models where no agent is a price-taker have been studied by Negishi (1961), Shubik (1973), Dubey and Shubik (1977), Shapley and Shubik (1977), Roberts (1987), Weretka (2011), and Carvajal and Weretka (2012). These models placed different assumptions on the *off-equilibrium* market clearing and trader optimization (e.g., arbitrary exogenously given behavior in Negishi (1961), optimization and market clearing given equilibrium monetary spending in Shapley and Shubik (1977, see Appendix D.1), optimization and market clearing given a linear approximation in Weretka (2011)). While this article reviews theory based on game-theoretic models, which accommodate asymmetric information, it is important to acknowledge that many ideas and mathematical properties of equilibria in the general equilibrium models relate to those of (Bayesian) Nash equilibria in the game-theoretic models based on demand submission games.

The competitive model tends not to be analyzed as one where the traders' choice variables are price-contingent demand or supply functions. However, this is to some extent a byproduct of the way in which the solution concept of the competitive equilibrium is defined and the complete information assumption. Namely, the competitive equilibrium is typically defined as a restriction on the profile of trades and prices — *levels rather than functions*. With uncertainty, price is a random variable and the competitive model is one in which traders submit price-contingent schedules.[11] Uncertainty about what price will clear the market motivates traders to submit price-contingent functions rather than quantities (i.e., price-inelastic schedules). With strategies that allow the players to condition on the outcomes to-be-realized, a demand submission game sidesteps the much-discussed difficulties of the general-equilibrium formulation concerning how equilibrium prices and trades are determined, how the equilibrium is implemented, or how asymmetric information can be incorporated (see Section 4.3).

Unrelated to work toward a theory of imperfectly competitive markets, the idea of accounting for the residual market's response in the Cournot context inspired the game-theoretic literature and the conduct parameter approach in industrial organization of the late 1970s and early '80s (see Berry, Gaynor, and Morton (2019)). Empirical studies provided part of the motivation: outcomes in many industries appeared to fall between those that the Cournot and Bertrand models would predict; i.e., they appeared to be imperfectly competitive but more competitive than in the Cournot model. The common explanation for these outcomes was that when making decisions, a firm anticipates and accounts for other firms' adjustments. The demand submission model captures this logic and produces outcomes between the Cournot and competitive outcomes (see "Comparison with the Cournot model" in Section 4.1.2).[12]

[11] With quasilinear utilities, for instance, adding the quantifier "for all price realizations" in the first-order condition of each trader would trace his competitive price-contingent demand, which coincides with the trader's primitive marginal utility.

[12] Early literature explored the *conjectural variation* approach to endogenizing a player's reaction function,

The 2000s brought to light how crucially market fragmentation, frictions, and design details each factor in the functioning of financial institutions.[13] An active literature on fragmented markets is now emerging at the intersection of finance, macro, and market design.[14]

Modeling strategic behavior in games with demand or supply functions as strategies has run into two kinds of challenges. Early literature recognized the conceptual challenge (how to formulate the problem so that equilibrium is well defined, i.e., it exists and is at least locally unique) and equilibrium multiplicity. During late 1970s and the 1980s, the literature made clear that uncertainty refines the set of equilibria (Klemperer and Meyer (1989) and Kyle (1989)). Tractability remained a challenge in developing a general theory — a useful perspective is that even competitive models with heterogeneous traders are typically analyzed using numerical methods. The setting with quadratic payoffs has enabled researchers to gain insights into the key forces of imperfect competition, and the advances in this setting have led to a common analytic framework for static, dynamic, centralized, and decentralized markets. Specifically, dispensing with symmetry assumptions on traders' preferences, information, and the market enabled analysis of market structures and problems that require flexible information structures. Additionally, steps taken to address the challenges presented by trader heterogeneity in dynamic models with private information (e.g., forecasting others' forecasts, the curse of dimensionality, non-Markovian learning) have enabled analysis of equilibrium with *dynamic* inference and trading.

One strand of the literature popular among industry practitioners and in mathematical finance research takes the perspective of an agent trading against an exogenously given supply function, the functional form of which is motivated empirically (Almgren and Chriss (2000), Huberman and Stanzl (2004), and Almgren et al. (2005)). An increasingly popular approach to the modeling of trading behavior assumes quadratic transaction costs (Gârleanu and Pedersen (2013)) or inventory costs — which apply, respectively, per trade and allocation. This

which was not based on a Nash equilibrium (see Hahn (1977), Hart (1979), Bresnahan (1981), Robson (1983), Sabourian (1992), and the book by Figuières, Jean-Marie, Quérou, and Tidball (2004)). Endogenizing the reaction of a player's strategic counterparties in the conjectural variation approach presented difficulties for the same reason that, in the Cournot model, the off-equilibrium outcome is not well defined in the absence of price-taking traders whose choices define an elastic demand: When strategies are quantity levels rather than price-contingent schedules, accounting for (and determining) the adjustment in other players' behavior following a unilateral deviation necessitates a *dynamic* interaction logic in a static game. A demand submission game endogenizes the response by letting the players condition their quantity demanded on prices that both in and off equilibrium are required to respect optimization and market clearing. See Klemperer and Meyer (1989), Section 4.1.1, and ft. 37.)

[13]The paper by Duffie, Gârleanu, and Pedersen (2005) — based on a search model rather than a demand submission framework — played an important role in stimulating the literature, as did the financial crisis.

[14]E.g., Antill and Duffie (2017), Du and Zhu (2017a), Hendershott and Madhavan (2015), Hortaçsu, Kastl and Zhang (2018), Kyle and Lee (2017), Kyle, Obizhaeva, and Wang (2018), Menkveld, Yueshen and Zhu (2017), Bergemann, Heumann, and Morris (2019), Sannikov and Skrzypacz (2016), Duffie (2018), Zhu (2018a,b), Malamud and Rostek (2017), Babus and Kondor (2018), Babus and Hatchem (2019), Babus and Parlatore (2017), Rostek and Yoon (2018a,b, 2020), Yoon (2018), and Zhang (2018); see also Kastl (2017) and Milgrom (2019). The literature on divisible good auctions is growing as well (see Section 4.2.3).

literature focuses on the single-agent rather than the equilibrium problem, taking a trader's residual supply, and hence his price impact, as given.[15] The game-theoretic framework based on demand submission games is a multi-agent counterpart of these models and endogenizes price impact in equilibrium.

3 Setting

Notation. We use the following notation: $(x_i)_i$ is a vector in which the i^{th} element is x_i, and $(y_{ij})_{i,j}$ is a matrix such that the $(i,j)^{\text{th}}$ element is y_{ij}; sets of the respective elements are denoted by $\{x_i\}_i$ and $\{y_{ij}\}_{i,j}$. In addition, $diag(x_1, \cdots, x_N)$ is a diagonal matrix in $\mathbb{R}^{N\times N}$ in which the i^{th} diagonal element is x_i. The $(k,\ell)^{\text{th}}$ element of matrix M is denoted by $m_{k\ell}$ and the k^{th} row of M is denoted by M_k. We denote the transpose of matrix M by M'. Lastly, $\mathbf{1}_K \in \mathbb{R}^K$ and $\mathbf{1}_{K\times K} \in \mathbb{R}^{K\times K}$ are a vector and a matrix, all elements of which are one, and Id is the identity matrix.

Traders and assets. $I \geq 3$ traders trade K risky assets. Traders are indexed by i and assets by k.

The payoffs of the K risky assets are jointly Normally distributed $R = (r_k)_k \sim \mathcal{N}(\theta, \Sigma)$ with vector of expected payoffs $\theta \in \mathbb{R}^K$ and a positive semi-definite covariance matrix $\Sigma \in \mathbb{R}^{K\times K}$. There is also a riskless asset with a zero interest rate (which serves as a numéraire). Each trader initially holds $q_0^i = (q_{k,0}^i)_k \in \mathbb{R}^K$ units of risky assets. Unless stated otherwise, endowment q_0^i is privately known to trader i. Endowments $\{q_0^i\}_i$ are jointly Normally distributed and can be correlated (see Equation (3)).

Double auction. The market for the risky assets operates as a uniform-price double auction. Each trader i submits K strictly downward-sloping price-contingent (net) demand functions $q^i(\cdot) : \mathbb{R}^K \to \mathbb{R}^K$ specifying the quantity of each of the K assets demanded for any price vector $p = (p_1, ..., p_K) \in \mathbb{R}^K$; i.e., demand for each asset k is $q_k^i(p_1, ..., p_K)$. For $q_k^i > 0$, trader i is a buyer for asset k; for $q_k^i < 0$, he is a seller (Fig. 1A). Trade of the K assets clears simultaneously: the equilibrium price vector is determined by $\sum_i q^i(p_1, \cdots, p_K) = 0 \in \mathbb{R}^K$.[16] Trader i trades $\{q_k^i\}_k$ and pays $p \cdot q^i = \sum_k p_k q_k^i$.

Each trader i has a utility that is quasilinear in the riskless asset and quadratic (mean-variance) in the quantity of risky assets:

$$u^i(q_0^i + q^i) = \theta \cdot (q^i + q_0^i) - \frac{\alpha^i}{2}(q^i + q_0^i) \cdot \Sigma(q^i + q_0^i) - p \cdot q^i. \tag{1}$$

[15] Incidentally, the same behavioral assumption (i.e., of trading against a conjecture about the price response of the residual market) was employed in early studies based on the general-equilibrium approach (e.g., Negishi (1961), Hahn (1977, 1978)).

[16] The definition of the game can be completed in the usual way: If there is no such price or if multiple prices exist, then no trade takes place. The assumption that bids are strictly downward-sloping rules out trivial equilibria with no trade.

All traders are strategic.[17] Gains from trade come from risk sharing and diversification: endowments $\{q_0^i\}_i$ and risk preferences $\{\alpha^i\}_i$ are heterogeneous.

Equilibrium. All results for static markets are based on the Bayesian Nash equilibrium, and those for dynamic markets are based on Perfect Bayesian equilibrium, both in linear demand schedules.[18]

Definition 1 *A profile of (net) demand schedules $\{q^i(\cdot)\}_i$ is a Bayesian Nash Equilibrium if, for each i, $q^i(\cdot)$ maximizes the expected payoff*

$$\max_{q^i(\cdot):\mathbb{R}^K\to\mathbb{R}^K} E[\theta\cdot(q^i+q_0^i)-\frac{\alpha^i}{2}(q^i+q_0^i)\cdot\Sigma(q^i+q_0^i)-p\cdot q^i|q_0^i], \tag{2}$$

subject to

$$q^i+\sum_{j\neq i}q^j(p)=0,$$

given strategies of other traders $\{q^j(\cdot)\}_{j\neq i}$.

Benchmark models. We consider two benchmark models in the same setting: the competitive model and the Cournot model. The latter will also permit the comparison with the transaction costs approach.

In keeping with the literature, to ensure that the per capita aggregate endowment (equivalently, price) is random in the limit large market ($I\to\infty$), we allow for the common value component q_0^{cv} in traders' privately known endowments. The subsequent analysis accommodates an arbitrary joint Gaussian distribution for privately known endowments, in particular, the independent private value model, which is of separate interest ($Var(q_{k,0}^{cv})=0$).

Endowments $\{q_0^i\}_i$ are distributed according to $q_{k,0}^i\sim\mathcal{N}(E[q_{k,0}],\sigma_{cv}^2+\sigma_{pv}^2)$ for all i and k. For each asset k, endowments $\{q_{k,0}^i\}_i$ are correlated across players through a common value component $q_{k,0}^{cv}\sim\mathcal{N}(E[q_{k,0}^{cv}],\sigma_{cv}^2)$:[19]

$$q_{k,0}^i=q_{k,0}^{cv}+q_{k,0}^{i,pv};\qquad q_{k,0}^{i,pv}\overset{iid}{\sim}\mathcal{N}(0,\sigma_{pv}^2). \tag{3}$$

Trader i observes his endowment q_0^i but not its components q_0^{cv} or $q_0^{i,pv}$. The endowments $\{q_{k,0}^i\}_i$ and the common value $q_{k,0}^{cv}$ are independent across assets k. The endowment vector $(q_0^i)_i$ is distributed according to $\mathcal{N}(E[q_0^{cv}]\otimes\mathbf{1}_K,\Omega\otimes Id)$, where $\Omega=\sigma_{cv}^2\mathbf{1}_{I\times I}+\sigma_{pv}^2Id$ for all k.

Definition 2 (Competitive Market, Competitive Equilibrium) *Consider a market with $I<\infty$ traders. The* competitive market *is the limit game as $I\to\infty$, holding fixed all other*

[17] Noise traders could be incorporated with price-inelastic or -elastic random supply.

[18] Equilibrium is *linear* if schedules have the functional form of $q^i(p)=\alpha_0+\alpha_q^iq_0^i+\alpha_p^ip$.

[19] The common value component in $\{q_0^i\}_i$ affects the magnitude of inference coefficient in Appendices A.2, B, and C.2, but it does not affect any results qualitatively.

primitives. Letting $\{q^{i,I}(\cdot)\}_i$ *be the equilibrium in the market with* $I < \infty$ *traders, the* competitive equilibrium *is the limit of equilibria* $\{q^{i,I}(\cdot)\}_i$ *as* $I \to \infty$*:*

$$q^i(\cdot) = \lim_{I\to\infty} q^{i,I}(\cdot) \qquad \forall i.$$

Definition 3 (Cournot Model) [20] *There are* $I < \infty$ *traders with preferences (1). Each trader submits a quantity vector* $q^i \in \mathbb{R}^K$ *that represents (net) demand for the assets. There is also a continuum of price-taking investors whose (net) demand is described by a linear (inverse) supply:*

$$p = H + G\sum_i q^i,$$

where $H \in \mathbb{R}^K$ *is a vector and* $G \in \mathbb{R}^{K\times K}$ *is a non-singular matrix, which are exogenously given.*

4 Centralized Markets

We first lay out the implications of imperfect competition. Relaxing the price-taking assumption in trader optimization (i.e., allowing a finite number of traders) is the sole departure from the assumptions of the competitive equilibrium. All traders are strategic; no results rely on the presence of noise traders, nonstrategic or not-fully-optimizing agents (who could be easily incorporated).

4.1 Equilibrium

Trader i submits a (net) demand schedule $q^i(\cdot) : \mathbb{R}^K \to \mathbb{R}^K$ as a function of the price vector, i.e., $q^i(p) \in \mathbb{R}^K$ for each $p = (p_1, \cdots, p_K) \in \mathbb{R}^K$ to maximize expected utility:

$$\max_{q^i(p)=(q^i_1(p),\cdots,q^i_K(p))} E[\theta\cdot(q^i+q^i_0) - \frac{\alpha^i}{2}(q^i+q^i_0)\cdot\Sigma(q^i+q^i_0) - p\cdot q^i | q^i_0, p] \qquad \forall\, p\in\mathbb{R}^K. \quad (4)$$

In the *pointwise* optimization problem, the payoff of trader i in (4) is measurable with respect to conditioning and contingent variables $\{q^i_0, p\}$; thus, the expected payoff is the same as the *ex post* payoff:

$$\max_{q^i(p)=(q^i_1(p),\cdots,q^i_K(p))} \{\theta\cdot(q^i+q^i_0) - \frac{\alpha^i}{2}(q^i+q^i_0)\cdot\Sigma(q^i+q^i_0) - p\cdot q^i\} \qquad \forall\, p\in\mathbb{R}^K.$$

[20] E.g., Bayesian Cournot-Nash model of Vives (2002). Related models include Grossman and Miller (1988), Brunnermeier and Pedersen (2005), Pritsker (2005), DeMarzo and Urošević (2006), extended by Urošević (2005).

4.1.1 Equilibrium as a fixed point in price impacts

It is useful to adopt the perspective of an individual trader who optimizes against a residual market, represented by his residual supply function $S^{-i}(\cdot) \equiv -\sum_{j\neq i} q^j(\cdot)$. This leads to an equilibrium characterization in terms of two simple conditions (Proposition 1 below), which separate individual optimization from equilibrium.[21]

Equilibrium $\{q^i(\cdot)\}_i$ is characterized in the following two steps.

Step 1 (Optimization of trader i, given price impact) The first-order condition with respect to the demand for each asset q^i_k is: for any $q^i_0 \in \mathbb{R}^K$ and $p \in \mathbb{R}^K$,

$$0 = \underbrace{\theta_k - \alpha^i(\sigma_{kk}(q^i_k + q^i_{k,0}) + \sum_{\ell\neq k}\sigma_{k\ell}(q^i_\ell + q^i_{\ell,0}))}_{\text{Marginal utility w.r.t. } q^i_k} - \underbrace{(p_k + \frac{dp_k}{dq^i_k}q^i_k + \sum_{\ell\neq k}\frac{dp_\ell}{dq^i_k}q^i_\ell)}_{\text{Marginal payment w.r.t. } q^i_k} \qquad \forall k. \tag{5}$$

Written in matrix form, the first-order condition becomes:

$$\theta - \alpha^i\Sigma(q^i + q^i_0) = p + \Lambda^i q^i, \tag{6}$$

where matrix $\Lambda^i \equiv \frac{dp}{dq^i} \in \mathbb{R}^{K\times K}$ is the *price impact* of trader i:[22]

$$\Lambda^i \equiv \begin{bmatrix} \frac{dp_1}{dq^i_1} & \cdots & \frac{dp_K}{dq^i_1} \\ \vdots & \ddots & \vdots \\ \frac{dp_1}{dq^i_K} & \cdots & \frac{dp_K}{dq^i_K} \end{bmatrix}.$$

Its $(k,\ell)^{\text{th}}$ element $\frac{dp_\ell}{dq^i_k}$ represents the price change in asset ℓ following a deviation in asset k. The inverse of price impact is a common measure of *liquidity*: the lower the price impact, the smaller the price concession a trader must accept to trade, the more liquid the market.[23] Equation (6) is the standard first-order condition of a monopolist — required to hold for all prices. In a multilateral oligopoly model — the demand submission game — it applies for all traders.

Equalization of marginal utility with marginal payment pointwise for each price in Equa-

[21] The idea of considering the optimization of a single trader, given the residual market, goes back to Kyle (1989). Rostek and Weretka (2015a) introduce the equilibrium definition in terms of optimization and the fixed point in price impacts, showing the equivalence of the equilibrium characterization.

[22] (i.e., 'Kyle's lambda')

[23] The literature developed various notions of illiquidity due to exogenous transaction costs, demand (inventory) shocks, private information, and search frictions that capture the difficulty of finding a counterparty. Amihud, Mendelson, and Pedersen (2005) survey empirical measures of liquidity. Being a counterfactual (i.e., off-equilibrium) object, price impact is *by definition* associated with asymmetric information. In any demand submission model, the price impact can be decomposed into the effect of the decreasing marginal utility and asymmetric information induced off equilibrium. See Equation (38).

tion (6) gives the best-response schedule of trader i:

$$q^i(p) = (\alpha^i\Sigma + \Lambda^i)^{-1}(\theta - p - \alpha^i\Sigma q_0^i) \qquad \forall p \in \mathbb{R}^K, \tag{7}$$

given his price impact Λ^i that will be characterized in Step 2.

Step 2 (Price impacts are correct) In equilibrium, the price impact in the pointwise optimization (6) of trader i (which could have been arbitrary in the logic so far) must be correct: It must equal the $K \times K$ Jacobian matrix of the inverse residual supply function of trader i, $S^{-i}(\cdot) = -\sum_{j\neq i} q^j(\cdot)$, determined by aggregation of the other traders' submitted schedules $\{q^j(\cdot)\}_{j\neq i}$ in (7). Taking a derivative of the market-clearing condition[24] $q^i + \sum_{j\neq i} q^j(p) = 0$ with respect to $q^i = (q^i_k)_k$:

$$Id + \sum_{j\neq i} \frac{\partial q^j(p)}{\partial p}\frac{dp}{dq^i} = 0$$

gives the required condition:

$$\Lambda^i \equiv \frac{dp}{dq^i} = -(\sum_{j\neq i} \frac{\partial q^j(p)}{\partial p})^{-1}. \tag{8}$$

The best-responses (7) of traders $j \neq i$ determine demand Jacobians $\frac{\partial q^j(p)}{\partial p}$ for $j \neq i$:

$$\frac{\partial q^j(p)}{\partial p} = -(\alpha^j\Sigma + \Lambda^j)^{-1} \qquad \forall j \neq i. \tag{9}$$

Substituting demand Jacobians (9) in Equation (8) characterizes the price impact:

$$\Lambda^i \equiv \frac{dp}{dq^i} = (\sum_{j\neq i} (\alpha^j\Sigma + \Lambda^j)^{-1})^{-1}. \tag{10}$$

Proposition 1 (Equilibrium: Static Market) A profile of demand schedules $\{q^i(\cdot)\}_i$ is a Bayesian Nash equilibrium if and only if

(i) (Optimization of trader i, given price impact Λ^i) Demand schedule $q^i(\cdot) : \mathbb{R}^K \to \mathbb{R}^K$ is determined by pointwise equalization of marginal utility and marginal payment:

$$q^i(p) = (\alpha^i\Sigma + \Lambda^i)^{-1}(\theta - p - \alpha^i\Sigma q_0^i) \qquad \forall p \in \mathbb{R}^K, \tag{11}$$

given his price impact Λ^i, such that:

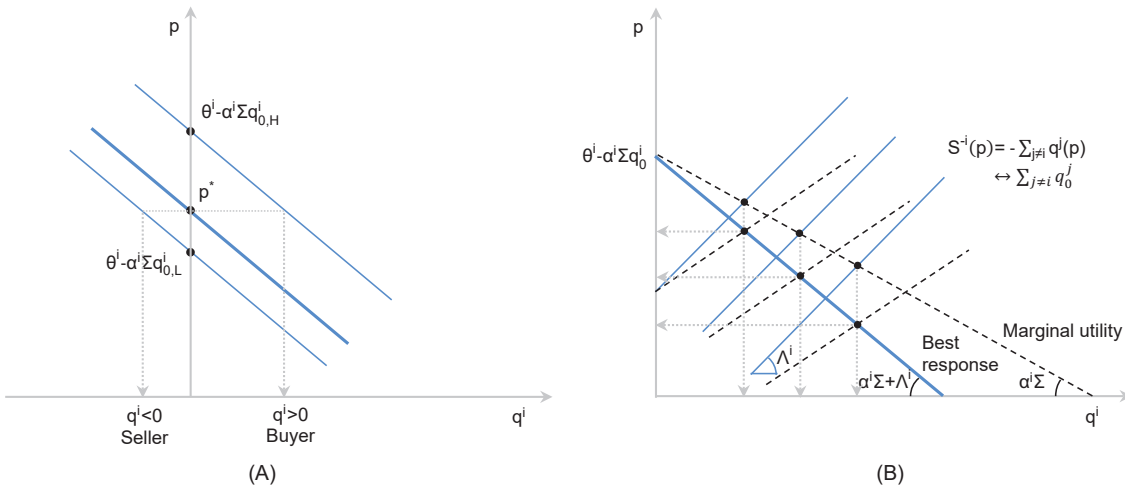
[24] Or, applying the implicit function theorem to $\{FOC^j\}_{j\neq i}$ where FOC^j is $lhs - rhs$ of Equation (6). The implicit function theorem is useful to derive price impacts in markets with richer contingent variables.

(ii) (Correct price impacts) The price impact of trader i equals the slope of his residual inverse supply function:

$$\Lambda^i = \left(\sum\nolimits_{j\neq i}(\alpha^j\Sigma + \Lambda^j)^{-1}\right)^{-1} \qquad \forall i. \tag{12}$$

Conditions (i) and (ii) jointly provide an *equivalent* characterization of the (Bayesian Nash) equilibrium in demand schedules (see Lemma 1 below): (i) traders optimize, given their assumed price impacts, (ii) which are correct. Analyzing price impact directly offers direct insights into the role of imperfectly competitive behavior and, as discussed in Section 4.3, inference. Counterparts of these conditions carry over to dynamic and decentralized markets (Propositions 3, 6, and 4).

FIGURE 1: POINTWISE OPTIMIZATION AND DEMAND REDUCTION



Notes: (A) The market-clearing price p^* is determined such that the aggregate net demand is zero: $\sum_i q^i(p^*) = 0$. Trader i trades $q^i = q^i(p^*)$ and is a buyer if $q^i > 0$ and a seller if $q^i < 0$. (B) Trader i's demand is optimal pointwise for each price realization against a family of the residual supply $S^{-i}(\cdot) = -\sum_{j\neq i} q^j(\cdot)$ with a deterministic slope (price impact) and random intercept (due to other traders' private information).

- *Pointwise optimization.*

Fig. 1 illustrates the method of finding equilibrium in games in demand schedules via pointwise optimization — known at least since Kyle (1989) — which applies even if equilibrium is not *ex post*. Given the linear best response functions submitted by traders $j \neq i$, the residual

supply function of trader i is:

$$S^{-i}(p) = \underbrace{-\sum_{j\neq i}(\alpha^j\Sigma+\Lambda^j)^{-1}(\theta-\alpha^j\Sigma q_0^j)}_{\text{Random intercept}} + \underbrace{\sum_{j\neq i}(\alpha^j\Sigma+\Lambda^j)^{-1}}_{\text{Slope }=(\Lambda^i)^{-1}} p. \tag{13}$$

With private information about other traders' endowments, trader i faces a family of residual supplies with a random intercept and constant slope (Λ^i). Equilibrium price p in (18):

$$p = (\sum_j(\alpha^j\Sigma+\Lambda^j)^{-1})^{-1}((\alpha^i\Sigma+\Lambda^i)^{-1}(\theta-\alpha^i\Sigma q_0^i) + \underbrace{\sum_{j\neq i}(\alpha^j\Sigma+\Lambda^j)^{-1}(\theta-\alpha^j\Sigma q_0^j)}_{\text{Intercept of } -S^{-1}(\cdot)}).$$

Provided that the aggregate slope $\Psi \equiv \sum_j(\alpha^j\Sigma+\Lambda^j)^{-1}$ is not zero, price realizations $p \in \mathbb{R}^K$ map one-to-one to a realization of the intercept of the residual supply function. Equilibrium is characterized by pointwise optimization for demand functions and fixed-point conditions on slopes (or slope functions) rather than on levels.

- *The market-clearing condition is accounted for by the equilibrium condition (ii) for price impacts.*

The price impact of each trader is determined by the requirement that optimization and market clearing hold in equilibrium *and* following a unilateral demand change: By (12), the price impact of trader i is characterized as the price change at which other traders are willing to sell the extra units demanded by i (given that traders $j \neq i$ optimize according to (11)) so that the market clears. Each trader's maximization problem (4) is subject to the market-clearing condition as a constraint. When the trader knows his price impact Λ^i, the market-clearing condition can be removed from the constraint maximization problem.

- *Equilibrium equivalence.*

To show that equilibrium conditions (i) and (ii) jointly give an equivalent characterization of the Bayesian Nash equilibrium, suppose that each trader submits $q^i(\cdot;\Lambda^i)$, as in Step 1. When does bidding $\{q^i(\cdot;\Lambda^i)\}_i$ correspond to a (Bayesian Nash) equilibrium? A necessary (and sufficient) condition for a profile $\{q^i(\cdot;\Lambda^i)\}_i$ to be an equilibrium is that the price impact assumed by each trader is correct; this gives condition (ii) in Step 2. This yields that the price impact which satisfies condition (ii) is the sufficient statistic for a trader's equilibrium demand.

Lemma 1 (Equivalence) A profile of demands $\{q^i(\cdot;\Lambda^i)\}_i$ is a Bayesian Nash equilibrium if and only if $\Lambda^i = -(\sum_{j\neq i}\frac{\partial q^j(p)}{\partial p})^{-1}$ for all i.

4.1.2 Price impact

The equilibrium condition (ii) is a fixed point in price impacts. This system can be solved. Letting the sum of the traders' demand slopes

$$\Psi \equiv \sum_j (\alpha^j \Sigma + \Lambda^j)^{-1} \tag{14}$$

be the *slope of aggregate (net) demand*, the price impacts of all traders can be solved explicitly as functions of the aggregate slope:

$$\Lambda^i = (\Psi - (\alpha^i \Sigma + \Lambda^i)^{-1})^{-1}, \quad \text{equivalently,} \quad (\Lambda^i)^{-1} + (\alpha^i \Sigma + \Lambda^i)^{-1} = \Psi \qquad \forall i.$$

Denoting the aggregate slope by $\Psi = \psi \Sigma^{-1}$, the above matrix equation for $\Lambda^i = \alpha^i \beta^i \Sigma$ ($\beta^i > 0$) is simplified into a scalar (quadratic) equation: for each i,

$$\frac{1}{\alpha^i \beta^i} + \frac{1}{\alpha^i + \alpha^i \beta^i} = \psi, \quad \text{therefore,} \quad \beta^i = \frac{2 - \alpha^i \psi + \sqrt{(\alpha^i \psi)^2 + 4}}{2 \alpha^i \psi}.$$

Substituting β^i into (14) gives an implicit equation for ψ,[25]

$$\sum_i (2 + \alpha^i \psi + \sqrt{(\alpha^i \psi)^2 + 4})^{-1} = \frac{1}{2}, \tag{15}$$

which has a unique positive solution for ψ, and thus, a unique solution for price impacts

$$\Lambda^i = (\sum_{j \neq i} (\alpha^j \Sigma + \Lambda^j)^{-1})^{-1} = \alpha^i \beta^i \Sigma \qquad \forall i. \tag{16}$$

The price impact characterization above provides a condition for equilibrium existence: When traders know θ, equilibrium with trade does not exist when $I = 2$;[26] then, by Equation (15), $\psi = 0$. The zero aggregate slope (i.e., inelastic aggregate demand) implies that price impacts become infinite, i.e., $\beta^i = \infty$ for all i, and hence each trader's best response demand is inelastic at his initial endowment.

The properties of the fixed point offer a glimpse into the properties and determinants of price impact.

- *Price impact derives from the payoff concavity.*

What is the origin of price impact? From Equation (16), at a primitive level, the price impact

[25] Introducing aggregate slope Ψ reduces the dimension of the fixed point variables — from $\{\Lambda^i\}_i$ to Ψ in the static centralized market. Even outside of centralized markets, this dimension reduction improves the (numerical) tractability of equilibrium characterization.

[26] With (negatively) interdependent values $\{\theta^i\}_i$, equilibrium exists also for $I = 2$ (see Section 4.3.1).

of trader i derives from the concavity of preferences of his residual market: it is increasing in $\{\{\alpha^j\}_{j\neq i}, \Sigma\}$. The riskier the asset (i.e., the larger Σ in the positive-definite order) and the more risk averse traders $j \neq i$ are, the less elastic those traders' marginal utilities, the greater the effect of a demand change by trader i on their marginal utility, the larger the price concession they require to absorb the new units so that the market clears. Trader i's own risk aversion affects his own price impact as well — through the fixed point (16) — as it enters the residual supplies that determine the price impacts of traders $j \neq i$.

Price impact is strictly decreasing in the number of traders I.[27] In a larger market, the effect of trader i's demand change on the average marginal utility of other traders is smaller. Each trader absorbs a smaller fraction of the deviation.

From Equation (16), the price impact of every trader is *proportional* to the fundamental covariance matrix.[28] This link between incentives (i.e., Λ^i) and risk Σ has important implications for equilibrium and design (some of which are discussed in Section 6). One consequence of the proportionality of Λ^i to Σ is that the correlation $corr(p_k, p_\ell)$ between equilibrium prices follows the fundamental correlation between assets $\sigma_{k\ell}$, as in the competitive model. Letting $\gamma^i \equiv \frac{\alpha^i}{\alpha^i + \alpha^i\beta^i} \in \mathbb{R}$, the equilibrium price is:

$$p = \theta - (\sum_j \frac{1}{\alpha^j + \alpha^j\beta^j})^{-1} \sum_j \frac{\alpha^j}{\alpha^j + \alpha^j\beta^j} \Sigma q_0^j = \theta - (\sum_j \frac{\gamma^j}{\alpha^j})^{-1} \sum_j \gamma^j \Sigma q_0^j. \tag{18}$$

The cross-asset price impacts $\lambda^i_{k\ell} \equiv \frac{dp_\ell}{dq^i_k}$ for $\ell \neq k$ and for k inherit the sign from asset correlation $\sigma_{k\ell}$.

- *Risk preferences determine the ranking of price impacts.*

Traders' equilibrium price impacts are ranked inversely to their risk preferences:

$$\alpha^1 < ... < \alpha^I \Rightarrow \Lambda^1 > ... > \Lambda^I. \tag{19}$$

[27] Bertrand-style models with perfectly elastic marginal payoffs/costs suggest that two traders suffice for perfect competition. A version of that result can be seen here in the equilibrium condition for price impacts (16) if at least two traders have a constant marginal utility; i.e., $\alpha^i \to 0$. Then, in the limit, all traders' equilibrium price impacts are zero and traders' equilibrium demands coincide with their marginal utilities. With multiunit demands, however, the constant marginal utility is a rather strong assumption, as it implies that the trader's marginal valuation of every additional unit stays the same irrespective of how many units are already held.

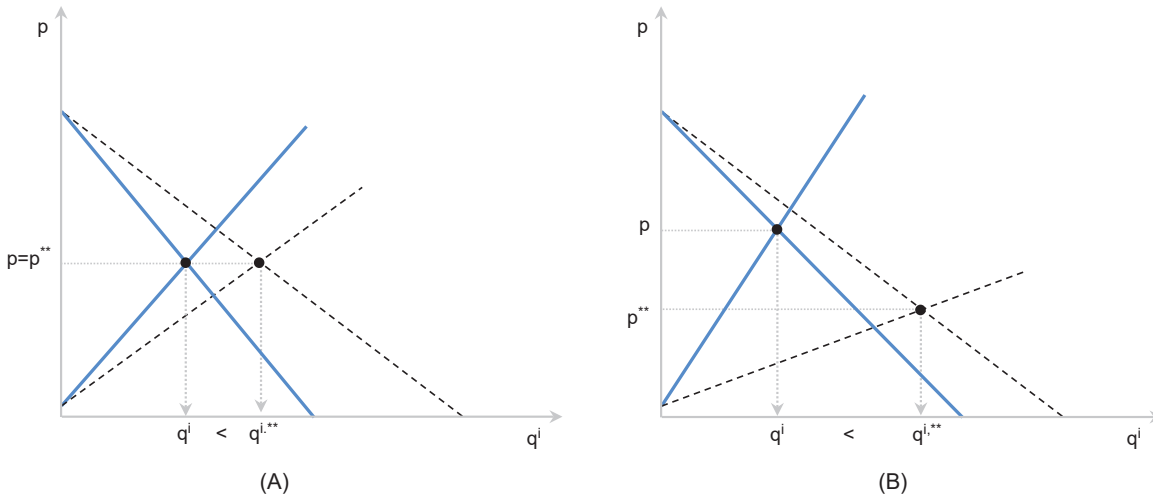
[28] Multiplying both sides of Equation (16) by $\Sigma^{-1/2}$ and letting $\hat{\Lambda}^i = \Sigma^{-1/2}\Lambda^i\Sigma^{-1/2}$, we have:

$$\hat{\Lambda}^i = (\sum_{j\neq i} (\alpha^j Id + \hat{\Lambda}^j)^{-1})^{-1} \qquad \forall i. \tag{17}$$

Any solution to (17) is of the form of $\hat{\Lambda}^i = \alpha^i\beta^i Id$ for some $\beta^i > 0$, and consequently, $\Lambda^i = \alpha^i\beta^i\Sigma$. The uniqueness of the linear equilibrium is proved in Malamud and Rostek (2017).

Traders' price impact matrices can be ranked linearly (i.e., asset by asset), given the proportionality of price impact in the covariance. Which counterparties are more important in the determination of a given trader's price impact? The less risk averse traders are. Mathematically, this result and the ranking (19) are implied by the harmonic mean structure of the fixed point: Price impact Λ^i is a harmonic mean of inverse demand slopes $\{\alpha^j\Sigma + \Lambda^j\}_{j\neq i}$. The harmonic mean[29] appears from aggregation in market models and assigns higher weights to smaller elements. The economic implication is that the less risk averse counterparties of trader i are more influential determinants of his price impact. These traders have more elastic marginal utility and thus absorb more of the demand change, given the price change.

FIGURE 2: TRADER HETEROGENEITY AND EQUILIBRIUM PRICE



Notes: Consider two markets — one in which all traders have the same risk aversion (panel A) and one in which some traders (sellers) are less risk averse than others (panel B). In both markets, traders reduce their demands and supplies in response to price impact. However, when traders are symmetric, the buyers and sellers reduce their demands by the same fraction (i.e., $\gamma^i = \gamma$ for all i), and the market clears at the competitive price even though trade can be significantly reduced relative to the competitive outcome. When traders are heterogeneous, the sellers face a less elastic residual market than the buyers and thus reduce their schedules more. Equilibrium price differs from the one in the competitive market.

- *Relative price impacts matter.*

In models of monopoly, market power increases the monopolist's revenue. In a financial market context, however, price impact is seen as a transaction *cost*. These intuitions can be reconciled when we observe that the *relative price impacts* determine whether a trader benefits from the ability to move prices in the sense that the equilibrium price is higher

[29]The harmonic mean of $\{\alpha^j\}_j$ is the reciprocal of the arithmetic mean of the reciprocals: $(\sum_j(\alpha^j)^{-1})^{-1}$.

or lower than if traders were price-takers: With heterogeneous traders, the risk premium $Q \equiv \theta - p$ in price depends on price impacts among traders who buy and sell in equilibrium, as determined by their relative endowments (Equation (18)). The fact that the relative price impacts (or relative convexities of buyers and sellers) determine the equilibrium properties is one of the enduring lessons of the imperfectly competitive markets. (See Fig. 2.)

A central insight of the competitive model was the link between concavity and allocation (22) (see, e.g., Duffie and Sonnenshein (1989)). Imperfect competition qualifies the role of concavity of the primitive utility functions as the key structure underlying equilibrium. Given the concavity, the fixed point among the interacting agents' price impact alters the qualitative properties of equilibrium. With interdependent values (Section 4.3), there is an additional fixed point for traders' inference coefficients, given their price impacts.

4.2 Classical Results Revisited: Static Markets

The following properties of competitive equilibrium will serve as a reference in the characterization of allocations and prices in imperfectly competitive markets:

- Traders are *price-takers.* Equation (16) gives the following inequality for price impact:

$$\frac{\min_j \alpha^j}{I-2}\Sigma < \Lambda^i = (\sum_{j\neq i}(\alpha^j\Sigma + \Lambda^j)^{-1})^{-1} < \frac{\max_j \alpha^j}{I-2}\Sigma. \tag{20}$$

  Taking the limit of the inequality (20) as $I \to \infty$, $\Lambda^i \to 0$ for all i.

- The competitive equilibrium demand coincides with the inverse marginal utility, given the quasilinearity of the utility function. Taking the limit of Equations (11) and (18), the competitive demand schedules and price are determined by:

$$q^i(p) \;\to\; (\alpha^i\Sigma)^{-1}(\theta - p) - q_0^i \quad \forall p \in \mathbb{R}^K \;\forall i; \qquad p \to \theta - (\sum_j \frac{1}{\alpha^j})^{-1}\Sigma \sum_j q_0^j. \tag{21}$$

  When price impact is positive, $\Lambda^i > 0$, trader i demands (or sells) less relative to his competitive schedule (Fig. 1B).

The competitive and imperfectly competitive theories of financial markets were developed using different tools and solution concepts. Some of the apparent dichotomies between the results based on the general-equilibrium approach to modeling competitive markets and those based on the game-theoretic approach to modeling markets in general are due to the differences in the solution concept itself (Nash equilibrium vs. not). These include the determination of equilibrium prices and trades, the implementation of equilibrium demands,[30] and

[30]In the general equilibrium model, how agents get to equilibrium/price formation is often justified by

the possibility of incorporating asymmetric information.[31] In the game-theoretic framework of demand submission, the dichotomies between competitive and imperfectly competitive results disappear. Those differences that remain illuminate why certain properties of the competitive equilibrium are special.

4.2.1 Prices and allocations

- *Imperfect competition severs the link between equilibrium prices and the fundamental value (i.e., the average marginal utility $\theta - \alpha\Sigma\bar{q}_0$).*

In contrast to the competitive model, equilibrium prices differ from the fundamental values, except when the utility functions of all traders are the same (Equation (18)). When traders are equally risk averse, the market clears at competitive prices despite the allocation being imperfectly competitive (Fig. 2A and Example 2 in Appendix A.1). This result is independent of how many traders will be buyers and sellers in equilibrium: aggregating demands (11) with $\alpha^i = \alpha$ for all i, we have that $p = \theta - \alpha\Sigma\bar{q}_0$ where $\bar{q}_0 \equiv \frac{1}{I}\sum_j q_0^j$. One implication is that measures of market competitiveness or liquidity should not be based on price levels (or markups) but price changes. One cannot rely on price levels "canceling out" when analyzing equilibrium outcomes or their changes. The variation in prices — price impact — is what determines the outcome.

- *Imperfect competition shifts the focus from aggregate risk to idiosyncratic risk; for aggregate risk, it shifts the focus from systematic to systemic risk.*

In the competitive market, equilibrium allocation coincides with the *efficient allocation*, i.e.,

$$q^i + q_0^i \;\rightarrow\; \frac{1}{\alpha^i}\Big(\sum_j \frac{1}{\alpha^j}\Big)^{-1} \sum_j q_0^j \qquad \forall i, \tag{22}$$

the interpretation based on a competitive auctioneer. It is unclear why agents should take the prices as given (for any auction rules, manipulative behavior may be possible). In the demand submission game: (i) The auctioneer/algorithm story is precisely how the market-clearing price is determined in one round by aggregating demand and supply schedules. (ii) Traders do not take prices as given unless price impact is zero and how prices are determined is part of the description of the rules of the game. (iii) No manipulation can occur, since the informational asymmetry that arises from a deviation is accounted for. Because all realizations of contingent variables can occur in equilibrium, Bayes' rule holds on and off equilibrium (i.e., following a unilateral demand change by a player). See the discussion below Proposition 2.

[31] In the demand schedule game (competitive or not), equilibrium incorporates asymmetric information — the players submit schedules contingent on the realization of a random variable. Player-specific asymmetric information (and not aggregate uncertainty alone) affects equilibrium: when a player's impact on the price is not negligible, equilibrium can impacted by the inference about the variables that describe a single player (as opposed to aggregate random variables). Additionally, in the definition of the competitive equilibrium in the general equilibrium model, an agent does not take into account what other agents know; hence, in equilibrium, some traders may be willing to exchange units in events that other traders know will not occur.

which maximizes total welfare over the set of all feasible allocations $\{\{q^i+q_0^i\}_i : \sum_i(q^i+q_0^i) = \sum_i q_0^i\}$, or equivalently, minimizes the total utility loss associated with risk exposure, and is given by

$$q^{i,**} = \frac{1}{\alpha^i}q^{**}, \quad \text{where } q^{**} = (\sum_j \frac{1}{\alpha^j})^{-1}\sum_j q_0^j. \tag{23}$$

In an imperfectly competitive equilibrium, equilibrium trades and allocations are, respectively,

$$q^i = \gamma^i(\frac{1}{\alpha^i}q^* - q_0^i) \quad \text{and} \quad q^i + q_0^i = \gamma^i\frac{1}{\alpha^i}q^* + (1-\gamma^i)q_0^i, \tag{24}$$

when $\gamma^i \equiv \frac{1}{1+\beta^i}$ for each i. The vector of market clearing prices is given by $p = \theta - Q$. The common portfolio allocated to all traders

$$q^* \equiv \Sigma^{-1}Q = (\sum_j \frac{\gamma^j}{\alpha^j})^{-1}\sum_j \gamma^j q_0^j = \bar{\alpha}\frac{1}{I}\sum_j q_0^j + (\frac{1}{I}\sum_j \frac{\gamma^j}{\alpha^j})^{-1}\operatorname{Cov}((\gamma^i)_i,(q_0^i)_i), \tag{25}$$

represents the aggregate risk portfolio, where $\bar{\alpha} \equiv (\sum_j \frac{\gamma^j}{\alpha^j})^{-1}\sum_j \gamma^j$ and $\operatorname{Cov}((\gamma^i)_i,(q_0^i)_i) \equiv \frac{1}{I}\sum_j \gamma^j q_0^j - (\frac{1}{I}\sum_j \gamma^j)(\frac{1}{I}\sum_j q_0^j)$.

Definition 4 (Aggregate (Systemic) Risk, Systematic Risk) Aggregate risk (systemic risk)

$$Q \equiv \theta - p = (\sum_j(\alpha^j\Sigma + \Lambda^j)^{-1})^{-1}\sum_j(\alpha^j\Sigma+\Lambda^j)^{-1}\alpha^j\Sigma q_0^j = (\sum_j \frac{\gamma^j}{\alpha^j})^{-1}\sum_j \gamma^j\Sigma q_0^j \tag{26}$$

represents the risk that cannot be diversified in equilibrium and corresponds to the risk premium in prices relative to the mean return θ. In the centralized market, q^* is the aggregate risk portfolio (common portfolio)*: $\theta - p = \Sigma q^*$.*

Systematic risk *is $Q^{**} \equiv \Sigma q^{**}$, where q^{**} is the efficient allocation (23), is independent of the market structure, and represents risk nondiversifiable in the market.*

In the competitive market, the aggregate (systemic) and systematic risk coincide ($Q = Q^{**}$) and depend on endowments only through the aggregate endowment $\sum_j q_0^j$.

When traders have price impact:

(i) Traders are exposed to idiosyncratic risk: a trader retains a fraction $1-\gamma^i$ of his initial endowment q_0^i. The First Welfare Theorem does not hold.

(ii) The aggregate (systemic) risk to which traders are exposed differs from systematic risk in two ways. First, in imperfectly competitive markets, risk that is nondiversifiable in equilibrium depends on price impact. Moreover, aggregate risk depends on the *distribution* of endowments among agents rather than merely aggregate endowment $\sum_i q_0^i$.

Moreover, the way the distribution of endowments determines aggregate risk depends on the market participants' risk preferences; i.e., the joint distribution of $\{q_0^i\}_i$ and $\{\alpha^i\}_i$ matters.

(iii) The common portfolio q^* that gets allocated to all agents in equilibrium differs from the efficient portfolio except when all traders' risk preferences $\{\alpha^i\}_i$ are the same (then, $\gamma^i = \frac{I-2}{I-1}$ for all i) or $I \to \infty$.

(iv) The ranking of agents' risk aversion still impacts who holds more risk in equilibrium (Equation (19)). However, the risk itself changes: it depends on price impact.

- *Noise trade is not essential for equilibrium to be well defined.*

Equilibrium in the demand submission game does not rely on the presence of noise traders or not-fully-optimizing traders. Thus, welfare analysis is well defined. Early literature (e.g., Kyle (1989)) assumed pure common values, which requires noise traders in order to lower the informativeness of price and ensure the existence of equilibrium with trade (i.e., finite price impacts).

- *Strategic complementarity.*

In centralized markets, price impacts are strategic complements; however, the demand submission game is not supermodular. More generally, price impacts are strategic complements when schedules are contingent even when the market is decentralized (see Section 6).

4.2.2 Method: the demand submission games approach

The following features of the characterization in terms of pointwise optimization and fixed point in price impact are worth highlighting. We will note their suitable variants in dynamic and decentralized market models.

- *Equilibrium as a fixed point in price impacts.*

Given individual optimization (condition (i)), the equilibrium characterization as a fixed point in traders' strategies $\{q^i(\cdot)\}_i$ is equivalent to the fixed point problem in price impacts $\{\Lambda^i\}_i$ in (10). In particular, the equilibrium representation in Proposition 1 endogenizes all the demand coefficients, including inference coefficients in the conditional expectation, as functions of price impact (see Section 4.3).

- *Tractability.*

The comparative statics of equilibrium and welfare can be characterized through the properties of equilibrium price impacts; the comparison of traders' beliefs is not required separately. Relying on the demand function representation to characterize equilibrium and prove the comparative statics result would be much less direct than the approach for characterizing price impacts.

The characterization in terms of conditions (i) and (ii) is more tractable. It reduces the number of endogenous variables to I price impact matrices. In particular, demand coefficients and inference coefficients do not need to be solved for as part of the fixed point problem, because they are functions of price impacts (see Section 4.3).

- *Direct link to the competitive model via price impact.*

The demand functions model allows a direct comparison with the competitive model by setting $\Lambda^i = 0$ for all i modulo scaling of risk aversion with market size.[32] Thus, the equilibrium representation in terms of price impact facilitates the development of imperfectly competitive (and decentralized market) analogues of the techniques commonly used in competitive analysis, such as the role of spanning, state-contingent pricing.

- *Price impact as a sufficient statistic.*

From Equation (11), when values θ are private, knowing own price impact suffices for a trader who knows his marginal utility (and the prior) to compute his best response. Price impact is the sufficient statistic for all the payoff-relevant information about the residual market against which he trades. No information about the primitive or endogenous variables — like the strategies of others, their utility functions, identities, or the number of market participants — is required. If the model were common knowledge, the equilibrium condition (ii) for price impact would be superfluous.

With interdependent values (Section 4.3), a trader needs to know the equilibrium price distribution and price impact.[33]

The sufficient statistic for the best response is the same in a competitive and imperfectly competitive market when cast as demand submission games — if the equilibrium price impact

[32] The common value component σ^2_{cv} in traders' endowments $\{q_0^i\}_i$ (equation (3)) ensures that the price (equivalently, the intercept of a trader's residual supply or the per capita aggregate endowment) is random in the limit large market ($I \to \infty$). To make the price variance $Var(p|q_0^i) = \Sigma Var((\sum_j \frac{\gamma^j}{\alpha^j})^{-1} \sum_j \gamma^j q_0^j | q_0^i)\Sigma'$ independent of the number of traders I, the risk aversion $\{\alpha^i\}_i$ can be scaled according to $\alpha^{i,I} = \alpha^i \sqrt{(\sigma^2_{cv} + \frac{1}{I}\sigma^2_{pv})^{-1}\sigma^2_{cv}}$. As $I \to \infty$, $\alpha^{i,I} \to \alpha^i > 0$ for all i.

[33] If information contained in price is needed to infer θ^i, then a trader needs to know the price distribution, as well. Market fragmentation does not change what a trader needs to know: the equilibrium price distribution and price impact.

is zero, traders submit their competitive demands.[34] The demand submission game weakens the role of prices (price *level*) in guiding equilibrium behavior, and hence in the determination of outcomes. What matters is how agents respond to *changes* in the variables whose realizations they condition their demands upon. (See Equation (11) and Fig. 1B.)

In the Cournot model, equilibrium is not equivalent to a fixed point in price impact and even with private values, the trader needs to know the equilibrium price distribution to compute his best response/equilibrium demand.

The sufficient statistic also captures the anonymity of investors akin to the competitive equilibrium model: (Nash) equilibrium is a result of optimization by anonymous investors whose information about the market is summarized by their price impact and the price distribution, which they correctly estimate.

- *Link to industry practice.*

Optimization given price impact corresponds to the practice of institutional traders who estimate their price impact functions, treating them as sufficient statistics for the payoff-relevant information about the residual market against which they each trade.

The logic of optimization against the residual market (Equation (13)) underlies the software for price impact estimation in the financial industry. These models adopt the perspective of an agent who trades against a residual market with an exogenously given price impact function the functional form of which is motivated empirically. A number of stylized facts about the shape of the price impact function have been established.[35] A similar model underlies the algorithms used to compute implementation shortfall and Gârleanu and Pedersen's (2013) transaction costs.

The linear-in-trade, possibly time-dependent price impact function is the predominant

[34]Early game-theoretic literature on imperfect competition has raised concerns about greater requirements for rationality and expectations in game-theoretic models relative to the general equilibrium approach; see, e.g., Arrow (1951) and Wilson (1987). An influential example of this critique has come to be known as the "Wilson's doctrine": "Game theory has a great advantage in explicitly analyzing the consequences of trading rules that presumably are common knowledge; it is deficient to the extent it assumes other features to be common knowledge, such as one agent's probability assessment about another's preferences or information. I foresee the progress of game theory as depending on successive reductions in the base of common knowledge required to conduct useful analyses of practical problems." Wilson presented this concern in a chapter devoted to the game-theoretic method of the theory of exchange, calling for progress using game theory via unification and generalization but also building more complex models of markets. This chapter has motivated the robust approach to economic analysis (surveyed, e.g., in Bergemann and Morris (2012), see also Morris (2019 Schwarz lecture)). A related criticism concerning the implementation of the Second Welfare Theorem motivated developments in the public finance literature with private information.

[35]For example, Almgren and Chriss (2000), Almgren et al. (2005) implemented by Citigroup, Huberman and Stanzl (2004); see also a recent contribution by Gârleanu and Pedersen (2013). While these models focus on single-agent optimization, they are closely related to the general equilibrium literature on price impact (see, e.g., an overview by Hart (1985)), which considers an exchange economy with the same optimization assumption as ours; each agent trades against a residual supply with an exogenously given price impact function.

assumption among practitioners (the "quadratic cost model").[36] For an overview, see, e.g., Almgren and Chriss (2000) and Almgren et al. (2005), whose models are the basis of Citigroup's Best Execution Consulting Services software, or the Handbook chapter by Vayanos and Wang (2013).

- *Multilateral oligopoly.*

Unlike in goods markets, in financial markets, it is often not the case that a trader always buys or sells. The demand submission game is a model of multilateral market power and captures precisely that: whether a trader takes a buying or a selling position depends not only on his own demand (e.g., signal, endowment shock, risk preference) but also those of his residual market (Equation (13)). Setting expected endowment $E[q_0^i]$ to be sufficiently low for some traders and high for others gives a model with buyers and sellers. Likewise, in the Cournot model, both price-takers and strategic traders can be buyers and sellers. As explained in Section 4.2.4 below, the key difference is whether all traders or only the price-taking traders provide liquidity for strategic traders (in demand submission and Cournot, respectively).

- *The demand submission approach weakens the criticism of rational expectations.*

Among the critiques of the rational expectations *assumption* (in the game-theoretic and general equilibrium approach alike), two have stood out: traders need to know the equilibrium price to act optimally, and — perhaps, the main difficulty of the rational expectations concept — the players must hold the same (and correct) expectations about where the equilibrium price will form ("communism of models", a term due to Thomas Sargent). Similar sentiments were echoed by Milgrom (1981), Dubey, Geanakoplos, and Shubik (1987), Aumann and Dreze (2008), and Woodford (2013).

The demand submission game mitigates these difficulties:

- Traders do not need to know (or compute) the equilibrium price (or outcome) to play equilibrium — their demand schedules allow them to condition on the *realizations of equilibrium variables* to-be-determined. In *equilibrium*, a trader takes as given (the payoff-relevant statistic of) others' demands to ensure the mutual best response property. As noted, knowledge of price impact (i.e., changes in market-clearing price) rather than price levels suffices for a player who knows his own marginal utility to play equilibrium.

- There's no requirement that all players have the same expectations about the equilibrium play. Instead, traders *independently* optimize knowing only *own* price impact and — with interdependent values — subjective distribution.

[36] The empirically documented concavity of price impact function (as a function of trade size) is implied by a dynamic equilibrium in a model with nonstationary price impact — larger trades are associated with smaller price impact.

- Rational expectations about equilibrium price and trades (i.e., levels) are a result, not an assumption. In the general equilibrium model of the competitive equilibrium, tatonnement need not converge; an alternative is omniscience (a variant of common knowledge: agents will solve for equilibrium and the prices will emerge).

4.2.3 Methods: Divisible assets

- *Results for divisible- and indivisible-asset auctions generally differ.*

Even for traditional auction-theoretic questions (e.g., which format maximizes revenue or efficiency), some of the most important markets are those for divisible (multi unit) settings. Examples include those for Treasury bills, spectrum, electricity, emission permits, refinancing and natural resources (e.g., water, diamonds).

It is well understood that the results from auction theory (i.e., the centralized market) for single objects in which bidders have unit demands or have a constant marginal utility do not hold in multiunit settings with decreasing marginal utility. The key to why, relative to auctions for single objects, auction theory and mechanism design for divisible good settings is relatively less developed is that a bidder valuation and strategy is not a point but a demand schedule that may specify a different willingness to pay depending on the number of units. For the relationship and differences between divisible and indivisible good markets, see Wilson (1979), Ausubel et al. (2014), and Pycia and Woodward (2018). The work of Wilson, Ausubel and Cramton (see, e.g., Ausubel et al. (2014)), Back and Zender (1993), and Wang and Zender (2002), Kremer and Nyborg (2004) has developed many of the key properties and principles for market design of divisible goods or assets. More recent contributions include McAdams (2007), Milgrom and Strulovici (2009), Kastl (2012), Holmberg, Newbery, and Ralph (2013), Rostek and Yoder (2017), Vives and Yang (2017), Andreyanov and Sadzik (2018), Burkett and Woodward (2018), Even, Tahbaz-Salehi, and Vives (2018), Rahi and Zigrand (2018), Burkett and Baisa (2018, 2019), Aldrich and Friedman (2019), Cespa and Vives (2019), and Rahi (2019).

4.2.4 Comparison with the Cournot model: bilateral vs. one-sided market power

Trader i submits a quantity demanded $q^i \in \mathbb{R}^K$ that maximizes his expected payoff:

$$\max_{q^i=(q^i_1,\cdots,q^i_K)\in\mathbb{R}^K} E[\theta\cdot(q^i+q^i_0)-\frac{\alpha^i}{2}(q^i+q^i_0)\cdot\Sigma(q^i+q^i_0)-p\cdot q^i|q^i_0], \tag{27}$$

where the inverse (net) supply of price-taking investors is given by:

$$p = H + G\sum_i q^i.$$

The best-response quantity of trader i is:

$$q^{i,C} = (\alpha^i \Sigma + \Lambda^i)^{-1}(\theta - E[p|q_0^i] - \alpha^i \Sigma q_0^i) \qquad \forall i,$$

where the price impact is defined by the slope of the inverse supply function defined by the price-taking investors:

$$\Lambda^i \equiv \frac{dp}{dq^i} = G \qquad \forall i.$$

When expected price is $E[p|q_0^i] = C_\theta^i \theta + C_q^i q_0^i$, where the inference coefficients $C_\theta^i, C_q^i \in \mathbb{R}^{K \times K}$ are determined by the projection theorem, the price in the Cournot model is:

$$p = H + G \sum_j (\alpha^j \Sigma + G)^{-1}(\theta - E[p|q_0^i] - \alpha^j \Sigma q_0^j) = H + G \sum_j (\alpha^j \Sigma + G)^{-1}((Id - C_\theta^j)\theta - (\alpha^j \Sigma + C_q^j) q_0^j).$$

Equilibrium allocations are:

$$q^{i,C} + q_0^i = (\alpha^i \Sigma + G)^{-1}((Id - C_\theta^i)\theta + (G - C_q^i) q_0^i) \qquad \forall i.$$

In contrast to the demand submission game, in the Cournot model:

- Equilibrium is not *ex post*: the best response of trader i depends on the expected price $E[p|q_0^i]$. Traders do not condition their demand on price realizations but only on expected price.
- Price impact is exogenous: the price impact of any trader is independent of the behavior (hence, of the concavity in the objective functions) of other strategic traders.
- A trader's allocation $q^{i,C} + q_0^i$ does not depend on the aggregate risk (i.e., on the risk due to the endowments of traders $j \neq i$).

To see whether the demand submission game or the Cournot game is more suitable in a given market context, consider a unified model with $I < \infty$ large traders described above and the (net) demand of price-taking investors. Suppose first that the large traders submit demand schedules rather than quantities. Then, the price impact of a large trader i is:

$$\Lambda^i \equiv \frac{dp}{dq^i} = G(Id + \sum_{j \neq i} \frac{\partial q^j(p)}{\partial p} \Lambda^i) = (Id - G \sum_{j \neq i} \frac{\partial q^j(p)}{\partial p})^{-1} G. \tag{28}$$

Contrasting the counterfactual that defines a strategic trader's price impact identifies which counterparties determine his liquidity. Suppose that one of the large traders (the players), i, increases his demand at a margin. Who sells these extra units?

- In the Cournot model, holding the quantity strategies of other players fixed — i.e., $\frac{\partial q^j}{\partial p} = 0$ for all $j \neq i$ in Equation (28) (a *Nash* equilibrium counterfactual) — the unilateral demand change of trader i is absorbed entirely by the price-taking traders. That is, the Cournot price impact $\Lambda^i = G$ of a large trader is defined by the (net) supply function derived from the marginal utilities of price-taking traders alone.

- In a demand submission game with no price-takers ($G^{-1} \to 0$ in Equation (28)), the price impact of every large trader is determined by the elastic demands of other *large* (strategic) traders $j \neq i$; $\frac{\partial q^j(p)}{\partial p} \neq 0$.[37]

- In the model with both large traders who submit schedules and price-taking investors (Equation (28)),[38] the liquidity of large traders comes from other large traders as well as price-taking investors.[39]

- Price correlation is the same as asset correlations if and only if G is proportional to Σ (and endowments are independent across assets).

In summary, in the Cournot model, market power is "one-sided": the demand change of a large (strategic) trader is absorbed by price-taking (nonstrategic) traders. In the model in which strategic traders submit demand schedules (i.e., the demand submission game or Equation (28)), market power is "two-sided": liquidity for large traders comes from all other large traders rather than small investors.

4.3 Interdependent Values

Inference from prices interacts with price impact and equilibrium behavior even in a static market: Demand schedule $q^i(p)$ allows a trader to specify the quantity demanded depending on the price to-be-determined, and hence, depending on signal realizations. To consider inference effects and the problem of information aggregation, in the market for a divisible good described in Section 4.1.1, we do not assume that the values $\{\theta^i\}_i$ are private. Traders are uncertain about how much the asset is worth to them: the intercepts of marginal utilities $\{\theta^i\}_i$ are random. This uncertainty may represent shocks to preferences or initial endowments (or inventories).

[37] Without the demand of price-taking traders, the outcome of the Cournot game — or the counterfactual that specifies the traders' price impact — is not well defined. See ft. 12. In the market structure of *multilateral oligopoly* (i.e., in which strategic traders submit schedules), market clearing among large investors does not require the presence of price-taking traders.

[38] Substituting the price impact from Equation (28) into the supply function model without price-takers characterizes equilibrium in this model. Price impact is amended by the liquidity of price-takers.

[39] In the competitive equilibrium, the counterfactual following a trader's unilateral demand change assumes no change in contingent variables. In a model with no uncertainty (e.g., Grossman (1981)), the counterfactual is not specified.

4.3.1 Price inference and equilibrium

Trader i observes a signal $s^i = \theta^i + \varepsilon^i$ of his true value θ^i. Random vector $(\theta^i, \varepsilon^i)_i$ is jointly Normally distributed according to the cdf $F((\theta^i, \varepsilon^i)_i)$, where noise ε^i_k has mean-zero and variance σ^2_ε and is independent across traders i and across assets k; $\{\varepsilon^i\}_i$ and $\{\theta^i\}_i$ are independent. Value θ^i_k has expectation $E[\theta^i_k]$ and variance σ^2_θ for all i and k, and is independent across assets k. Let $\sigma^2 \equiv \sigma^2_\varepsilon/\sigma^2_\theta$ denote the variance ratio.[40] The information structures studied in the literature can be mapped into the $I \times I$ variance-covariance matrix of the joint distribution of values $\{\theta^i_k\}_i$ for each k, normalized by variance σ^2_θ (i.e., the matrix of correlations of the joint distribution of $\{\theta^i_k\}_i$): for each asset k,

$$\begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1I} \\ \rho_{21} & 1 & \cdots & \rho_{2I} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{I1} & \rho_{I2} & \cdots & 1 \end{bmatrix}. \tag{29}$$

- The *independent (private) value* model corresponds to $\rho_{ij} = 0$ for all i and $j \neq i$.
- In the *fundamental value* model, the correlations are the same for all trader pairs $\rho_{ij} = \rho$ for all $j \neq i$ (e.g., Vives (2011)); e.g., all traders' values are affected by only common (e.g., Kyle (1989)) or idiosyncratic shocks.
- In the *equicommonal* model, each trader's value θ^i is *on average* correlated with other traders' values θ^j, $j \neq i$, in the same way: for each i, $\frac{1}{I-1}\sum_{j\neq i}\rho_{ij} = \bar{\rho}$ for some $\bar{\rho} \in [-1, 1]$ (Rostek and Weretka (2012)).
- Arbitrary Gaussian information structures with correlations $\{\rho_{ij}\}_{i,j\neq i}$ among $\{\theta^i\}_i$.

We assume that for each i, there exists $j \neq i$ such that $\rho_{ij} < 1$. Also, to simplify equations, we assume that traders know the endowment vector $(q^i_0)_i$ in this section.[41]

Trader i maximizes his utility (1) net of payment $p \cdot q^i$:

$$\max_{q^i(\cdot):\mathbb{R}^K \to \mathbb{R}^K} E[\theta^i \cdot (q^i + q^i_0) - \frac{\alpha^i}{2}(q^i + q^i_0) \cdot \Sigma(q^i + q^i_0) - p \cdot q^i | s^i]. \tag{30}$$

[40] The variances σ^2_ε and σ^2_θ, and more generally the priors $F((\theta^i, \varepsilon^i)_i)$, can be heterogenous; we impose the symmetry for the simplicity of notation.

[41] Equilibrium characterization generalizes to uncertain endowments $(q^j_0)_j$ that are independent across assets (as specified in Section 3) and independent of values $(\theta^j)_j$; results do not change qualitatively. Kawakami (2017) and Kyle and Lee (2017) generalize the model (in an imperfectly competitive version of Diamond and Verrecchia (1981)) to allow for different correlation among endowments and values (i.e., $Corr(q^i_0, q^j_0) \neq Corr(\theta^i, \theta^j)$). Then, equilibrium price endogenously weighs aggregate value and aggregate risk, potentially creating a trade-off between risk sharing of endowment risk and information aggregation.

The first-order condition is

$$\underbrace{E[\theta^i|s^i,p]-\alpha^i\Sigma(q^i+q_0^i)}_{\text{Expected marginal utility}}=\underbrace{\frac{d(q^i\cdot p(\cdot))}{dq^i}}_{\text{Marginal payment}}\qquad\forall p\in\mathbb{R}^K, \tag{31}$$

given the posterior expectations $E[\theta^i|s^i,p]=C_\theta^iE[\theta^i]+C_s^is^i+C_p^ip$ with inference coefficients C_s^i, C_p^i, and C_θ^i in $\mathbb{R}^{K\times K}$. The first-order condition (31) becomes:

$$E[\theta^i|s^i,p]-\alpha^i\Sigma(q^i+q_0^i)=p+\Lambda^iq^i\qquad\forall p\in\mathbb{R}^K.$$

In the method of finding equilibrium via pointwise optimization, there's an additional step: endogenizing posterior expectations (see Appendix A.2).

Proposition 2 (Equilibrium: Interdependent Values) A profile of (net) demand schedules $\{q^i(\cdot)\}_i$ is a linear Bayesian Nash Equilibrium if and only if for each trader i,

(i) (Optimization of trader i, given posterior expectations $E[\theta^i|s^i,p]$ and price impact Λ^i) Schedule $q^i(\cdot)$ satisfies the first-order condition:

$$q^i(p)=(\alpha^i\Sigma+\Lambda^i)^{-1}(E[\theta^i|s^i,p]-p-\alpha^i\Sigma q_0^i)\qquad\forall p\in\mathbb{R}^K, \tag{32}$$

given i's posterior expectations $E[\theta^i|s^i,p]=C_\theta^iE[\theta^i]+C_s^is^i+C_p^ip$ and i's price impact Λ^i such that:

(ii) (Bayesian posterior expectations) $E[\theta^i|s^i,p]=C_\theta^iE[\theta^i]+C_s^is^i+C_p^ip$ is determined by the Bayes' rule applied to a Gaussian price distribution $F(p|s^i)$;[42] and

(iii) (Correct price impacts) Λ^i equals the slope of i's residual inverse supply function:

$$\Lambda^i=-(\sum_{j\neq i}\frac{\partial q^j(p)}{\partial p})^{-1}=(\sum_{j\neq i}(\alpha^j\Sigma+\Lambda^j)^{-1}(Id-C_p^j))^{-1}. \tag{33}$$

Price impacts $\{\Lambda^i\}_i$ determine the equilibrium price distribution: applying the market-clearing condition to the profile of best responses $\{q^i(\cdot)\}_i$ from (i), treated as functions of price impacts and distribution of price:

$$p=(\sum_j(\alpha^j\Sigma+\Lambda^j)^{-1})^{-1}\sum_j(\alpha^j\Sigma+\Lambda^j)^{-1}(E[\theta^j|s^j,p]-\alpha^j\Sigma q_0^j). \tag{34}$$

Regarding the role of inference, it is worth noting the following:

[42] In a linear equilibrium, the price distribution induced by other traders' demands, held fixed, is Gaussian.

- With multiunit demands, submitting a demand contingent on price affects bidding, and hence efficiency, in two ways: (i) it improves learning about values: a bidder's demand is contingent on the information about his own value from other traders' signals, aggregated by price (there is a price inference coefficient C_p^i in the posterior expectations $E[\theta^i|s^i,p] = C_\theta^i E[\theta^i] + C_s^i s^i + C_p^i p$), and (ii) it changes price impact, which is a function of C_p^i.

- A trader's own price impact Λ^i and the price distribution $F(p|s^i)$ conditional on the trader's private information are the sufficient statistics for the demands $\{q^j(\cdot)\}_{j\neq i}$ of other players when computing his *best response* demand. This demand schedule is an *equilibrium* demand if and only if the price impacts are correct (analogously to Lemma 1). Thus, as with private values, equilibrium in demand schedules is equivalent to the fixed point in price impacts, given optimization (condition (i)). In particular, the equilibrium price distribution is in closed form in price impacts $\{\Lambda^i\}_i$.[43]

- Bayes' rule applies for all price realizations because all could occur in equilibrium for some realizations of signals, given downward-sloping demands and unbounded support of uncertainty. When equilibrium exists, it is fully separating.

- Because all price realizations can occur in equilibrium, Bayes' rule applies in and off equilibrium. However, a deviation of trader i gives rise to asymmetric information off equilibrium:[44] Following a demand change by trader i, traders $j \neq i$, who assume that all others including i play equilibrium, attribute the price realization induced by i's deviation to a different realization of signals instead. Consequently, all traders' demands off equilibrium are different: they are determined by optimization, given information. In contrast, in the competitive and the Cournot model, traders' $j \neq i$ demands off equilibrium are the same as in equilibrium because, respectively, the deviation of trader i is negligible in price or traders' strategies do not condition on price realizations but only expected price and are thus unaffected by the price change.

- When $I = 2$, a (linear) equilibrium with downward-sloping demands exists for any $\rho < 0$. Intuitively, with negatively correlated asset values, then, traders have opposite trading

[43] This is because price impacts are sufficient statistics for the weight on a trader's signal in the equilibrium price. When traders are heterogeneous, price aggregates their signals with heterogeneous weights:

$$p = (\sum_j (\alpha^j \Sigma + \Lambda^j)^{-1}(Id - C_p^j))^{-1}(\sum_j (\alpha^j \Sigma + \Lambda^j)^{-1}(C_\theta^j E[\theta^j] + C_s^j s^j - \alpha^j \Sigma q_0^j)), \tag{35}$$

where $E[\theta^i|s^i,p] = E[\theta^i] + Cov(\theta^i,(s^i,p))Var((s^i,p))^{-1}(s^i - E[s^i], p - E[p])'$ for all i. In symmetric markets (Example 1 below), equilibrium price is a function of the unweighted average of traders' signals (Equation (37)) with a coefficient that depends on price impact.

[44] With independent private values (i.e., $\rho_{ij} = 0$ for all i and $j \neq i$), such off-equilibrium asymmetric information does not affect traders' behavior.

needs. This can be seen in Equations (38)-(39): traders' demands are downward-sloping (i.e., $\frac{1}{\alpha+\lambda} > 0$) if and only if the inference coefficient $c_p < \gamma$. When trade is bilateral (i.e., $I = 2$ and $\gamma = 0$), this condition becomes $c_p < 0$, which requires $\rho < 0$, i.e., a trader's asset value is negatively correlated with his counterparty's value. See Equation (39).[45]

Example 1 illustrates the key effect of interdependent values: the fixed point between inference and price impact.

Example 1 (Inference and Equilibrium Behavior) Consider a double auction with price-contingent schedules for one asset ($K = 1$). In this example, for simplicity, we assume that the utility is symmetric, i.e., $\alpha^i\sigma = \alpha$ for all i, and that the information structure is equicommonal, i.e., $\frac{1}{I-1}\sum_{j\neq i}\rho_{ij} = \bar{\rho}$ for all i, for some $\bar{\rho} \in [-1, 1]$; this allows for symmetric equilibrium bid functions. In the symmetric linear Bayesian Nash Equilibrium, the demand schedules and market clearing price are given by

$$q^i(p) = \frac{1}{(\alpha^i + \lambda^i)}(E[\theta^i|s^i, p] - p - \alpha^i q_0^i) = \frac{\gamma - c_p}{1 - c_p}\frac{c_\theta}{\alpha}E[\theta^i] + \frac{\gamma - c_p}{1 - c_p}\frac{c_s}{\alpha}s^i - \frac{\gamma - c_p}{\alpha}p - \gamma q_0^i, \quad (36)$$

$$p^* = \frac{c_\theta}{1 - c_p}E[\theta^i] + \frac{c_s}{1 - c_p}\bar{s} - \frac{1}{1 - c_p}\alpha\bar{q}_0, \quad (37)$$

where $\gamma \equiv 1 - \frac{1}{I-1} \in [0, 1]$ is the index of market size and $\bar{s} \equiv \frac{1}{I}\sum_i s^i$ and $\bar{q}_0 \equiv \frac{1}{I}\sum_i q_0^i$ are the average signal and average endowment, respectively.

The price impact and price inference coefficients are:

$$\lambda^i = -(\sum_{j\neq i}\frac{\partial q^j(p)}{\partial p})^{-1} = (\underbrace{\sum_{j\neq i}\frac{1}{\alpha^j + \lambda^j}}_{\text{Direct effect}} - \underbrace{\sum_{j\neq i}\frac{c_p^j}{\alpha^j + \lambda^j}}_{\text{Inference effect}})^{-1} = \frac{(1 - \gamma)\alpha}{\gamma - c_p}, \quad (38)$$

where

$$c_p = \frac{(2 - \gamma)\bar{\rho}}{1 - \gamma + \bar{\rho}}\frac{\sigma^2}{1 - \bar{\rho} + \sigma^2}; \qquad c_s = \frac{1 - \bar{\rho}}{1 - \bar{\rho} + \sigma^2}. \quad (39)$$

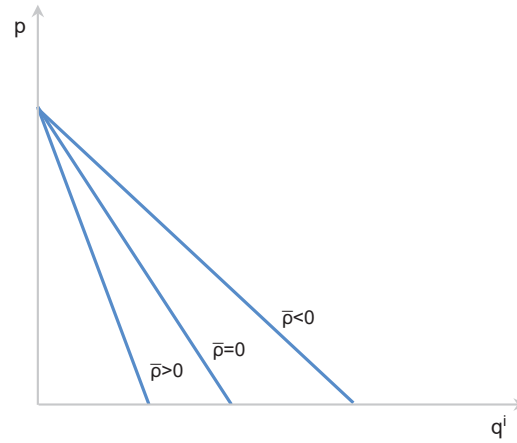
When values $\{\theta^i\}_i$ are independent ($\rho_{ij} = 0$ for all $j \neq i$, i.e., $c_p^i = 0$ for all i in Equation (39)), price impacts are independent of inference: price impact of trader i is determined by the concavity of utilities of the residual market (i.e., the direct effect in Equation (38)). When values are interdependent, however, price inference interacts with price impact: following a demand change by trader i, other traders who assume that all others including i play equilibrium, attribute the price realization to a different realization of the average signal

[45] Du and Zhu (2017b) points to a class of nonlinear *ex post* equilibria in divisible double auctions with two traders.

instead, and revise their posterior expectations about their θ^j (Equations (38) and (39) and Fig. 3).

When the inference effect is strong ($\bar{\rho}$ is large), the market becomes illiquid, i.e., price impact $\lambda^i \to \infty$ (Equation (38)). Due to this adverse selection effect, price cannot be too informative for equilibrium with trade to exist (see Appendix A.2). □

FIGURE 3: DEMAND REDUCTION AND PRICE INFERENCE



Notes: From Equation (38), inference from prices increases (decreases, has no impact on) price impact when the price inference coefficient $c_p > 0$ ($c_p < 0$, $c_p = 0$). By Equation (39), this occurs when traders' values are positively correlated (negatively correlated, uncorrelated) on average; i.e., $\bar{\rho} > 0$ ($\bar{\rho} < 0$, $\bar{\rho} = 0$). As a result, demands are less elastic than (more elastic than, as elastic as) they would be if values were independent. When $\bar{\rho} < 0$, demands are more elastic than with independent values.

As with private values, when traders have price impact, the aggregate risk portfolio q^* that gets allocated to traders in equilibrium differs from the (constrained, given private information) efficient portfolio except when all traders have symmetric payoffs.

$$q^i + q_0^i = (\alpha^i\Sigma + \Lambda^i)^{-1}(E[\theta^i|s^i, p] - \overline{v}) + (\alpha^i\Sigma + \Lambda^i)^{-1}Q + (\alpha^i\Sigma + \Lambda^i)^{-1}\Lambda^i, \tag{40}$$

where the aggregate valuation $\overline{v}$ and aggregate risk Q are:

$$\overline{v} = (\sum_j (\alpha^j\Sigma + \Lambda^j)^{-1})^{-1} \sum_j (\alpha^j\Sigma + \Lambda^j)^{-1} E[\theta^j|s^j, p], \tag{41}$$

$$Q = (\sum_j (\alpha^j\Sigma + \Lambda^j)^{-1})^{-1} \sum_j (\alpha^j\Sigma + \Lambda^j)^{-1}\alpha^j\Sigma q_0^j. \tag{42}$$

4.3.2 Information aggregation

A question explored by many researchers is whether price can aggregate traders' private information. Informational efficiency is commonly understood as a privately revealing price.

Definition 5 *The equilibrium price is* privately revealing *if, for each bidder i, the conditional cdfs of the posterior of θ^i satisfy $F(\theta^i|s^i, p^*) = F(\theta^i|\{s^j\}_j)$ for every state $\{s^j\}_j$, given the corresponding equilibrium price $p^* = p^*(\{s^j\}_j)$.*

A privately revealing price allows a Bayesian player i, who also observes his signal s^i, to learn about value θ^i as much as he would if he had access to all the information available in the market, $\{s^j\}_j$.

Information aggregation (equivalently, the *ex post* property) requires joint symmetry assumptions. For instance, in the market with symmetric traders and equicommonal information structure (Example 1), price is fully privately revealing if and only if $\rho_{ij} = \rho$ for all i and $j \neq i$ (the fundamental value model).[46]

In markets with $\rho_{ij} = \rho$ for all i and $j \neq i$, even if learning through market does not suffice for traders to learn their exact values, in markets such as those in which uncertainty is driven by fundamental shocks, traders learn all available information. One lesson from the "if" part of the result (and small-market literature) is that strategic behavior and the non-negligibility of individual signals in price in finite markets can be consistent with informational efficiency. The "only if" part qualifies the aggregation prediction by underscoring how (in)efficiency is affected by heterogeneity in interdependence among trader values plays for (in)efficiency.

4.3.3 Market power and price informativeness

In practical market design, the encouragement of market participation used to be seen as an enhancement to liquidity, inference, and efficiency. Early industrial organization and auction theory suggested that in markets with greater participation (1) price impact is lower; (2) traders learn more from prices through trading; and (3) price-taking behavior obtains in large markets.

Results (1) and (2) rely on the joint symmetry assumptions on risk preferences and information structure (more precisely, the fundamental value model). With heterogeneous traders, price informativeness may exhibit essentially arbitrary nonmonotone behavior. Informational *in*efficiency severs the link between market size and price informativeness and, in turn, welfare.[47]

[46] Vives (2011) and Rostek and Weretka (2012).

[47] Early studies of information aggregation focused on the information structure in which the values of the asset are determined by an underlying *common* shock (i.e., the fundamental value assumption). With heterogeneity in (income, endowment or preference) shocks among market participants, such as spatial dependence, price informativeness may not increase monotonically with market size and can be maximized for intermediate market size or markets with bilateral transactions.

- Strategic interactions among heterogeneous traders qualify the robustness of predictions regarding price-taking behavior in large markets and the ability of prices to aggregate information (in both small and large markets). Price-taking is predicted robustly in large markets and independent of details of information structure other than a uniform bound on the large-market commonality: $lim_{I\to\infty} sup[\bar{\rho}^I] \to 1$. On the other hand, even in large markets, aggregation of information is not generic and holds only in markets described by the fundamental value model.[48]

- Learning from prices (and, hence, information acquisition) need not improve welfare, given the market size: Improved estimation of agent values improves individual decisions, but may also increase market power. Policies that maximize learning from prices need not correspond to the maximization of welfare or liquidity. Policies that minimize market power need not maximize welfare.

- The demand submission game has offered a resolution of the Grossman and Stiglitz (1980) paradox through a partially revealing equilibrium (e.g., Kyle (1989), Vives (2011), and Kyle and Lee (2017)).

4.3.4 Heterogeneity in information structure and risk preferences

When traders are heterogeneous, the central difficulty involves dealing with the fixed point between heterogeneous traders' inference and price impacts. Apart from the papers that allow for the heterogeneity in the information structure mentioned in the previous section, the literature can now accommodate the heterogeneity in risk preferences — primitive and/or induced by the heterogeneity in the market structure — with independent private values (Malamud and Rostek (2017) and Babus and Kondor (2018). Finally, a couple of papers have taken the steps towards modeling markets that can both accommodate heterogeneity in traders' risk preference and information structure. Manzano and Vives (2017) consider a market with two groups of identical bidders whose bid functions are asymmetric. Combined with the symmetry of bidders within a group, the two-group assumption gives tractability by permitting fully revealing (*ex post*) equilibria. Inference from prices (more precisely, the fact that the players' values are interdependent) still matters. Glebkin (2019) examines a two-group model in which price is not fully revealing, thus allowing the informational inefficiency and price impact to interact; the assumption that traders in one of the groups are price-takers gives tractability. Rostek and Yoon (2018b) characterize equilibrium with heterogeneous risk preferences and general Gaussian information structures.

[48] Early literature (based on the competitive rational expectations equilibrium or double auctions) studied price-taking and information aggregation in large markets. The question of monotonicity of price impact in finite markets with interdependent preferences received attention more recently. Many financial institutions, including the CFTC have been concerned with how growth in market participation, which has increased trader diversity, affects the traditional roles of markets in price discovery and efficiency.

The classic model by Kyle (1989) had informed and uninformed traders. These papers can accommodate informed and uninformed traders, allow traders' strategies to differ, and allow for information structures richer than the fundamental value model.

5 Dynamic Markets

Dynamic models are tied directly to data — most markets are dynamic. Three themes have emerged from the body of research on dynamic trading. First, imperfect competition critically affects the equilibrium properties of prices and allocations relative to competitive consumption-trading models. In a seminal paper, Vayanos (1999) introduces the first infinite horizon model with large traders, private information, and no noise traders. Subsequent literature has identified additional implications of market power brought by nonstationarity of equilibrium, the second theme, and trader heterogeneity, a third theme.

We review the theory of dynamic trading in a unified framework that preserves the core structure of the models in the literature.[49] The framework allows us to relate the results for equilibrium models of strategic trading to the transaction-costs approach and to relate demand submission and Cournot games.

We first review models with symmetric traders, on which the literature has focused, applying recursive methods in stationary and nonstationary equilibrium. We then discuss the new effects brought forth by trader heterogeneity and how they change the structure of equilibrium and the solution methods. Among the modeling decisions that arise are (1) the choice of the relevant source of randomness (e.g., endowment or valuation shocks or exogenous supply shocks) and (2) the choice of trading cost (e.g., an exogenous transaction or inventory cost, or an endogenous transaction cost that derives from the concavity of the value function at consumption). We discuss how alternative assumptions affect the methods to characterize equilibrium and induce different dynamic properties — e.g., which among the popular assumptions to model imperfectly competitive dynamic markets induce the martingale or Markov property of prices.

5.1 Equilibrium

To fix ideas, there are T trading rounds $t = 1, \cdots, T$ after which asset payment occurs. While we focus on the dynamic-trading results, we also touch on the consumption-trading framework for imperfectly competitive markets. That is, to analyze jointly the (nonstationary) trading problem and the (stationary) consumption problem, we allow trade to be more frequent than asset payments by embedding the T-round trading problem in the standard infinite-horizon

[49] Including Vayanos (1999), Du and Zhu (2017a), Rostek and Weretka (2015a), Sannikov and Skrzypacz (2016), Duffie and Zhu (2017), Antil and Duffie (2017), Kyle, Obizhaeva, and Wang (2018), Ollar, Rostek, and Yoon (2017), and Rostek and Yoon (2019).

model with asset payments (and consumption) occurring every T rounds. This permits a direct mapping of the infinite-horizon competitive consumption-trading model.

Before a trade in each round t a trader receives a privately known endowment shock $\delta_t^i \in \mathbb{R}^K$. Endowment shocks $(\delta_{k,t}^i)_i \in \mathbb{R}^I$ are jointly Normally distributed $\mathcal{N}(0,\Omega)$ and are independent across assets k and time t.[50] The quantity traded by trader i in round t is $q_t^i \in \mathbb{R}^K$. Payoffs (1) are realized after T rounds. Trader i's payoff equals the utility from final allocation $q_T^{i,0} + q_T^i$ net of the trading costs:

$$\underbrace{\theta^i \cdot (q_T^{i,0} + q_T^i) \; - \; \frac{\alpha^i}{2}(q_T^{i,0} + q_T^i) \cdot \Sigma(q_T^{i,0} + q_T^i)}_{\text{Utility from allocation after the last round } T} \quad - \quad \underbrace{\sum_{t=1}^{T}(p_t \cdot q_t^i + c(q_t^i))}_{\text{Trading costs in rounds } t=1,\cdots,T}, \tag{43}$$

where $q_t^{i,0}$ is the post-trade allocation (inventory) in the beginning of round t.[51] The trading costs include the payment $p_t \cdot q_t^i$ for the (net) purchase of q_t^i units of risky assets in each round $t = 1, \cdots, T$, as well as a trading cost with the concavity parameter κ:[52]

$$c(q_t^i) = \frac{\kappa}{2} q_t^i \cdot \Sigma q_t^i. \tag{44}$$

In round t, a trader i submits a demand schedule $q_t^i(\cdot) : \mathbb{R}^K \to \mathbb{R}^K$ as a function of current price p_t that maximizes his expected payoff (43) conditional on private information $\mathcal{I}_t^i$:

$$\max_{q_t^i(\cdot):\mathbb{R}^K \to \mathbb{R}^K} E[\theta^i \cdot (q_T^{i,0} + q_T^i) - \frac{\alpha^i}{2}(q_T^{i,0} + q_T^i) \cdot \Sigma(q_T^{i,0} + q_T^i) - \sum_{\tau=1}^{T}(p_\tau \cdot q_\tau^i + c(q_\tau^i))|\mathcal{I}_t^i].$$

To indicate the key properties of the value function, we assume symmetric risk preference $\alpha^i = \alpha$, θ^i is certain, and information $\mathcal{I}_t^i = \{\{\delta_s^i\}_{s \le t}, \{p_s\}_{s<t}\}$ for all i.

[50] In demand submission games, uncertainty in a residual market in all rounds is necessary for equilibrium existence. Endowment shocks $\{\delta_t^i\}_i$ (or, alternatively, an exogenous supply shock H_t) in each round t ensure that the residual supply function (equivalently, p_t) is uncertain to trader i.

[51] $q_t^{i,0} \equiv \sum_{s=1}^{t-1}(\delta_s^i + q_s^i) + \delta_t^i$, for all $t \le T$ and all i. The first-round endowment shock is the initial endowment $q_0^i \equiv \delta_1^i$.

[52] Thus, the costs comprise execution costs and opportunity costs — a distinction commonly made in computing the implementation shortfall. Introducing the additional concavity κ to the payoff function (value function) at each round t guarantees equilibrium existence.

5.1.1 Value function

The value function at t is defined as:[53]

$$V_t^i(q_t^i;\mathcal{I}_t^i,p_t) \equiv E[\theta^i\cdot(q_t^{i,0}+q_t^i+\sum_{\tau>t}^{T}q_\tau^i(p_\tau))-\frac{\alpha}{2}(q_t^{i,0}+q_t^i+\sum_{\tau>t}^{T}q_\tau^i(p_\tau))\cdot\Sigma(q_t^{i,0}+q_t^i+\sum_{\tau>t}^{T}q_\tau^i(p_\tau)$$
$$-\sum_{\tau>t}^{T}(p_\tau\cdot q_\tau^i(p_\tau)+c(q_\tau^i(p_\tau)))|\mathcal{I}_t^i,p_t], \tag{45}$$

given schedules $\{q_\tau^i(\cdot)\}_{\tau>t}$. The best response $q_t^i(\cdot)$ at t is determined by the *pointwise* optimization as in the static market (Section 4.3).[54] The first-order condition of $q_t^i(p_t)$ is:

$$\frac{dV_t^i}{dq_t^i}\equiv E[v_t^i|\mathcal{I}_t^i,p_t]-\alpha_t\Sigma(q_t^{i,0}+q_t^i)=p_t+\Lambda_t q_t^i+\kappa\Sigma q_t^i \qquad \forall p_t\in\mathbb{R}^K, \tag{46}$$

where the expected asset valuation is characterized by $E[v_t^i|\mathcal{I}_t^i,p_t]=C_{\theta,t}\theta^i+C_{s,t}\mathcal{I}_t^i+C_{p,t}p_t$.[55] In round T, $v_T^i=\theta^i$ and $\alpha_T=\alpha$.

A key observation here is that the intercept $v_t^i\equiv\frac{dV_t^i}{dq_t^i}$ and slope $\alpha_t\Sigma\equiv\frac{d^2V_t^i}{(dq_t^i)^2}$ of the marginal value function are both endogenous for $t<T$, and depend on the price impact in future rounds $s>t$. Given the link between concavity and price impact (Equation (52) below), one thus expects the liquidity to be affected by the number of trading opportunities T. The concavity parameter represents *effective risk aversion* (sometimes called "risk bearing capacity"):

$$\alpha_t=\underbrace{(\prod_{s>t}^{T}(1-\gamma_s)^2)\alpha}_{\text{"Discounted" risk aversion }\alpha}+\underbrace{\sum_{\tau>t}^{T}\gamma_\tau^2\prod_{s>t}^{\tau-1}(1-\gamma_s)^2\kappa}_{\text{"Discounted" future cost convexity }\kappa}, \tag{47}$$

where $1-\gamma_s$ represents the (net) demand reduction and $(1-\gamma_s)^2\equiv(\frac{\kappa+\lambda_s}{\alpha_s+\kappa+\lambda_s})^2$ acts as the endogenous discount factor in round s. Analogously to the static market, γ_t is derived, given the current-round effective risk aversion α_t and price impact $\Lambda_t=\lambda_t\Sigma$.[56] Dynamic trading also alters a trader's asset valuation v_t^i. Except in the last round T, v_t^i is not equal to the primitive asset return θ^i, as it would be with price-taking traders, but is instead a convex

[53] The mean-zero independent endowment shocks does not affect the value function. Therefore, we suppress notation for endowment shocks $\{\delta_\tau^i\}_{\tau>t}$.

[54] Even when traders know their values θ^i as in Section 4.1.1, uncertainty about future prices is what matters for the effect of dynamic inference on the value function.

[55] The pojection theorem determines inference coefficients $C_{\theta,t}\in\mathbb{R}^{K\times K}$ on the primitive asset value θ^i, $C_{s,t}\in\mathbb{R}^{K\times|\mathcal{I}_t^i|}$ on (private and public) history $\mathcal{I}_t^i$, and $C_{p,t}$ on current price p_t.

[56] Price impacts Λ_t^i are proportional to the asset payoff covariance Σ in the dynamic centralized market. Furthermore, as in Section 4.3, the inference coefficients $C_{p,t}^i$ is proportional to the identity matrix. For simplicity, we denote $\Lambda_t^i=\lambda_t^i\Sigma$ and $C_{p,t}^i=c_{p,t}^i Id$.

combination of θ^i and future prices $\{p_\tau\}_{\tau>t}$

$$v_t^i = \prod_{s>t}^{T}((1-\gamma_s)^2 + \gamma_s^2\frac{\kappa}{\alpha_s})(\theta^i - p_T) + \sum_{\tau>t}^{T-1}\prod_{s>t}^{\tau}((1-\gamma_s)^2 + \gamma_s^2\frac{\kappa}{\alpha_s})(p_{\tau+1} - p_\tau) + p_{t+1}. \quad (48)$$

The more competitive the market (i.e., the larger $\{\gamma_s\}_{s>t}$ are) and the more trading opportunities (T) present, the more the trader's valuation v_t^i depends on the price path $\{p_\tau\}_{\tau>t}$ relative to the fundamental asset value θ^i.[57] When traders have price impact, since only a fraction of gains from trade is realized in any given round ($\gamma_s < 1$), a trader's valuation of the asset at t depends on future prices: a trader's valuation of the asset v_t^i is endogenous.[58]

5.1.2 Prices and allocations

Given the endogenous valuation v_t^i and concavity α_t, the equilibrium price and allocations in round t can be decomposed analogously into their static counterparts. With symmetric traders, the equilibrium price in round t is:

$$p_t = \overline{v}_t - Q_t \quad (49)$$

where

$$\overline{v}_t \equiv \frac{1}{I}\sum_j E[v_t^j|\mathcal{I}_t^j, p_t] \quad \text{and} \quad Q_t \equiv \alpha_t\Sigma\frac{1}{I}\sum_j q_t^{j,0}.$$

Trader i's post-trade allocation at t can be decomposed into a convex combination of an aggregate risk portfolio and initial portfolio, and a component due to heterogeneous valuations:

$$q_t^{i,0} + q_t^i = \gamma_t \underbrace{\frac{1}{\alpha_t}\Sigma^{-1}(E[v_t^i|\mathcal{I}_t^j, p_t] - \overline{v}_t)}_{\text{Valuation adjustment}} + \gamma_t \underbrace{\frac{1}{\alpha_t}\Sigma^{-1}Q_t}_{\text{Aggregate risk portfolio}} + (1-\gamma_t)\underbrace{q_t^{i,0}}_{\text{Initial portfolio}} \quad (50)$$

In the competitive model (i.e., $\gamma_t = 1$) with no exogenous or strategic trading costs (i.e., $\kappa = 0$, $\Lambda_t = 0$), the valuation adjustment is zero in all rounds $t < T$, even when the fundamental asset values are not common.[59] Writing the allocation (50) as

$$q_t^{i,0} + q_t^i = \gamma_t \underbrace{q_t^{i,*}}_{\text{Target portfolio}} + (1-\gamma_t)\underbrace{q_t^{i,0}}_{\text{Initial portfolio}}, \quad (51)$$

[57] More precisely, the more a trader's valuation depends on the fundamental value through prices rather than directly, the more it depends on the aggregate risks $\{Q_s\}_{s>t}$ in prices.

[58] Furthermore, when traders are heterogenous, the traders' valuation of the asset at t differs across traders because of the asymmetry of $\{\gamma_s^i\}_i$, even when the fundamental asset value $\theta^i = \theta$ is common for all traders.

[59] The value adjustment is non-zero if either $\{\theta^i\}_i$ or $\{\alpha^i\}_i$ differ (so that the traders' relative exposure to future aggregate risks $\{Q_s\}_{s>t}$ differs across rounds).

trader i's allocation is as if he traded towards a *target portfolio*:

$$q_t^{i,*} \equiv \frac{1}{\alpha_t}\Sigma^{-1}(E[v_t^i|\mathcal{I}_t^j, p_t] - \overline{v}_t + Q_t),$$

the quantity that i would buy if his *own* price impact at t were zero (i.e., $\gamma_t = 1$). Gârleanu and Pedersen (2013) introduce the notion of a "target portfolio" in the context of a single-agent trading problem with exogenous transaction costs. The notion of target portfolio — which differs across traders and time (because asset valuation $E[v_t^i|\mathcal{I}_t^i, p_t]$ and aggregate risk Q_t differ) — appears endogenously from the equilibrium fixed point in decomposition (51). In the symmetric market and the competitive market, the target portfolio and efficient risk portfolio coincide.

In response to their price impact, traders reduce their demands and supplies relative to the demand they would submit as price-takers. That is, in a dynamic market with imperfect competition, traders trade slowly; they break up their orders into blocks placed sequentially on the market. The inefficiency associated with slow trading is central to the difference between equilibrium properties in dynamic imperfectly competitive and competitive markets. Section 5.2 reviews how the equilibrium structure of slow trading and the associated inefficiency give rise to endogenous limits to arbitrage (with price impact, profits from buy-sell orders are finite) and the non-neutrality of several instruments (e.g., trading frequency, public information releases) for welfare.

5.1.3 Price impact in dynamic markets

In equilibrium, each trader's price impact at t must match the slope of his residual inverse supply function; this is characterized by the condition familiar in the static model: the harmonic mean of inverse demand slopes of traders $j \neq i$,

$$\Lambda_t^i = -(\sum_{j\neq i}\frac{\partial q_t^j(\cdot)}{\partial p_t})^{-1} = (\sum_{j\neq i}((\alpha_t^j + \kappa)\Sigma + \Lambda_t^j)^{-1}(Id - C_{p,t}^j))^{-1}. \tag{52}$$

- In a dynamic imperfectly competitive market, the inference coefficient $C_{p,t}$ is present at $t < T$ even when equilibrium is *ex post*, because asset valuation v_t^i depends on future prices, which are correlated with the current-round price (Equation (48)).

- Dynamic models provide important insights into the determinants of market power: Market power in a static market (or at T) is due solely to the traders' decreasing marginal utility. The imperfect competitiveness of dynamic markets is, in turn, due to a dynamic mechanism: *Future market thinness induces current market thinness.* Suppose $\kappa = 0$; if (and only if) the market at t is competitive, so too would be the

market in any round $s < t$.[60]

- Price impact is bounded away from zero in all rounds, even as the number of trading rounds between asset payments grows to infinity ($T \to \infty$) and even without private information. Due to demand reduction in future rounds, traders are effectively risk averse in any round (Equation (47)).

- In markets in which trade is more frequent than payments (i.e., $T > 1$), equilibrium is not stationary and price impact is generally not constant but can be essentially arbitrary and nonmonotone across trading rounds $1, \cdots, T$.

  The dynamics of equilibrium price impact results from two countervailing effects: With symmetric traders:

  (i) The effective risk aversion increases between rounds 1 and T: The further away from the payment after round T, the more opportunities to diversify and re-trade, the less costly it is for the investors to depart from their current holdings. The value function makes this mechanism apparent as it becomes more concave over time, i.e., traders are effectively more risk averse and less willing to trade risky assets at given price concessions (Equation (47)). The dynamics of price impact implies the dynamics of trade: the investors choose to trade more when their price impacts are smallest (Fig. 4). The effective risk aversion introduces an exponential component into price impact.

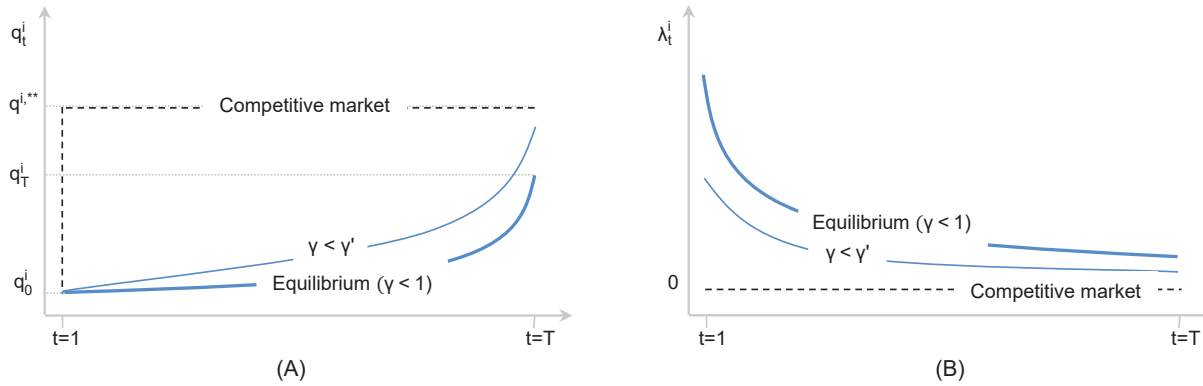
  (ii) $C_{p,t}$ decreases between rounds 1 and T: In earlier rounds, a trader's asset valuation v_t^i depends more on aggregate risks $\{Q_\tau\}_{\tau>t}$ in future prices $\{p_\tau\}_{\tau>t}$ (Equation (48)). The stronger the dependence on aggregate risks, the stronger the interdependence among $\{v_t^i\}_i$. Thus, holding fixed the effective risk aversions, price impacts are larger in earlier rounds (Fig. 4B).

- The model with time-varying price impact (and hence demand elasticity) gives rise to the representation of a market as a sequence of short- and long-run demands, as used in empirical studies (e.g., Greenwood (2005)).

[60] Suppose that traders had no price impact at t. Then, the traders' endogenous price impact at $s < t$ would be zero. Knowing that at t they can trade arbitrary amounts without price concessions, traders would become effectively risk neutral at $s < t$ (i.e., $\alpha_s \to 0$) and thus willing to trade arbitrary amounts at the s-round price. The marginal value function at $s < t$, would be perfectly elastic and equal to the s-round price. (See ft. 66.)

[61] With no discounting between rounds, traders are indifferent in which round to trade in the competitive market. With an arbitrarily small discount factor, all trade will take place in the first round.

FIGURE 4: DYNAMIC ALLOCATION AND PRICE IMPACT



Notes: In markets with more frequent trade than asset payments (i.e., in a nonstationary equilibrium), by Equation (52), price impact decreases over time when the effect of decreasing $c_{p,t}$ (Equation (48)) dominates the effect of increasing α_t^i (Equation (47)). Trade q_t^i is larger in rounds with lower price impact. In the competitive equilibrium, price impact is zero ($\lambda_t^i = 0$ in all rounds t) and all gains from trade can be realized in one round.[61]

5.2 Classical Results Revisited: Dynamic Markets

With heterogeneous traders, the counterpart of Equation (49) is:

$$p_t = \overline{v}_t - Q_t \tag{53}$$

where the aggregate valuation and aggregate risk are:

$$\overline{v}_t \equiv (\sum_j \frac{\gamma_t^j}{\alpha_t^j})^{-1} \sum_j \frac{\gamma_t^j}{\alpha_t^j} E[v_t^j | \mathcal{I}_t^j, p_t] \quad \text{and} \quad Q_t \equiv (\sum_j \frac{\gamma_t^j}{\alpha_t^j})^{-1} \sum_j \gamma_t^j \Sigma q_t^{j,0} \tag{54}$$

(see Example 4 in Appendix B for details). The decomposition of trader i's post-trade allocation at t becomes:

$$q_t^{i,0} + q_t^i = \gamma_t^i \underbrace{\frac{1}{\alpha_t^i}\Sigma^{-1}(E[v_t^i|\mathcal{I}_t^i, p_t] - \overline{v}_t)}_{\text{Valuation adjustment}} + \gamma_t^i \underbrace{\frac{1}{\alpha_t^i}\Sigma^{-1}Q_t}_{\text{Aggregate risk portfolio}} + (1-\gamma_t^i)\underbrace{q_t^{i,0}}_{\text{Initial portfolio}}. \tag{55}$$

The asymmetric value function parameters $\{v_t^i, \alpha_t^i\}_{i,t}$ are derived by the steps in Example 4.

5.2.1 Allocations

- *Slow trading.*

In the competitive market (i.e., $\gamma_t^i \to 1$ for all t), one trading round suffices to realize all gains

from trade:[62]

$$q_t^{i,0} + q_t^i \;\to\; \frac{1}{\alpha^i}\Big(\sum_j \frac{1}{\alpha^j}\Big)^{-1} \sum_j q_t^{j,0} \qquad \text{as} \quad I \to \infty.$$

When traders have price impact ($\Lambda_t^i > 0$ for all i and t), the realization of gains from trade requires *multiple* rounds of trading, even if no new gains are generated by endowment shocks (other than the initial source of the gains form trade) and no information about the asset value is revealed between rounds.[63] With non-zero price impact, the equilibrium trading strategy is uniquely pinned down without introducing any preference for urgency: traders trade whenever they have the opportunity.

- *Imperfect competition changes the efficiency properties of equilibrium.*

Pursuant to the First Welfare theorem, equilibrium allocation can attain Pareto efficiency in one round of trading in a competitive market. All traders hold the aggregate portfolio — idiosyncratic risk is fully diversified irrespective of the initial distribution of endowments among the traders. With imperfect competition, the post-trade allocation is inefficient in all rounds. From the first-order condition (46), in every round, a trader i closes only a fraction $\gamma_t^i \equiv \frac{\alpha_t^i}{\alpha_t^i+\kappa+\lambda_t^i} < 1$ of the difference between his endowment $q_t^{i,0}$ and the target portfolio $q_t^{i,*} = \frac{1}{\alpha_t^i}\Sigma^{-1}(E[v_t^i|\mathcal{I}_t^i, p_t] - \overline{v}_t + Q_t)$ (see Equation (55)). Thus, in imperfectly competitive markets, traders generally do not hold an efficient ("market") portfolio in any round.

Moreover, unless the risk preferences $\{\alpha^i\}_i$ of all traders are the same, the aggregate (target) portfolio $q_t^{i,*}$, which traders trade towards at t, differs from the efficient portfolio $q_t^{i,**}$:[64]

$$q_t^{i,*} = \frac{1}{\alpha_t^i}\Sigma^{-1}(E[v_t^i|\mathcal{I}_t^i, p_t] - \overline{v}_t) + \frac{1}{\alpha_t^i}\Big(\sum_j \frac{\gamma_t^j}{\alpha_t^j}\Big)^{-1} \sum_j \gamma_t^j q_t^{j,0} \quad \neq \quad \frac{1}{\alpha^i}\Big(\sum_j \frac{1}{\alpha^j}\Big)^{-1} \sum_j q_t^{j,0}.$$

The discrepancy between the efficient and aggregate portfolios is due to heterogeneous demand reduction.[65]

- *Changes in the frequency of trading relative to the frequency of asset payment or consumption are not neutral for outcomes.*

[62] Unless an update of a trader's expected valuation $E[v_t^i|\mathcal{I}_t^i, p_t]$ renews the gains from trade.

[63] The result might seem reminiscent of — but differs from — the Coase conjecture type of results for the classical durable-good monopoly. With bilateral (buyer and seller) price impact, slow trading ("delay in trade") is optimal, even in the absence of discounting or heterogeneous beliefs about fundamentals and even when the equilibrium price is constant (more generally, deterministic).

[64] Efficient portfolio $q_t^{i,**}$ is subject to past and current endowment shocks: $q_t^{i,**} = \frac{1}{\alpha^i}\big(\sum_j \frac{1}{\alpha^j}\big)^{-1} \sum_j q_t^{j,0} = \frac{1}{\alpha^i}\big(\sum_j \frac{1}{\alpha^j}\big)^{-1} \sum_j (q_0^j + \sum_{s>1}^t \delta_s^j)$ for i and t.

[65] This can be seen in equilibrium models with heterogeneous traders and the single-agent models of Gârleanu and Pedersen (2013) with an anticipated price process.

The assumption that both trade and asset payments occur in each round, as in the standard competitive setup, allows for stationary price impact and trading. The inefficiency due to slow trading motivates the study of separation of trading frequency ($t = 1, \cdots, T$) relative to the frequency of asset payments (T), which would be neutral with price-taking agents. When traders have price impact, the unequal frequency gives rise to nonstationary equilibrium dynamics of prices and liquidity, consumption-savings behavior, and welfare. Likewise, the relative frequency of *consumption* to trading matters, because it affects the concavity of the value function, as well.[66] See Appendix B.

Models that separate these frequencies depart from stationarity at the trading frequency while still allowing for stationarity at the payment (consumption) frequency.[67] Equilibrium coincides with that in the standard competitive infinite-horizon consumption-trade framework if traders are price-takers.

Such models have made a case for the potential use of new — relative to competitive markets — instruments to affect consumption and welfare. In particular, the trading frequency, decisions regarding how to disclose public information about fundamentals (e.g., asset return), and the choice of which statistics of past outcomes should be available to traders all determine traders' price impact. The link between price impact and concavity (Equation (52)) is key to understanding the implications of imperfect competition in dynamic markets. Given the inefficiency (slow trading), welfare is not neutral to instruments that affect that link. The non-neutrality applies even when there is no discounting and no shocks that renew the gains from trade after the first round.

In imperfectly competitive markets, the instruments that affect price impact affect the endogenous concavity $\overline{\alpha}^i$ of value function $V^i_{lT}(\cdot)$ in consumption at $lT = T, 2T, \cdots$. Due to slow trading, a trader's value function is more concave than it would be if the market were competitive. Letting r be the return of the riskless asset in consumption rounds, the effective risk aversion is:

$$\bar{\alpha}^i = \xi^i \frac{\alpha^i + r\kappa}{r^2} + \frac{\alpha^i}{r},$$

where the coefficient $\xi^i = \xi(\{\gamma^i_t\}_t, \kappa, T) \in [0, 1]$ can be interpreted as the speed of trading in a

[66] In an imperfectly competitive market with symmetric traders, the last-round allocation

$$q^{i,0}_T + q^i_T = \sum_{t=1}^{T}(\prod_{s>t}^{T}(1-\gamma_s))\frac{\gamma_t}{\alpha_t}\Sigma^{-1}(E[v^i_t|\mathcal{I}^i_t, p_t] - \overline{v}_t + Q_t) + (\prod_{s=1}^{T}(1-\gamma_s))q^i_0 + \sum_{t=1}^{T}\prod_{s=t}^{T}(1-\gamma_s)\delta^i_t \qquad \forall i,$$

depends on T because $\gamma_s = \frac{\alpha_s}{\alpha_s+\kappa+\lambda_s} < 1$ for all s. When traders have no price impact, they can attain the efficient allocation within one round ($q^{i,0}_T + q^i_T = \frac{1}{\alpha}\Sigma^{-1}Q_T = \frac{1}{I}\sum_j q^{j,0}_T$); thus, the trading frequency T is neutral.

[67] Vayanos (1999) and Du and Zhu (2017a) allow the frequencies of trading (equal to that of consumption) and asset payment to differ while maintaining the stationarity of equilibrium. Rostek and Weretka (2015a), Duffie and Zhu (2017), Antil and Duffie (2017), and Rostek and Yoon (2019) separate the frequencies of trading and consumption (equal to that of asset payment).

T-round market. With $\alpha^i = \alpha$ for all i, effective risk aversion is increasing in price impact (ξ^i is decreasing in γ_t^i and T). The trading speed ξ^i decreases to zero as $T \to \infty$ in an imperfectly competitive market, and $\xi^i = 0$ in the competitive market for any T. (See Appendix B for details.)

Budish, Cramton, and Shim (2015) advocate for frequent batch (discrete) uniform-price auctions to improve upon the continuous limit-order book design. The uniform-price design, the authors argue, transforms competition on speed into competition on price. Imperfectly competitive models complement the insights of that paper, in which traders have no price impact. The seminal paper on dynamic imperfectly competitive trading by Vayanos (1999) demonstrates that when endowment information is private (and independent across traders), a higher frequency may decrease liquidity and lower welfare. A higher frequency lowers the amount of information per trading round, thus increasing adverse selection and price impact. In their study of optimal trading frequencies, Du and Zhu (2017a) point out that frequent trading induces a trade-off: it allows for more immediate asset reallocation after new information arrives at the cost of a lower volume of beneficial trades in each double auction. The authors show that with the arrival of private (interdependent) information, a low trading frequency is optimal for scheduled, but not continuous, arrival. With symmetric information disclosures (Rostek and Weretka (2015a)), allowing for a higher trading frequency increases welfare. Kyle and Lee (2017) argue that, if feasible to implement, a continuous time exchange would eliminate the need to break up orders, thus saving time and resources of large investors.

Disclosing asset-return information at t lowers Σ, and hence price impact in all rounds $s \leq t$. In contrast, the effects of changes in transparency — observing signals about *private* information (e.g., endowment $q_t^{i,0}$ or asset value θ^i) — on price impact and welfare depend on the information structure and which statistics of past outcomes become available. This is due to the fixed point between inference coefficients and price impact (see Section 4.3).

5.2.2 Prices

- *Equilibrium prices in imperfectly competitive markets do not generally satisfy the martingale property.*

With price-making traders, the martingale property of equilibrium prices does not hold when traders have heterogeneous risk preferences ($\alpha^i \neq \alpha^j$ for some $j \neq i$) or heterogeneous time-varying transaction costs. The heterogeneity across traders gives asymmetric weights $\{\frac{\gamma_t^i}{\alpha_t^i}\}_i$ in the average asset valuation $\overline{v}_t$ and aggregate risk Q_t in price formula (53). In nonstationary equilibria, the time-varying weights induce deterministic changes in expected prices ($E[p_t] \neq E[p_\tau]$ for any $\tau \neq t$).

In addition, even with symmetric traders, prices are not martingales with supply shocks

$\{H_t\}_t$, because each-round price p_t is a noised signal of fundamentals $\overline{v}_t - Q_t$:

$$p_t = \overline{v}_t - Q_t - (\sum_i \frac{\gamma_t^i}{\alpha_t^i})^{-1} \Sigma H_t \qquad \forall t. \tag{56}$$

The noise $(\sum_j \frac{\gamma_t^j}{\alpha_t^j})^{-1} \Sigma H_t$ induces an endogenously time-varying variance, and thus, the price distribution (i.e., both its mean and variance) changes over time. In the competitive model with a stationary distribution of shocks, the martingale property holds.

- *Aggregate risk (risk premium) depends on trading frequency, releases of public information, and inference about private information over time.*

In Equation (54), the aggregate risk Q_t depends on the effective risk aversion $\{\alpha_t^i\}_t$ and price impacts $\{\Lambda_t^i\}_t$, which are endogenously determined by trading frequency, releases of public information, and inference about private information over time. (See Equation (53).) If $I \to \infty$, then prices (and outcomes) are subject to systematic risk.

- *Imperfect competition changes the arbitrage properties of equilibrium, compared to the competitive model.*

In the competitive model, shocks have only a negligible effect on prices; if traders were price-takers, anticipated price differentials would create infinite profit opportunities. With non-price-taking behavior, the argument behind no-arbitrage differs in two ways from that in the competitive model. First, profits from arbitrage are *bounded* for any round-trip trade — large enough buying and selling positions would change prices and yield strictly negative profits. Second, an additional order (e.g., a round-trip) placed by a trader who participates in a market has an externality on other units held by the trader; thus, when traders have price impact, the incentives to arbitrage differ between the participating traders and outside investors.

If all the positions of the participating traders are held in the market, these traders do not have any incentive to arbitrage anticipated price differentials since they optimize. For outside investors, profit opportunities from arbitraging anticipated price differentials (such as those associated with an anticipated shock H_τ in future price p_τ in Equation (56) for $\tau > t$) are only *finite* in an imperfectly competitive market. Taking an unbounded position or placing a round-trip order involving a purchase at τ of more shares than the shock H_τ sold in the next round yields a strictly negative profit. Hence, unlike the competitive model, sufficient fixed-entry costs can discourage outside investors from arbitraging the liquidity effect. In practice, entry costs include not only explicit trading costs, but also costs associated with learning and monitoring the characteristics of particular stocks.[68]

[68] Mitchell, Pedersen and Pulvino (2007) examine price behavior in the convertible bond market in 2005

If the participating traders hold positions outside the market, the possibility to influence prices in an imperfectly competitive market creates an incentive to manipulate prices in order to profit from the prices of correlated assets. This incentive has been at the core of the LIBOR scandal and is a problem for designing a robust benchmark.

5.2.3 Methodology

- *Imperfect competition changes the scope for recursive analysis.*

With the standard strategies (demands contingent on current prices and conditioned on past outcomes), equilibrium in imperfectly competitive markets is not recursive when traders are heterogeneous and $T < \infty$ — the state variable is *(endogenously) non-Markovian* for any state variables that do not fully disclose information after every round. As traders learn over time, the entire history matters and cannot generally be summarized by the previous round's outcome.

With rich enough strategies (i.e., demands contingent on past demand functions rather than current prices), one can characterize equilibrium recursively.[69] However, such strategies would require the player and the modeler to compute the joint distribution of all possible future prices, which is not necessary for demand at t, only conditioning on actual realizations is necessary.

A non-recursive equilibrium draws a distinction between the agent's ability to make decisions based on past outcomes and the ability to assess the impact of current behavior on future outcomes. It also allows analysis of the effects of changes in transparency (i.e., making available new statistics of past outcomes). In markets with symmetric traders (more generally, in a stationary equilibrium), equilibrium is invariant to changes in transparency — prices are fully revealing and thus learning is static. When traders are heterogeneous, availability of new statistics of past outcomes, and more generally the design of conditioning statistics, affect the traders' price impact. Furthermore, changes in transparency alter the composition

and surrounding the collapse of LTCM in 1998, and merger targets in the 1987 market crash. During these events, "natural" liquidity providers were forced to liquidate their holdings, which depressed prices below the fundamental values, despite the fact there was little change in the overall fundamentals. In the convertible bond markets, the prices deviated from the fundamental values, reaching the maximum discount of 2.7% in 2005 (2.5 standard deviations from the historical average) and 4% in 1998 (4 standard deviations from the average). During the 1987 crash, the median merger-arbitrage deal spreads increased to 15.1%. In all episodes, it took several months for traders to increase their capital or for better-capitalized traders to enter. The authors attribute the slow entry to information barriers and the costs of maintaining dormant financial and human capital in a state of readiness when arbitrage opportunities arise.

[69] Suppose that in each round t trader i submits a (net) demand that is a complete contingent plan, i.e., specifies demand at t for all of his possible past demand functions in addition to past price realizations, instead of submitting a (net) demand contingent on the current- and past-round price. Characterizing equilibrium with contingent plans involves a variant of the curse of dimensionality. First, contingent plans (recursive updating) require the posterior distribution $F^i_\tau(\theta^i, p_{\tau+1}, \cdots, p_T | \mathcal{I}^i_t, p_t, \{p_s\}_{t<s\leq\tau})$ for all $\tau > t$ at t to be computed for *all possible future prices* $\{p_s\}_{t<s\leq\tau}$ *rather than actual realizations.* Second, the dimension of input variables in contingent plans (i.e., past schedules) is large and increases over time.

of risk to which traders are exposed in equilibrium as well as the distribution of risk among the traders.[70]

As we noted in previous sections, in demand submission games, all price realizations are in the equilibrium support of the price distribution. Consequently, like in static markets, Bayes' rule and optimization determine beliefs in and off equilibrium, as well as traders' price impact, which completely describes how off equilibrium affects equilibrium behavior in a dynamic game.

5.2.4 Impact of news and supply shocks

It is routine for price-impact-estimation software in the financial industry to distinguish between "permanent" and "temporary" price-impact effects of trades.[71] Price behavior in financial markets has been interpreted as a temporary departure of prices from their fundamental values as a reaction to market events that correspond to exogenous shocks in supply (or demand). Such behavior has long been documented by event studies.[72]

Notably, even when supply shocks have been publicly preannounced, the temporary price changes relative to the long-run level still occur and take place on the actual event date, not the announcement date.[73] Thus, predictable changes in market outcomes are observed.

[70]The effects of increasing the number of traders on price informativeness, price impact, and welfare are non-monotone (Rostek and Weretka (2012, 2015b)), as are the effects on inference coefficients for price and one's own signal in traders' posterior expectations (Rostek and Yoon (2018b)). The effects of changes in transparency depend on the market.

[71]The program for price impact estimation implemented by Citigroup separately estimates a permanent price-impact component ("reflects the information transmitted to the market by the buy/sell imbalance"), which is believed to be roughly independent of trade scheduling, and a temporary price-impact component ("reflects the price concession needed to attract counterparts within a specified short time interval"), which is sensitive to trade scheduling (Almgren et al. (2005)). "The temporary impact affects only the execution price and has no effect on the 'fair value' or fundamental price. In contrast, the permanent impact directly affects the fair value of the security while having no direct effect on the execution price. Thus we can think of the temporary impact as connected to the liquidity cost faced by the agent ..." (Li and Almgren (2011), p.2). "The temporary impact component of cost is interpreted as the additional premium that must be paid for execution in a finite time above a suitably prorated fraction of the permanent cost" (Almgren (2009), p. 1).

[72]The transitory price effects have been known in the academic literature as early as Kraus and Stoll (1972). Numerous studies for various securities examined (net) supply shocks, such as stock inclusions into or deletions from market indices, index weight changes, forced liquidations, issuance of new debt, selling initial public offerings (IPOs), and fire sales. For example, weight changes are administrative decisions and data on ownership used for reweighting is publicly available before events. Hence, an index weight change induces a demand shock that is not associated with new information about the fundamental value of an asset.

[73]Events such as preannounced weight changes in stock market indices facilitate control of the informational component of the price change. Lou, Yan, and Zhang (2013) show that prices in the secondary market for Treasury security decrease significantly during the few days before Treasury auctions and recover shortly thereafter, even though the timing and size of each auction are announced in advance. The study shows that the price effects are significant even in very liquid markets. In one of the first event studies, Newman and Rierson (2004) find that new bond issuance in the European telecommunication sector increased the yield spreads of other firms in the sector. The effect was transitory, significant, and peaked on the day of issuance, not on the day of announcement. For a review of related event studies, see Duffie (2010, the AFA Presidential

Deterministic departures of prices from fundamental values is inconsistent with the competitive model — price-taking traders could make infinite profits. Irrespective of when the shock occurs, the competitive equilibrium price adjusts fully at the shock announcement. With imperfect competition, transitory departures of prices from the fundamental values are an equilibrium response to shocks.

- *Slow trading fundamentally changes the way markets respond to shocks — shocks to information, (net) supply, and endowments. First, any supply shock has two effects on prices; apart from the permanent price change, price exhibits temporary changes.*

Any exogenous supply or demand shock has two effects on prices: a *fundamental effect*, which is permanent, and a *liquidity effect*, which is temporary. These two effects differ in their origin, persistence, timing, and magnitude dynamics.

The fundamental effect, which is permanent, captures the change in the average market holdings (aggregate risk) and is also present in markets with price-taking traders.

In an imperfectly competitive market, the fundamental effect is amplified by a temporary liquidity effect, which results from traders' order reduction in response to price impact (and occurs even when information is symmetric); see Fig. 5A. Suppose that a supply shock H_t arrives in round t in the symmetric market. The current price p_t becomes

$$p_t = \overline{v}_t - \alpha_t \Sigma \frac{1}{I} \sum_j q_t^{j,0} - \underbrace{\alpha_t \Sigma \frac{1}{I} H_t}_{\text{Fundamental effect}} - \underbrace{(\kappa + \lambda_t) \Sigma \frac{1}{I} H_t}_{\text{Liquidity effect}},$$

while future prices p_τ for all $\tau > t$ shift by the change in fundamentals due to additional supply H_t in the market:

$$p_\tau = \overline{v}_\tau - \alpha_\tau \Sigma \frac{1}{I} \sum_j (q_\tau^{j,0} + \frac{1}{I} H_t) = \overline{v}_\tau - \alpha_\tau \Sigma \frac{1}{I} \sum_j q_\tau^{j,0} - \underbrace{\alpha_\tau \Sigma \frac{1}{I} H_t}_{\text{Fundamental effect}} .$$

While price exhibits the permanent effect upon the announcement of the shock, the liquidity effect always occurs at the moment of trade.

- *Second, whether the shock is pre-announced affects temporary price changes. Thus, announcements about fundamentals have real effects (i.e., induce changes in price).*

  In imperfectly competitive markets, price effects induced by a supply shock do not occur fully in rounds in which the information about the shock becomes available. Rather, shock-induced price effects occur on the date of the shocks as well as in rounds other than event dates (i.e., other than shock announcement and occurrence).

Address), Gromb and Vayanos (2010), and Lou, Yan, and Zhang (2013).

If anticipated, the liquidity effect of the shock induces additional temporary effects, which are present in all periods between the announcement t and the supply shock event $t' > t$; see Fig. 5B. This is because the price $p_\tau = \overline{v}_\tau - Q_\tau$ in round τ between t and t' (more precisely, the asset valuation $\overline{v}_\tau$) is affected by the anticipated future price change $p_{t'}$ in round t': from Equation (48), at each $\tau < t'$, the change in $\overline{v}_\tau$ by shock $H_{t'}$ is

$$\Delta\overline{v}_\tau = \sum_{s>\tau}^{t'} \prod_{l>\tau}^{s-1} ((1-\gamma_l)^2 + \gamma_l^2 \frac{\kappa}{\alpha_l})(1-(1-\gamma_s)^2 - \gamma_s^2 \frac{\kappa}{\alpha_s})\Delta p_s \qquad \forall t \leq \tau < t',$$

where Δp_s is the change in p_s for $s > t$ by shock $H_{t'}$ at t'. The temporary adjustment is maximal at the moment of trade t'.[74]

Thus, price impact gives rise to separation between trade announcements and trade-induced price effects in time *on the equilibrium path.* The price path is thus characterized by *anticipated* price changes. In the standard competitive model, these features of price behavior are ruled out by no-arbitrage, which equates the equilibrium price to the fundamental value. Any effects of announcements of future trades are reflected in price fully upon the announcement. With imperfect competition, no-arbitrage (implied by equilibrium) does not imply that the equilibrium price must equal the fundamental value.

In the competitive market, long-run prices are not affected by whether the trade is divided into smaller orders or by the time at which the trades take place, whether the shocks are anticipated or not. The partition and timing of trades do affect the price path in the imperfectly competitive market. This occurs because a future sale depresses prices during the period between the announcement and the shock occurrence; moreover, the effects of multiple sales on prices are cumulative.[75]

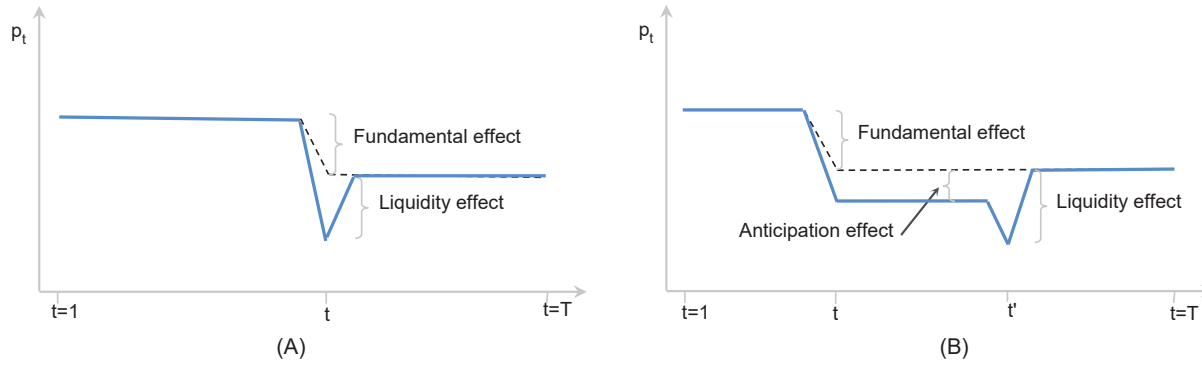
- *In contrast to competitive markets, when traders have price impact, price responds differently to endowment shocks than to shocks in demand or supply.*

In Equation (53), if the endowment of a strategic trader i increases by δ_t^i, then equilibrium price dynamics would follow that of the fundamental value; only the fundamental effect, and

[74] Bond, Edmans, and Goldstein (2012) review models that feature a feedback effect from financial markets to the real economy due to the informational role of prices. The authors advance two general conclusions. First, the concept of price efficiency should account both for the extent to which prices reflect information that is useful for the efficiency of real decisions and the extent to which prices forecast future cash flows. Second, incorporating the feedback effect into financial market models can explain evidence on limits to arbitrage, manipulative short selling, information-based trade, and incentives to disclose private information.

[75] Sannikov and Skrzypacz (2016) study temporary price impact off equilibrium; the intertemporal price impacts are constant since equilibrium is stationary (see Section 5.3.2). Rostek and Yoon (2019) examine temporary changes in equilibrium prices.

FIGURE 5: UNANTICIPATED AND ANTICIPATED SUPPLY SHOCK



Notes: (A) Suppose that an unanticipated supply shock H_t arrives at t. In a competitive market, equilibrium price adjusts to the new fundamental value (i.e., fundamental effect) within one round. In an imperfectly competitive market, in addition to the permanent (fundamental) effect, there is a temporary price change due to price impacts (i.e., liquidity effect). (B) When a shock $H_{t'}$ at t' is pre-announced at $t < t'$, the competitive-market price adjusts to the new fundamental value of the assets upon news arrival. In the imperfectly competitive market, the equilibrium price also changes in all rounds $t < t < t'$ between news and shock arrivals.

no liquidity effect, would be observed:[76]

$$p_t = (\sum_j \frac{\gamma_t^j}{\alpha_t^j})^{-1} \sum_j \frac{\gamma_t^j}{\alpha_t^j} E[v_t^j | \mathcal{I}_t^j, p_t] - (\sum_j \frac{\gamma_t^j}{\alpha_t^j})^{-1} \sum_j \gamma_t^j \Sigma q_t^{j,0} - \underbrace{(\sum_j \frac{\gamma_t^j}{\alpha_t^j})^{-1} \gamma_t^i \Sigma \delta_t^i}_{\text{Fundamental effect}}.$$

Hence, with endowment shocks, price dynamics are not affected by the timing of shocks but a full price adjustment occurs solely on the date of the shock announcement. Shocks of size δ_t^i to both endowments and (net) demands change *per capita* holdings, as the net trade of liquidity providers is zero in every round, and the price impact does not create a wedge between the equilibrium price and the average marginal utility.

This suggests that traders who liquidate risky positions have incentives to bypass centralized markets and exchange shares over the counter to avoid price concessions resulting from the liquidity effect. The elimination of the liquidity effect with the direct exchange of shares also motivates alternative clearing arrangements, such as netting.

- *In contrast to competitive markets, preannouncing the shock has a redistributive impact on consumption and wealth.*

[76] With endowment shocks, in every round the price equals the fundamental value, which varies over time (e.g., Vayanos (1999)). If the new supply enters the holdings directly (i.e., an endowment shock) instead of being absorbed by the liquidity providers through trading (i.e., a supply shock), the temporary liquidity effect is present.

Anticipated supply shocks redistribute wealth toward buyers, whereas demand shocks redistribute wealth toward sellers. In contrast, private endowment shocks do not necessarily benefit one side of the market in all subsequent rounds.

5.3 Techniques to Characterize Equilibrium

5.3.1 Symmetric traders (static inference)

In markets with symmetric traders, the value function parameters v_t^i and α_t can be characterized by backward induction. Given the value function $V_t^i(\cdot)$, equilibrium demand $\{q_t^i(\cdot)\}_i$ at t can be characterized through the same steps as in the static model. Even with private information, in markets with shocks to private endowments or valuation, but not to supply, when traders are symmetric, inference effects are static without loss of generality; i.e., current price p_t summarizes the history $\{p_s\}_{s<t}$.[77]

5.3.2 Heterogeneous traders (dynamic inference)

Incorporating trader heterogeneity (e.g., in risk preferences) in a dynamic model is challenging because the effect of each trader on the residual market (i) differs across traders, (ii) impacts current and *all* future rounds, and (iii) is time-varying. Jointly, these effects make recursive methods inapplicable.

Methodologically, trader heterogeneity gives rise to dynamic inference and allocation effects. In markets with symmetric traders, a demand increase by trader i in round t (i.e., a unilateral deviation) impacts only the current-round price.[78] When traders are heterogeneous — either because their risk preferences differ or their demands condition on different statistics — *equilibrium price aggregates traders' valuations with different weights* $\{\frac{\gamma_t^i}{\alpha_t^i}\}_i$ *(because traders value assets differently), which change over time (because effective risk aversion* α_t^i *changes with* t*)* (see Equation (53)). With time-dependent weights $\{\frac{\gamma_t^i}{\alpha_t^i}\}_i$, a unilateral demand change at t affects the current price in round t and, in contrast to markets with symmetric traders, (the distribution of) all future prices. Hence, equilibrium characterization has to take into account the resulting change in all bidders' *inference* and *allocations* in rounds $s \geq t$, and hence demand functions and valuations in all future rounds $s > t$ and — through a fixed point — their value function in the current round. Consequently, the best response at t is

[77] With supply shocks (e.g., Vayanos (1999, 2001)), history matters even with symmetric traders when equilibrium is recursive. Price is not fully revealing but provides a noisy signal on the aggregate risk $Q_t = \frac{1}{I}\sum_j q_t^{j,0}$, which is constant in all rounds. Time-invariant aggregate risk in prices implies that, even when prices are not fully revealing, the history is summarized by a one-dimensional statistics, i.e., an *average* of past and current prices, which allows the inference coefficients to be expressed in a closed form.

[78] When traders are symmetric, equilibrium price coincides with the competitive price in all rounds, and may vary over time only due to arrival of new shocks. The competitive price assigns equal weight to all traders' private information and is a function only of traders' aggregate inventory in this round and not the distribution across traders.

defined as a fixed point between the value function at t and the best response demands at $s > t$.

As a result, in contrast to markets with symmetric traders, the dynamic inference and allocation effects challenge equilibrium recursivity as follows:

(i) Learning is (endogenously) gradual over time rather than Markovian. This is because prices in different rounds are different functions of private information $\{q_0^i\}_i$ and thus contain different information even without arrival of new signals. That is, the entire history matters.

(ii) Additionally, how traders value assets in any round depends not only on the past price history but also on the future price path, of which traders are uncertain. Because of the uncertainty about future prices, the payoff-relevant random variable that determines a trader's marginal valuation changes over time. (With symmetric traders, the endogenous valuation v_t^i depends only on the current price, to which all expected future prices are equal.)

Dynamic inference effects are present even in markets with independent private values (when $T > 2$, except when traders are equally risk averse). *Uncertainty about future prices* gives rise to intertemporal inference effects in and off equilibrium (Example 4 in the Appendix makes this explicit); inference effects due to interdependence among valuation $\{\theta^i\}_i$ amplify these effects.

For these reasons, many dynamic models of imperfectly competitive trading — models of financial markets or more generally, dynamic oligopoly — have studied markets in which information is disclosed fully after each round, prices are fully revealing in all rounds (Vayanos (1999), Rostek and Weretka (2015a), Antill and Duffie (2017), Du and Zhu (2017a), Kyle, Obizhaeva, and Wang (2018), and Sannikov and Skrzypacz (2016)) or private information is Markovian and players are symmetric (Vayanos (1999, 2001)).[79]

With symmetric traders, a demand increase in t does not affect future prices and traders' inference is static for all rounds. Moreover, after the unilateral deviation at t, the value functions, v_t^i and α_t, stay the same for all i, and thus, equilibrium is recursively characterized by the standard backward induction.

Two recent papers introduce methods that allow for the characterization of equilibrium with heterogeneous traders. Sannikov and Skrzypacz (2016) let traders submit schedules in each round t contingent on the $(I-1)$-dimensional vector of other traders' quantities to be traded at t. Then, (i) equilibrium is *ex post* in each round (i.e., history does not matter and

[79] An alternative approach is to relax the assumptions about the state variable of which the players keep track; e.g., oblivious equilibrium (Weintraub, Benkard, and Van Roy (2006)) and variants of self-confirming equilibrium used in theoretical and empirical models (e.g., Fudenberg and Levine (1993, 2009), Dekel, Fudenberg, and Levine (1999), Cho and Sargent (2008), Pakes (2016)).

the dynamic inference effect disappears) and (ii) even though a demand increase of trader i at t affects all future prices through changes in allocations, future demand functions in $s > t$ do not change after the deviation (hence, no fixed point for the value function). This allows a recursive characterization of equilibrium for heterogeneous traders.[80]

Rostek and Yoon (2019) characterize a nonstationary equilibrium in a finite horizon model in which trade is more frequent than asset payments and which encompasses fully transparent markets (i.e., private information is revealed in every round) and those that are less transparent (i.e., non *ex post* equilibrium). With the presence of both dynamic inference and allocation effects, a deviation at t affects prices in *current and all future rounds* $\tau \geq t$ and traders' inference conditional on past price history. The key observation is that a trader's current-round price impact $\Lambda_t^i \equiv \frac{dp_t}{dq_t^i}$ and intertemporal price impacts $\Lambda_t^{i,\tau} \equiv \frac{dp_\tau}{dq_t^i}$ $(\tau > t)$ — which, respectively, capture the impact of a trader's demand change at t on the current and future round prices — are a sufficient statistic for the behavior of other players after a unilateral deviation at t.[81] Moreover, the fixed point for the dynamic non-recursive equilibrium in demand schedules $\{q_t^i(\cdot)\}_{i,t}$ is equivalent to the fixed point in price impacts $\{\Lambda_t^i, \{\Lambda_t^{i,\tau}\}_{\tau>t}\}_{i,t}$. In particular, and analogously to the static model (Proposition 2), all of the coefficients of the best response demand, including the inference coefficients (i.e., conditional expectations), can be endogenized as functions of current-round and intertemporal price impacts (see Appendix B).[82]

Financial models with private information commonly assume that there is a fundamental value of the asset and investors receive signals of that fundamental value with additive noise. In dynamic trading models, receiving signals of the fundamental value renders the model nonstationary even when traders have symmetric risk preferences. Du and Zhu (2017a) solved a dynamic double auction with interdependent values, thus taking an important step beyond Vayanos (1999). Du and Zhu (2017a) propose that the fundamental value itself varies over time; every time traders receive signals, there is an innovation in the fundamental value and one in private values. Importantly, investors' signals are about the innovations rather than the level of the fundamental value. This approach lends itself to a stationary equilibrium and prices that are martingales.

These models capture different types of adverse selection due to the within-round inference (Du and Zhu (2017a)) or both static and intertemporal inference (Vayanos (1999) and Rostek

[80] Sannikov and Skrzypacz (2016) consider an infinite horizon model with trade and asset payments in every round. Their approach is useful to characterize *ex post* equilibria both in infinite- and finite-horizon models (respectively, stationary and nonstationary).

[81] In particular, the current-round price impacts $\{\Lambda_t^i\}_{i,t}$ are not sufficient for equilibrium in the dynamic market (except when, in each round t, traders can condition on an $(I-1)$-dimensional statistic of the outcome at t) because of the fixed-point between the value function and demands in a nonstationary equilibrium.

[82] Moreover, the joint distribution of all future prices is *unnecessary* for a trader's equilibrium demand at t. Thus, a non-recursive equilibrium representation as a fixed point in price impacts both reduces the dimension of the variables that characterize equilibrium and eliminates the information on future equilibrium variables that is unnecessary for optimization in a given round.

and Yoon (2019)).

5.3.3 Comparison with the Cournot model

With *price-taking* liquidity providers (e.g., Grossman and Miller (1988), the Cournot model), the implications of slow trading — for welfare, savings, and the effects of announcements — differ from those with price-making liquidity providers (i.e., demand competition game). Namely, unless information happens to be revealed contemporaneously with the shocks' occurrence, shocks only have permanent and no transitory price impact; hence, the liquidity effect of shocks only depends on contemporaneous uncertainty. With bilateral market power, price impact in any given round depends on *future and current* uncertainty about asset payoffs. This is why, with bilateral price impact, the effects of information disclosure about an event and the event itself are separated in time.[83]

Unlike with price-contingent demands, in the Cournot model a demand change at t affects current *and future* prices even with symmetric traders. This is because a trader's strategy at t depends on the expected (rather than to-be-realized) price in the current round. Consequently, there is an intertemporal price impact even with symmetric traders. The technique to solve dynamic demand submission games also applies to the dynamic Cournot model with heterogeneous traders (and/or private information): equilibrium can be characterized as a fixed point in traders' intertemporal price impacts. The price impacts within round are exogenous; thus, in contrast to games with price contingent schedules, there is intertemporal, but not within-round, adverse selection.[84] Demand coefficients, including inference coefficients, are functions of the intertemporal price impacts. See Appendix B.

Bonatti, Cisternas, and Toikka (2017) consider a dynamic Cournot model with heterogeneous traders and private information. Traders consume in each round; consequently, equilibrium bids are stationary and prices have a Markov property even with heterogeneous traders. Since traders' inference depends only on the previous-round price, the effect of a demand change (i.e., intertemporal price impact) lasts one round. Thus, equilibrium can be recursively characterized even though traders are heterogeneous and traders' inference is not static.

5.3.4 Related macroeconomic models

In parallel to financial market studies, a number of recent papers have proposed deviations from the benchmark of the competitive equilibrium or full-information rational-expectations

[83] In contrast to markets with price-taking liquidity providers (or all traders), in markets with strategic traders, price impact is strictly positive even without any price risk. In the Grossman and Miller (1988) model, the assumption $Var_1(E_2\bar{P}_2) \equiv \sigma_2^2 > 0$ is necessary for a round-trip trade to affect the equilibrium price.

[84] A unilateral demand increase only affects the future-round, but not the current-round, expected price for other traders.

to improve the predictive power of the standard new-Keynesian model motivated by newly available evidence from macro and labor data. Some contributions include Gabaix (2016), Fahri and Werning (2017), Angeletos and Huo (2018), Angeletos and Lian (2018), Werning and Wang (2018), Woodford (2018), and Molavi (2019). See also Bianchi, Lettau, and Ludvigson (2020), Lagos and Zhang (2020), and Malamud and Schrimpf (2017, 2018).

Underlying the implication of imperfect competition described above is inefficiency due to an inability to fully realize gains from trade in one round.[85] Suitable counterparts of the implications of "slow trading" due to (bilateral) price impact will be shared by models based on alternative departures from the competitive, frictionless model.

Like the demand submission approach, these alternatives take steps towards alleviating the difficulties associated with backward induction by modeling forward-looking behavior. These ideas go back to the notion of temporary equilibrium (Hicks (1939), Grandmont (1977, 1982); see also Grandmont (1988)).[86]

5.3.5 Comparison with (single-agent) transaction costs approach

The transaction cost approach (Gârleanu and Pedersen (2013, 2016), Collin-Dufresne, Daniel, Moallemi, and Sağlam (2015)) enables the characterization of the best response trading strategy for any constant price impact matrix, which is taken as given. (A time-varying exogenous cost can be accommodated.) Multi-trader equilibrium models (such as those reviewed in previous sections) endogenize price impact — and the time-varying effects of traders and prices — and allow for the incorporation of inference effects. By endogenizing the relationships between price, volume, and other variables, they therefore provide additional structure to map ideas from the transaction cost approach to data. Because price impact is determined in equilibrium, one can consider how it responds endogenously to changes in the environment (e.g., the market structure, financial innovation).

5.3.6 Sources of predictability in imperfectly competitive markets

There is more to learn about the impact of frequencies on equilibrium and welfare, particularly with heterogeneous traders and dynamic inference effects. Investigating equilibrium dynamics more systematically when the relative frequencies of trading, payment, and consumption differ can help uncover the time-series properties of prices and volatility.

When traders have price impact, differences in relative frequencies, announcements of events that represent supply shocks, and heterogeneity in risk aversion (risk-bearing capacity)

[85] Naturally, the inefficiency is not specific to the uniform-price mechanisms and would be present with other market-clearing rules in markets with two-sided private information.

[86] Temporary equilibrium allows for an arbitrary closure of the model. The additional structure of more recent approaches to modeling dynamic equilibrium allows one to demonstrate whether the comparative statics is, in some sense, invariant to the corresponding closure of the model.

all lead to deterministic equilibrium price changes.

The existence of the liquidity effect of supply shocks gives rise to excess volatility, which when combined with the nonstationarity of price impact, induces changes in volatility unrelated to changes in fundamentals.

The representation of equilibrium as a fixed point in price impacts makes explicit how market impact models and market data improve predictability. The separation of the forward problem (characterization of inference) and the backward problem (characterization of the value function) described below in Section B can be useful in importing the results from the equilibrium approach to the agent-based approach and vice versa.

6 Decentralized Markets

Apart from assuming price-taking behavior, the standard equilibrium model of financial markets assumes a particular market structure: all traders trade all assets in a single exchange (or the equilibrium is as if that were the case). In the Arrow-Debreu model, the assumption that there is a single market clearing is, in fact, explicit (see also Section 6.3.2). While this assumption fits certain marketplaces, actual trading environments are considerably more complex. With imperfect competition, such differences in the organization of trade can play a crucial role when considering design and regulation. Indeed, in practice, dealing with price impact often serves as a primary motivation for creating an alternative exchange or introducing a new financial product.[87] The fact that agents are non-negligible in trading helps to explain why trading is not centralized even if doing so were feasible — we review the arguments below.

6.1 Financial Market Structure

The number of alternative trading venues and the volume of trade outside open exchanges have increased severalfold over the past couple of decades.[88] New types of electronic marketplaces have emerged to offer different forms of market clearing (e.g., direct matching with an intermediary, trading in a dealer network, or an online exchange) to different types of traders, institutional and retail.[89] During this time, a significant fraction of transactions occurring outside open exchanges — particularly, customer-to-dealer and interdealer transactions — has shifted to such electronic auctions available to large institutional traders and dealers, as

[87] E.g., Biais and Green (2007), Biais, Bisière, and Spatt (2010), Knight Capital Group (2010), and Angel, Harris, and Spatt (2011).

[88] Regulatory changes have also contributed to the trend of the increasing number of venues. Before 2007, European equity markets were characterized by dominant exchanges in each domestic market. The MiFID reform in 2007 created more than 200 new trading venues in which equities, bonds, and derivatives are traded.

[89] E.g., BondDesk, BrokerTec, eSpeed, MarketAxess, and TradeWeb.

well as to retail investors. This transition in the way markets clear is sometimes called the "call to electronic."[90]

Markets are also fragmented because investors trade different assets, either by choice or regulation. Pension funds cannot trade many types of derivatives, while banks are allowed to hold but not trade loans, and hedge funds tend to specialize in trading a limited number of securities. Even large financial institutions trade a small subset of securities and participate in only a few trading venues.

Moreover, many products traded over the counter (OTC) are hard to standardize.[91] Increasing the standardization of assets has been considered a key challenge in regulating OTC markets (Financial Stability Board (2010) and Duffie (2018).[92] Even for standardized products traded in large volumes, the mandated use of trade platforms (the Dodd-Frank Act and MiFID II) to encourage direct trading among investors has had limited success (e.g., Duffie (2018)).

A growing body of empirical studies evaluates the impact that advances in technology (e.g., new market-clearing arrangements, high speeds) and the availability of new detailed market data have had on liquidity and the functioning of markets more generally. To quote from a recent paper by O'Hara (2015): "What is particularly intriguing is the new role played by microstructure. One might have expected that when things are fast the market structure becomes irrelevant — the opposite is actually the case. At very fast speeds, only the microstructure matters." (See also Pagano (1989), Biais and Foucault (2014), Budish, Cramton, and Shim (2015), Budish, Lee, and Shim (2019), Foucault and Moinas (2018), Linton and Mahmoodzadehy (2018), Pagnotta and Philippon (2018), Cespa and Vives (2019), and Baldauf and Mollner (2020).)

6.2 Theory of Decentralized Trading

To identify the implications of market fragmentation, it is useful to first consider what exactly the *centralized market assumption* entails. What does an off-the-shelf model of centralized trading assume about the market? The essence of centralized trading, in the classical theory, is not a matter of centralized or decentralized implementation but rather that a single market clearing applies for all traders and assets. In particular, two assumptions are implicit in the centralized market assumption. The first is that *trader participation in the market is com-*

[90] E.g., Hendershott and Madhavan (2015) examine the transition toward auctions in the US corporate bonds market. Innovations in trading technology allow traders to engage in multilateral trading as opposed to sequentially contacting dealers.

[91] While this article focuses on standardized products, observe that the framework allows for K arbitrary assets described by $\mathcal{N}(\theta, \Sigma)$, standardized or bespoke.

[92] Market participants looking to hedge specific risks may not find a standardized product that would effectively match their exposure and instead may prefer to use a bespoke product. Bespoke products often do not have the level of standardization required for trading on organized exchanges.

plete in the following sense: Each trader trades all assets with all other market participants; however, actual market structures combine private exchanges with restricted participation, public exchanges in which all traders can participate, and a variety of dealer networks, for the same or different assets.

The second assumption implicit in the centralized market assumption concerns the *type of demand schedules allowed.* Consider a market with two exchanges, each for one asset, possibly the same asset. Should the demands that traders submit in one exchange allow them to condition on the price from the other exchange as well? In other words, should a trader submit *contingent schedules* $q_1^i(\cdot) : \mathbb{R}^2 \to \mathbb{R}$ and $q_2^i(\cdot) : \mathbb{R}^2 \to \mathbb{R}$, each specifying the quantity demanded of each asset for any price vector (p_1, p_2), or *uncontingent schedules* $q_1^i(\cdot)$ and $q_2^i(\cdot)$, each $q_k^i(\cdot) : \mathbb{R} \to \mathbb{R}$ specifying the quantity demanded of asset k as a function of price p_k. With uncontingent schedules, the market clears *asset by asset.* With contingent schedules — assumed in the standard competitive model of centralized markets — the market clears *jointly* for all assets. To reiterate, the assumption of fully contingent schedules requires a single aggregation for all traders and assets in the classical model. While uncontingent schedules are more prevalent, contingent schedules are employed by (or equivalent to[93]) some trading protocols in practice.[94]

More generally, once one departs from the assumption that all units are cleared via a single aggregation, the assumptions implicit in the centralized market assumption become apparent. When reviewing the literature, we focus on two in particular:

Definition 6 (Centralized Market, Decentralized Market) *A market is* centralized *if*

(i) (Complete participation) All traders $i \in I$ trade all assets $k \in K$; and

(ii) (Complete demand conditioning) Each trader's demand schedules are contingent on the price vector $p \in \mathbb{R}^K$, i.e., $q^i(\cdot) = (q_k^i(\cdot))_k : \mathbb{R}^K \to \mathbb{R}^K$.

A market is decentralized *otherwise.*

In Section 6.3, we review two corresponding classes of decentralized market models that relax these assumptions.

6.2.1 Comparison with search models

An influential tradition in modeling decentralized markets is based on search and matching models for over-the-counter trading via bilateral transactions.[95] By relaxing the assumption of

[93] E.g., RegNMS and UTS in US stock exchanges de facto induce contingent demand schedules.

[94] Given that most markets are dynamic, traders can condition their demands in each exchange on past outcomes from other exchanges. Even if past shocks are aggregated, incomplete conditioning will still matter for how the current-round shocks affect equilibrium.

[95] E.g., Gale (1986a,b), Duffie, Gârleanu, and Pedersen (2005), Vayanos and Weill (2008), Weill (2008), Duffie, Malamud, and Manso (2009, 2014), Lagos and Rocheteau (2009), Lagos, Rocheteau, and Weill (2011),

complete participation, these models have led to substantial progress in explaining equilibrium dynamics and shifted the focus to a *distribution* of equilibrium prices.

The random search approach examines trading environments in which trade occurs among a continuum of traders who are small relative to the market. The relationships form via random meetings among traders, and the outcome of transactions is determined by bilateral bargaining. The search friction captures that it takes time to find a counterparty who is willing to trade. In turn, the framework based on demand submission games fits trading environments with complementary features: It accommodates markets with an arbitrary finite number of traders, all of whom can be non-negligible in the market and have price impact, and who interact through relationships that are not random. Terms of trade are determined for any subsets of interacting agents (e.g., an exchange).

For many markets, a more realistic description would neither impose that the probability of two counterparties that trade at date t meeting again in the future is almost surely zero (random search models) nor that trading relationships are fixed (demand submission models). There is scope for integrating the ideas and techniques from search models and demand submission games to develop joint equilibrium implications of traders' price impact and frictions (e.g., search cost in the time needed to find a counterparty).[96] Such a framework would elucidate how trading relationships emerge.[97] It would help separate the implications of market fragmentation in the cross section from those in time series.

The demand submission framework contributes equilibrium analysis methods for:

(1) General market structures.

Demand submission games accommodate (i) any form of limited participation that encompasses market structures "between" centralized and bilateral trading, including centralized markets (Definition 6), intermediated markets (vertically or horizontally integrated trading arrangements), networked markets on graphs (i.e., bilateral links) and market structures that cannot be described by bilateral links without loss, and are more appropriately represented by a generalization of graphs known as hypergraphs;[98] and (ii) arbitrary restrictions on demand conditioning "between" contingent $q^i(p) = (q^i_k(p))_k : \mathbb{R}^K \to \mathbb{R}^K$ and uncontingent $\{q^i_k(p_k) : \mathbb{R} \to \mathbb{R}\}_k$.

Afonso and Lagos (2015), Huggonier, Lester, and Weill (2014), and Atkeson, Eisfeldt, and Weill (2015). For a book treatment, see Duffie (2012).

[96] Recent contributions by Lester, et al. (2018, 2019) introduce market power in a search models.

[97] An average bank trades with a small number of counterparties, and most banks form stable relationships with at least one lending counterparty; e.g., the US Federal Funds market (Bech and Atalay (2010), Afonso, Kovner, and Schoar (2013)), interbank markets (Craig and Peter (2014), Cocco, Gomes, and Martins (2009)), and US municipal bonds market (Li and Schürhoff (2012)).

[98] A hypergraph generalizes a graph by allowing an edge to connect any number of nodes, beyond just two (e.g., Berge (1973)). Hypergraph representations of complex trading relationships allow the study of interactions among groups and scenarios in which aspects other than the traders — trading venues and trader or asset characteristics — are essential.

The framework can thus incorporate some of the increasingly diverse set of trading venues and trading protocols used in practice, such as markets with multiple coexisting venues for the same or distinct assets and investors taking positions in multiple venues.

(2) Multilateral trading relations.

Just as the classical equilibrium theory assumes that there is a single market clearing for all traders and all assets (i.e., the centralized market assumption), early literature on decentralized trading made the opposite assumption — i.e., that all transactions are bilateral — modeled as random or fixed graphs.[99] Often, assuming that relations are only between pairs of agents when groups can interact as well provides an incomplete description of complex trading relationships (e.g., coexisting exchanges and bilateral interactions) and, indeed, introduces biases in inference and possibly artifacts in the data in the study of equilibrium as well as connectivity, clustering, and other topological properties.[100]

Allowing for relations other than bilateral also facilitates a systematic comparison of centralized and decentralized markets. In some settings, one can identify the centralized market with a complete graph; however, equilibria in complete graphs generally differ from those in which agents interact as a group, as in exchanges. For the purpose of developing a systematic comparison of centralized and decentralized markets, working with purely bilateral relations may not be without loss. Moving beyond bilateral links also allows the analysis of when and why a decentralized market can function like the centralized market as well as the conditions under which equilibrium behavior differs.

(3) The uniform-price double auction with private information.

The standard uniform-price auction enables explicit treatment of imperfectly competitive behavior in decentralized markets. Buyers and sellers act strategically, the quantity traded affects the price at which a transaction takes place. Market clearing based on the uniform price is a precise counterpart of the market clearing in *centralized market* trading models — in the competitive (general equilibrium) model and its imperfectly competitive (game-theoretic) counterpart (e.g., Wilson (1979), Kyle (1989), Klemperer and Meyer (1989), and Vives (2011)). It thus permits a direct comparison with classical results.

In order to avoid the complexities of having to model bargaining with incomplete information, random- and fixed-graph models typically derive the terms of trade from bargaining (e.g.,

[99] Outside the demand submission framework, exceptions include Corominas-Bosch (2004) and Elliott (2015), who allow for multilateral bargaining with search. Rahi and Zigrand (2013) study trade of price-taking investors intermediated by arbitrageurs.

[100] Atalay, Hortaçsu, Roberts, and Syverson (2011) made this observation in the context of production networks. The demand submission games are also suitable for the trading environments where transactions are bilateral (see Babus and Kondor (2018) and Yoon (2018)).

take-it-or-leave-it offers) or posted prices, and have efficient surplus sharing on each link.[101] Accounting for the inefficiency due to price impact and private information (two-sided, i.e., buyer and seller) allows one to establish how decentralized trading affects inefficiency and to identify which changes in market structure are conducive to efficiency.

Due to the difficulties of characterizing equilibrium in bargaining with externalities, models of bargaining among more than two players, or bargaining among players involved in multiple relationships, often rely on concepts from cooperative game theory or adopt a separability assumption. The (noncooperative) demand submission game based on the uniform-price auction determines the outcome in multilateral relationships and/or multiple relationships with externalities.

6.3 Equilibrium in Decentralized Markets

We review two classes of decentralized market models, which depart from the complete participation and complete conditioning assumptions of centralized trading, respectively. Relaxing each assumption separately allows a systematic delineation of the functioning of decentralized markets.[102] What economic mechanisms in decentralized markets have no centralized market counterparts?

Consider a market with I traders who trade K risky assets in N exchanges. An *exchange* $n \in N$ is identified by a subset of agents $I(n) \subseteq I$ who trade there and the subset of assets traded $K(n) \subseteq K$. The *market structure* is described by the set of exchanges $\{(I(n), K(n))\}_n$ (a hypergraph). As in the centralized market model based on demand submission, in each class of decentralized market models reviewed below, equilibrium can be represented as a fixed point in price impacts, and price impact is a sufficient statistic for the market (and the game). Knowing the details of the market structure $\{(I(n), K(n))\}_n$ other than price impact (and, in the second model, the joint distribution of equilibrium prices of the assets a trader trades) would not affect the optimal behavior when values are independent private.

6.3.1 Incomplete participation

Agents can participate in many different types of trading venues for possibly non-disjoint subsets of traders with the same or different counterparties (e.g., a public exchange, in which all traders participate, a private exchange, which restricts participation to a subset of traders, and intermediation; relationships need not be exclusive). Each trader i participates in $N(i) \subseteq N$ exchanges (and trades $K(i) \equiv \sum_{n \in N(i)} K(n)$ assets).

[101] I.e., models of decentralized trading with bargaining often assume that agents reveal their private types upon the match; information is symmetric in trade.

[102] Babus and Kondor (2018), Babus and Parlatore (2017), and Malamud and Rostek (2017) study markets with incomplete participation and contingent contracts. Chen and Duffie (2020), Rostek and Yoon (2020), and Wittwer (2020) relax the assumption of contingent schedules.

Equilibrium in any decentralized market $\{(I(n), K(n))\}_n$ can be characterized by two simple conditions that mimic those for centralized trading. This follows from two observations, which lead to Proposition 3. First, it is useful to *treat assets traded in different exchanges as different assets.* This is in order not to impose that the same assets (in the sense of the asset payoff distribution $\mathcal{N}(\theta, \Sigma)$) trade at the same prices in different exchanges. Accordingly, a market with K assets traded in N exchanges is analyzed as a market with $\sum_n K(n)$ (replicas of) assets. The covariance matrix $V \in \mathbb{R}^{(\sum_n K(n))\times(\sum_n K(n))}$, induced by covariance Σ and the set of exchanges, and the vector in $\mathbb{R}^{\sum_n K(n)}$ induced by mean return $\theta \in \mathbb{R}^K$ (we do not use additional notation for that vector) describe the interconnectedness among the exchanges via traders and assets.

The distribution of the assets $K(i)$ that trader i trades is $\mathcal{N}(\theta_{K(i)}, V_{K(i)})$, where $\theta_{K(i)} \in \mathbb{R}^{K(i)}$ and $V_{K(i)} \in \mathbb{R}^{K(i)\times K(i)}$ are a subvector of θ and a submatrix of V, respectively, corresponding to assets $K(i)$. Trader i maximizes

$$\max_{q^i(\cdot):\mathbb{R}^{K(i)}\to\mathbb{R}^{K(i)}} E[\theta_{K(i)} \cdot (q^i + q_0^i) - \frac{\alpha^i}{2}(q^i + q_0^i) \cdot V_{K(i)}(q^i + q_0^i) - p_{K(i)} \cdot q^i | q_0^i], \tag{57}$$

and submits his demand schedules $q^i(\cdot) : \mathbb{R}^{K(i)} \to \mathbb{R}^{K(i)}$ as a function of the prices $p_{K(i)} \in \mathbb{R}^{K(i)}$ of the assets he trades.

With limited participation, traders' price impacts are of different dimensionality, correspond to different assets, and generally are not independent across exchanges, so the market-clearing condition cannot be written exchange by exchange. *Lifting* traders' demands allows one to apply market clearing to all assets in all exchanges in order to determine the market clearing prices $p \in \mathbb{R}^{\sum_n K(n)}$: $\sum_i \mathbf{q}^i(\mathbf{p}_{K(i)}) = 0$. Lifting restores common dimensionality.

Any symmetric matrix A that acts on space $\mathbb{R}^{\sum_n K(n)}$ can be decomposed into a block form

$$A = \begin{bmatrix} A_{i,i} & A_{i,-i} \\ A'_{i,-i} & A_{-i,-i} \end{bmatrix}, \tag{58}$$

where $A_{i,i} = A_{K(i)}$, $A_{-i,-i} = A_{\sum_n K(n)\setminus K(i)}$, and $A_{i,-i}$ is a rectangular block. Let $\mathbf{A}_{i,i}$ denote the *lifted* matrix, which acts on $\mathbb{R}^{\sum_n K(n)}$, and with a slight abuse of notation, let $\mathbf{A}_{i,i}^{-1}$ denote its inverse:

$$\mathbf{A}_{i,i} = \mathbf{A}_{K(i)} \equiv \begin{bmatrix} A_{i,i} & 0 \\ 0 & 0 \end{bmatrix}; \quad \mathbf{A}_{i,i}^{-1} = \mathbf{A}_{K(i)}^{-1} \equiv \begin{bmatrix} A_{i,i}^{-1} & 0 \\ 0 & 0 \end{bmatrix}. \tag{59}$$

We use capital bold notation for objects defined directly in a decentralized market.

If assets traded in different exchanges are treated as distinct assets and aggregation is tackled by lifting, one can characterize equilibria in a decentralized market (represented by any hypergraph) by two conditions that mirror those for the centralized market (Proposition 1).

Proposition 3 (Equilibrium: Incomplete Participation) A profile of schedules $\{q^i(\cdot) : \mathbb{R}^{K(i)} \to \mathbb{R}^{K(i)}\}_i$ is an equilibrium if and only if

(i) Each trader submits a schedule that equalizes his marginal utility and marginal payment, given his price impact:

$$\theta_{K(i)} - \alpha^i V_{K(i)}(q^i + q_0^i) = p_{K(i)} + \Lambda^i q^i \qquad \forall p_{K(i)} \in \mathbb{R}^{K(i)}; \tag{60}$$

(ii) The price impact $\Lambda^i \in \mathbb{R}^{K(i)\times K(i)}$ is correct (i.e., equals the slope of the residual inverse supply function resulting from the aggregation of other traders' schedules, projected on the assets relevant for trader i after lifting):

$$\Lambda^i = -((\sum_{j\neq i} \frac{\partial q^j(p_{K(j)})}{\partial p})^{-1})_{K(i)} = ((\sum_{j\neq i}(\alpha^j \mathbf{V}_{K(j)} + \mathbf{\Lambda}^j)^{-1})^{-1})_{K(i)} \qquad \forall i. \tag{61}$$

- *With independent private values, equilibrium is* ex post *for any market structure* $\{(I(n), K($

- *Price impact is not proportional to but concave in* Σ.

The equilibrium price impact tuple $\{\Lambda^i\}_i$ is increasing and concave in risk aversion $\{\alpha^i\}_i$ and the covariance matrix Σ, and is decreasing in the number of traders in the market $\{I(n)\}_n$. Mathematically, the concavity follows from Gaussian conditioning. Denoting by $\Psi \equiv \sum_j(\alpha^j \mathbf{V}_{K(j)} + \mathbf{\Lambda}^j)^{-1} \in \mathbb{R}^{(\sum_n K(n))\times(\sum_n K(n))}$ the slope of the aggregate (net) market demand, the condition that price impact is the harmonic mean of (lifted) inverse demand slopes can be written as:

$$\Lambda^i = (((\Psi^{-1})_{K(i)})^{-1} - (\alpha^i V_{K(i)} + \Lambda^i)^{-1})^{-1}, \tag{62}$$

where $((\Psi^{-1})_{K(i)})^{-1} = \Psi_{i,i} - \Psi_{i,-i}\Psi_{-i,-i}^{-1}\Psi_{i,-i}'$.

- *When a market becomes more decentralized (i.e., when some agents trade fewer assets or trade with fewer other traders), traders' price impacts weakly increase in all exchanges.*

This holds irrespective of the asset structure in the more decentralized market (see Definition 9 in Appendix C.1) and for any number of traders and any risk aversion. The centralized market thus minimizes price impact. Moreover, a general complementarity property holds: any change in the market structure that lowers price impact "locally" in one exchange also lowers price impact in all others as long as they are indirectly connected. For instance, creating new private exchanges (i.e., in which participation is restricted) weakly improves liquidity in the market.

6.3.2 Incomplete demand conditioning

When schedules are contingent and values are independent private, one can study the effects of incomplete participation itself without having to tackle the effects of information. With limited cross-asset conditioning, equilibrium is not *ex post*, even with independent private values. As a result, dispensing with the assumption that schedules are contingent requires techniques different from those used with models that are based on contingent demands.

Consider a market with the same traders I and assets K as in the previous section. For simplicity, each asset is traded in a separate exchange. To isolate the decentralized market effects stemming from limited conditioning and not limited participation, assume that all traders participate in each of the K exchanges. $(N = K, K(k) = \{k\}, I(k) = I.)$ A trader submits an uncontingent demand schedule $q_k^i(p_k) : \mathbb{R} \to \mathbb{R}$ in each exchange $\{k\}$ to maximize his expected payoff pointwise for any price $p_k \in \mathbb{R}$:

$$\max_{\{q_k^i(\cdot):\mathbb{R}\to\mathbb{R}\}_k} E[\theta \cdot (q^i + q_0^i) - \frac{\alpha^i}{2}(q^i + q_0^i) \cdot \Sigma(q^i + q_0^i) - p \cdot q^i | q_0^i]. \tag{63}$$

Equilibrium prices of assets are determined exchange by exchange (here, asset by asset): setting the aggregate net demand for each asset k equal to zero, $\sum_i q_k^i(p_k) = 0$, determines the equilibrium price p_k.[103] The following changes in equilibrium (Proposition 6, Appendix C.2) relative to complete conditioning are crucial for efficiency and design.

- *Equilibrium is not* ex post.

With uncontingent trading, the demand submitted by a trader in the exchange for asset k cannot condition on realized trades and instead depends on his *expected trades* $E[q_\ell^i | p_k, q_0^i]$ and expected prices $E[p_\ell | q_0^i, p_k]$ of other assets $\ell \neq k$. Therefore, equilibrium is not *ex post* except when asset payoffs are independent (i.e., the utility Hessian Σ is separable; a trader's marginal utility for asset k is independent of the quantities traded of other assets). The dependence of the demand for asset k on expected trade of other assets induces cross-asset inference effects (and adverse selection).[104] The inference coefficient $c_{p,k\ell}^i \equiv \frac{Cov(p_\ell, p_k | q_0^i)}{Var(p_k | q_0^i)}$ is non-zero even when traders' privately known endowments $q_{k,0}^i$ are independent across assets; it captures correlated price risk in prices p_ℓ and p_k, given trader i's private information. Inference plays a key role in determining the equilibrium price impacts.

- *Price impact within and across exchanges differs from that with contingent trading.*

[103] With contingent demands, the K assets clear *simultaneously*: the equilibrium price vector is determined by $\sum_i q^{i,c}(p_1, \cdots, p_K) = 0 \in \mathbb{R}^K$. With either type of schedule, trader i receives trade $\{q_k^i\}_k$ and pays $\sum_k p_k q_k^i$.

[104] Following a trader's demand increase for asset k, other traders $j \neq i$ attribute the increase in price p_k to a low realization of endowments instead (for all assets, as long as asset payoffs are not independent, i.e., Σ is not diagonal). Expecting higher prices of positively correlated assets and lower prices of negatively correlated assets, traders $j \neq i$ adjust the price at which they are willing to sell units of asset k.

Because traders submit demands for each asset conditioned only on the price of that asset, the cross-asset price impacts are zero: $\lambda^i_{k\ell} \equiv \frac{dp_\ell}{dq^i_k} = 0$ for all $\ell \neq k$ and k, for all i. The price impact λ^i_{kk} (simply, λ^i_k) $\equiv \frac{dp_k}{dq^i_k}$ in each exchange $\{k\}$ is the slope of the residual inverse supply function for asset k:

$$\lambda^i_k \equiv \frac{dp_k}{dq^i_k} = -\left(\sum_{j\neq i} \frac{\partial q^j_k(p_k)}{\partial p_k}\right)^{-1} \qquad \forall i. \tag{64}$$

In equilibrium, the per-unit price impact within exchanges differs from its contingent-demand counterpart unless asset payoffs are independent (i.e., inference does not matter) or perfectly correlated (i.e., inference is perfect). This is due to the inference effect across assets present when the same traders participate in different exchanges. (In markets with K assets, price impact increases maximally K-fold — when asset payoffs are perfectly correlated (cf. Equation (67))).

Even though the cross-exchange price impact is zero, equilibrium prices and allocations are not independent across exchanges — then, the market would operate as a collection of independent exchanges — unless traders' utility Hessian Σ is separable (i.e., asset payoffs are independent).

6.4 Classical Results Revisited: Decentralized Markets

When trading is decentralized, traders are not negligible in the trading relationships in which they participate. In market structures characterized by incomplete participation or incomplete conditioning, price impact is positive under general conditions. Therefore, imperfectly competitive strategic behavior (i.e., price impact) is inherent to the analysis of equilibrium and the design of decentralized trading environments. Allowing (not assuming) noncompetitive behavior turns out to be important for considering the implications of decentralized trading.

6.4.1 Equilibrium

- *Decentralized trading changes the link between the fundamental risk (covariance) and demands.*

In centralized markets, traders' equilibrium price impact Λ^i is *proportional* to the assets' fundamental covariance $V = \Sigma$, because the aggregate inverse demand slope Ψ^{-1} is (Equation (14)). When the market is decentralized (i.e., participation or demand conditioning is incomplete), price impact is generally not proportional to fundamental risk. The non-proportionality has important implications for thinking about equilibrium and design.[105]

[105] More precisely, what matters is not that the link between the covariance Σ and price impact is proportional but that the relation which holds when trading is centralized does not hold in decentralized markets.

With limited participation, since price impact depends on the *projected risk* in $((\Psi^{-1})_{K(i)})^{-1}$ rather than the fundamental risk Σ (Equation (62)), it is not proportional to but *concave* in Σ.[106] Intuitively, traders' incentives depend on the riskiness of the assets in an exchange in which trader i participates through the *residual risk* of these assets (via $\Psi_{i,-i}\Psi_{-i,-i}^{-1}\Psi_{i,-i}'$), net of the risk that can be diversified in other exchanges by i and/or his counterparties. Notably, the projection that defines the residual risk of the assets $K(i)$ traded by agent i in Equation (62) is endogenous, as it depends on price impacts. Essentially, trading relationships transform the risk that traders are subject to. This observation further qualifies the role of concavity (of the primitive utility functions or due to primitive risk) as the key structure underlying equilibrium.

- *Demand substitutability for assets is endogenous and heterogeneous across traders.*

Consequently, demand substitutability for the same assets (defined by the cross-asset elements of the demand Jacobian: $\frac{\partial q^i}{\partial p_{K(i)}} = -(\alpha^i V_{K(i)} + \Lambda^i)^{-1}$ for limited participation or $\frac{\partial q^i}{\partial p} = diag(s_k^i)$ for limited demand conditioning; see Appendix C.2) is endogenous and differs from the payoff substitutability of the assets (defined by the exogenous covariance). Moreover, it is heterogeneous among traders who participate in different exchanges. For instance, assets that are payoff substitutes can act as demand complements; assets that are demand substitutes for some traders can act as demand complements for others (see Example 5 in Appendix C.1, part (1)). The endogeneity of traders' demand substitutability with respect to market structure indicates new (relative to the centralized markets) possibilities for impacting efficiency or profits by introducing alternative types of trading venues or assets. For instance, an exchange with restricted participation can Pareto dominate an innovation via a public exchange in which all traders participate.

- *Aggregate (systemic) risk depends on the market structure.*

This is another consequence of the lack of proportionality: the weights $\{(\alpha^j \mathbf{V}_{K(j)} + \mathbf{\Lambda}^j)^{-1}\alpha^j \mathbf{V}_{K(j)}$ on traders' endowments in the aggregate endowment are not proportional across traders in markets with incomplete participation. Equilibrium price and allocations are:

$$p = \theta - Q, \tag{65}$$

$$q^i + q_0^i = (\alpha^i V_{K(i)} + \Lambda^i)^{-1} Q_{K(i)} + (\alpha^i V_{K(i)} + \Lambda^i)^{-1}\Lambda^i q_0^i \qquad \forall i, \tag{66}$$

where the aggregate risk Q is determined equivalently to Section 4:

$$Q = (\sum_j (\alpha^j \mathbf{V}_{K(j)} + \mathbf{\Lambda}^j)^{-1})^{-1} \sum_j (\alpha^j \mathbf{V}_{K(i)} + \mathbf{\Lambda}^j)^{-1}\alpha^j \mathbf{V}_{K(j)} \mathbf{q}_0^j.$$

[106] Equation (62) is equivalent to $((\Lambda^i)^{-1} + (\alpha^i V_{K(i)} + \Lambda^i)^{-1})^{-1} = (\Psi^{-1})_{K(i)}$. Since the harmonic mean function $f(x) = (x^{-1} + (v+x)^{-1})^{-1}$ is convex, its functional inverse is concave; hence, Λ^i is (weakly) concave as a function of Ψ^{-1}.

The corresponding decompositions with incomplete demand conditioning depend on inference coefficients (analogously to the model with interdependent values, Section 4.3); see Appendix C.2.[107]

- *A decentralized market can function like the centralized market (i.e., under conditions, equilibrium is as if trading were centralized).*

Theorem 1 provides sufficient conditions on the market structure itself under which the equivalence holds.[108]

Theorem 1 (Equilibrium Equivalence) Consider a market $\{(I(n), K(n))\}_n$ for traders I and assets K. Traders' equilibrium utilities are the same as in the centralized market if one of the following conditions holds:

(i) (Incomplete participation) Demands are contingent, i.e., each trader submits a demand contingent on the prices of assets in the exchanges in which he participates, and each pair of exchanges has at least two common participating traders: $|I(n) \cap I(n')| > 1$.

(ii) (Incomplete conditioning) For every pair of assets k, ℓ, there is a derivative whose payoff r_d is a linear combination of the payoffs r_k and r_ℓ and does not replicate r_k or r_ℓ.[109]

Even if the market structure does not contain a single exchange for all traders and assets, or if trading of all assets clears independently, the equilibrium prices, trades, and price impacts are as if there were a single exchange for all traders and assets in which the trading of all assets clears jointly. Consider first markets with incomplete participation: When two exchanges are connected by one trader (a monopolist, i.e., $I(n) \cap I(n') = \{i\}$), prices will generally differ between these two exchanges unless the slopes of the trader's residual supplies are the same in exchanges n and n'. By Equation (62), the cross-exchange price impacts are zero, because the off-diagonal elements of $(\alpha^j \mathbf{V}_{K(j)} + \mathbf{\Lambda}^j)^{-1}$ corresponding to $(K(n), K(n'))$ are zero for all $j \neq i$ (see Example 5 in Appendix C.1, part (2)). When two agents trade in both exchanges n and n', they equalize the within-exchange with across-exchange price impacts for *all* traders in n and n':

$$\Lambda^i = \Lambda^{i,c} \otimes \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$$

[107] Corresponding weights $\{b_k^i\}_{i,k}$ in the aggregate risk are proportional neither across traders nor across assets (Equation (143), Appendix C.2).

[108] Necessary and sufficient conditions for the equivalence of equilibria between the centralized and decentralized markets is given by Malamud and Rostek (2017) for incomplete participation and by Rostek and Yoon (2018b, 2020) for incomplete demand conditioning.

[109] I.e., there are at least $\frac{K(K-1)}{2}$ linearly independent non-replicating derivatives whose payoffs are linear combinations of those of the K assets: $R_d = (r_d)_d = W \cdot R$ for a rectangular matrix $W \in \mathbb{R}^{K \times (K(K-1)/2)}$ whose rows are not mutually proportional.

where $\Lambda^{i,c}$ is the equilibrium price impact if all traders from n and n' participate in a single exchange for assets $K(n) \cap K(n')$ instead and $\otimes$ denotes the Kronecker product. Thus, equilibrium price impacts per unit of each asset in $K(n) \cap K(n')$ are the same as in the centralized market for the same traders and assets. Equilibrium prices of the asset are equalized between exchanges and are the same as if these assets were traded in a single exchange, irrespective of what other exchanges exist (see Example 5 in Appendix C.1, part (3)).

The argument differs from standard Bertrand-style price equalization: *prices equalize because price impacts do.*[110] *In general, prices do not coincide with competitive ones,* as they would in the (centralized) exchange for the corresponding traders and assets. The Bertrand result for competitive (i.e., zero-price-impact) games asserts that two traders suffice for the price and the outcome to be competitive. In demand submission games, two common participating traders in different exchanges suffice for price impacts (and outcomes) to coincide with the *centralized* trading price impacts for the traders in these exchanges. If equilibrium price impacts are zero, the classic result obtains: prices equalize and are competitive.

With incomplete demand conditioning, when the same asset (say, 1) is traded as part of $K-1$ derivatives whose payoffs bundle those of other assets as well, the prices of these derivatives yield different linear combinations of the random variables (i.e., aggregate endowments, $\bar{q}_{k,0}, k \neq 1$ in addition to $\bar{q}_{1,0}$, since a trader is uncertain about other traders' endowments). With at least K such conditioning variables fo all assets, inference about all K random variables $\{\bar{q}_{k,0}\}_k$ is perfect. Equilibrium is as if traders could condition their demand for each asset on the price vector even if in each exchange, the demand conditions only on a subset of prices of all assets (i.e., $K(n) < K$ for all n), and so no expectations of trade $\{E[q^i_\ell | q^i_0, p_k]\}_{\ell \neq k}$ are perfect.[111]

The equivalence conditions suggest which changes in the market or asset structure would be neutral for equilibrium. For example, condition (ii) puts a bound on the number of derivatives that can be introduced in a market and be nonredundant.

[110] In noncompetitive markets, the relevant notion of an absence of arbitrage is based on equalization of price impact rather than prices; the linearity of prices is implied.

[111] With perfectly correlated assets k and $\ell \neq k$, traders' inference on prices and trades are perfect: $E[p_\ell | q^i_0, p_k] = p_\ell$ and $E[q^i_\ell | q^i_0, q^i_k] = q^i_\ell$, and the uncontingent market implement *ex post* equilibrium. In this case, price impacts mimic cross-asset price impact with contingent trading. A unilateral demand increase in asset k by a trader is absorbed by other traders at a price concession based on those traders' (expected) demand changes in both asset k and ℓ. Thus, the price impact is twice that in the contingent market:

$$\Lambda^i = \Lambda^{i,c} \otimes \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}.$$

The price impact per unit of each asset traded in equilibrium (Equation (67)) is the same as in the centralized market. More generally, any market with uncontingent demands that implements *ex post* equilibrium yields the same utilities as the contingent demands do. See Example 6 in Appendix C.2.

6.4.2 Financial innovation, innovation in market design

- *Decentralized trading weakens the role of spanning (in the sense of representing risk through state-contingent securities) for asset pricing and financial innovation.*[112]

In the standard competitive theory, which is based on contingent bidding, securities introduced in the span of the existing assets (i.e., derivatives, whose returns are linear combinations of the returns of existing assets) are redundant: their introduction does not change equilibrium utilities. Diversification of risk does not depend on the asset structure, only on the asset span. When demands are uncontingent, spanning does not hold. Introducing a new asset with payoffs in the span of the existing assets is never redundant. A new asset impacts traders' payoffs even if it is in zero net supply and even if the returns of the underlying assets are independent.[113]

The equivalence between the equilibrium fixed point in demand schedules and price impacts implies the following general observation (cf. Equation (67)).

Corollary 1 (Nonredundancy: Condition on Price Impact) In imperfectly competitive markets, derivatives (or other changes in the market or asset structure) are nonredundant if and only if they change the equilibrium price impact per unit of the existing assets.

Suppose that new securities (derivatives) K' in the span of assets K are introduced to be traded along with the existing assets, each in a separate exchange. The return of each security $r_{k'}$ for all $k' \in K'$ is a linear combination of those of assets $R = (r_k)_k$; i.e., $r_{k'} = \omega_{k'} \cdot R$ for some $\omega_{k'} \in \mathbb{R}^K$. A new security thus either replicates an asset $k \in K$ or is a bundle of assets in K. Equilibrium price impact $\Lambda \in \mathbb{R}^{(K+K')\times(K+K')}$ can be decomposed into impacts for assets K and new securities K':

$$\Lambda = \begin{bmatrix} \Lambda_K & 0 \\ 0 & \Lambda_{K'} \end{bmatrix}.$$

The *projected price impact* $\widetilde{\Lambda}_K \in \mathbb{R}^{K\times K}$ determines whether derivatives are non-redundant and how they affect trading costs (see Definition 10 in Appendix C.2):

$$\widetilde{\Lambda}_K = (\Lambda_K^{-1} + W\Lambda_{K'}^{-1}W')^{-1} = \Lambda_K - (\Lambda_K W(\Lambda_{K'} + W'\Lambda_K W)^{-1}W'\Lambda_K), \tag{67}$$

[112] For a general reference on financial innovation, see, e.g., Allen and Gale (1994) and Tufano (2003).

[113] In centralized trading theory, a large literature has studied asset innovation outside the span of existing assets K. When trading is decentralized, introducing a security *in* the span is generally not redundant. It is well understood that in imperfectly competitive markets, introducing an asset in non-zero net supply would change prices and utilities. The argument underlying the results for decentralized markets is new — it holds for zero net supply.

where $W \equiv (\omega_1, \cdots, \omega_{K'}) \in \mathbb{R}^{K \times K'}$ is the weight matrix that defines the securities, each column $\omega_{k'}$ of which represents how the return of a security is correlated with underlying assets.

Decentralized trading thus motivates a theory of financial market design with instruments that would be redundant if trading were centralized.

Theorem 2 (Spanning: Conditions on the Primitives) Consider a market $\{(I(n), K(n))$ and suppose that equilibrium payoffs are not equivalent to those in the centralized market for the same traders and assets. Introducing derivatives in the span of the existing assets does not change traders' payoffs if:

(i) The payoffs of all assets are perfectly correlated, (i.e., $\sigma_{k\ell} = 1$ or -1 for all k, $\ell \neq k$); or the payoff of all assets are independent (i.e., $\sigma_{k\ell} = 0$ for all k, $\ell \neq k$) and each derivative k' is perfectly correlated with an existing asset (i.e., $r_{k'} = r_k$ for some k); and

(ii) Traders' values $\{\theta^i_k\}_{k,i}$ and endowments $\{q^i_{k,0}\}_{k,i}$ are independent across traders and assets.

Under these conditions, the introduction of derivatives is neutral for equilibrium payoffs; the neutrality extends to other innovations such as new exchanges or changes in market clearing. Spanning holds because these innovations do not affect traders' inference or, *equivalently*, price impact.

Conditions (i) and (ii) ensure that prices of derivatives are not informative about the values of assets. In practice, other than for tax-related purposes, derivatives are traded to diversify payoffs in contingencies not covered by the existing assets (due to some form of market incompleteness) or to improve liquidity. When demands are contingent, the classical spanning results still hold: the introduction of securities in the span is redundant. Under conditions (i) and (ii), derivatives are redundant in non-competitive as well as competitive markets. In either one's absence, derivatives are traded to improve liquidity or reduce informational error.

If trading were centralized, the following classes of innovations would be redundant:

(1) The creation of new trading venues for *existing* assets (which with limited conditioning are not redundant even with participation of the same traders);

(2) The introduction of new assets in the span of existing assets;[114] and

(3) Innovation in the types of market clearing that decentralized trading enables (e.g., netting, CCPs, asset deconsolidation, ring-fencing of investment banking units).

[114] E.g., exchange-traded products (ETPs), exchange-traded funds (ETFs), derivatives, and other financial instruments whose rate of innovation and trading activity continue to be high.

These instruments have been explicitly promoted or discouraged by changes in regulation following the recent financial crisis (e.g., the Dodd-Frank Act and MiFID I and II).

- *Decentralized trading motivates new types of financial innovation that are not based on spanning, as well as new types of innovation that are.*

In the competitive centralized market, a financial innovation outside the span of the traded assets is always weakly Pareto improving with quasilinear utility; this result extends to imperfectly competitive markets. It is well known that in centralized markets with nonquasilinear utility, the introduction of an unspanned asset may increase or decrease welfare if it *changes relative prices* of the assets traded. When trading is decentralized, the welfare effects of financial innovation (i.e., introducing an unspanned asset or an exchange for a spanned asset) can increase or decrease welfare even with quasilinear utility if it *changes relative price impacts* for the traded assets.[115]

- *Financial-market design.*

The literature has already demonstrated that trader heterogeneity matters for bidding behavior and equilibrium outcomes, and that it can overturn implications for policy or design. For instance, increasing access to information and encouraging participation are no longer seen as tenets of design. The literature has only begun to explore the conditions under which efficiency aligns with particular forms of market decentralization, such as restricting trader participation, specialization in trading certain assets, requiring that an asset be cleared OTC rather than in an exchange, breaking up exchanges, asset deconsolidation, and demergers.

New data has exposed a wide range of specialized and unique challenges that shape market design — design of information sharing, market structure, contracts, and market mechanisms. Understanding of the properties of such arrangements is rapidly evolving. For instance, Duffie and Zhu (2017) investigate the impact of adding a fixed-price trading mechanism called 'workup' (a form of size discovery) to a double auction. Although the fixed price may be biased, it mitigates the strategic avoidance of price impact. A suitably arranged size discovery trading mechanism can improve efficiency. Antill and Duffie (2017) show that overall allocative efficiency is reduced by augmenting a double auction with size discovery, except when a size-discovery session is conducted before the exchange market opens, as shown by Duffie and Zhu (2017). Duffie and Zhu (2017) analyze auction design for CDS settlement. Hortaçsu and Kastl (2012) and Boyarchenko, Lucca, and Veldkamp (2018) examine the effects of information sharing arrangements among financial intermediaries.

[115]This observation was first made by Malamud and Rostek (2017) regarding innovations of new exchanges for existing assets based on spanning. Given the equivalence of the equilibrium fixed points in demand schedules and price impacts, the non-neutrality of any type of innovation (e.g., new assets, exchanges or other types of venues, market clearing protocols) can be verified and evaluated in terms of changes in the relative per-unit (e.g., projected) price impacts.

Recent events consistent with the manipulation of foreign exchange rates and LIBOR have raised concerns about the design of financial benchmarks. The relevant framework for evaluating the current design of benchmarks is one with imperfectly competitive and decentralized trading (see Duffie and Dworczak (2014), Duffie, Dworczak, Zhu (2017), Duffie (2018), and Zhang (2019)). Investors rely on benchmarks to improve valuation of the assets they trade; the traders who hold related assets and whose demand determines the value of the benchmark have an incentive to manipulate the value of the benchmark.

6.4.3 Welfare: centralized vs. decentralized markets

Many regulators have raised concerns about the effects of market fragmentation, particularly in the wake of the financial crisis. Until recently, the academic literature on financial markets has emphasized important frictions associated with decentralization of trade: a search friction in the time needed to find a counterparty, counterparty risk externality, and a friction in aggregating information. A typical comparative statics in the literature asserts that as the friction vanishes, the limit market is the efficient centralized market.[116] Thus, the very way decentralization is introduced — as a friction in a competitive market — presupposes that a decentralized market is inefficient.[117]

Now, an alternative view of decentralized trading is suggested by several arguments. If feasible, would centralized trading be desirable, given the traders and assets, taking efficiency as the objective? Decentralized trading may improve traders' learning about the asset value (Rostek and Weretka (2015b), Babus and Kondor (2018),[118] Glode and Opp (2016), Even, Tahbaz-Salehi, and Vives (2018), Yoon (2018), and Glebkin (2019)) or asset price (Zhu (2014))[119]; redistribute risk towards less risk averse traders (Malamud and Rostek (2017)); improve risk sharing and diversification of risk across assets (Rostek and Yoon (2018b, 2020)); or be more stable than the centralized market (Peivandi and Vohra (2017)).[120]

[116]Exceptions include Farboodi, Jarosch, and Menzio (2017) and Farboodi, Jarosch, and Shimer (2017).

[117]In the early academic literature, "the fact that many markets are fragmented and remain so for a long period of time, despite strong economic arguments for consolidation" became known as "the network externality puzzle" (Madhavan (2000)). As Cantillon and Yin (2011) pointed out, what makes fragmentation even more puzzling is the many unsuccessful attempts by exchanges to organize trading in assets already traded at other exchanges. Another early version of the argument against decentralized trading is that decentralized bargaining may produce inefficient outcomes because local incentives of traders do not necessarily align with global efficiency. The literature reviewed in this section has shown that decentralized trading may instead improve efficiency precisely because local incentives for trade can better align individual traders' needs and behavior with global efficiency. (See also Cantillon and Yin (2011).)

[118]In the model of Babus and Kondor (2018), information is aggregated in the decentralized market. This allows the authors to investigate the relative efficiency of centralized vs. decentralized trading for assets whose values are closer to independent private or the common values.

[119]Nava (2015) makes a related point in the Cournot model.

[120]Results from the theory of *centralized* markets with asymmetric information can be interpreted as suggesting reasons for market decentralization when trade is motivated not only by diversification but also by asymmetric information (e.g., in Akerlof (1970)). Adverse selection may incentivize trade in a separate market.

Imperfectly competitive models of decentralized trading (e.g., those in Sections 6.3.1 and 6.3.2) provide two general observations:

- Market fragmentation may improve welfare relative to centralized trading by lowering the total trading costs that stem from price impact[121] within and/or across exchanges. Because of the equivalence of the equilibrium fixed point in demand schedules and price impacts (Proposition 1 in Section 4), welfare comparative statics can be analyzed through changes in the structure of price impact itself (projected price impacts, see Section 6.4.5), which, in particular, accounts for changes in the aggregate risk.

- In fact, if exchanges are suitably designed, decentralized markets are weakly more efficient than centralized markets. This follows from the equivalence of equilibrium across market structures (Theorem 1).

Why might one expect markets with multiple exchanges for assets with *homogeneous* units to strictly increase welfare relative to the centralized market? Restrictions on trader participation and conditioning (i.e., models in Sections 6.3.1 and 6.3.2) isolate the corresponding welfare effects. Restricting trader participation to trading only some assets or interacting with some classes of investors (while allowing contingent schedules) (1) can improve *allocation of risk across traders* provided that traders' risk preferences differ. Moreover, (2) given that a more decentralized market always lowers liquidity (increases price impact, see Section C.1), the redistribution of risk towards less risk averse traders is a *necessary* condition for a more decentralized market to improve welfare (with contingent demands). The welfare impact of changes in the market structure can thus be understood through the trade-off between price impact and aggregate risk (cf. Equation (68)). Namely, substituting the equilibrium allocation (66) into the utility function (125) gives the indirect utility function as a function of price impact and aggregate risk:

$$V^i(q^i + q_0^i) = \underbrace{U^i(q_0^i)}_{\text{Utility from no trade}} + \underbrace{(Q_{K(i)} - \alpha^i V_{K(i)} q_0^i) \cdot \Upsilon^i(\Lambda^i)(Q_{K(i)} - \alpha^i V_{K(i)} q_0^i)}_{\text{Utility surplus from trade}}, \tag{68}$$

where

$$\Upsilon^i(\Lambda^i) \equiv (\alpha^i V_{K(i)} + \Lambda^i)^{-1}(\frac{1}{2}\alpha^i V_{K(i)} + \Lambda^i)(\alpha^i V_{K(i)} + \Lambda^i)^{-1}.$$

In the centralized market, the proportionality of price impact $\Lambda^{i,c} = \alpha^i \beta^{i,c} \Sigma$ to the payoff

[121] This is not to say the price impact itself decreases. Due to the non-proportionality between price impact and risk in decentralized markets, the relationship between price impact and the indirect utility function (which is monotone in the centralized markets) is not necessarily monotone when trading is decentralized. See Equation (68) below. In the centralized market, the proportionality of price impact and risk (i.e., $\Lambda^i = \alpha^i \beta^i \Sigma$) implies that the aggregate risk $Q = (\sum_j (\alpha^j)^{-1})^{-1} \sum_j \Sigma q_0^j$ is proportional to Σ, and that the indirect utility is decreasing in price impact β^i.

covariance Σ (i.e., the counterpart of $V_{K(i)}$) simplifies $\Upsilon^i(\Lambda^i)$ to:

$$\Upsilon^i(\Lambda^{i,c}) = (\alpha^i\Sigma + \Lambda^{i,c})^{-1}(\frac{1}{2}\alpha^i\Sigma + \Lambda^{i,c})(\alpha^i\Sigma + \Lambda^{i,c})^{-1} = \frac{\frac{1}{2} + \beta^{i,c}}{(1+\beta^{i,c})^2}(\alpha^i\Sigma)^{-1}.$$

(3) Trader heterogeneity matters: When all traders' risk preferences are symmetric, the centralized market provides higher welfare than any decentralized market structure. Welfare gains from decentralization exist even if the number of traders is arbitrarily large — only for a few market participants need to be sufficiently heterogeneous (e.g., in terms of their risk preferences), and endowments need not become more extreme as the market grows. In turn, restrictions on demand conditioning (4) can lower the cost of trading due to either the within- or across-exchange price impact, or both. Accordingly, such restrictions *can improve both risk sharing and diversification of risk across assets* in ways not feasible with contingent trading or heterogeneous participation, *even if traders' risk preferences are symmetric.*

When viewed through the lens of game-theoretic analysis, the fact that a decentralized market may lead to higher welfare is not in itself "surprising." Developing principles of market design, regulation, and financial market antitrust requires an understanding of *how* market structure affects incentives. There is substantial scope for exploring the comparative advantage of different types of innovation — e.g., financial innovation, exchange design, and alternative market-design arrangements — in improving liquidity or markets' effectiveness in facilitating risk sharing and diversification. Apart from shedding light on the effects of different types of innovations, such results could contribute to the evaluation of (de)mergers or specialization in trading certain assets.

6.4.4 Endogenous market structure

- *Decentralized trading suggests a theory to explain which markets are open and which are not.*[122]

This issue was not resolved in the general equilibrium with incomplete markets (see, e.g., the surveys by Geanakoplos (1990), Duffie and Rahi (1995), and Bisin (1998)). A general theory of endogenous market structures — based on equilibrium (i.e., how the market structure evolves) or stability (i.e., what is the market structure in which no further adjustments occur) — awaits further advances. For instance, accommodating complementary agreements in matching models alongside substitutable ones remains an open problem. The econometric issues associated with endogeneity and identification have long been studied (see, e.g., the survey by Durlauf, Blume, Brock and Ioannides (2011)). Much remains to be understood

[122]Missing markets and the resulting exposure to systemic risk have been central to discussions of the 2008 financial crisis. To monitor financial innovation, in September 2009, the Security and Exchange Commission (SEC) created the Division of Risk, Strategy and Financial Innovation (its first new division in 37 years).

regarding precisely which theoretical results concerning equilibrium or design will change *qualitatively.*

Existing welfare results for decentralized markets offer several insights.

- When markets are imperfectly competitive *and* decentralized in some sense, there are individual or social incentives to limit participation of certain traders in some exchanges (e.g., the model from Section 6.3.1) or not to introduce certain assets (e.g., the model from Section 6.3.2). Thus, although a general characterization of endogenous market structures has yet to be provided, it follows from the existing literature that if traders can choose which assets they trade and with whom, the endogenous market structure will not be centralized in general. For instance, some traders have an incentive to establish a new venue to accept higher price impact for improved aggregate risk (i.e., a trade-off between Λ^i and Q (or $q^{i,*}$), e.g., Babus and Parlatore (2017)) or to benefit from lower price impact or improved price informativeness, or both (e.g., Yoon (2018)).

- The study of endogenous formation of exchanges in decentralized markets — with respect to welfare and other objectives — should not be separated from innovation in exchange design, assets and market clearing. Babus and Hachem (2019) examine how the securities issued and the market structure are endogenously determined to explain why standardized securities are frequently traded in decentralized markets. They show that financial intermediaries have an incentive to issue debt in markets with fewer investors (larger price impact) and issue equity in larger markets. Investors choose to trade at a larger price impact cost in order to benefit from less risky securities from issuers (i.e., there is a trade-off between price impact Λ^i — impacted by the number of traders — and Σ).

- The sub-problem of choosing contingent vs. uncontingent schedules in endogenizing the market structure is straightforward. If traders themselves were choosing whether to submit an uncontingent or contingent schedule, the choice of the contingent schedule follows by individual optimization, taking as given the schedules submitted by others. Submission of contingent schedules by all traders would be the unique outcome. Therefore, the welfare enhancements possible with limited conditioning require that the cross-exchange condition be restricted. In practice, implementation of contingent trading would require coordination of market clearing for all traders and assets (e.g., RegNMS and UTS rules in the US stock market).

Given that the trader's allocation (and, hence, indirect utility) depends on equilibrium behavior through price impact Λ^i and the aggregate risk Q (Equation (66)), the welfare effects of changes in the market structure can be understood through the effects on these variables.

6.4.5 Methods

- *Classical methods for equilibrium analysis, which are based on spanning, do not apply in decentralized markets.*

Pricing of derivatives is typically done (and taught) either using arbitrage — i.e., assuming the existence of a replicating portfolio (the payoff of a derivative can be created synthetically by trading available assets), hence assuming that a derivative is redundant — or in a partial-equilibrium model with a monopolistic trader who innovates. Derivatives markets are often seen by regulators as particularly challenging, partly due to the difficulty of assessing the impact and the scope for innovating nonredundant derivatives or their efficiency benefits. The development of an *equilibrium* theory of *nonredundant* derivatives can help guide the regulation of various types of derivatives markets. An equilibrium framework helps address questions such as this: Given the structure of the underlying assets, which derivatives will be introduced in markets?

The classical techniques based on spanning have laid the foundation for analysis of equilibrium and asset pricing; these include the representation of uncertainty through the implied state space over which state-contingent securities are defined. These methods do not generally apply in decentralized markets. There is room to develop new theory for the pricing and valuation of financial innovation — new assets, exchanges, or contracts — based on game-theoretic tools. The approach to characterizing equilibrium in demands as a fixed point in price impacts allows for the introduction of such alternative techniques. These methods can be extended to the comparative design analysis for different instruments, based on the *projected price impacts* of the matrix representation of the corresponding fixed points.

- *Corresponding techniques for equilibrium analysis differ from those in the linear best response models with actions rather than schedules as choices.*

With nonlinear effects (due to the harmonic mean structure of the fixed point in price impacts as well the fixed point between price impact and conditional expectations) and nonlocal effects, the techniques from the linear best response models are inapplicable. For instance, summary statistics, such as the smallest eigenvalue (e.g., the *Handbook* chapter by Jackson and Zenou (2014)), do not suffice to capture the network's properties. Methods from the theory of shorted operators (e.g., Anderson (1971)) — in particular, the concavity and monotonicity properties of shorted operators — and monotone matrix functions (e.g., Donoghue (1974)) can be applied to develop comparative statics of equilibrium and welfare for decentralized markets with incomplete participation and contingent schedules.

7 Conclusions

Financial markets have not been centralized in modern history — except when regulated — nor have they been perfectly competitive. As the literature review has revealed, some of the classic results from the competitive equilibrium theory do not apply when markets are decentralized and/or imperfectly competitive. Throughout the survey, we have indicated research areas and problems where room remains for the innovation of methods to enrich equilibrium analysis of imperfectly competitive markets. Such methods will help to advance new areas of market design that are of direct practical and policy relevance.

Demand submission games offer a flexible framework for equilibrium analysis of imperfectly competitive financial markets within which one can integrate details of market design and market structure. At the same time, the framework described in this article is special in three respects: the restriction to financial markets, the use of the quadratic payoffs to make the analysis tractable, and the restriction to the uniform-price mechanism. We conclude by discussing these modeling features.

Other market mechanisms. The survey has focused on the uniform-price mechanism, since it is a prevalent mechanism in markets (as well as the mechanism on which aggregation is based in the general equilibrium model). From the theory of games with (two-sided) private information, one expects other pricing mechanisms to lead to inefficiency — which is central to why imperfect competition and market fragmentation are generally not neutral. The multitude of variants of the uniform-price market clearing mechanism itself, as well as other mechanisms used in practice, suggest that the scope for efficiency broadens as a result of mitigating specific types of inefficiency in various settings (e.g., Antil and Duffie (2017)).

As bargaining and contracting with externalities are difficult to model, a strand of the structural literature has relied on the Nash-in-Nash approach. By its virtue of allowing for different forms of conditioning in a player's strategy, the language of demand submission games has the potential to account for externalities — both when characterizing equilibrium and endogenizing the market structure.

General utilities. Essentially all of the predictive results in the literature based on demand games have come from models with quadratic payoffs. The qualitative conclusions about the impact of imperfect competition or market fragmentation which this article tends to highlight do not hinge on the assumption of quadratic utilities.

The scarcity of analytic results for general utilities[123] is one reason why, like in analysis of competitive models with heterogenous traders, econometric and computational methods

[123]With non-quadratic utilities, the marginal utility of the traders is nonlinear, i.e., the concavity of the utility depends on quantity. Hence, equilibrium price impact is not constant but a function of quantity. The characterization of equilibrium and price impact for general utilities is an open problem (see Klemperer and Meyer (1989), Glebkin (2015), Du and Zhu (2017b), Malamud and Teguia (2017), Breon-Drish (2015), Rostek and Yoon (2018d), and Holmberg, Newbery, and Ralph (2013).

have been employed to quantify the importance of imperfect competition and market fragmentation. Beginning with Hortaçsu (2002) and Hortaçsu and McAdams (2010), a growing empirical literature applies structural methods to characterize bidding behavior and market outcomes in divisible-good settings. This literature is reviewed in Hortaçsu and McAdams (2010) and Kastl (2017). The quadratic model invites CAPM-style analysis — for decentralized as well as centralized markets.

Beyond financial markets. Imperfect competition and fragmentation are increasingly recognized as important not only in financial markets. A number of recent studies have documented decreasing concentration in virtually every industry of the economy (e.g., Furman (2016), De Loecker and Eeckhout (2018), and De Loecker, Eeckhout, and Unger (2018)). The ideas, results, and methods developed in the literature on imperfectly competitive financial markets are not specific to the markets for financial assets and carry over to other contexts. Among these results are slow trading, persistence, transitory and permanent effects of shocks, separation between the real effects of events' announcements (about shocks to trades or information) and the events themselves, and the role of bilateral (buyer and seller) vs. one-sided market power. Observations about the effects pertinent to demand submission games and the Cournot model of market competition apply to other canonical models, which tend to be used in markets other than financial, including Bertrand Nash and Dixit-Stiglitz frameworks.

References

Afonso, G. and R. Lagos (2015): "Trade Dynamics in the Market for Federal Funds," Econometrica, 83, 1, 263-313.

Afonso, G., A. Kovner, and A. Schoar (2013): "Trading Partners in the Interbank Lending Market," Federal Reserve Bank of New York Staff Report 620.

Akerlof, G. A. (1970): "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," Quarterly Journal of Economics, 84, 3, 488-500.

Aldrich, E. M. and D. Friedman (2019): "Order Protection through Delayed Messaging," Working Paper.

Allen, F. and D. Gale (1994): Financial Innovation and Risk Sharing, Cambridge, MA: MIT Press.

Almgren, R. (2009): "Execution Costs," In: Encyclopedia of Quantitative Finance, Ed. R. Cont. Hoboken, NJ: Wiley.

Almgren, R. and N. Chriss (2000): "Optimal Execution of Portfolio Transactions," Journal of Risk, 3, 5-39.

Almgren, R., C. Thum, E. Hauptman, and H. Li (2005): "Equity Market Impact," Risk, July, 57-62.

Amihud, Y., H. Mendelson, and L. H. Pedersen (2005): "Liquidity and Asset Prices," Foundations and Trends in Finance, 1, 4, 269-364.

Anderson, W. N. Jr. (1971): "Shorted Operators," SIAM Journal of Applied Mathematics, 20, 3, 520-525.

Andreyanov, P. and T. Sadzik (2018): "Robust Mechanisms of Exchange," Working Paper.

Angel, J., L. Harris, and C. Spatt (2011): "Equity Trading in the 21st Century," Quarterly Journal of Finance, 1, 1-53.

Angeletos, G.-M. and Z. Huo (2018): "Myopia and Anchoring," Working Paper.

Angeletos, G.-M. and C. Lian (2018): "Forward Guidance without Common Knowledge," American Economic Review, 108, 9, 2477-2512.

Antill, S. and D. Duffie (2017): "Augmenting Markets with Mechanisms," Working Paper.

Arrow, K. (1951): "An Extension of the Basic Theorems of Classical Welfare Economics," Proceedings of the Second Berkley Symposium on Mathematical Statistics and Probability, J. Neyman (ed.), University of California Press, 507-532.

Arrow, K. and G. Debreu (1954): "Existence of an Equilibrium for a Competitive Economy," Econometrica, 22, 265-90.

Arrow, K. and F. H. Hahn (1971): General Competitive Analysis, Holden-Day, Inc., San Francisco.

Atalay, E., A. Hortaçsu, J. Roberts, and C. Syverson (2011): "Network Structure of Production," Proceedings of the National Academy of Sciences, 108, 13, 5199-5202.

Atkeson, A. G., A. L. Eisfeldt, and P.-O. Weill (2015): "Entry and Exit in OTC Derivatives Markets," Econometrica, 83, 2231-2292.

Aumann, R. and J. Dreze (2008): "Rational Expectations in Games," American Economic Review, 98, 1, 72-86.

Ausubel, L., P. Cramton, M. Pycia, M. Rostek, and M. Weretka (2014): "Demand Reduction and Inefficiency in Multiunit Auctions," Review of Economic Studies, 81, 4, 1366-1400.

Babus, A. and K. Hachem (2019): "Markets for Financial Innovation," Working Paper.

Babus, A. and P. Kondor (2018): "Trading and Information Diffusion in Over-The-Counter Markets," Econometrica, 86, 5, 1727-1769.

Babus, A. and C. Parlatore (2017): "Strategic Fragmented Markets," Working Paper.

Back, K. and Zender, J. F. (1993): "Auctions of Divisible Goods: On the Rationale for the Treasury Experiment," Review of Financial Studies, 6, 733-764.

Baldauf, M. and J. Mollner (2020): "High-Frequency Trading and Market Performance," Journal of Finance, 75, 3, 1495-1526.

Bech, M. and E. Atalay (2010): "The Topology of the Federal Funds Market," Physica A: Statistical Mechanics and its Applications, 389, 22, 5223-5246.

Berge, C. (1973): Graphs and Hypergraphs, North-Holland Mathematical Library.

Bergemann, D., T. Heumann, and S. Morris (2018): "Information and Market Power," Cowles Foundation Discussion Paper No. 2017.

Bergemann, D., T. Heumann, and S. Morris (2019): "Information, Market Power and Price Volatility," Working Paper.

Bergemann, D. and S. Morris (2012): Robust Mechanism Design: The Role of Private Information and Higher Order Beliefs, World Scientific.

Berry, S., M. Gaynor, and F. S. Morton (2019): "Do Increasing Markups Matter? Lessons from Empirical Industrial Organization," Journal of Economic Perspectives, 33, 3, 44-68.

Biais, B., C. Bisière, and C. Spatt (2010): "Imperfect Competition in Financial Markets: An Empirical Study of Island and Nasdaq," Management Science, 56, 12, 2237-2250.

Biais, B. and T. Foucault (2014): "HFT and Market quality," Bankers, Markets & Investors No 128.

Biais, B. and R. Green (2007): "The Microstructure of the Bond Market in the 20th Century," IDEI Working Paper, 482.

Bianchi F., M. Lettau, and S. Ludvigson (2020): "Monetary Policy and Asset Valuation," Working Paper.

Bisin, A. (1998): "General Equilibrium and Endogenously Incomplete Financial Markets," Journal of Economic Theory, 82, 1, 19-45.

Bonano, G. (1990): "General Equilibrium Theory with Imperfect Competition," Journal of Economic Surveys, 4, 4, 297-328.

Bonatti, A., G. Cisternas, and J. Toikka (2017): "Dynamic Oligopoly with Incomplete Information," Review of Economic Studies, 84, 2, 503-546.

Bond, P., A. Edmans, and I. Goldstein (2012): "The Real Effects of Financial Markets," Annual Reviews of Financial Economics, 4, 339-360.

Boyarchenko, N., D. Lucca, and L. Veldkamp (2018): "Taking Orders and Taking Notes: Dealer Information Sharing in Financial Markets," Working Paper.

Bramoullé, Y., A. Galeotti, and B. Rogers (2016): The Oxford Handbook of the Economics of Networks, Oxford University Press.

Breon-Drish, B. (2015): "On Existence and Uniqueness of Equilibrium in a Class of Noisy Rational Expectations Models," Review of Economic Studies, 82, 3, 868-921.

Bresnahan, T. F. (1981): "Duopoly Models with Consistent Conjectures," American Economic Review, 71, 934-945.

Brunnermeier, M. and L. Pedersen (2005): "Predatory Trading," Journal of Finance, 4, 1825-1863.

Budish, E., P. Cramton, and J. Shim (2015): "The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response," Quarterly Journal of Economics, 130, 4, 1547-1621.

Burkett, J. and B. Baisa (2018): "Large Multiunit Auctions with a Large Bidder," Journal of Economic Theory, 174, 2018, 1-15.

Burkett, J. and B. Baisa (2019): "Optimal Quantity Caps in Discriminatory Auctions with Resale," Theoretical Economics, Forthcoming.

Burkett, J. and K. Woodward (2018): "Uniform Price Auctions with a Last Accepted Bid Pricing Rule," Working Paper.

Cantillon, E. and P.-L. Yin (2011): "Competition between Exchanges: A Research Agenda," International Journal of Industrial Organization, 29, 3, 329-336.

Carvajal, A. and M. Weretka (2012): "No-arbitrage, State Prices and Trade in Thin Financial Markets," Economic Theory, 50, 1, 223-268.

Budish, E., R. Lee, and J. Shim (2019): "A Theory of Stock Exchange Competition and Innovation: Will the Market Fix the Market?" Working Paper.

Cespa, G. and X. Vives (2019): "Exchange Competition, Entry, and Welfare," Working Paper.

Chan, L. and J. Lakonishok (1993): "Institutional Trades and Intraday Stock Price Behavior," Journal of Financial Economics, 33, 2, 173-199.

Chan, L. and J. Lakonishok (1995): "The Behavior of Stock Prices around Institutional Trades," Journal of Finance, 50, 4, 1147-1174.

Chen, D. and D. Duffie (2020): "Market Fragmentation," Working Paper.

Cho, I.-K. and T. Sargent (2008): "Self-confirming equilibrium," In: S. Durlauf, and L. Blume (eds.), The New Palgrave Dictionary of Economics.

Cocco, J., F. Gomes, and N. Martins (2009): "Lending Relationships in the Interbank Market," Journal of Financial Intermediation, 18, 1, 24-48.

Collin-Dufresne, P., K. Daniel, C. Moallemi, and M. Sağlam (2015): "Dynamic Asset Allocation with Predictable Returns and Transaction Costs," Working Paper.

Corominas-Bosch, M. (2004): "Bargaining in a Network of Buyers and Sellers," Journal of Economic Theory 115, 1, 35-77.

Craig, B. and G. Peter (2014): "Interbank Tiering and Money Center Banks,"Journal of Financial Intermediation, 23, 3, 322-347.

Debreu, G. (1951): "The Coefficient of Resource Utilization," Econometrica, 19, 273-292.

Dekel, E., D. Fudenberg, and D. Levine (1999): "Payoff Information and Self-Confirming Equilibrium," Journal of Economic Theory, 89, 165-185.

De Loecker, J. and J. Eeckhout (2018): "Global Market Power," Working Paper.

De Loecker, J., J. Eeckhout, and G. Unger (2018): "The Rise of Market Power and the Macroeconomic Implications," Working Paper.

DeMarzo, P. M. and B. Urošević (2006): "Ownership Dynamics and Asset Pricing with a Large Shareholder," Journal of Political Economy, 4, 145-174.

Diamond, R. and E. Verrecchia (1981): "Information Aggregation in a Noisy Rational Expectations Economy," Journal of Financial Economics, 9, 3, 221-235.

Donoghue, W. F. Jr. (1974): Monotone Matrix Functions and Analytic Continuation, New York: Springer-Verlag.

Du, S. and H. Zhu (2017a): "What is the Optimal Trading Frequency in Financial Markets?" Review of Economic Studies, 84, 1606-1651.

Du, S. and H. Zhu (2017b): "Bilateral Trading in Divisible Double Auctions," Journal of Economic Theory, 167, 285-311.

Dubey, P., J. Geanakoplos, and M. Shubik (1987): "The Revelation of Information in Strategic Market Games: A Critique of Rational Expectations Equilibrium," Journal of Mathematical Economics, 16, 2, 105-138.

Dubey, P. and M. Shubik (1977): "A Closed Economic System with Production and Exchange Modelled as a Game of Strategy," Journal of Mathematical Economics, 4, 3, 253-288.

Duffie, D. (2010): "Asset Price Dynamics with Slow-Moving Capital," American Finance Association Presidential Address, Journal of Finance, 65, 1238-1268.

Duffie, D. (2012): Dark Markets, Princeton Lectures in Finance, Y.Aït-Sahalia (ed.), Princeton University Press.

Duffie, D. (2018): "Notes on LIBOR Conversion," Working Paper.

Duffie, D. (2018): "Post-Crisis Bank Regulations and Financial Market Liquidity," Thirteenth Paolo Baffi Lecture on Money and Finance, Banca d'Italia, Eurosystem.

Duffie, D. and P. Dworczak (2014): "Robust Benchmark Design," Working Paper.

Duffie, D., P. Dworczak, and H. Zhu (2017): "Benchmarks in Search Markets," Journal of Finance, 72, 5, 1983-2044.

Duffie, D., N. Gârleanu, and L. H. Pedersen (2005): "Over-the-Counter Markets," Econometrica, 73, 1815-1847.

Duffie, D., S. Malamud, and G. Manso (2009): "Information Percolation with Equilibrium Search Dynamics," Econometrica, 77, 1513-1574.

Duffie, D., S. Malamud, and G. Manso (2014): "Information Percolation in Segmented Markets," Journal of Economic Theory, 153, 1-32.

Duffie, D. and R. Rahi (1995): "Financial Market Innovation and Security Design: An Introduction," Journal of Economic Theory, 65, 1-42.

Duffie, D. and H. Sonnenshein (1989): "Arrow and General Equilibrium Theory," Journal of Economic Literature, 27, 2, 565-598.

Duffie, D. and H. Zhu (2017): "Size Discovery," Review of Financial Studies, 30, 4, 1095-1150.

Durlauf, S., L. Blume, W. Brock, and Y. Ioannides (2011): "Identification of Social Interactions," in Handbook of Social Economics, J. Benhabib, A. Bisin, and M. Jackson, eds., Amsterdam: North Holland.

Elliott, M. (2015): "Inefficiencies in Networked Markets," American Economic Journal: Microeconomics, 7, 4, 43-82.

Even, Y., A. Tahbaz-Salehi, and X. Vives (2018): "Information and Learning in Heterogeneous Markets," Working Paper.

Farhi, E. and I. Werning (2017): "Monetary Policy, Bounded Rationality and Incomplete Markets," Working Paper.

Farboodi, M., G. Jarosch, and M. G. Menzio (2017): "Intermediation as Rent Extraction," Working Paper.

Farboodi, M., G. Jarosch, and R. Shimer (2017): "The Emergence of Market Structure," Working Paper.

Figuières, C., A. Jean-Marie, N. Quérou, and M. Tidball (2004): Theory of Conjectural Variations, World Scientific.

Financial Stability Board (2010): "Implementing OTC Derivatives Market Reforms," Report, 25 October 2010.

Foucault, T. (2010): "Limit order markets," in Encyclopedia of Quantitative Finance, Vol. III, p.1057: Wiley.

Foucault, T. and S. Moinas (2018): "Is trading fast dangerous," Global Algorithmic Capital Markets: High Frequency Trading, Dark Pools, and Regulatory Challenges, Forthcoming.

Frazzini, A., R. Israel, and T. Moskowitz (2018): "Trading Costs," Working Paper.

Furman, J. (2016): "Beyond Antitrust: The Role of Competition Policy in Promoting Inclusive Growth," Searle Center Conference on Antitrust Economics and Competition Policy, Chicago, IL, September 16, 2016.

Fudenberg, D. and D. Levine (1993): "Self-Confirming Equilibrium," Econometrica, 61, 523-546.

Fudenberg, D. and D. Levine (2009): "Self Confirming Equilibrium and the Lucas Critique," Journal of Economic Theory, 144, 2354-2371.

Gabaix, X. (2019): "A Behavioral New Keynesian Model," Working Paper.

Gale, D. (1986a): "Bargaining and Competition Part I: Characterization," Econometrica, 54, 785-806.

Gale, D. (1986b): "Bargaining and Competition Part II: Existence," Econometrica, 54, 807-818.

Gale, D. (2000): Strategic Foundations of General Equilibrium: Dynamic Matching and Bargaining Games, Cambridge and New York: Cambridge University Press.

Gârleanu, N. and L. Pedersen (2013): "Dynamic Trading with Predictable Returns and Transaction Costs," Journal of Finance, 68, 6: 2309-2340.

Gârleanu, N. and L. Pedersen (2016): "Dynamic Portfolio Choice with Frictions," Journal of Economic Theory, 165, 487-516.

Gary-Bobo, R. J. (1989): "Cournot-Walras and Locally Consistent Equilibria," Journal of Economic Theory, 49, 1, 10-32.

Geanakoplos, J. (1990): "An Introduction to General Equilibrium with Incomplete Asset Markets," Journal of Mathematical Economics, 19, 1-2, 1-38.

Glebkin, S. (2015): "Strategic Trading without Normality," Working Paper.

Glebkin, S. (2019): "Liquidity versus Information Efficiency," Working Paper.

Glode, V. and C. Opp (2016): "Asymmetric Information and Intermediation Chains," American Economic Review, 106, 9, 2699-2721.

Glosten, L. and P. Milgrom (1985): "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders," Journal of Financial Economics, 14, 1, 71-100.

Grandmont, J.-M.. (1977): "Temporary General Equilibrium Theory," Econometrica, 45, 535-572.

Grandmont, J.-M.. (1982): "Temporary General Equilibrium Theory," In: Handbook of Mathematical Economics, K.J. Arrow and M. D. Intriligator (ed.), Vol. 2, 879-922: Elsevier.

Grandmont, J.-M. (ed.) (1988): Temporary equilibrium: Selected readings, Economic Theory, Econometrics, and Mathematical Economics series, San Diego, London, Sydney and Toronto: Harcourt Brace Jovanovich, Academic Press.

Greenwood, R. (2005): "Short- and Long-term Demand Curves for Stocks: Theory and Evidence on the Dynamics of Arbitrage," Journal of Financial Economics, 75, 607-649.

Gromb, D. and D. Vayanos (2010): "Limits of Arbitrage," Annual Review of Financial Economics, 2, 251-275.

Grossman, S. J. (1981): "Nash Equilibrium and the Industrial Organization of Markets with Large Fixed Costs," Econometrica, 49, 1149-1172.

Grossman, S. J. and M. H. Miller (1988): "Liquidity and Market Structure," Journal of Finance, 43, 617-637.

Grossman, S. J. and J. E. Stiglitz (1980): "On the Impossibility of Informationally Efficient Markets," American Economic Review, 70, 3, 393-408.

Hahn, F. H. (1977): "Exercises in Conjectural Equilibria," Scandinavian Journal of Economics, 79, 2, 210-226.

Hahn, F. H. (1978): "On Non-Walrasian Equilibria," Review of Economics Studies, 45, 1, 1-17.

Hart, O. (1979): "Monopolistic Competition in a Large Economy with Differentiated Commodities," Review of Economics Studies, 46, 1, 1-30.

Hart, O. (1985): "Imperfect Competition in General Equilibrium: An Overview of Recent Work," Frontiers of Economics, K. Arrow and S. Honkapohja (eds.): Basil Blackwell, Oxford.

Hendershott, T. and A. Madhavan (2015): "Click or Call? Auction versus Search in the Over-the-Counter Market," Journal of Finance, 70, 1, 419-447.

Hicks, J. R. (1939): Value and Capital, Oxford: Clarendon.

Holmberg, P., D. Newbery, and D. Ralph (2013): "Supply Function Equilibria: Step Functions and Continuous Representations," Journal of Economic Theory, 148, 1509-1551.

Holthausen, R., R. Leftwich, and D. Mayers (1987): "The Effect of Large Block Transactions on Security Prices: A Cross-Sectional Analysis," Journal of Financial Economics, 19, 2, 237-267.

Horn, R. and C. Johnson (2013): Matrix Analysis, Cambridge University Press.

Hortaçsu, A. (2002): "Bidding Behavior in Divisible Good Auctions: Theory and Evidence from the Turkish Treasury Auction Market," Working Paper.

Hortaçsu, A., J. Kastl, and A. Zhang (2018): "Bid Shading and Bidder Surplus in the US Treasury Auction System," American Economic Review, 108, 1, 147-169.

Hortaçsu, A. and D. McAdams (2010): "Mechanism Choice and Strategic Bidding in Divisible Good Auctions: An Empirical Analysis of the Turkish Treasury Auction Market," Journal of Political Economy, 118, 5, 833-865.

Huberman, G. and W. Stanzl (2004): "Price Manipulation and Quasi-Arbitrage," Econometrica, 72, 4, 1247-1275.

Huggonier, J., B. Lester, and P.-O. Weill (2014): "Heterogeneity in Decentralized Asset Markets," Working Paper.

Jackson, M. and Y. Zenou (2014): "Games on Networks," In: Handbook of Game Theory, Vol. 4, P. Young and S. Zamir (eds.): Elsevier Science.

Kastl, J. (2012): "On the Properties of Equilibria in Private Value Divisible Good Auctions with Constrained Bidding," Journal of Mathematical Economics, 48, 6, 339-352.

Kastl, J. (2017): "Recent Advances in Empirical Analysis of Financial Markets: Industrial Organization Meets Finance," In B. Honore, A. Pakes, M. Piazzesi, and L. Samuelson (Eds.), Advances in Economics and Econometrics: Eleventh World Congress, Econometric Society Monographs, 231-270, Cambridge: Cambridge University Press.

Kawakami, K. (2017): "Welfare Consequences of Information Aggregation and Optimal Market Size," American Economic Journal: Microeconomics, 9, 4, 303-323.

Keim, D. and A. Madhavan (1995): "Anatomy of the Trading Process Empirical Evidence on the Behavior of Institutional Traders," Journal of Financial Economics, 37, 3, 371-398.

Keim, D. and A. Madhavan (1996): "The Upstairs Market for Large-Block Transactions: Analysis and Measurement of Price Effects," Review of Financial Studies, 9, 1, 1-36.

Keim, D. and A. Madhavan (1998): "The Cost of Institutional Equity Trades: An Overview," Rodney L. White Center for Financial Research Working Paper Series.

Klemperer, P. and M. Meyer (1989): "Supply Function Equilibria in Oligopoly under Uncertainty," Econometrica, 57, 1243-1277.

Knight Capital Group (2010): Current Perspectives on Modern Equity Markets: A Collection of Essays by Financial Industry Experts: Knight Capital Group, Inc.

Koutsougeras, L. C. (2003): "Non-Walrasian Equilibria and the Law of One Price," Journal of Economic Theory, 108, 1, 169-175.

Kraus, A. and H. Stoll (1972): "Price Impacts of Block Trading on the New York Stock Exchange," Journal of Finance, 27, 3, 569-588.

Kremer, I. and K. G. Nyborg (2004): "Divisible-Good Auctions: The Role of Allocation Rules," The RAND Journal of Economics, 35, 1, 147-159.

Kyle, A. S. (1985): "Continuous Auctions and Insider Trading," Econometrica, 53, 1315-1335.

Kyle, A. S. (1989): "Informed Speculation and Imperfect Competition," Review of Economic Studies, 56, 517-556.

Kyle, A. S. and J. Lee (2017): "Information and Competition with Speculation and Hedging," Working Paper.

Kyle, A. S., A. A. Obizhaeva, and Y. Wang (2018): "Smooth Trading with Overconfidence and Market Power," Review of Economic Studies, 85, 1, 611-662.

Lagos, R. and S. Zhang (2020): "The Limits of onetary Economics: On Money as a Latent Medium of Exchange" NBER Working Paper 26756.

Lagos, R. and G. Rocheteau (2009): "Liquidity in Asset Markets with Search Frictions," Econometrica, 77, 2, 403-426.

Lagos, R., G. Rocheteau, and P.-O. Weill (2011): "Crises and Liquidity in Over-the-Counter Markets," Journal of Economic Theory, 146, 6, 2169-2205.

Lester, B., A. Shourideh, V. Venkateswaran, and A. Zetlin-Jones (2018): "Market-Making with Search and Information Frictions," Working Paper.

Lester, B., A. Shourideh, V. Venkateswaran, and A. Zetlin-Jones (2019): "Screening and Adverse Selection in a Frictional Market," Journal of Political Economy, 127, 1, 338-377.

Li, D. and N. Schürhoff (2012): "Dealer Networks," Working Paper.

Li, T. M., and R. Almgren (2011): "A Fully-Dynamic Closed-Form Solution for Hedging with Market Impact," Working Paper.

Linton, O. and S. Mahmoodzadeh (2018): "Implications of High-Frequency Trading for Security Markets," Working Paper.

Lou, D., H. Yan, and J. Zhang (2013): "Anticipated and Repeated Shocks in Liquid Markets," Review of Financial Studies, 26, 1891-1912.

Madhavan, A. (2000): "Market Microstructure: A Survey," Journal of Financial Markets, 3, 3, 205-258.

Malamud, S. and M. Rostek (2017): "Decentralized Exchange," American Economic Review, 107, 11, 3320-3362.

Malamud, S. and A. Schrimpf (2017): "Intermediation markups and monetary policy pass through," Working paper, EPFL and BIS.

Malamud, S. and A. Schrimpf (2018): "An intermediation-based model of exchange rates," Working paper, EPFL and BIS.

Malamud, S. and A. Teguia (2017): "Asset Pricing with Large Investors," Working Paper.

Manzano, C. and X. Vives (2017): "Market Power and Welfare in Asymmetric Divisible Good Auctions," Working Paper.

McAdams, D. (2007): "Adjustable Supply in Uniform Price Auctions: Non-Commitment as a Strategic Tool," Economics Letters, 95, 1, 48-53.

McLennan, A. and H. Sonnenschein (1991): "Sequential Bargaining as a Noncooperative Foundation of General Equilibrium," Econometrica, 59, 1395-1424.

Menkveld, A., B. Z. Yueshen, and H. Zhu (2017): "Shades of Darkness: A Pecking Order of Trading Venues," Journal of Financial Economics, 124, 3, 503-534.

Milgrom, P. (1981): "Rational Expectations, Information Acquisition, and Competitive Bidding," Econometrica, 49, 921-943.

Milgrom, P. (2019): "Auction Market Design: Recent Innovations," Annual Review of Economics, 11, 383-405.

Milgrom, P. and B. Strulovici (2009): "Substitute Goods, Auctions, and Equilibrium," Journal of Economic Theory, 144, 212-247.

Mitchell, M., L. H. Pedersen, and T. Pulvino (2007): "Slow Moving Capital," American Economic Review, 97, 2, 215-220.

Molavi, P. (2019): "Macroeconomics with Learning and Misspecification: A General Theory and Applications," Working Paper.

Nava, F. (2015): "Efficiency in Decentralized Oligopolistic Markets," Journal of Economic Theory, 157, 315-348.

Negishi, T. (1961): "Monopolistic Competition and General Equilibrium," Review of Economics Studies, 28, 3, 196-201.

Newman, Y. and M. Rierson (2004): "Illiquidity Spillovers: Theory and Evidence from European Telecom Bond Issuance," Working Paper.

O'Hara, M. (2015): "High Frequency Market Microstructure," Journal of Financial Economics, 116, 2, 257-270.

Ollar, M., M. Rostek, and J. H. Yoon (2017): "Privacy in Markets," Working Paper.

Pagano, M. (1989): "Trading Volume and Asset Liquidity," Quarterly Journal of Economics, 104, 2: 255-274.

Pagnotta, E. and T. Philippon (2018): "Competing on Speed," Econometrica, 86, 3: 1067-1115.

Pakes A. (2016): "Methodological Issues in Analyzing Market Dynamics," In: Advances in Dynamic and Evolutionary Games: Theory, Applications, and Numerical Methods, F. Thuijsman and F. Wagener (eds.), 43-75, Switzerland: Springer International Publishing.

Peck, J. and K. Shell (1989): "On the Nonequivalence of the Arrow-Securities Game and the Contingent Commodities Game," In: Economic Complexity: Chaos, Sunspots, Bubbles, and Nonlinearity, J. G. William Barnett and K. Shell (eds.), 61-85, Cambridge University Press.

Peivandi, A. and R. Vohra (2017): "Instability of Centralized Markets," Working Paper.

Pesendorfer, W. and J. Swinkels (2000): "Efficiency and Information Aggregation in Auctions," American Economic Review, 90, 3, 499-525.

Pritsker, M. (2005): "Large Investors: Implications for Equilibrium Asset Returns. Shock Absorption, and Liquidity," Federal Reserve Board, Washington, D.C. Staff Working Paper 2005-36, the Finance and Economics Discussion Series.

Pycia, M. and K. Woodward (2018): "Auctions of Homogeneous Goods: A Case for Pay-as-Bid," Working Paper.

Rahi, R. (2019): "Information Acquisition with Heterogeneous Valuations," Working Paper.

Rahi, R. and J.-P. Zigrand (2013): "Market Quality and Contagion in Fragmented Markets," Working Paper.

Rahi, R. and J.-P. Zigrand (2018): "Information Acquisition, Price Informativeness, and Welfare," Journal of Economic Theory, 177, 558-593.

Reny, P. and M. Perry (2006): "Toward a Strategic Foundation of Rational Expectations Equilibrium," Econometrica, 49, 655-678.

Roberts, D. J. (1987): "General Equilibrium Analysis of Imperfect Competition: an Illustrative Example," In: G. Feiwel, Ed., Arrow and the ascent of economic theory, New York, MacMillan.

Robson, A. (1983): "Existence of Consistent Conjectures: Comment," American Economic Review, 73, 3, 454-456.

Rostek, M. and M. Weretka (2012): "Price Inference in Small Markets," Econometrica, 80, 2, 687-711.

Rostek, M. and M. Weretka (2015a): "Dynamic Thin Markets," Review of Financial Studies, 28, 10, 2946-2992.

Rostek, M. and M. Weretka (2015b): "Information and Strategic Behavior," Journal of Economic Theory, 158, 536-557.

Rostek, M. and N. Yoder (2017): "Core Selection in Auctions and Exchanges," Working Paper.

Rostek, M. and J. H. Yoon (2018a): "Supply Function Games with General Gaussian Information Structures," Working Paper.

Rostek, M. and J. H. Yoon (2018b): "Innovation in Decentralized Markets," Working Paper.

Rostek, M. and J. H. Yoon (2018c): "Demand Submission Games with General Utilities," Working Paper.

Rostek, M. and J. H. Yoon (2019): "Dynamic Imperfectly Competitive Markets with Private Information," Working Paper.

Rostek, M. and J. H. Yoon (2020): "Exchange Design and Efficiency," Working Paper.

Rubinstein, A. and A. Wolinsky (1985): "Equilibrium in a Market with Sequential Bargaining," Econometrica, 53, 295-328.

Rubinstein, A. and A. Wolinsky (1990): "Decentralized Trading, Strategic Behavior and the Walrasian Outcome," Review of Economic Studies, 57, 63-78.

Sabourian, H. (1992): "Rational Conjectural Equilibrium and Repeated Games," In: Economic Analysis of Markets and Games, P. Dasgupta, D. Gale, O. Hart, and E. Maskin (eds.), 228-257: M.I.T. Press.

Salanie, B. (2000): The Microeconomics of Market Failures, MIT Press.

Sannikov, Y. and A. Skrzypacz (2016): "Dynamic Trading: Price Inertia and Front-Running," Working Paper.

Shapley, L. and M. Shubik (1977): "Trade Using One Commodity as a Means of Payment," Journal of Political Economy, 85, 5, 937-968.

Shleifer, A. (1986): "Do Demand Curves for Stocks Slope Down?" Journal of Finance, 41, 3, 579-590.

Shubik, M. (1973): "Commodity Money, Oligopoly, Credit and Bankruptcy in a General Equilibrium Model," Western Economic Journal, 11, 1, 24-38.

Tufano, P. (2003): "Financial Innovation," In: Handbook of the Economics of Finance, Vol. 1, G. M. Constantinides, M. Harris, and R. M. Stulz (eds.), 307-335, Amsterdam: Elsevier.

Urošević, B. (2005): "Moral Hazard and Dynamics of Insider Ownership Stakes," Universitat Pompeu Fabra, Economics Working Papers, 787.

Vayanos, D. (1999): "Strategic Trading and Welfare in a Dynamic Market," Review of Economic Studies, 66, 219-254.

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

Vayanos, D. and J. Wang (2013): "Market Liquidity: Theory and Empirical Evidence," In: Handbook of the Economics of Finance, G. Constantinides, M. Harris, and R. Stulz (eds.), Chapter 19.

Vayanos, D. and P.-O. Weill (2008): "A Search-based Theory of the On-the-run Phenomenon," Journal of Finance, 63, 1351-1389.

Vives, X. (2002): "Private Information, Strategic Behavior, and Efficiency in Cournot Markets," The RAND Journal of Economics, 33, 3, 361-376.

Vives, X. (2008): Information and Learning in Markets: The Impact of Market Microstructure, Princeton: Princeton University Press.

Vives, X. (2011): "Strategic Supply Function Competition with Private Information," Econometrica 79, 6, 1919-1966.

Vives, X. and L. Yang (2017): "Costly Interpretation of Asset Prices," Working Paper.

Wang, J. J. D. and Zender, J. F. (2002): "Auctioning Divisible Goods," Economic Theory, 19, 673-705.

Weill, P.-O. (2008): "Liquidity Premia in Dynamic Bargaining Markets," Journal of Economic Theory, 140, 66-96.

Weintraub, G., L. Benkard, and B. Van Roy (2006): "Oblivious Equilibrium: A Mean Field Approximation for Large-Scale Dynamic Games," Working Paper.

Werning, I. and O. Wang (2018): "Dynamic Oligopoly and Price Stickiness," Working Paper.

Weretka, M. (2011): "Endogenous Market Power," Journal of Economic Theory, 146, 6, 2281-2306.

Wilson, R. (1977): "A Bidding Model of Perfect Competition," Review of Economic Studies, 44, 3, 511-518.

Wilson, R. (1979): "Auctions of Shares," Quarterly Journal of Economics, 94, 675-689.

Wilson, R. (1987): "Game-Theoretic Analyses of Trading Processes," In: Advances in Economic Theory: Fifth World Congress, Truman Bewley (ed.), Chapter 2, pp. 33-70, Cambridge UK: Cambridge University Press.

Wittwer, M. (2020): "Connecting Disconnected Financial Markets?," Working Paper.

Woodford, M. (2013): "Macroeconomic Analysis Without the Rational Expectations Hypothesis," Annual Review of Economics, 5, 303-346.

Woodford, M. (2018): "Monetary Policy Analysis When Planning Horizons Are Finite," Working Paper.

Yoon, J. H. (2018): "Endogenous Market Structure: Over-the-Counter versus Exchange Trading," Working Paper.

Zhang, A. L. (2018): "Auctions with Liquidity Subsidies," Working Paper.

Zhang, A. L. (2019): "Competition and Manipulation in Derivative Contract Markets," Working Paper.

Zhu, H. (2014): "Do Dark Pools Harm Price Discovery?" Review of Financial Studies, 27, 3, 747-789.

Zhu, H. (2018a): "A Multi-Maturity Clock Auction Design," Market Design Proposal.

Zhu, H. (2018b): "Design of CCP Default Management Auctions," Market Design Proposal.

Appendix

The appendix provides a sketch of the main steps in the equilibrium characterization, comparative statics results, and additional results for static, dynamic, centralized, and decentralized markets. The presentation highlights the common arguments across settings as well as model-specific observations. These include the comparison with the competitive and Cournot models. For simplicity, we assume two risky assets, $K = 2$.

Appendix A: Centralized markets with (A.1) private and (A.2) interdependent valuations.

Appendix B: Dynamic (centralized) markets.

Appendix C: Decentralized markets with (C.1) incomplete participation and (C.2) incomplete demand conditioning.

Appendix D: Relation to Other Models and Results.

We begin by describing the setting and the results that apply in all sections.

Assets and traders. There are two risky (divisible) assets and a riskless asset (numéraire). The return of the numéraire is normalized to one. The returns of the risky assets are jointly Normally distributed $R = (r_1, r_2) \sim \mathcal{N}(\theta, \Sigma)$, where $\theta = (\theta_1, \theta_2) \in \mathbb{R}^2$ and $\Sigma \in \mathbb{R}^{2\times 2}$. The mean value θ is uncertain, unless stated otherwise, and is heterogeneous across traders $\theta^i \in \mathbb{R}^2$ with the jointly Normal distribution $\mathcal{N}((E[\theta^i])_i, \Sigma_\theta)$.[124] Each trader i has an expected payoff:

$$E[u^i(q^i)|q_0^i] = E[-\exp(-\alpha^i(R\cdot(q^i+q_0^i) - p\cdot q^i))|q_0^i], \tag{69}$$

where α^i is the risk aversion, $q^i \in \mathbb{R}^2$ is quantity of assets traded, and $p \in \mathbb{R}^2$ is market prices. A trader's initial endowment $q_0^i \in \mathbb{R}^2$ is privately known. Trader i is uncertain of other trader's endowments $\{q_0^j\}_{j\neq i}$. In later sections, we will generalize the information structure of traders on $\{\theta^i, q_0^i\}_i$. The asset return R is independent of traders' endowments $\{q_0^i\}_i$.

Maximization of the expected payoff (69) is equivalent to maximization of the quasilinear-quadratic payoff by a monotonic transformation:

$$U^i(q^i) = E[\theta^i\cdot(q^i+q_0^i) - \frac{\alpha^i}{2}(q^i+q_0^i)\cdot\Sigma(q^i+q_0^i) - p\cdot q^i|q_0^i] \qquad \forall i. \tag{70}$$

A unilateral demand change in $q^i(\cdot)$ is understood (without loss of generality) as the change in demand's intercept.

Techniques for equilibrium derivation. Equilibrium in demand submission games can be characterized through the following steps:

Step 1 (Optimization) *The demand schedule $q^i(\cdot)$ maximizes the expected payoff (70), given price impact Λ^i and expected asset valuation $E[\theta^i|s^i, p]$.*

Step 2 (Conditional expectations) *Price distribution $\mathcal{N}(E[p], Var(p))$ is a function of demand coefficients. The projection theorem determines inference coefficients $\{C_\theta^i, C_s^i, C_p^i\}_i$ in $E[\theta^i|s^i, p] = C_\theta^i E[\theta^i] + C_s^i s^i + C_p^i p$ for i, given price impacts $\{\Lambda^i\}_i$.*

Step 3 (Price impact) *By substituting inference coefficients $\{C_p^i\}_i$ as functions of price impacts, the equilibrium condition for price impacts (e.g., (74) below) becomes a fixed point problem for $\{\Lambda^i\}_i$ alone.*

The following observations are used in the characterization of equilibrium. Variants of the corresponding equations hold in all settings.

[124] In Appendix A.2, individual asset valuations (mean return) $(\theta^i)_i$ are described by an arbitrary Normal distribution. Appendix A.1 and Appendix C.1 assume a constant mean of asset return θ^i, but the derivation and the results in these sections extend to unknown independent θ^i.

- *Equilibrium outcomes are characterized by demand reduction parameters* $\{1-\gamma^i\}_i$.

In response to his price impact, each trader i reduces his (net) demand by a fraction $\gamma^i = \frac{\alpha^i}{\alpha^i+\lambda^i}$. The demand reduction parameters $\{1-\gamma^i\}_i$ determine how traders' expected asset valuations $\{E[\theta^i|s^i,p]\}_i$ and idiosyncratic risks $\{\alpha^i\Sigma q_0^i\}_i$ are aggregated in equilibrium price:

$$p = \overline{v} - Q, \tag{71}$$

where the aggregate asset valuation $\overline{v}$ and the aggregate risk Q are defined by:

$$\overline{v} \equiv (\sum_j \frac{\gamma^j}{\alpha^j})^{-1} \sum_j \frac{\gamma^j}{\alpha^j} E[\theta^j|s^j,p]; \qquad Q \equiv (\sum_j \frac{\gamma^j}{\alpha^j})^{-1} \sum_j \gamma^j \Sigma q_0^j. \tag{72}$$

When θ^i is uncertain, substituting $E[\theta^i|s^i,p] = C_\theta^i E[\theta^i] + C_s^i s^i + C_p^i p$ into (71), one can write the equilibrium price equation:

$$p = (\sum_j \frac{\gamma^j}{\alpha^j}(1-C_p^j))^{-1} \sum_j \frac{\gamma^j}{\alpha^j}(C_\theta^j E[\theta^j] + C_s^j s^j) - (\sum_j \frac{\gamma^j}{\alpha^j}(1-C_p^j))^{-1} \sum_j \gamma^j \Sigma q_0^j. \tag{73}$$

Price Equation (73) gives the price distribution $\mathcal{N}(E[p], Var(p))$ and $Cov(s^i,p)$.

- *The projection theorem characterizes the expected asset valuation* $E[\theta^i|s^i,p]$.

Applying the projection theorem to the price equation (73) characterizes the inference coefficients $\{C_\theta^i, C_s^i, C_p^i\}_i$:

$$E[\theta^i|s^i,p] = E[\theta^i] - \begin{bmatrix} Cov(\theta^i,s^i) \\ Cov(\theta^i,p) \end{bmatrix} \cdot \begin{bmatrix} Var(s^i) & Cov(s^i,p) \\ Cov(p,s^i) & Var(p) \end{bmatrix}^{-1} \begin{bmatrix} s^i - E[s^i] \\ p - E[p] \end{bmatrix} \qquad \forall i.$$

The fixed point (sub)problem for inference coefficients is a system of *third-order* polynomial equations for $\{C_\theta^i, C_s^i, C_p^i\}_i$, given price impacts $\{\Lambda^i\}_i$. That is, the profile of price impacts $\{\Lambda^i\}_i$ is a sufficient statistic for equilibrium $\{q^i(\cdot)\}_i$.

- *Price impact is the harmonic mean of the utility functions' Hessians.*

Price impact $\Lambda^i \equiv \frac{dp}{dq^i}$ is characterized by the slope of inverse residual supply function $S^{-i}(\cdot) \equiv -\sum_{j\neq i} q^j(\cdot)$:[125]

$$\Lambda^i = -(\sum_{j\neq i} \frac{\partial q^j(\cdot)}{\partial p})^{-1} = (\sum_{j\neq i} (\alpha^j\Sigma + \Lambda^j)^{-1}(Id - C_p^j))^{-1} \qquad \forall i. \tag{74}$$

[125] If θ^i is known (to all) as in Section 4, then $C_p^i = 0$ for all i and the price impact simplifies as in Equation (10).

When values $\{\theta^i\}$ are independent private (i.e., $C_p^i = 0$ for all i), the system of Equations (74) for $\{\Lambda^i\}_i$ can be reduced to a fixed point problem for a single variable: the aggregate slope of all traders' demands $\Psi \equiv \sum_i (\alpha^i \Sigma + \Lambda^i)^{-1}$. See Section 4.1.2.

This article considers equilibrium with downward sloping demands — equivalently, positive semi-definite price impact matrices $\Lambda^i \geq 0$ for all i — which is a sufficient condition for the second-order conditions to hold for all traders.

A Centralized Markets

A.1 Independent Private Values

Throughout, we identify the symmetric market with the symmetric equilibrium.

Definition 7 (Symmetric Market) *Equilibrium in a market is symmetric if equilibrium demands of all traders have the same coefficients on price and private information: Letting a trader's demand be*

$$q^i(p) = a^i - b^i q_0^i - s^i p \qquad \forall p \in \mathbb{R}^2 \quad \forall i,$$

equilibrium is symmetric if $b^i = b$ and $s^i = s$ for all i.

Equilibrium in the centralized market with private values is symmetric if and only if traders have symmetric risk preferences $\alpha^i = \alpha$.

Example 2 (Symmetric Market) In the symmetric market (Definition 7) where $\theta^i = \theta$, traders' price impacts are the same: $\Lambda^i = \alpha\beta\Sigma$ for all i. By Equation (18), the symmetric-equilibrium price p is:

$$p = \theta - (\sum_j \frac{\gamma}{\alpha})^{-1} \sum_j \gamma \Sigma q_0^j = \theta - \alpha\Sigma \frac{1}{I} \sum_j q_0^j \ = p^{**},$$

where $\gamma \equiv \frac{\alpha}{\alpha+\alpha\beta}$. The price p is independent of the price impact. In the symmetric market, the price coincides with the competitive price p^{**} (Equation (21)). Thus, the aggregate risk coincides with systematic risk, $Q = \alpha\Sigma\frac{1}{I}\sum_j q_0^j = Q^{**}$.

Each trader's allocation $q^i + q_0^i$ is a convex combination of the unweighted average endowment (aggregate risk portfolio) and his own endowment (idiosyncratic risk portfolio):

$$q^i + q_0^i = \gamma \frac{1}{I} \sum_j q_0^j + (1-\gamma) q_0^i \qquad \forall i. \tag{75}$$

In the competitive market, $\Lambda \to 0$ and $\gamma \to 1$, and hence, all traders' allocations (75) coincide with the average endowment (i.e., the efficient portfolio).

For the comparative statics, it is useful to write trader i's indirect utility as a function of price impact and aggregate risk:

$$V^i(q_0^i) = (\theta \cdot q_0^i - \frac{\alpha}{2}(q_0^i) \cdot \Sigma q_0^i) + (Q - \alpha \Sigma q_0^i) \cdot (\alpha\Sigma + \Lambda)^{-1} (\frac{1}{2}\alpha\Sigma + \Lambda)(\alpha\Sigma + \Lambda)^{-1}(Q - \alpha\Sigma q_0^i). \quad \square$$

A.2 Interdependent Values

Proposition 2 (Section 4.3) characterizes equilibrium with interdependent values. Relative to equilibrium in markets with independent values (Proposition 1), there is a new condition for conditional expectations and a corresponding new step in the characterization.

Step 1 (Optimization) The first-order condition is:

$$E[\theta^i | s^i, p] - \alpha^i \Sigma (q^i + q_0^i) = p + \Lambda^i q^i \qquad \forall p \in \mathbb{R}^2, \tag{76}$$

given price impact Λ^i (characterized in Step 3) and the conditional expectations $E[\theta^i | s^i, p]$ (characterized in Step 2 given the distribution of price p).[126]

Step 2 (Conditional expectations) Fix the best responses (76) of traders $j \neq i$. Equilibrium price distribution is determined by the market-clearing condition:

$$p = (\sum_j (\alpha^j \Sigma + \Lambda^j)^{-1})^{-1} ((\alpha^i \Sigma + \Lambda^i)^{-1}(E[\theta^i | s^i, p] - \alpha^i \Sigma q_0^i) + \underbrace{\sum_{j \neq i} (\alpha^j \Sigma + \Lambda^j)^{-1}(E[\theta^j | s^j, p] - \alpha^j \Sigma q_0^j)}_{\text{Intercept of residual supply function}} \tag{77}$$

- *Expected value $E[\theta^i | s^i, p]$ is conditional on the price to be realized.*

By the price equation (77), the realization of price p maps one-to-one to the intercept of the residual supply function $S^{-i}(\cdot) = -\sum_{j \neq i} q^j(\cdot)$, which is an aggregate statistic of other traders' private information ($E[\theta^j | s^j, p] - \alpha^j \Sigma q_0^j$ for $j \neq i$). Given the linear best-response schedules of traders $j \neq i$, trader i optimizes against a family of residual supplies with a deterministic slope Λ^i and random intercept determined by the signal realizations of traders $j \neq i$ (Fig. 1B).

In the quadratic-Gaussian setting, given the equilibrium price equation (77), the conditional expectations of a trader's asset valuation θ^i are characterized by the projection theorem as a linear function of the conditioning variables:

$$E[\theta^i | s^i, p] = C_\theta^i E[\theta^i] + C_s^i s^i + C_p^i p \qquad \forall i, \tag{78}$$

[126] A unilateral demand change in $q^i(\cdot)$ is understood (without loss of generality) as the change in demand's intercept.

where $C^i_\theta, C^i_s, C^i_p \in \mathbb{R}^{2\times 2}$ are the inference coefficients of trader i.

Trader i's expectations $E[\theta^i|s^i,p]$ are correct, i.e., the projection theorem is applied to the equilibrium price distribution. Using the parameterization in (78), the equilibrium price (77) becomes:

$$p = (\sum_j (\alpha^j\Sigma + \Lambda^j)^{-1}(Id - C^j_p))^{-1}(\sum_j (\alpha^j\Sigma + \Lambda^j)^{-1}(C^j_\theta E[\theta^j] + C^j_s s^j - \alpha^j\Sigma q^j_0)). \tag{79}$$

The traders' conditional expectations $\{E[\theta^i|s^i,p]\}_i$ in Equation (78) and the price distribution p from (79) determine the inference coefficients $\{C^i_\theta, C^i_s, C^i_p\}_i$ as a fixed point, given price impacts $\{\Lambda^i\}_i$.

Step 3 (Price impact) Differentiating the first-order conditions (76) of traders $j \neq i$ with respect to p gives $\frac{\partial q^j(p)}{\partial p}$:

$$\frac{\partial q^j(p)}{\partial p} = -(\alpha^j\Sigma + \Lambda^j)^{-1}(Id - \frac{\partial}{\partial p}E[\theta^j|s^j,p]) = -(\alpha^j\Sigma + \Lambda^j)^{-1}(Id - C^j_p) \qquad \forall j \neq i. \tag{80}$$

Substituting for $\frac{\partial q^j(p)}{\partial p}$ into $\Lambda^i = -\sum_{j\neq i}\frac{\partial q^j(p)}{\partial p}$ characterizes trader i's price impact: Equation (74).

The profile of demand schedules $\{q^i(\cdot)\}_i$, each given by

$$q^i(p) = (\alpha^i\Sigma + \Lambda^i)^{-1}(E[\theta^i|s^i,p] - p - \alpha^i\Sigma q^i_0) \qquad \forall p \in \mathbb{R}^2, \tag{81}$$

is an equilibrium if and only if price impacts $\{\Lambda^i\}_i$ and conditional expectations $\{E[\theta^i|s^i,p]\}_i$ are endogenized by the fixed point problem defined by (78) and (74).

Sufficient statistics for a best response and equilibrium. Assuming individual optimization (i.e., (76)), the profile of price impacts $\{\Lambda^i\}_i$ is sufficient for *equilibrium* schedules $\{q^i(\cdot)\}_i$, because the equilibrium price distribution (and hence inference coefficients) is determined as a function of price impacts $\{\Lambda^i\}_i$ (Equation (79))). The sufficient statistic for a trader i's *best response* is his price impact Λ^i *and* the price distribution $p \sim \mathcal{N}(E[p], Var(p))$.

Interdependence between price impacts and inference coefficients. Example 1 (Section 4.3) characterizes the comparative statics of price impacts and conditional expectations in symmetric markets with interdependent values (the equicommonal model) and one asset. The logic of equilibrium characterization with one asset extends to markets with multiple assets. Assuming independent private values across assets, we discuss the results that hold for an arbitrary number of assets: equilibrium existence, welfare implications, and information aggregation.

Example 1 - Cont'd (Symmetric Information Structure) Consider a two-asset version of the market from Example 1: all traders are symmetric in their risk preferences and

information structure. Asset valuations $\theta^i = (\theta^i_1, \theta^i_2)$ are *interdependent* across traders i and *independent* and identically distributed across assets k: for each $k = 1, 2$, $(\theta^i_k)_i \in \mathbb{R}^I$ is distributed according to $\mathcal{N}(0, \Omega)$.[127]

In the symmetric markets, equilibrium price

$$p = \frac{1}{I}\sum_j (E[\theta^j | s^j, p] - \alpha \Sigma q^j_0) = (Id - C_p)^{-1}\frac{1}{I}\sum_j (C_\theta E[\theta^j] + C_s s^j - \alpha \Sigma q^j_0) \tag{82}$$

is Normally distributed. The independence of asset values θ^i_k (hence, signals s^i_k) and endowments $q^i_{k,0}$ across assets implies that equilibrium prices p_1 and p_2 are independent. It follows that the inference coefficients are proportional to the identity matrix: $C_\theta = c_\theta Id$, $C_s = c_s Id$, and $C_p = c_p Id$, where $c_\theta, c_s, c_p \in \mathbb{R}$ are identical to the coefficients in the one-asset market from Example 1.

The inference coefficient c_p and the price impact Λ are given by:

$$c_p = \frac{I\sigma^2\bar{\rho}}{(1+\sigma^2-\bar{\rho})(1+(I-1)\bar{\rho})}; \quad \Lambda = \lambda\Sigma = \frac{\alpha}{(I-1)(1-c_p)-1}\Sigma. \tag{83}$$

The price impact in Equation (83) gives a necessary and sufficient condition for the existence of equilibrium with downward sloping demands. The price impact λ is positive, i.e., the demand schedule is downward sloping (with respect to p), if and only if $c_p < 1 - \frac{1}{I-1}$ or, equivalently, $\bar{\rho} < \bar{\rho}^+$, where $\bar{\rho}^+$ is the unique positive solution to:

$$0 = I(I-1)\sigma^2\bar{\rho}^+ - (I-2)(1+\sigma^2-\bar{\rho}^+)(1+(I-1)\bar{\rho}^+).$$

Fixing the number of traders I:

- An increase in $\bar{\rho}$ increases $Var(\theta^i|s^i, p)$, but it also increases price impact λ and decreases total welfare.
- Price may not be fully revealing (see Definition 5) even when traders' risk preferences are symmetric and information structure is equicommonal for each asset, unless an additional symmetry condition holds. The inequality

  $$Var(\theta^i|s^i, p) = Var(\theta^i|s^i, \bar{s}) \geq Var(\theta^i|(s^j)_j)$$

  becomes an equality for all i (i.e., price is fully revealing) if and only if $\rho_{ij} = \rho$ for all i and $j \neq i$. Then, the unweighted average of others' signals is a sufficient statistics of the signal vector $(s^j)_{j\neq i} \in \mathbb{R}^{I-1}$ in each trader's conditional expectation: $E[\theta^i|s^i, (s^j)_{j\neq i}] = E[\theta^i|s^i, \bar{s}]$ for all i. □

[127] More precisely, the asset valuation θ^i of trader i is distributed $(\theta^i_1, \theta^i_2)_i \sim \mathcal{N}(\theta \otimes \mathbf{1}, \Omega \otimes Id)$ in the space of $\mathbb{R}^{KI}$. Ω captures the interdependence among values of traders.

Trader i submits his demand schedule $q^i(\cdot)$ as a function of the price vector p to be realized:

$$q^i(p) = \frac{\gamma^i}{\alpha^i}\Sigma^{-1}(E[\theta^i|s^i,p] - p) - \gamma^i q_0^i \qquad \forall p \in \mathbb{R}^2, \tag{84}$$

Parameter $1-\gamma^i \equiv \frac{\lambda^i}{\alpha^i+\lambda^i}$ represents the demand reduction of trader i. Each trader's allocation $q^i + q_0^i$ can be decomposed as follows:

$$q^i + q_0^i = \gamma^i \underbrace{\frac{1}{\alpha^i}\Sigma^{-1}(E[\theta^i|s^i,p] - \overline{v})}_{\text{Valuation adjustment}} + \gamma^i \underbrace{\frac{1}{\alpha^i}\Sigma^{-1}Q}_{\text{Aggregate risk portfolio}} + (1-\gamma^i)\underbrace{q_0^i}_{\text{Initial portfolio}} \qquad \forall i, \tag{85}$$

where the aggregate asset valuation $\overline{v}$ and aggregate risk Q are defined in Equation (72).

Comparison with the competitive market. As $I \to \infty$, the limit of Equation (74) gives:

$$\Lambda^i = (\sum_{j\neq i}(\alpha^j\Sigma + \Lambda^j)^{-1}(Id - C_p^j))^{-1} \leq \frac{1}{I-1}\max_j \alpha^j\Sigma(Id - C_p^j)^{-1} \to 0 \qquad \forall i.$$

By Steps 1 and 2, using that $\Lambda^i = 0$ (equivalently, $\gamma^i = 1$), the competitive demand $q^i(\cdot) : \mathbb{R}^2 \to \mathbb{R}^2$ is:

$$q^i(p) = (\alpha^i\Sigma)^{-1}(E[\theta^i|s^i,p] - p - \alpha^i\Sigma q_0^i) \qquad \forall p \in \mathbb{R}^2.$$

The conditional expectations $E[\theta^i|s^i,p] = C_\theta^i E[\theta^i] + C_s^i s^i + C_p^i p$ are determined by applying the projection theorem to the distribution of the competitive price:

$$p = (\sum_j \frac{1}{\alpha^j})^{-1}\sum_j(\frac{1}{\alpha^j}E[\theta^j|s^j,p] - \Sigma q_0^j) = (\sum_j \frac{1-c_p^j}{\alpha^j})^{-1}\sum_j(\frac{c_\theta^j}{\alpha^j}E[\theta^j] + \frac{c_s^j}{\alpha^j}s^j - \Sigma q_0^j) \tag{86}$$

Equilibrium inference coefficients $\{C_\theta, C_s, C_p\}$ and price (Equation (73)) in imperfectly competitive markets differ from those in competitive markets, except when traders are symmetric. In the symmetric market, the equilibrium price (82) does not depend on price impact and so coincides with the competitive price (Equation (86)).

Comparison with the Cournot model. In the Cournot model with supply $p = H + G\sum_j q^j$, trader i submits a quantity demanded q^i (not contingent on price) that maximizes the same expected payoff as (30):

$$\max_{q^i=(q_1^i,q_2^i)} E[\theta^i \cdot (q^i + q_0^i) - \frac{\alpha^i}{2}(q^i + q_0^i)\cdot\Sigma(q^i + q_0^i) - p\cdot q^i|s^i].$$

Trader i's best response is:

$$q^i = (\alpha^i\Sigma + \Lambda^i)^{-1}(E[\theta^i|s^i] - E[p|s^i] - \alpha^i\Sigma q_0^i) \qquad \forall i.$$

A trader i's quantity demanded depends on the expected price $E[p|q_0^i]$. This contrasts with the demand submission game, where the schedule $q^i(\cdot)$ specifies the quantity demanded for each realization of price p. The price impact is determined by the slope of the exogenous supply, $\Lambda^i = G$ for all i. The inference coefficients in $E[\theta^i|s^i] = C_\theta^i E[\theta^i] + C_s^i s^i$ still differ from those in the competitive model.

B Dynamic Centralized Market

In the symmetric market, the intercept $v_t^i \in \mathbb{R}^2$ and slope $\alpha_t^i \in \mathbb{R}$ of the marginal value function $\frac{dV_t^i}{dq_t^i} \equiv E[v_t^i|\mathcal{I}_t^i, p_t] - \alpha_t^i \Sigma(q_t^{i,0} + q_t^i)$ are determined by backward induction. Given the value function $V_t^i(\cdot)$, equilibrium $\{q_t^i(\cdot)\}_i$ is then characterized following Steps 1-3 for each t.

Proposition 4 (Equilibrium: Dynamic Market with Symmetric Traders) Suppose th $\alpha^i = \alpha$ for all i. In a dynamic centralized market, equilibrium $\{\{q_t^i(\cdot) : \mathbb{R}^K \to \mathbb{R}^K\}_{t=1,\dots,T}\}_i$ is characterized by the following conditions:

(i) Each trader i submits a demand schedule $q_t^i(\cdot)$ such that:

$$\frac{dV_t^i}{dq_t^i} \equiv E[v_t^i|\mathcal{I}_t^i, p_t] - \alpha_t \Sigma(q_t^{i,0} + q_t^i) = p_t + \Lambda_t q_t^i + \kappa \Sigma q_t^i \qquad \forall p_t \in \mathbb{R}^2, \tag{87}$$

where the intercept $v_t^i \equiv \frac{dV_t^i}{dq_t^i}$ and slope $\alpha_t \Sigma \equiv \frac{d^2 V_t^i}{(dq_t^i)^2}$ of the marginal value function are characterized as follows: for $t < T$,

$$\begin{aligned} v_t^i &= \prod_{l>t}^{T}((1-\gamma_l)^2 + \gamma_l^2 \frac{\kappa}{\alpha_l})(\theta^i - p_T) + \sum_{\tau>t}^{T-1}\prod_{l>t}^{\tau}((1-\gamma_l)^2 + \gamma_l^2 \frac{\kappa}{\alpha_l})(p_{\tau+1} - p_\tau) + p_{t+1} \qquad (88) \\ &= ((1-\gamma_{t+1})^2 + \frac{\kappa}{\alpha_{t+1}}\gamma_{t+1}^2)(v_{t+1}^i - p_{t+1}) + p_{t+1}, \\ \alpha_t &= (\prod_{l>t}^{T}(1-\gamma_l)^2)\alpha + \sum_{\tau>t}^{T}\gamma_\tau^2 \prod_{l>t}^{\tau-1}(1-\gamma_l)^2\kappa = (1-\gamma_{t+1})^2\alpha_{t+1} + \gamma_{t+1}^2\kappa, \qquad (89) \end{aligned}$$

and $v_T^i = \theta^i$, $\alpha_T = \alpha$. Here, $\gamma_t = \frac{\alpha_t}{\alpha_t + \kappa + \lambda_t}$ and $\Lambda_t = \lambda_t \Sigma$ for all t.

(ii) The conditional expectations $E[v_t^i|\mathcal{I}_t^i, p_t] = C_{\theta,t}\theta^i + C_{p,t}p_t$ are characterized by $C_{p,t} = c_{p,t} Id \in \mathbb{R}^{2\times 2}$, where

$$c_{p,t} = 1 - (\alpha_t + \sum_{s>t}^{T}\prod_{l>t}^{s-1}((1-\gamma_l)^2 + \gamma_l^2\frac{\kappa}{\alpha_l})((1-(1-\gamma_s)^2)\alpha_s - \gamma_s^2\kappa)(1-c_{p,s})^{-1})^{-1}\alpha_t \tag{90}$$

in rounds $t < T$ and $C_{p,T} = 0$ at $t = T$.

(iii) Price impact $\Lambda_t = \lambda_t \Sigma$ is characterized by:

$$\Lambda_t = -(\sum_{j \neq i} \frac{\partial q_t^j(p_t)}{\partial p_t})^{-1} = \frac{\alpha_t + \kappa}{(I-1)(1-c_{p,t}) - 1}\Sigma. \tag{91}$$

Example 3 illustrates the backward-recursive characterization of the value function. We discuss the Markov property of learning (i.e., the conditional expectations) which holds in the symmetric market.

Example 3 (Dynamic Market ($T < \infty$): Symmetric Traders) Consider a market with $\alpha^i = \alpha$ for all i. Each trader knows privately his own endowment shocks and all traders observe past prices: $\mathcal{I}_t^i = \{q_0^i, \{\delta_s^i\}_{s \leq t}, \{p_s\}_{s<t}\}$. We assume that traders' values θ^i are known (to all) for simplicity of the derivation. Suppose that the value function parameters v_t^i and α_t in round t are given by Equations (88)-(89). Equilibrium demands $\{q_t^i(\cdot)\}_i$ in each round t can then be derived following Steps 1-3. The demand schedule of trader i is:

$$q_t^i(p_t) = ((\alpha_t + \kappa)\Sigma + \Lambda_t)^{-1}(E[v_t^i|\mathcal{I}_t^i, p_t] - p_t - \alpha_t \Sigma q_t^{i,0}) = \frac{\gamma_t}{\alpha_t}\Sigma^{-1}(E[v_t^i|\mathcal{I}_t^i, p_t] - p_t) - \gamma_t q_t^{i,0}, \tag{92}$$

given price impact $\Lambda_t = \lambda_t \Sigma$ and expected asset valuation $E[v_t^i|\mathcal{I}_t^i, p_t]$. The inference coefficients in $E[v_t^i|\mathcal{I}_t^i, p_t] = C_{\theta,t}\theta^i + C_{p,t}p_t$ are derived by applying the projection theorem, using that the equilibrium price at t is:

$$p_t = \frac{1}{I}\sum_j (E[v_t^j|\mathcal{I}_t^j, p_t] - \alpha_t \Sigma q_t^{j,0}) = (Id - C_{p,t})^{-1}(C_{\theta,t}\theta^j - \alpha_t \Sigma \frac{1}{I}\sum_j q_t^{j,0}). \tag{93}$$

- *The conditional expectations $E[v_t^i|\mathcal{I}_t^i, p_t]$ do not depend on price history $\{p_s\}_{s<t}$ in the symmetric market.*

By Equation (93) there is a one-to-one mapping between the price p_t and the aggregate risk Q_t, which — in the symmetric market — is an unweighted average of inventories $\{q_t^{i,0}\}_i$. Because past prices $\{p_s\}_{s<t}$ do not provide additional information beyond the current price p_t, $E[v_t^i|\mathcal{I}_t^i, p_t] = E[v_t^i|p_t]$ for each t.

Equations (88) and (93) determine the joint distribution of v_t^i and p_t:

$$E[v_t^i] = \theta^i - \sum_{l>t}^{T}((1-(1-\gamma_l)^2)\alpha_l - \gamma_l^2\kappa)\Sigma E[\bar{q}_l^0],$$

$$E[p_t] = E[v_t^i] - \alpha_t \Sigma E[\bar{q}_t^0] = \theta^i - \sum_{l>t}^{T}((1-(1-\gamma_l)^2)\alpha_l - \gamma_l^2\kappa)\Sigma E[\bar{q}_l^0] - \alpha_t \Sigma E[\bar{q}_t^0],$$

$$Cov(v_t^i, p_t) = \sum_{s>t}^{T}\prod_{l>t}^{s-1}((1-\gamma_l)^2 + \gamma_l^2\frac{\kappa}{\alpha_l})(1-(1-\gamma_s)^2 - \gamma_s^2\frac{\kappa}{\alpha_s})\frac{\alpha_s}{1-c_{p,s}}\Sigma Cov(\bar{q}_s^0, \bar{q}_t^0)\Sigma\frac{\alpha_t}{1-c_{p,t}}.$$

The inference coefficient $C_{p,t} = Cov(v_t^i, p_t)Var(p_t)^{-1}$, as a function of the current and future demand coefficients, is then characterized by Equation (90).

Differentiating the demand function (92) with respect to p and substituting for $\frac{\partial q^j(p)}{\partial p}$ into $\Lambda^i = -\sum_{j\neq i} \frac{\partial q^j(p)}{\partial p}$, the price impact Λ_t is characterized by Equation (91). Given equilibrium price impact Λ_t (equivalently, the demand reduction parameter $(1-\gamma_t)$), equilibrium allocations in round t are:

$$q_t^{i,0} + q_t^i = \frac{\gamma_t}{\alpha_t}(E[v_t^i|p_t] - \frac{1}{I}\sum_j E[v_t^j|p_t]) + \gamma_t \frac{1}{I}\sum_j q_t^{j,0} + (1-\gamma_t)q_t^{i,0} \qquad \forall i.$$

Next, taking the value function $V_t^i(\cdot)$ and demands $\{q_t^i(\cdot)\}_i$ as given, we show that the recursive characterization (88)-(89) for the value function parameters holds for round $t-1$. In round $t-1$, the value function $V_{t-1}^i(\cdot)$ is characterized recursively:

$$V_{t-1}^i(q_{t-1}^{i,0}+q_{t-1}^i) \equiv E[V_t^i(q_{t-1}^{i,0}+q_{t-1}^i+\delta_t^i+q_t^i(p_t)) - \frac{\kappa}{2}q_t^i(p_t)\cdot\Sigma q_t^i(p_t) - p_t\cdot q_t^i(p_t)|\mathcal{I}_{t-1}^i, p_{t-1}]. \quad (94)$$

Substituting demand (92) at t into the rhs of Equation (94) gives:

$$\begin{aligned} rhs \;= E[const. + v_t^i\cdot(1-\gamma_t)q_t^{i,0} - \gamma_t(v_t^i - p_t)\cdot(1-\gamma_t)q_t^{i,0} - \frac{\alpha_t}{2}(1-\gamma_t)q_t^{i,0}\cdot\Sigma(1-\gamma_t)q_t^{i,0} \\ +\kappa\gamma_t^2\frac{v_t^i - p_t}{\alpha_t}\cdot q_t^{i,0} - \frac{\kappa}{2}\gamma_t^2 q_t^{i,0}\cdot\Sigma q_t^{i,0} + p_t\cdot\gamma_t q_t^{i,0}|\mathcal{I}_{t-1}^i, p_{t-1}]. \end{aligned}$$

Letting $\frac{dV_{t-1}^i}{dq_{t-1}^i} \equiv E[v_{t-1}^i|\mathcal{I}_{t-1}^i, p_{t-1}] - \alpha_{t-1}\Sigma(q_{t-1}^{i,0} + q_{t-1}^i)$, Equations (88)-(89) then give the recursive characterization of the intercept v_{t-1}^i and slope α_{t-1} of the marginal value function. □

Proposition 5 generalizes the characterization of equilibrium in Proposition 4 with symmetric traders to asymmetric equilibrium with heterogeneous traders. We first discuss the differences in the derivation of the symmetric and asymmetric equilibrium.

Differences in the derivation of the symmetric vs. asymmetric equilibrium.

- *Time-varying price distribution.*

When risk preferences are heterogeneous or the joint distribution of $\{\theta_k^i\}_i$ is not equicommonal (see the definition below Equation (29)), equilibrium price aggregates traders' marginal values with asymmetric weights $\{\gamma_t^j\}_j$:

$$p_t = (\sum_j \frac{\gamma_t^j}{\alpha_t^j})^{-1}\sum_j \frac{\gamma_t^j}{\alpha_t^j}(E[v_t^j|\mathcal{I}_t^j, p_t] - \alpha_t^j\Sigma q_t^{j,0}) \qquad \forall t. \quad (95)$$

The weights $\{\gamma_t^j \equiv \frac{\alpha_t^j}{\alpha_t^j+\kappa+\lambda_t^j}\}_j$ change over t. Steps 2 and 3 characterize the conditional expectations and price impacts, which deterministically *vary over time*. Two differences with respect to symmetric markets follow. First, past and current prices convey different information (i.e., "history matters") and receive non-zero weights in trader i's inference:

$$E[v_t^i|\mathcal{I}_t^i, p_t] = C_{\theta,t}^i E[v_t^i] + C_{s,t}^i \mathcal{I}_t^i + C_{p,t}^i p_t, \tag{96}$$

where $C_{\theta,t}^i, C_{p,t}^i \in \mathbb{R}^{2\times 2}$, and $C_{s,t}^i \in \mathbb{R}^{2\times|\mathcal{I}_t^i|}$. Second, a trader's *intertemporal price impact* — the price change at at $\tau > t$ due to a unilateral demand change at t — is not zero, $\Lambda_t^{i,\tau} \equiv \frac{dp_\tau}{dq_t^i} \neq 0$ for $\tau > t$.

Lemma 2 (Intertemporal Price Impacts in Symmetric Markets) Suppose that traders are symmetric, $\alpha^i = \alpha$ for all i, and there are no supply shocks.[128] Then, the intertemporal price impacts $\Lambda_t^{i,\tau} \equiv \frac{dp_\tau}{dq_t^i}$ are zero for all i and $\tau > t$.

This is because with symmetric traders, equilibrium price $p_\tau = \overline{v}_\tau - Q_\tau$ (Equation (93), $\tau > t$) does not depend on outcomes in the current-round t or in rounds between t and τ other than through the unweighted average of traders' inventories $\frac{1}{I}\sum_j q_\tau^{j,0}$. In particular, history $\mathcal{I}_\tau^j$ does not affect the conditional expectations ($E[v_\tau^j|\mathcal{I}_\tau^j, p_\tau] = E[v_t^j|p_\tau]$) and hence, $\overline{v}_\tau$. By the market-clearing condition, the unweighted average of the inventories $\frac{1}{I}\sum_j q_\tau^{j,0}$ is equal to *per capita* supply in the market *in and off* equilibrium (i.e., following a demand increase of trader i at t). It follows that a trader's deviation does not affect Q_τ at $\tau > t$.

- *Nonrecursive backward and forward characterization.*

Step 1 mimics the symmetric market. Steps 2' and 3' below define a fixed point problem between price impacts $\{\Lambda_t^i, \{\Lambda_t^{i,\tau}\}_{\tau>t}\}_{i,t}$ (characterized backward) and inference coefficients $\{C_{p,t}^i\}_{i,t}$ (characterized forward).

(i) The backward characterization of price impacts is not recursive, i.e., the parameters at t cannot be characterized by those at $t+1$ if (and only if) traders are heterogeneous. Similarly to Equation (88), the intercept of value function v_t^i is a convex combination of θ^i and future prices $\{p_\tau\}_{\tau>t}$:

$$v_t^i = \xi_t^{i,T}(\theta^i - p_T) + \sum_{\tau>t}^{T-1} \xi_t^{i,\tau}(p_{\tau+1} - p_\tau) + p_{t+1}. \tag{97}$$

However, in contrast to the market with symmetric traders, the weights $\{\xi_t^{i,\tau}\}_{i,\tau>t}$ (or equivalently intertemporal price impacts $\{\Lambda_t^{i,\tau}\}_{i,\tau>t}$) are determined non-recursively.

[128] See ft. 78.

(ii) The forward characterization of inference coefficients is not recursive unless each trader's demand conditions on the outcomes of other players (or their $I-1$-dimensional linear combination) in every round or traders are symmetric. Then, history $\mathcal{I}_t^i$ does not provide additional information beyond that in the current-round conditioning variables or the current-round price p_t. The nonrecursivity is due to two reasons: past prices include different information and because current-round demand (or a deviation) affects each of the future prices differently.

- *In competitive markets with $\kappa = 0$, these differences in the characterization of symmetric and asymmetric equilibrium do not arise.*

Because $\gamma_t^j \to 1$ for all j and t, price p_t in Equation (95) maps one-to-one to the unweighted average of traders' inventories $\{q_t^{i,0}\}_i$. It follows that, irrespective of the traders' heterogeneity, their conditional expectations in every round only depend on the current-round price (and not all past prices) and their intertemporal price impacts are zero.

Proposition 5 (Equilibrium: Dynamic Market with Heterogeneous Traders) A profile of (net) demand schedules $\{\{q_t^i(\cdot) : \mathbb{R}^K \to \mathbb{R}^K\}_{t=1,\cdots,T}\}_i$ is an equilibrium if and only if, for each i,

(i) Trader i chooses his best response (99), given his marginal value function $\frac{dV_t^i}{dq_t^i} \equiv E[v_t^i|\mathcal{I}_t^i, p_t] - \alpha_t^i \Sigma(q_t^{i,0} + q_t^i)$, price impact Λ_t^i, and conditional expectations $E[v_t^i|\mathcal{I}_t^i, p_t]$;

(ii) Conditional expectations $\{E[v_t^i|\mathcal{I}_t^i, p_t]\}_{i,t}$ are characterized from price distributions (109)-(111), given the value function parameters $\{v_t^i, \alpha_t^i\}_{i,t}$ and price impacts $\{\Lambda_t^i, \{\Lambda_t^{i,\tau}\}_{\tau>t}\}_{i,t}$;

(iii) Value function parameters $\{v_t^i, \alpha_t^i\}_t$ and optimal directions $\{\frac{dq_\tau^i}{dq_t^i}\}_{t,\tau>t}$ solve the fixed point Equations (97)-(105); price impacts $\{\Lambda_t^i, \{\Lambda_t^{i,\tau}\}_{\tau>t}\}_t$ are characterized by Equations (106)-(107).

Subsequently, Example 4 provides the main steps that characterize the counterparts of equilibrium conditions in Proposition 4 for markets with heterogeneous traders.

Example 4 (Dynamic Market ($T < \infty$): Heterogeneous Traders) Consider the dynamic market from Example 3 but now assume that traders' risk aversion $\{\alpha^i\}_i$ is heterogeneous.

Step 1 (Optimization in round t) In each round t, the first-order condition of trader i is:

$$E[v_t^i|\mathcal{I}_t^i, p_t] - \alpha_t^i \Sigma(q_t^{i,0} + q_t^i) = p_t + \Lambda_t^i q_t^i + \kappa \Sigma q_t^i \qquad \forall p_t \in \mathbb{R}^2, \tag{98}$$

given the value function parameters v_t^i and α_t^i, the conditional expectations $E[v_t^i|\mathcal{I}_t^i, p_t]$ and price impact Λ_t^i in round t. The best response schedule of trader i at t is:

$$q_t^i(p_t) = ((\alpha_t^i+\kappa)\Sigma+\Lambda_t^i)^{-1}(E[v_t^i|\mathcal{I}_t^i, p_t]-p_t-\alpha_t^i\Sigma q_t^{i,0}) = \frac{\gamma_t^i}{\alpha_t^i}\Sigma^{-1}(E[v_t^i|\mathcal{I}_t^i, p_t]-p_t)-\gamma_t^i q_t^{i,0}. \quad (99)$$

Let $1-\gamma_t^i \equiv \frac{\kappa+\lambda_t^i}{\alpha_t^i+\kappa+\lambda_t^i}$ be the demand reduction parameter, using that $\Lambda_t^i = \lambda_t^i\Sigma$ for all i and t.

Step 3' (Price impacts and value function parameters) Price impacts $\{\Lambda_t^i\}_{i,t}$ and parameters v_t^i and α_t^i of value function $\{V_t^i(\cdot)\}_{i,t}$ are characterized backward, taking the conditional expectations $E[v_t^i|\mathcal{I}_t^i, p_t]$ for *all* i and *all* t as given.

- *(Nonrecursive) Derivation of the value function parameters.*

Fix the price impacts of trader i, $\{\Lambda_t^i, \{\Lambda^{i,\tau}\}_{\tau>t}\}$. The marginal value function $\frac{dV_t^i}{dq_t^i} = E[v_t^i|\mathcal{I}_t^i, p_t] - \alpha_t^i\Sigma(q_t^{i,0}+q_t^i)$ at t is:

$$\begin{aligned}\frac{dV_t^i}{dq_t^i} &= (1+\sum_{\tau>t}^{T}\frac{dq_\tau^i}{dq_t^i})(\theta^i-\alpha^i\Sigma(q_t^{i,0}+q_t^i+\sum_{\tau>t}E[q_\tau^i|\mathcal{I}_t^i,p_t]))-\sum_{\tau>t}^{T}\frac{dq_\tau^i}{dq_t^i}E[p_\tau|\mathcal{I}_t^i,p_t] \quad (100)\\ &\quad -\sum_{\tau>t}^{T}\frac{dq_\tau^i}{dq_t^i}\kappa\Sigma E[q_\tau^i|\mathcal{I}_t^i,p_t]-\sum_{\tau>t}^{T}\Lambda_t^{i,\tau}E[q_\tau^i|\mathcal{I}_t^i,p_t],\end{aligned}$$

where the Jacobian matrix $\frac{dq_\tau^i}{dq_t^i} = (\frac{dq_{\ell,\tau}^i}{dq_{k,t}^i})_{k,\ell} \in \mathbb{R}^{K\times K}$ is defined by the impact of the trade q_τ^i of trader i in each future round $\tau > t$. $(\frac{dq_\tau^i}{dq_t^i})_{\tau\geq t} \equiv (Id, \{\frac{dq_\tau^i}{dq_t^i}\}_{\tau>t}) \in \mathbb{R}^{K\times K\times(T-t+1)}$ represents the *optimal direction* in the first-order condition of trader i and is endogenized as part of the fixed point problem. The demand increase by trader i at t alters the distribution of future prices $\{p_\tau\}_{\tau>t}$ off equilibrium — unlike in the symmetric market.[129] Given his own deviation, trader i adjusts his conditional expectations $E[v_\tau^i|\mathcal{I}_\tau^i, p_\tau]$ (more precisely, inference coefficients) and demand schedules $q_\tau^i(\cdot)$ for all $\tau > t$ according to the off equilibrium price distribution.[130]

[129] A unilateral demand change in $q^i(\cdot)$ is understood as the change in demand's intercept; this allows for a linear equilibrium. When inference is dynamic (e.g., when traders are heterogeneous), arbitrary functional deviations $q^i(\cdot)$ may give rise to a nonlinear equilibrium.

[130] In the symmetric market, the impact of the demand increase at t on future *trade* $\frac{dq_\tau^i}{dq_t^i}$ is still present, but the future demand *schedule* $q_\tau^i(\cdot)$ of trader i is the same in equilibrium and off equilibrium. Moreover, the partials $\{\frac{dq_\tau^i}{dq_t^i}\}_{\tau>t}$ do not affect trader i's optimization in the current round t. These are followed by the envelope theorem: From the definition of value function (94), the marginal value function $\frac{dV_t^i}{dq_t^i}$ at t can be written in terms of the next-round value function $V_{t+1}^i(\cdot)$ as follows:

$$\frac{dV_t^i}{dq_t^i} = \frac{dV_{t+1}^i}{dq_{t+1}^i} + \frac{dq_{t+1}^i}{dq_t^i}\underbrace{(E[v_{t+1}^i|\mathcal{I}_{t+1}^i, p_{t+1}] - \alpha_{t+1}^i\Sigma q_{t+1}^i - p_{t+1} - (\Lambda_{t+1}^i+\kappa\Sigma)q_{t+1}^i)}_{=\, lhs-rhs \text{ of the first-order condition (98) at } t+1 \,\equiv\, FOC_{t+1}^i}. \quad (101)$$

With interdependence of endowment shocks $\{\delta_t^i\}_{i,t}$ across assets, the Jacobian matrix $\frac{dq_\tau^i}{dq_t^i}$ is proportional to the identity matrix for each $\tau > t$ and t.[131] With a slight abuse of notation, $\frac{dq_\tau^i}{dq_t^i}$ denotes both the matrix $\frac{dq_\tau^i}{dq_t^i} = \frac{dq_\tau^i}{dq_t^i} Id \in \mathbb{R}^{K\times K}$ and the scalar $\frac{dq_\tau^i}{dq_t^i} \in \mathbb{R}$. Taking these Jacobian matrices as given, we first characterize the intercept v_t^i and slope α_t^i of trader i's marginal value function. The expected future demand q_τ^i in Equation (99) are written as functions of the current allocation $q_t^{i,0} + q_t^i$ and future risk return $\{\frac{v_\tau^i - p_\tau}{\alpha_\tau^i}\}_{\tau>t}$: for each $\tau > t$,

$$E[q_\tau^i|\mathcal{I}_t^i, p_t] = \gamma_\tau^i \Sigma^{-1} E[\frac{v_\tau^i - p_\tau}{\alpha_\tau^i} - \sum_{s>t}^{\tau-1}(\prod_{l>s}^{\tau-1}(1-\gamma_l^i))\gamma_s^i \frac{v_s^i - p_s}{\alpha_s^i}|\mathcal{I}_t^i, p_t] - \gamma_\tau^i(\prod_{l>t}^{\tau-1}(1-\gamma_l^i))(q_t^{i,0} + q_t^i) \quad (102)$$

Substituting for $E[q_\tau^i|\mathcal{I}_t^i, p_t]$ from Equation (102) for all $\tau > t$ into Equation (100), v_t^i and α_t^i are characterized as follows: for given $\{v_\tau^i, \alpha_\tau^i, \gamma_\tau^i, \frac{dq_\tau^i}{dq_t^i}\}_{\tau>t}$,

$$\begin{aligned} v_t^i &= (1 + \sum_{\tau>t}^{T} \frac{dq_\tau^i}{dq_t^i})(\theta^i - p_T) + \sum_{s>t}^{T-1}(1 + \sum_{\tau>t}^{s} \frac{dq_\tau^i}{dq_t^i})(p_{s+1} - p_s) + p_{t+1} \quad (103)\\ &\quad -(1 + \sum_{\tau>t}^{T} \frac{dq_\tau^i}{dq_t^i})\alpha^i \sum_{s>t}^{T}(\prod_{l>s}^{T}(1-\gamma_l^i))\gamma_s^i \frac{v_s^i - p_s}{\alpha_s^i} - \sum_{\tau>t}^{T}(\frac{dq_\tau^i}{dq_t^i}\kappa + \lambda_t^{i,\tau})\gamma_\tau^i \frac{v_\tau^i - p_\tau}{\alpha_\tau^i}\\ &\quad + \sum_{\tau>t}^{T}(\frac{dq_\tau^i}{dq_t^i}\kappa + \lambda_t^{i,\tau})\gamma_\tau^i \sum_{s>t}^{\tau-1}(\prod_{l>s}^{\tau-1}(1-\gamma_l^i))\gamma_s^i \frac{v_s^i - p_s}{\alpha_s^i},\\ \alpha_t^i &= \underbrace{(1 + \sum_{\tau>t}^{T} \frac{dq_\tau^i}{dq_t^i})(\prod_{l>t}^{T}(1-\gamma_l^i))\alpha^i}_{\text{“Discounted” risk aversion}} - \underbrace{\sum_{\tau>t}^{T}\gamma_\tau^i(\prod_{l>t}^{\tau-1}(1-\gamma_l^i))(\frac{dq_\tau^i}{dq_t^i}\kappa + \lambda_t^{i,\tau})}_{\text{“Discounted” future cost convexity}}. \quad (104) \end{aligned}$$

In Equation (103), since v_τ^i is a convex combination of θ^i and future prices in rounds $s > \tau$, the intercept v_t^i is a convex combination of the same variables (see Equation (97)). In the two components in the effective risk aversion α_t^i in Equation (104), the "discounting" is endogenous and arises because of demand reduction $\{1-\gamma_l^i\}_{l>t}$. In contrast to the symmetric market, the discounting $(1 + \sum_{\tau>t}^{T} \frac{dq_\tau^i}{dq_t^i})(\prod_{l>t}^{T}(1-\gamma_l^i))$ is not recursive because the factor $(1 + \sum_{\tau>t}^{T} \frac{dq_\tau^i}{dq_t^i})$, which captures a change in the last-round allocation following a demand increase at t, is not recursive.

Further, taking the value function parameters $\{v_\tau^i, \alpha_\tau^i\}_{\tau\geq t}$ as given, the partials $\{\frac{dq_\tau^i}{dq_t^i}\}_{\tau>t}$

In the symmetric market, because the conditional expectations $E[v_{t+1}^i|\mathcal{I}_{t+1}^i, p_{t+1}]$ of trader i (and price p_{t+1}) do not change following his deviation at t, the optimization at $t+1$ (i.e., $FOC_{t+1}^i = 0$ as defined in Equation (101)) ensures that the round-t marginal value function is independent of $\frac{dq_{t+1}^i}{dq_t^i}$, i.e., $\frac{dV_t^i}{dq_t^i} = \frac{dV_{t+1}^i}{dq_{t+1}^i}$.

[131] This is due to the proportionality of Λ_t^i and $\Lambda_t^{i,\tau}$ to Σ: $\Lambda_t^i = \lambda_t^i \Sigma$ and $\Lambda_t^{i,\tau} = \lambda_t^{i,\tau}\Sigma$ for all i, t, and $\tau > t$. Furthermore, analogously to Example 1 (the static market with interdependent values), the independence of private information $\{q_k^{i,0}\}$ across k implies that inference coefficients $C_{p,t}^i$ are proportional to the identity matrix: $C_{p,t}^i = c_{p,t}^i Id$ for all i and t.

at t are characterized by differentiating $q^i_\tau(\cdot)$ in Equation (99) with respect to q^i_t:[132]

$$\frac{dq^i_\tau}{dq^i_t} = \frac{\gamma^i_\tau}{\alpha^i_\tau}\Sigma^{-1}\frac{d(v^i_\tau - p_\tau)}{dq^i_t} - \gamma^i_\tau(Id + \sum_{s>t}^{\tau-1}\frac{dq^i_s}{dq^i_t}) \qquad \forall \tau > t. \tag{105}$$

Using Equation (97), we have:

$$\frac{d(v^i_\tau - p_\tau)}{dq^i_t} = -\xi^{i,T}_t\Lambda^{i,T}_t + \sum_{s>\tau}^{T-1}\xi^{i,s}_t(\Lambda^{i,s+1}_t - \Lambda^{i,s}_t) + (\Lambda^{i,\tau+1}_t - \Lambda^{i,\tau}_t) \qquad \forall \tau > t, i.$$

From Lemma 2, *when traders are symmetric,* $\frac{d(v^i_\tau - p_\tau)}{dq^i_t} = 0$ *for all* i *and* $\tau > t$ because the intertemporal price impacts $\Lambda^{i,s}_t$ are zero for all $s > t$ and i.[133]

- *Derivation of price impacts.*

In each round t, the (within-round) price impact $\Lambda^i_t = \frac{dp_t}{dq^i_t}$ is characterized as in the static market with interdependent values (Section 4.3), taking as given the effective risk aversion α^i_t and the inference coefficients in $E[v^i_t|\mathcal{I}^i_t, p_t]$ in Equation (96) — characterized in Step 2' below — for all i and t:

$$\Lambda^i_t = -(\sum_{j\neq i}\frac{\partial q^j_t(p_t)}{\partial p_t})^{-1} = (\sum_{j\neq i}((\alpha^j_t + \kappa)\Sigma + \Lambda^j_t)^{-1}(Id - C^j_{p,t}))^{-1} \qquad \forall i. \tag{106}$$

The intertemporal price impact $\Lambda^{i,\tau}_t \equiv \frac{dp_\tau}{dq^i_t}$ can be decomposed into an *allocation effect* and an *inference effect*. Differentiating Equation (95) in round $t+1$ with respect to q^i_t gives the *intertemporal price impact of trader* i *in round* t:

$$\begin{aligned}\Lambda^{i,t+1}_t = &\underbrace{(\sum_j \frac{\gamma^j_{t+1}}{\alpha^j_{t+1}})^{-1}\sum_j \frac{\gamma^j_{t+1}}{\alpha^j_{t+1}}(E[\frac{dv^j_{t+1}}{dq^i_t}|\mathcal{I}^j_t, p_t, p_{t+1}] - \alpha^j_{t+1}\Sigma\frac{dq^j_t}{dq^i_t})}_{\text{Allocation effect (due to the change in future payments and allocations)}} \\ &+\underbrace{(\sum_j \frac{\gamma^j_{t+1}}{\alpha^j_{t+1}})^{-1}\sum_{j\neq i} \frac{\gamma^j_{t+1}}{\alpha^j_{t+1}}(\frac{dE[v^j_{t+1}|\mathcal{I}^j_t, \tilde{p}_t, p_{t+1}]}{d\tilde{p}_t}\Lambda^i_t + \frac{dE[v^j_{t+1}|\mathcal{I}^j_t, p_t, \tilde{p}_{t+1}]}{d\tilde{p}_{t+1}}\Lambda^{i,t+1}_t)}_{\text{Inference effect (due to the change in the inference by traders } j\neq i \text{ through conditioning variables)}}.\end{aligned} \tag{107}$$

[132] By the proportionality of price impacts to the payoff covariance, $\frac{d(v^i_\tau - p_\tau)}{dq^i_t}$ is proportional to Σ and hence, $\frac{dq^i_\tau}{dq^i_t}$ is proportional to the identity matrix.

[133] In the symmetric market, the allocation effect becomes zero, because price p_{t+1} aggregates the inventories $\{q^{j,0}_{t+1}\}_j$ with equal weights and the market-clearing condition implies $\sum_j q^{j,0}_{t+1} = 0$ in and off equilibrium. The inference effect in $\Lambda^{i,t+1}_t$ becomes zero when the conditional expectations of asset value are independent of p_t for $j \neq i$ ($E[v^j_{t+1}|\mathcal{I}^j_t, p_t, p_{t+1}] = E[v^j_{t+1}|p_{t+1}]$) — the case in symmetric markets.

The intertemporal price impacts $\Lambda_t^{i,\tau} \equiv \frac{dp_\tau}{dq_t^i}$ with respect to rounds $\tau > t$ can be similarly decomposed. The allocation effect in $\Lambda_t^{i,t+1}$ is present because an increase of demand schedule $q_t^i(\cdot)$ by trader i changes the inventory $q_{t+1}^{j,0} = q_t^{j,0} + q_t^j + \delta_{t+1}^j$ for all j through the price increase Λ_t^i. Because price aggregates (changes in) inventories $\{q_{t+1}^{j,0}\}_j$ with heterogeneous weights $\{\gamma_{t+1}^j\}_j$, the next-round price p_{t+1} changes (as do all future prices p_τ at $\tau > t$). In addition, the inference effect is present because following a unilateral deviation of trader i, the expectations $E[v_{t+1}^j|\mathcal{I}_t^j, p_t, p_{t+1}]$ of traders $j \neq i$ are conditioned on p_t (in fact, all history). Traders $j \neq i$ attribute the increase in p_t off equilibrium due to the change in trader i's demand to a lower aggregate risk instead and increase the intercepts of their demand $q_{t+1}^j(\cdot)$ by $\frac{\gamma_{t+1}^j}{\alpha_{t+1}^j} \frac{dE[v_{t+1}^j|\mathcal{I}_t^j, \tilde{p}_t, p_{t+1}]}{d\tilde{p}_t} \Lambda_t^i$.

- *In the symmetric market with no supply shocks, both components of $\Lambda_t^{i,\tau}$ — the allocation and information effects — are zero for all i, t, and $\tau > t$.*

See Lemma 2. Then, the future demand schedules $q_\tau^i(\cdot)$ at $\tau > t$ do not change following trader i's deviation at t. Therefore, in the symmetric market, equilibrium is characterized by recursive backward induction without loss. Outside of symmetric markets, equilibrium is not recursively characterized. Given the value function and the optimal direction for all i and t as functions of price impacts, Equations (106)-(107) define the price impacts $\{\Lambda_t^i, \{\Lambda^{i,\tau}\}_{\tau>t}\}_{i,t}$ for all i and t, by taking the conditional expectations $\{E[v_t^i|\mathcal{I}_t^i, p_t]\}_{i,t}$ as given.

Step 2' (Conditional expectations in round t) In each round t, Equation (95) characterizes equilibrium price p_t as a weighted average of the marginal value functions' intercepts $\{v_t^i\}_i$ and inventories $\{q_t^{i,0}\}_i$. Substituting

$$E[v_t^i|\{q_s^{i,0}\}_{s\leq t}, \{p_s\}_{s<t}, p_t] = C_{\theta,t}^i \theta^i + C_{q,t}^i (\Sigma q_s^{i,0})_{s\leq t} + C_{s,t}^i (p_s)_{s<t} + C_{p,t}^i p_t \qquad \forall t \tag{108}$$

into Equation (95) gives the equilibrium prices:

$$p_1 = \Big(\sum_j \frac{\gamma_1^j (1 - c_{p,1}^j)}{\alpha_1^j}\Big)^{-1} \sum_j \frac{\gamma_1^j}{\alpha_1^j} (c_{\theta,1}^j \theta^j + c_{q,1}^j \Sigma q_1^{j,0} - \alpha_1^j \Sigma q_1^{j,0}), \tag{109}$$

$$p_t = \Big(\sum_j \frac{\gamma_t^j (1 - c_{p,t}^i)}{\alpha_t^j}\Big)^{-1} \sum_j \frac{\gamma_t^j}{\alpha_t^j} (c_{\theta,t}^j \theta^j + c_{q,t}^j (\Sigma q_s^{j,0})_{s\leq t} + c_{s,t}^j (p_s)_{s<t} - \alpha_t^j \Sigma q_t^{j,0}) \quad 1 < \forall t < T \tag{110}$$

$$p_T = \Big(\sum_j \frac{\gamma_T^j}{\alpha^j}\Big)^{-1} \sum_j \frac{\gamma_T^j}{\alpha^j} (\theta^j - \alpha^j \Sigma q_T^{j,0}). \tag{111}$$

Price p_t in round t is a linear function of past prices $\{p_s\}_{s<t}$ and past inventories $\{q_s^{i,0}\}_{i,s<t}$. By substituting past inventories $q_s^{i,0} = \sum_{l<s}(\delta_l^i + q_l^i) + \delta_s^i$ evaluated at past trades $\{q_l^i\}_{l<s}$ from Equation (102) into Equations (109)-(111), price p_t can be written as a linear function of asset return θ^i, initial endowments $\{q_0^i\}_i$, and endowment shocks $\{\delta_s^i\}_{i,s\leq t}$.

The marginal value function intercept v_t^i at t is a convex combination of θ^i and $\{p_\tau\}_{\tau>t}$. Applying the projection theorem to the price distribution derived from Equations (109)-(111) defines the fixed point problem for inference coefficients, given price impacts. □

In an imperfectly competitive market, the equilibrium allocation of trader i in each round t is:

$$q_t^{i,0} + q_t^i = \frac{\gamma_t^i}{\alpha_t^i}\Sigma^{-1}(E[v_t^i|\mathcal{I}_t^i, p_t] - \overline{v}_t) + \frac{\gamma_t^i}{\alpha_t^i}\Sigma^{-1}Q_t + (1-\gamma_t^i)q_t^{i,0} \qquad \forall i,$$

where the aggregate asset valuation $\overline{v}_t$ and the aggregate risk Q_t are:

$$\overline{v}_t \equiv (\sum_j \frac{\gamma_t^j}{\alpha_t^j})^{-1}\sum_j \frac{\gamma_t^j}{\alpha_t^j}E[v_t^j|\mathcal{I}_t^j, p_t]; \qquad Q_t \equiv (\sum_j \frac{\gamma_t^j}{\alpha_t^j})^{-1}\sum_j \gamma_t^j \Sigma q_t^{j,0}.$$

Equilibrium price satisfies $p_t = \overline{v}_t - Q_t$ in each round t.

Comparison with competitive dynamic market. In the competitive market (i.e., $I \to \infty$) with $\kappa = 0$, the within-round and intertemporal price impacts become zero: $\Lambda_t^i \to 0$ for all i and t, and $\Lambda_t^{i,\tau} \to 0$ for all i, t, and $\tau > t$. In each round t, traders realize the gains from trade due to the current-round shocks $\{\delta_t^i\}_i$. Allocations are thus efficient in each round t for all traders i:

$$q_t^{i,0} + q_t^i = \frac{1}{\alpha_t^i}(\sum_j \frac{1}{\alpha_t^j})^{-1}\sum_j q_t^{j,0}; \qquad p_t = \frac{1}{I}\sum_j \theta^j - (\sum_j \frac{1}{\alpha_t^j})^{-1}\sum_j \Sigma q_t^{j,0}.$$

Relative frequency of asset payments (consumption) and trade in infinite-horizon models. Suppose that a trader receives (and consumes) the asset payments $R \cdot (q_{lT}^{i,0} + q_{lT}^i)$ every T rounds (i.e., in rounds $lT = T, 2T, 3T, ...$) and trades q_t^i in all rounds t (i.e., in rounds $t = 1, ..., T, T+1, ...$). Trader i maximizes a discounted sum of payoffs from asset payments at $lT = T, 2T, 3T, ...$:

$$U(\{q_{lT}^{i,0} + q_{lT}^i\}_l) = E\sum_l^\infty \beta^l\{rw_{lT}^i + \theta_l^i \cdot (q_{lT}^{i,0} + q_{lT}^i) - \frac{\alpha^i}{2}(q_{lT}^{i,0} + q_{lT}^i) \cdot \Sigma(q_{lT}^{i,0} + q_{lT}^i)\}, \qquad (112)$$

where $\beta \in (0,1)$ is the discount factor and r is the return on the risk-free asset. Equilibrium trading decisions in each round t maximize the utility from consumption that depends on holdings of riskless and risky assets (w_0^i, q_0^i) after the last round T.[134] Before assets pay, the value function is:

$$V_T^i(w_0^i, q_0^i) = w_0^i + \bar{\theta}^i \cdot \Sigma q_0^i - \frac{\bar{\alpha}^i}{2}(q_0^i) \cdot \Sigma q_0^i. \qquad (113)$$

The intercept $\bar{\theta}^i$ and slope $\bar{\alpha}^i$ of the marginal value function $V_T^i(\cdot)$ are endogenously determined and depend on the market structure. When traders do not participate in a financial market

[134] The convexity of trading cost κ is normalized by the trading frequency T.

(i.e., in autarky), the marginal value function from consumption is:

$$\frac{dV_T^{i,A}(q_0^i)}{dq_0^i} = v_0^{i,A} - (\frac{\alpha^i + r\kappa}{r^2} + \frac{\alpha^i}{r})\Sigma q_0^i.$$

A competitive trader anticipates trading all of his initial holdings q_0^i for the efficient portfolio $q^{i,**}$ at once. The trader is effectively less risk averse than the agent in autarky:

$$\frac{dV_T^{i,CM}(q_0^i)}{dq_0^i} = v_0^{i,CM} - \frac{\alpha^i + r\kappa}{r^2}\Sigma q^{i,**} - \frac{\alpha^i}{r}\Sigma q_0^i.$$

Due to slow trading, the risk aversion $\bar{\alpha}^i$ in the imperfectly competitive market is between those of the competitive and autarkic agents:

$$\bar{\alpha}^i = \xi^i \frac{\alpha^i + r\kappa}{r^2} + \frac{\alpha^i}{r},$$

where $\xi^i = \xi^i(\{\gamma_t^j\}_{j,t}, \kappa, T) \in (0,1)$ is determined by $q_T^{i,0} + q_T^i$ in Equation (102) at T as a function of the initial endowment q_0^i. The coefficient ξ^i converges to zero as $T \to \infty$ or as $I \to 0$.

Given the value function (113) at lT, each trader i's best response, the conditional expectations, and price impacts in trading rounds between consumption at $(l-1)T$ and lT are characterized in the same way as in the market with a one-time asset payment at T (see Example 3 and 4). Equilibrium is stationary (i.e., α_t^i is constant over all trading rounds t for all i) when the frequency of trading and consumption is the same (i.e., $T = 1$).[135] When trading is more frequent than consumption ($T > 1$), equilibrium is nonstationary in both symmetric and asymmetric markets.[136]

Price dynamics and the martingale property. In a nonstationary equilibrium, equilibrium price process $\{p_t\}_t$ is a martingale if and only if prices are fully revealing in all rounds and traders are symmetric. In a stationary equilibrium, the price process is a martingale if and only if prices are fully revealing.

Definition 8 (Martingale and Markov Process[137]) *Price process* $\{p_t\}_t$ *is*

(a) a martingale if $E[p_{t+1}|p_1, \cdots, p_t] = p_t$ *for all* t*;*

(b) a Markov process if $F(p_{t+1}|p_1, \cdots p_t) = F(p_{t+1}|p_t)$ *for all* t*.*

[135] Equilibrium would be stationary if the concavity of the consumption utility in (113) is zero (i.e., $\overline{\alpha}^i = 0$ for all i). This is ruled out by $\alpha^i > 0$ for all i.

[136] Collin-Dufresne, Daniel, Moallemi, and Sağlam (2015) characterize a nonstationary equilibrium in the transaction costs approach with a stochastic variance.

[137] The Markov property does not imply the martingale property. The converse is true in the Guassian setting.

With symmetric traders, trading costs with time-invariant convexity κ, and no supply shocks (Example 3), market price $\{p_t\}_t$ is a martingale and hence a Markov process:

$$p_{t+1} = p_t + (\alpha - \alpha_t)\Sigma(E[\vec{q}^0_{t+1}] - E[\vec{q}^0_t]) - \alpha\Sigma(\vec{q}^0_{t+1} - \vec{q}^0_t) \qquad \forall t.$$

With heterogeneous traders (Example 4), price process $\{p_t\}_t$ is not a martingale or a Markov process. In Equations (109)-(111), equilibrium price p_t aggregates the inventory $\{q_t^{j,0}\}_j$ with heterogeneous weights $\{\frac{\gamma_t^j}{\alpha_t^j}\}_j$ — the previous-round price is not a sufficient statistic for the information contained in past price history.

- *Different joint assumptions on the payoff concavity and shocks change equilibrium properties of prices.*

The concavity of the value function due to a convex cost per unit of round-t inventory $q_t^{i,0}$ or a supply shock H_t guarantees equilibrium existence. However, equilibrium properties differ depending on the source of the concavity and randomness. Table 1 summarizes some of the differences across models.

Vayanos (1999, 2001) considers a dynamic model with an i.i.d. supply shock $H_t \sim \mathcal{N}(0, \Sigma_H)$ in each round. When trading is more frequent than consumption (i.e., $T > 1$),[138] equilibrium price is:

$$p_t = (\sum_j \frac{\gamma_t^j}{\alpha_t^j})^{-1} \sum_j \frac{\gamma_t^j}{\alpha_t^j}(E[v_t^j|\mathcal{I}_t^j, p_t] - \alpha_t^j \Sigma q_t^{j,0}) - (\sum_j \frac{\gamma_t^i}{\alpha_t^j})^{-1} H_t. \tag{114}$$

With supply shocks, p_t is a function of a time-varying supply H_t. Expected future prices depend on the history $\{p_s\}_{s\le t}$ (not only the current-round price p_t) and, with symmetric traders, are functions of average prices $\frac{1}{t}\sum_{s\le t} p_s$. The price process $\{p_t\}_t$ is not a martingale (i.e., $E[p_{t+1}|p_1, \cdots, p_t] \neq p_t$) and not Markov (i.e., $E[p_{t+1}|p_1, \cdots, p_t]$ is not a function of p_t only).

Impact of public or private shocks. Imperfect competition changes the impact of public shocks (e.g., a supply shock H_t that is common for all traders) and private shocks (e.g., an endowment shock δ_t^i of some trader i) on equilibrium relative to markets with price-taking traders.

- *A public supply shock affects prices through liquidity and inference effects in addition to the fundamental effect.*

[138] In Vayanos (1999), traders have a concave utility from consumption in every round t. We separate the frequency of consumption (at T) from the frequency of trading (at $t = 1, ..., T$) and introduce a convex trading cost. Alternatively, we could have followed Du and Zhu (2017a), who introduce an inventory cost $-\frac{\kappa}{2}(q_t^{i,0}) \cdot \Sigma(q_t^{i,0})$, which is a function of the current-round inventory $q_t^{i,0}$ rather than trade q_t^i.

Table 1: Markov and martingale properties of equilibrium price

	$T<\infty$ Symmetric	$T<\infty$ Heterogeneous	$T=1$ or $T=\infty$ (Stationary) Symmetric/Heterogeneous
Trading cost	$\{p_t\}_t$ is Markov and a martingale	$\{p_t\}_t$ is not Markov Learning is not recursive	$\{p_t\}_t$ is Markov and a martingale
Supply shock	$\{p_t\}_t$ is not Markov Learning is recursive		$\{p_t\}_t$ is not Markov Learning is recursive

Notes: The Markov and martingale properties are defined in Definition 8. Traders' inference is recursive when v_t^i maps one-to-one to the aggregate risk Q_t of which weights $\{\frac{\gamma_t^i}{\alpha_t^i}\}_i$ on traders' inventories $\{q_t^{i,0}\}_i$ are time-invariant. T represents the number of trading rounds relative to the number of rounds in which consumption or asset payment takes place.
Models with trading costs: Sannikov and Skrzypacz (2016), Antil and Duffie (2017), Duffie and Zhu (2017), Du and Zhu (2017a), Examples 3 and 4.
Models with supply (public) shocks: Rostek and Weretka (2015a), Kyle, Obzihaeva, and Wang (2017), and Rostek and Yoon (2019). Vayanos (1999) studies a stationary equilibrium in the absence of supply shocks (i.e., $Var(H_t)\to 0$). Vayanos (2001) allows for supply shocks (i.e., $Var(H_t)>0$).

When a new shock $\hat{H}_t$ arrives at t, price p_t shifts due to two effects: a temporary liquidity effect and a permanent fundamental effects. Denoting price in equilibrium with the new shock $\hat{H}_t$ by $\hat{p}_t$, the price shift is given by:

$$\Delta p_t \equiv \hat{p}_t - p_t = \underbrace{-(\sum_j \frac{1}{\alpha_t^j})^{-1}\Sigma\hat{H}_t}_{\text{Fundamental effect}} \underbrace{-(\sum_j \frac{1}{\alpha_t^j})^{-1}(\sum_j \frac{1-\gamma_t^j}{\alpha_t^j})(\sum_j \frac{\gamma_t^j}{\alpha_t^j})^{-1}\Sigma\hat{H}_t}_{\text{Liquidity effect}}.$$

Recursively, the arrival of shock $\hat{H}_t$ has an inference effect and a fundamental effect on the change in future prices p_τ: $\Delta p_\tau \equiv \hat{p}_\tau - p_\tau = \Delta\overline{v}_\tau - \Delta Q_\tau$ for any $\tau > t$. For instance,

$$\Delta p_{t+1} = \underbrace{(\sum_j \frac{\gamma_{t+1}^j}{\alpha_{t+1}^j})^{-1}\sum_j \frac{\gamma_{t+1}^j}{\alpha_{t+1}^j}\Delta v_{t+1}^j}_{\text{Inference effect } \equiv \Delta\overline{v}_{t+1}} \underbrace{-(\sum_j \frac{\gamma_{t+1}^j}{\alpha_{t+1}^j})^{-1}\sum_j \gamma_{t+1}^j\frac{\gamma_t^j}{\alpha_t^j}(\sum_j \frac{\gamma_t^j}{\alpha_t^j})^{-1}\Sigma\hat{H}_t}_{\text{Fundamental effect } \equiv \Delta Q_{t+1}}.$$

The fundamental effect ΔQ_τ of shock $\hat{H}_t$, defined by the change in the aggregate risk, arises for future rounds $\tau > t$ because $\hat{H}_t$ shifts traders' inventories. The public shock $\hat{H}_t$ increases (or decreases) current and future prices: $\Delta Q_\tau > 0$ for all $\tau \geq t$ if and only if $\hat{H}_t > 0$. This is because the changes in Q_τ and $q_\tau^{i,0}$ are proportional to $\hat{H}_t$. The inference effect $\Delta\overline{v}_\tau$ arises only in the asymmetric market. The shock $\hat{H}_t$ reduces the informativeness of price p_t in inference $E[v_\tau^j|\mathcal{I}_t^j, p_t, p_{t+1}]$ for $\tau > t$.

- *A private endowment shock affects prices through inference and fundamental effects but*

not the liquidity effect.

Suppose that a trader i gets a new private endowment shock $\hat{\delta}_t^i$ in round t. The current price changes due to the fundamental effect, but the liquidity effect is absent:

$$\Delta p_t \equiv \hat{p}_t - p_t = \underbrace{-(\sum_j \frac{\gamma_t^j}{\alpha_t^j})^{-1}\gamma_t^i \hat{\delta}_t^i}_{\text{Fundamental effect}}.$$

Prices in future rounds $\tau > t$ shift due to the fundamental effect, and, in asymmetric markets, also due to the inference effect.

$$\Delta p_{t+1} = \underbrace{(\sum_j \frac{\gamma_{t+1}^j}{\alpha_{t+1}^j})^{-1}\sum_{j\neq i}\frac{\gamma_{t+1}^j}{\alpha_{t+1}^j}\Delta v_{t+1}^j}_{\text{Inference effect}} \underbrace{-(\sum_j \frac{\gamma_{t+1}^j}{\alpha_{t+1}^j})^{-1}\sum_j \gamma_{t+1}^j \Sigma \Delta q_{t+1}^{j,0}}_{\text{Fundamental effect}}.$$

In contrast to the effects of a public shock, a private shock in a trader's endowment can shift current and future prices differently, rather than increasing or decreasing all prices. Moreover, the current price change does not have any liquidity effect when a shock to a trader's endowment is private.

Comparison with the Cournot model (T=2). Suppose that traders are *symmetric*: $\alpha^i = \alpha$ for all i. Equilibrium in Cournot model is characterized non-recursively, as in the dynamic demand submission game for *heterogeneous* traders in Example 4. This is because traders optimize their quantity demanded conditional on expected prices, rather than to-be-realized prices, and hence, the history of past outcomes matters even with symmetric traders. Furthermore, intertemporal price impacts Λ_t^τ and optimal directions $\frac{dq_\tau^i}{dq_t^i}$ are not zero.

Each symmetric trader i optimizes his quantity bid $q_t^i \in \mathbb{R}^2$ to maximize

$$\max_{q_t^i \in \mathbb{R}^2} E[\theta^i \cdot (q_T^{i,0} + q_T^i) - \frac{\alpha}{2}(q_T^{i,0} + q_T^i)\cdot\Sigma(q_T^{i,0} + q_T^i) - \sum_{t\leq 1}^{T}(\frac{\kappa}{2}q_t^i \cdot \Sigma q_t^i + p_t \cdot q_t^i)|\mathcal{I}_t^i]. \tag{115}$$

Step 1 (Optimization in round t) At each round t, the marginal value function of trader i is given by $\frac{dV_t^i}{dq_t^i} \equiv E[v_t^i|\mathcal{I}_t] - \alpha_t\widehat{\Sigma}_t(q_t^{i,0} + q_t^i)$, where $v_t^i \equiv \frac{dV_t^i}{dq_t^i} \in \mathbb{R}^2$ and $\alpha_t\widehat{\Sigma}_t \equiv \frac{d^2V_t^i}{(dq_t^i)^2} \in \mathbb{R}^{2\times 2}$. The first-order condition of trader i is:

$$E[v_t^i|\mathcal{I}_t^i] - \alpha_t\widehat{\Sigma}_t(q_t^{i,0} + q_t^i) = E[p_t|\mathcal{I}_t^i] + \Lambda_t q_t^i + \kappa\Sigma q_t^i,$$

for given price impact Λ_t and *expected price* $E[p_t|\mathcal{I}_t^i]$ (rather than to-be-realized price p_t). Trader i's quantity demanded is:

$$q_t^i = (\alpha_t^i\widehat{\Sigma}_t + \kappa\Sigma + \Lambda_t)^{-1}(E[v_t^i - p_t|\mathcal{I}_t^i] - \alpha_t^i\widehat{\Sigma}_t q_t^{i,0}). \tag{116}$$

Equilibrium price is determined by supply function $p_t = H_t + G_t \sum_j q_t^j$:

$$p_t = H_t + G_t(\alpha_t \widehat{\Sigma}_t + \kappa\Sigma + \Lambda_t)^{-1} \sum_j (E[v_t^i - p_t | \mathcal{I}_t^i] - \alpha_t \widehat{\Sigma}_t q_t^{j,0}). \tag{117}$$

Step 3' (Price impacts and value function parameters) At the last round $T = 2$, the intercept and slope of the value function are the same as those of utility (115): $v_T^i = \theta^i$ and $\alpha_T \widehat{\Sigma}_T = \alpha\Sigma$ for all i.

At round $t = 1$, fix price impact Λ_2 and intertemporal price impact Λ_1^2. Similarly to Equation (100) in Example 4, the value function parameters v_1^i and $\alpha_1 \widehat{\Sigma}_1$ are characterized by differentiating utility (115) with respect to q_1^i:

$$E[(1 + \frac{dq_2^i}{dq_1^i})(\theta^i - \alpha\Sigma(q_2^{i,0} + q_2^i)) - (\kappa \frac{dq_2^i}{dq_1^i}\Sigma + \Lambda_1^2)q_2^i - \frac{dq_2^i}{dq_1^i} p_2 | q_1^{i,0}] = E[p_1 | q_1^{i,0}] + \Lambda_1 q_1^i + \kappa\Sigma q_1^i.$$

Substituting the expected future trade $E[q_2^i | q_1^{i,0}]$ from Equation (116) at $t = 2$, the value function parameters are characterized by:

$$\begin{aligned} v_1^i &= (1 + \frac{dq_2^i}{dq_1^i})\theta^i - \frac{dq_2^i}{dq_1^i} p_2 - ((1 + \frac{dq_2^i}{dq_1^i})\alpha\Sigma + \kappa \frac{dq_2^i}{dq_1^i}\Sigma + \Lambda_1^2)((\alpha + \kappa)\Sigma + \Lambda_2)^{-1}(v_2^i - p_2), \\ \alpha_t^i \widehat{\Sigma}_1 &= (1 + \frac{dq_2^i}{dq_1^i})\alpha\Sigma((\alpha + \kappa)\Sigma + \Lambda_2)^{-1}(\alpha + \kappa)\Sigma - (\kappa \frac{dq_2^i}{dq_1^i}\Sigma + \Lambda_1^2)((\alpha + \kappa)\Sigma + \Lambda_2)^{-1}\alpha\Sigma. \end{aligned}$$

In addition, differentiating future trade q_2^i from Equation (116) at $t = 2$ with respect to q_1^i, the Jacobian matrix $\frac{dq_2^i}{dq_1^i} \in \mathbb{R}^{2\times 2}$ of trader i is:

$$\frac{dq_2^i}{dq_1^i} = -((\alpha + \kappa)\Sigma + \Lambda_2)^{-1}\Lambda_1^2.$$

In general, the slope of the marginal value function $\alpha_1 \widehat{\Sigma}_1$ at $t = 1$ is not proportional to the primitive covariance matrix Σ, unlike in demand submission games. This is because price impacts Λ_2 and Λ_1^2 are not necessarily proportional to the asset covariance Σ.

Price impact at each round t is equal to the slope of the given supply function: $\Lambda_t = G_t$ for all t. Price impact Λ_t^i is independent of traders' best responses (116) and the concavity of their value functions $\alpha_t \widehat{\Sigma}_t$. The intertemporal price impact Λ_1^2 is characterized by differentiating price Equation (117) at $t = 2$ with respect to q_1^i:

$$\Lambda_1^2 = -G_2((\alpha + \kappa)\Sigma + \Lambda_2)^{-1}(\underbrace{I\Lambda_1^2 + \alpha\Sigma}_{\text{Allocation effect}} + \underbrace{\sum_{j\neq i} \frac{dE[p_2 | q_2^{j,0}, \tilde{p}_1]}{d\tilde{p}_1}\Lambda_1}_{\text{Inference effect}}).$$

The inference effect arises even for symmetric traders, because the expected price $E[p_2|q_2^{j,0}, p_1]$ depends on history p_1 in the Cournot model.

Step 2' (Conditional expectations) Price equation (117) at $t = 1, 2$ provides the joint distribution of (p_1, p_2) as a function of price impacts $\{\Lambda_t\}_t$ and intertemporal price impact Λ_1^2. Inference coefficients at $t = 1$ and $t = 2$:

$$\begin{aligned} E[p_2|q_1^{i,0}, q_2^{i,0}, p_1] &= C_{\theta,2}E[p_2] + C_{q,2}(q_1^{i,0}, q_2^{i,0}) + C_{p,2}p_1, \\ E[v_1^i - p_1|q_1^{i,0}] &= C_{\theta,1}E[v_1^i - p_1] + C_{q,1}q_1^{i,0}. \end{aligned}$$

are determined by the projection theorem. In contrast to Example 3 and 4, equilibrium is not *ex post* even at the last round T and even with symmetric traders.

- *Effects of shocks in the dynamic Cournot model.*

When an unexpected increase in supply $\hat{H}_1 = H_1 + \Delta H_1$ arrives at $t = 1$, price at t shifts by ΔH_1, i.e., there exists a (temporary) liquidity effect $\Delta p_1 = \Delta H_1$ as in the demand submission game. Even though p_1 shifts, the shock ΔH_1 does not affect traders' bids q_1^i or inventories $q_2^{i,0}$, because traders submit quantity bids q_1^i rather than contingent schedules $q_1^i(\cdot)$. Hence, the shock does not have a (permanent) fundamental effect on future prices, i.e., $\Delta p_2 = 0$, which is a key difference between the Cournot model and the demand submission game. A round-t announcement of an anticipated supply shock in round $t' > t$ changes all prices p_τ between t and t' in the Cournot model, because of the inference effect $\Delta\overline{v}_\tau$ as in the demand submission game.

C Decentralized Market

C.1 Incomplete Participation

We will work with projected and lifted demands, as described in Section 6.3.1. Using the definition of lifting (Equations (58) and (59)), for any vector $d \in \mathbb{R}^{\sum_n K(n)}$, $d_{K(i)}$ denotes a subvector in $\mathbb{R}^{K(i)}$ that corresponds to the assets traded by trader i $K(i)$ in exchanges $N(i)$. Similarly, for any matrix $M \in \mathbb{R}^{(\sum_n K(n))\times(\sum_n K(n))}$, $M_{K(i)}$ denotes a submatrix in $\mathbb{R}^{K(i)\times K(i)}$ that corresponds to assets $K(i)$ in exchanges $N(i)$.

Trader i chooses a demand schedule $q^i(\cdot) : \mathbb{R}^{K(i)} \to \mathbb{R}^{K(i)}$ as a function of prices $p_{K(i)} \in \mathbb{R}^{K(i)}$ of the assets he trades to maximize a quasilinear quadratic payoff:

$$\max_{q^i(\cdot):\mathbb{R}^{K(i)}\to\mathbb{R}^{K(i)}} E[\theta_{K(i)} \cdot (q^i + q_0^i) - \frac{\alpha^i}{2}(q^i + q_0^i) \cdot V_{K(i)}(q^i + q_0^i) - p_{K(i)} \cdot q^i|q_0^i]. \tag{118}$$

The mean asset value $\theta \in \mathbb{R}^{\sum_n K(n)}$ is known to all traders and the initial endowment $q_0^i \in \mathbb{R}^{K(i)}$ is trader i's private information.

Proposition 3 in Section 6.3.1 characterizes equilibrium in decentralized markets in two steps: Step 1 (Optimization) and Step 3 (Price impact).

Market structure and price impact. Given Propositions 1 and 3, we can characterize the equilibrium properties in decentralized markets by comparing equilibrium price impact $\Lambda^i \in \mathbb{R}^{K(i)\times K(i)}$:

$$\Lambda^i = -((\sum_{j\neq i} \frac{\partial \mathbf{q}^j(p_{K(j)})}{\partial p})^{-1})_{K(i)} = ((\sum_{j\neq i}(\alpha^j \mathbf{V}_{K(j)} + \mathbf{\Lambda}^j)^{-1})^{-1})_{K(i)} \qquad \forall i \tag{119}$$

with that in centralized markets.

- *Price impact Λ^i is not proportional to the asset covariance $V_{K(i)}$.*

By Equation (119), price impacts $\{\Lambda^i\}_i$ are determined by the traders' utility Hessians, $\{\alpha^j V_{K(j)}\}_j$. In decentralized markets, equilibrium price impacts are always concave in the primitive covariance (see Equation (62)) and do not commute in general — with the covariance matrix or across agents.[139] The exception is when the same assets are traded in all exchanges (see Example 5, part (2)). Commutativity is weaker than the proportionality property, which holds in centralized markets (Equation (16)).

- *Trader i's price impact Λ^i in (119) depends on the traders' utilities and assets traded in exchanges $N \setminus N(i)$ via other traders' participation $\{N(j)\}_{j\neq i}$.*

A change in trader j's demand affects not only prices in exchanges $N(j)$ but also prices in all other exchanges $N \setminus N(j)$ that are directly or indirectly connected to $N(j)$. Let $\Psi \equiv \sum_j (\alpha^j \mathbf{V}_{K(j)} + \mathbf{\Lambda}^j)^{-1} \in \mathbb{R}^{(\sum_n K(n))\times(\sum_n K(n))}$ be the slope of the aggregate lifted demand. By the Schur complement equation,[140] Equation (119) becomes:

$$\Lambda^i = (\Psi_{i,i} - \Psi_{i,-i}\Psi_{-i,-i}\Psi_{-i,i} - (\alpha^i V_{K(i)} + \Lambda^i)^{-1})^{-1} \qquad \forall i. \tag{120}$$

The strategic behavior of trader $j \neq i$ in $N(j) \setminus N(i)$ (as well as $N(j) \cap N(i)$) affects trader i's price impact Λ^i through the projection $\Psi_{i,-i}\Psi_{-i,-i}\Psi_{-i,i}$. Traders' price impacts in two exchanges are independent (i.e., the cross-exchange price impacts are zero) when $\Psi_{i,-i} = 0$ — e.g., when the payoffs of assets traded in these exchanges are independent. See also Example 5 below.

[139] Matrices A and B commute if $AB = BA$. Diagonalizable matrices A and B commute if, and only if, they can be simultaneously diagonalized (e.g., Horn and Johnson (2013), Theorem 1.3.12). Commutativity captures a symmetry property of centralized market equilibria that naturally does not hold in decentralized markets, given the traders' heterogeneous participation in exchanges.

[140] $((\Psi^{-1})_{K(i)})^{-1} = \Psi_{i,i} - \Psi_{i,-i}\Psi_{-i,-i}^{-1}\Psi_{i,-i}'$.

The following example illustrates the derivation of price impacts and the effects of incomplete participation in a market where two exchanges are connected through a subset of participating traders.

Example 5 (Indirectly Connected Markets) Consider a market with two exchanges, $N = \{n_1, n_2\}$, and assume that there exist nonempty sets I_1 and I_2, such that $I_1 = \{i : N(i) = \{n_1\}\}$ and $I_2 = \{i : N(i) = \{n_2\}\}$. If we define another set of traders, $I_3 = I \setminus (I_1 \cup I_2)$, then $\{I_1, I_2, I_3\}$ is a partition of the set of traders I. Traders I_3 are intermediaries who trade assets in both exchanges, and two sets of clients I_1 and I_2 trade disjoint subsets of assets. We assume that all traders have the same risk aversion $\alpha^i = \alpha$, and hence, price impact Λ^ℓ is the same for all traders in I_ℓ for $\ell = 1, 2, 3$.

The slope of the aggregate (lifted) market demand is:

$$\Psi \equiv \sum_{\ell=1}^{3} I_\ell \left(\alpha \mathbf{V}_{K(\ell)} + \mathbf{\Lambda}^\ell\right)^{-1} = \begin{bmatrix} I_1 S^1 + I_3 S^3_{11} & I_3 S^3_{12} \\ I_3 (S^3_{12})' & I_2 S^2 + I_3 S^3_{22} \end{bmatrix},$$

where $S^\ell \equiv -\frac{\partial q^\ell(p_{K(\ell)})}{\partial p_{K(\ell)}} = (\alpha V_{K(\ell)} + \Lambda^\ell)^{-1} \in \mathbb{R}^{K(\ell)\times K(\ell)}$ is the demand slope of a trader in group $\ell = 1, 2, 3$. From Equation (120), the intermediaries' price impacts are:

$$\Lambda^3 = \left((I_3 - 1) \begin{bmatrix} S^3_{11} & S^3_{12} \\ (S^3_{12})' & S^3_{22} \end{bmatrix} + \begin{bmatrix} I_1 S^1 & 0 \\ 0 & I_2 S^2 \end{bmatrix} \right)^{-1}, \tag{121}$$

and the price impacts of traders in I_1 are:

$$\Lambda^1 = ((I_1 - 1)S^1 + I_3 S^3_{11} - (I_3 S^3_{12})(I_2 S^2 + I_3 S^3_{22})^{-1}(I_3 S^3_{12})')^{-1}. \tag{122}$$

(1) *Cross-exchange effects:* When $I_3 > 0$, by Equation (121), the intermediaries' price impacts Λ^3 in exchanges n_1 and n_2 are not independent as long as the assets traded in n_1 and n_2 are not independent. Depending on whether the assets in n_1 and n_2 are payoff substitutes or complements in the sense of $\sigma_{k\ell}$ for $k \in K(n_1)$, $\ell \in K(n_2)$, the intermediary's liquidity is higher or lower, relative to the case when the asset payoffs in n_1 and n_2 are independent (Λ^3_{12} is proportional to αV_{12}). The price impact of a trader i in I_ℓ for $\ell = 1, 2$ does not depend on the price impacts, risk aversion, or number of traders in $N \setminus \{n_\ell\}$ if and only if $I_3 = 0$ (i.e., the network is disconnected) or $V_{12} = 0$ (i.e., the payoffs of the assets traded in n_1 and n_2 are independent).

(2) *The same assets in n_1 and n_2:* Suppose that the same assets are traded in both exchanges n_1 and n_2: $V_{K(1)} = V_{K(2)} = V_{12} = \Sigma$. Then, price impacts Λ^l are proportional to the asset covariance matrix Σ and slopes S^l are proportional to Σ^{-1}. Letting

$$\Lambda^\ell \equiv \alpha\beta^\ell\Sigma; \qquad \Lambda^3_{k\ell} \equiv \alpha\beta^3_{k\ell}\Sigma \qquad \forall k, \ell = 1, 2,$$

Equations (121)-(122) are simplified to:

$$\begin{bmatrix} \beta^3_{11} & \beta^3_{12} \\ \beta^3_{12} & \beta^3_{22} \end{bmatrix} = \big((I_3 - 1) \begin{bmatrix} (1+\beta^3_{11}) & (1+\beta^3_{12}) \\ (1+\beta^3_{12}) & (1+\beta^3_{22}) \end{bmatrix}^{-1} + \begin{bmatrix} I_1(1+\beta^1)^{-1} & 0 \\ 0 & I_2(1+\beta^2)^{-1} \end{bmatrix} \big)^{-1}, \quad (123)$$

$$\beta^1 = \Big(\frac{(I_1-1)}{(1+\beta^1)} + \frac{I_3}{(1+\beta^3_{11})} - \frac{(I_3)^2}{(1+\beta^3_{12})^2}\Big(\frac{I_2}{(1+\beta^2)} + \frac{I_3}{(1+\beta^3_{22})}\Big)^{-1}\Big)^{-1}. \quad (124)$$

When $I_3 = 1$ (i.e., a single intermediary), the price impact of the intermediary is block diagonal:

$$\Lambda^3 = \begin{bmatrix} \frac{1}{I_1}(1+\beta^1)\alpha\Sigma & 0 \\ 0 & \frac{1}{I_2}(1+\beta^2)\alpha\Sigma \end{bmatrix}.$$

(3) *One vs. multiple intermediating traders:* When $I_3 > 1$, the system of Equations (123)-(124) gives $\beta^\ell = \beta = \frac{1}{I_1+I_2+I_3-2}$ for $\ell = 1, 2$ and $\beta^3_{k\ell} = \beta = \frac{1}{I_1+I_2+I_3-2}$ for $k, \ell = 1, 2$:

$$\Lambda^1 = \Lambda^2 = \alpha\beta\Sigma; \quad \Lambda^3 = \begin{bmatrix} \alpha\beta\Sigma & \alpha\beta\Sigma \\ \alpha\beta\Sigma & \alpha\beta\Sigma \end{bmatrix}.$$

Since equilibrium in demand schedules is equivalent to the fixed point in price impacts, it follows that *equilibrium (i.e., allocations and payoffs) in the decentralized market* $N = \{n_1, n_2\}$ *are the same as in the centralized market* with the traders and assets from exchanges n_1 and n_2. □

- *When a market becomes more decentralized, traders' price impacts weakly increase in all exchanges.*

A more decentralized market restricts the participation of some traders — with respect to traders or assets:

Definition 9 (More Decentralized Than) *Fix the set of traders and assets* (I, K). *Market* $\{(I(n'), K(n'))\}_{n'}$ *is more decentralized than* $\{(I(n), K(n))\}_n$ *if for any exchange* n' *in* $\{(I(n'), K(n'))\}_{n'}$, *there exists an exchange* n *in* $\{(I(n), K(n))\}_n$ *such that* $I(n') \subseteq I(n)$ *and* $K(n') \subseteq K(n)$.

Given the set of traders and assets (I, K), if market $\{(I(n'), K(n'))\}_{n'}$ is more decentralized than market $\{(I(n), K(n))\}_n$, then the price impacts of all traders are weakly larger: $\{\Lambda^{i,N'}\}_i \geq \{\Lambda^{i,N}\}_i$.

- *Incomplete participation may increase total welfare relative to the centralized market, despite the weakly larger price impacts for all traders.*

Prices and allocations. The best response demand schedule of trader i, $q^i(\cdot) : \mathbb{R}^{K(i)} \to \mathbb{R}^{K(i)}$, is:

$$q^i(p_{K(i)}) = (\alpha^i V_{K(i)} + \Lambda^i)^{-1}(\theta_{K(i)} - p_{K(i)} - \alpha^i V_{K(i)} q_0^i) \qquad \forall p_{K(i)} \in \mathbb{R}^{K(i)}.$$

Equilibrium price vector $p \in \mathbb{R}^{\sum_n K(n)}$ solves the market-clearing condition in all exchanges, $\sum_i \mathbf{q}^i(p_{K(i)}) = 0$,

$$p = \theta - (\sum_j (\alpha^j \mathbf{V}_{K(j)} + \mathbf{\Lambda}^j)^{-1})^{-1} \sum_j (\alpha^j \mathbf{V}_{K(j)} + \mathbf{\Lambda}^j)^{-1} \alpha^j \mathbf{V}_{K(j)} \mathbf{q}_0^j \equiv \theta - Q.$$

Equilibrium allocation of each trader i is:

$$q^i + q_0^i = (\alpha^i V_{K(i)} + \Lambda^i)^{-1} Q_{K(i)} + (\alpha^i V_{K(i)} + \Lambda^i)^{-1} \Lambda^i q_0^i \qquad \forall i.$$

Example 5 Cont'd (Indirectly Connected Markets) Suppose that a single intermediary participates in both exchanges n_1 and n_2, i.e., $I_3 = 1$, and the same assets are traded in both exchanges, i.e., $V_{K(1)} = V_{K(2)} = \Sigma$. We assume that the intermediary is endowed with $-q$ units of the asset. (The split of the intermediary's endowment $-q$ between the two exchanges (i.e., $q_0^3 \in \mathbb{R}^2$) does not affect equilibrium.) Traders of class I_1 are endowed with q/I_1 units of the asset, and class I_2 traders have zero initial endowment. If the market were centralized, the aggregate risk would be $Q^c = \alpha \Sigma \bar{q}_0 = 0$. In the decentralized market $\{n_1, n_2\}$, the aggregate risk Q is given by:

$$Q = \begin{bmatrix} Q_{K(1)} \\ Q_{K(2)} \end{bmatrix} = \begin{bmatrix} I_1 S^1 + S^3_{11} & S^3_{12} \\ (S^3_{12})' & I_2 S^2 + S^3_{22} \end{bmatrix}^{-1} \left(\begin{bmatrix} S^1 \alpha \Sigma q \\ 0 \end{bmatrix} + S^3 \begin{bmatrix} -\alpha \Sigma q \\ -\alpha \Sigma q \end{bmatrix} \right).$$

Observe that the aggregate risk is common to all traders in a given exchange n_ℓ, $\ell = 1, 2$. However, $Q_{K(\ell)}$ is a *weighted* average of the endowments of *all* traders in the exchanges n_1 and n_2. In addition, the exchange-specific aggregate risks $Q_{K(1)}$ and $Q_{K(2)}$ generally differ because the weights are heterogeneous across exchanges; moreover, they differ from $Q^c = 0$. Therefore, the same asset trades at different prices in different exchanges.

The equilibrium allocation in this intermediated market is determined analogously to the centralized market (Equation (24)):

$$\begin{aligned} q^1 + q_0^1 &= S^1 Q_{K(1)} + S^1 \Lambda^1 q, \\ q^2 + q_0^2 &= S^2 Q_{K(2)}, \\ q^3 + q_0^3 &= (1,1) \left(S^3 \begin{pmatrix} Q_{K(1)} \\ Q_{K(2)} \end{pmatrix} + S^3 \Lambda^3 \begin{pmatrix} -q/2 \\ -q/2 \end{pmatrix} \right), \end{aligned}$$

where the multiplication by $(1, 1)$ in the last line sums the holdings of the intermediary in

exchanges n_1 and n_2. The intermediary trades both components $Q_{K(1)}$ and $Q_{K(2)}$ of aggregate risk, whereas all other traders buy or sell its distinct components. □

Comparison with competitive markets. The equilibrium in a competitive market is characterized by traders' optimization (60) in Proposition 3 as $\Lambda^i \to 0$ for all i. A price-taking trader submits a demand schedule:

$$q^i(p_{K(i)}) \to (\alpha^i V_{K(i)})^{-1}(\theta_{K(i)} - p_{K(i)} - \alpha^i V_{K(i)} q_0^i) \qquad \forall p_{K(i)} \in \mathbb{R}^{K(i)}.$$

The price impact $\Lambda^i = 0$ is proportional to $V_{K(i)}$. The competitive prices and allocations are:

$$\begin{aligned} p &= \theta - \Big(\sum_j (\alpha^j \mathbf{V}_{K(j)})^{-1}\Big)^{-1} \sum_j \mathbf{q}_0^j \;\equiv\; \theta - Q^{**}, \\ q^i + q_0^i &= (\alpha^i V_{K(i)})^{-1} Q^{**}_{K(i)} \qquad \forall i; \end{aligned}$$

i.e., equilibrium is efficient given the market structure $\{(I(n), K(n))\}_n$.

Comparison with the Cournot model. Consider a decentralized market in the Cournot model with an exogenous supply in $\mathbb{R}^{\sum_n K(n)}$: $p = H + G\sum_i \mathbf{q}^i$ for a non-singular matrix $G \in \mathbb{R}^{(\sum_n K(n))\times(\sum_n K(n))}$ and a vector $H \in \mathbb{R}^{(\sum_n K(n))}$. Trader i chooses a quantity $q^i \in \mathbb{R}^{K(i)}$ to maximize

$$\max_{q^i \in \mathbb{R}^{K(i)}} E[\theta_{K(i)} \cdot (q^i + q_0^i) - \frac{\alpha^i}{2}(q^i + q_0^i) \cdot V_{K(i)}(q^i + q_0^i) - p_{K(i)} \cdot q^i | q_0^i].$$

The first-order condition of trader i is:

$$\theta_{K(i)} - \alpha^i V_{K(i)}(q^i + q_0^i) = E[p_{K(i)}|q_0^i] + \Lambda^i q^i \qquad \forall i.$$

As in the centralized market in the Cournot model, equilibrium is not *ex post* and, by the definition of the Cournot game, the price impact Λ^i is independent of inference coefficients and traders' demands: $\Lambda^i = G_{K(i)}$ for all i.

C.2 Incomplete Demand Conditioning

Consider a model with two risky assets, each traded in a separate exchange. Each trader participates in both exchanges and submits a demand schedule $q_k^i(\cdot) : \mathbb{R} \to \mathbb{R}$ in each exchange $k = 1, 2$ to maximize his expected payoff:

$$\max_{\{q_k^i(\cdot):\mathbb{R}\to\mathbb{R}\}_{k=1,2}} E[\theta \cdot (q^i + q_0^i) - \frac{\alpha^i}{2}(q^i + q_0^i) \cdot \Sigma(q^i + q_0^i) - p \cdot q^i | q_0^i]. \tag{125}$$

Equilibrium prices are determined by market clearing exchange by exchange: $\sum_i q_k^i(p_k) = 0$ for $k = 1, 2$. The *uncontingent* schedule $q_k^i(p_k)$ for asset k is not measurable with respect to prices of other assets p_ℓ for $\ell \neq k$ (cf. contingent schedule $q_k^i(p_1, p_2)$, $k = 1, 2$). Due to the non-measurability of demands across assets, equilibrium derivation for the market with incomplete demand conditioning in Steps 1"-3" below differs from Steps 1-3 in the market with contingent schedules.

Step 1" (Optimization) Fix price impacts $\{\lambda_k^i \equiv \frac{dp_k}{dq_k^i} \in \mathbb{R}\}_k$ and conditional expectations $\{E[p_\ell|q_0^i, p_k]\}_{k,\ell\neq k}$ for all traders. The first-order conditions of trader i for *uncontingent* demand schedules for assets $k = 1, 2$ are:

$$0 = E[\underbrace{\theta_1 - \alpha^i(\sigma_{11}(q_1^i + q_{1,0}^i) + \sigma_{12}(q_2^i + q_{2,0}^i)) - p_1 - \lambda_1^i q_1^i}_{\partial u^i/\partial q_1^i} |q_0^i, p_1] \quad \forall p_1 \in \mathbb{R}, \tag{126}$$

$$0 = E[\underbrace{\theta_2 - \alpha^i(\sigma_{22}(q_2^i + q_{2,0}^i) + \sigma_{21}(q_1^i + q_{1,0}^i)) - p_2 - \lambda_2^i q_2^i}_{\partial u^i/\partial q_2^i} |q_0^i, p_2] \quad \forall p_2 \in \mathbb{R}. \tag{127}$$

Due to the non-measurability of demands across exchanges, trader i's best response for asset k depends on expected trades $E[q_\ell^i|q_0^i, p_k]$ for assets $\ell \neq k$. Equilibrium is not *ex post* in general even when traders are symmetric and θ is common for all traders.

Step 2" (Conditional expectations) Conditional expectations $\{E[p_\ell|q_0^i, p_k]\}_{k,\ell\neq k}$ are characterized by the projection theorem, given price impacts $\{\lambda_k^i\}_{i,k}$.

We endogenize the demand coefficients and conditional expectations as functions of price impacts. To this end, we treat a trader's demand for asset k as a linear function of q_0^i and p_k,

$$q_k^i(p_k) \equiv a_k^i - b_k^i \cdot q_0^i - s_k^i p_k \qquad \forall p_k \in \mathbb{R}, \tag{128}$$

and characterize all demand coefficients — the demand intercept $a_k^i \in \mathbb{R}$, the inference coefficients $b_k^i \in \mathbb{R}^2$, and the demand slope $s_k^i \in \mathbb{R}$ — as functions of price impact alone. Taking the conditional expectation of (128) for asset ℓ gives the expected trade as a function of the expected price:

$$E[q_\ell^i|q_0^i, p_k] = a_\ell^i - b_\ell^i \cdot q_0^i - s_\ell^i E[p_\ell|q_0^i, p_k] \qquad \forall p_k \in \mathbb{R}. \tag{129}$$

By market clearing for each asset k applied to schedules (128), the equilibrium price is:

$$p_k = (\sum_j s_k^j)^{-1} \sum_j (a_k^j - b_k^j \cdot q_0^j). \tag{130}$$

Coefficients b_k^j determine the variance of the price distribution upon which expected trades depend. Substituting for expected prices from (130) into Equation (129) gives expected trades

as linear functions of expected prices: by the projection theorem,

$$E[q_\ell^i|q_0^i,p_k]=a_\ell^i-b_\ell^i\cdot q_0^i-s_\ell^i(E[p_\ell|q_0^i]+\frac{Cov(p_\ell,p_k|q_0^i)}{Var(p_k|q_0^i)}(p_k-E[p_k|q_0^i])). \tag{131}$$

The inference coefficient $\frac{Cov(p_\ell,p_k|q_0^i)}{Var(p_k|q_0^i)}$ captures correlated price risk in p_ℓ and p_k, given trader i's private information q_0^i:

$$\frac{Cov(p_\ell,p_k|q_0^i)}{Var(p_k|q_0^i)}=\frac{(\sum_j s_\ell^j)^{-1}\sum_{j,h\neq i}b_k^j\cdot Cov(q_0^j,q_0^h|q_0^i)b_\ell^h\ (\sum_j s_k^j)^{-1}}{(\sum_j s_k^j)^{-1}\sum_{j,h\neq i}b_k^j\cdot Cov(q_0^j,q_0^h|q_0^i)b_k^h\ (\sum_j s_k^j)^{-1}}. \tag{132}$$

Substituting (131) into the first-order condition (126)-(127) defines a fixed point problem for $\{a_k^i,b_k^i,s_k^i\}_{i,k}$, given price impacts $\{\lambda_k^i\}_{i,k}$.

Step 3" (Price impact) In each exchange $k=1,2$, the price impact $\lambda_k^i\equiv\frac{dp_k}{dq_k^i}\in\mathbb{R}$ is the slope of the inverse residual supply function:

$$\lambda_k^i\equiv\frac{dp_k}{dq_k^i}=-(\sum_{j\neq i}\frac{\partial q_k^j(p_k)}{\partial p_k})^{-1}=(\sum_{j\neq i}s_k^j)^{-1}\qquad\forall i. \tag{133}$$

Because market clearing is independent across exchange, the cross-asset price impacts are zero: $\lambda_{k\ell}^i\equiv\frac{dp_\ell}{dq_k^i}=0$ for all $k,\ell\neq k$, and all i.

Proposition 6 (Equilibrium: Incomplete Demand Conditioning) The profile of uncontingent schedules $\{q_k^i(\cdot):\mathbb{R}\to\mathbb{R}\}_{k,i}$ is an equilibrium if and only if

(i) Each trader i chooses his best response by (126)-(127);

(ii) Expected trades $\{E[q_\ell^i|q_0^i,p_k]\}_{k,\ell\neq k}$ are given by (131), given price distribution (130);

(iii) Price impacts $\{\lambda_k^i\}_i$ are characterized by the fixed point (133) for each asset k.

Example 6 (Symmetric Traders and Equilibrium Derivation in Matrices) Suppose that $\alpha^i=\alpha$ and $E[q_0^i]=E[\bar{q}_0]$ for all i. Without loss, we assume that the assets have the same variance:

$$\Sigma=\begin{bmatrix}1&\rho\\\rho&1\end{bmatrix}.$$

By the symmetry across traders and assets, the price impact is symmetric across traders and assets: $\lambda_k^i\equiv\lambda$ for all $k,\ell\neq k$ and all i. Traders' demand coefficients in (128) are symmetric across traders: $a_k^i=a_k$, $b_k^i=b_k$ for all i and $s_k^i=s$ for all i and all k. Observe that the vector b_k is also symmetric across k up to a permutation: i.e., if $b_1=(x,y)$, then $b_2=(y,x)$.

The equilibrium price (130) simplifies in the symmetric market to:

$$p_k=s^{-1}(a_k-b_k\cdot\bar{q}_0).$$

The conditional distribution of price, given private information q_0^i is:

$$E[p_k|q_0^i] = s^{-1}(a_k - b_k \cdot E[\bar{q}_0]) + \frac{Cov(p_k, q_0^i)}{Var(q_0^i)}(q_0^i - E[q_0^i]) = s^{-1}(a_k - b_k \cdot (\sigma_0^2 q_0^i + (1-\sigma_0^2)E[\bar{q}_0])), \tag{134}$$

where $\sigma_0^2 \equiv \frac{cov(\bar{q}_0^j, q_0^i)}{var(q_0^i)} = \frac{\sigma_{cv}^2 + \frac{1}{I}\sigma_{pv}^2}{\sigma_{cv}^2 + \sigma_{pv}^2}$. In addition, the inference coefficient (132) in $E[p_\ell|q_0^i, p_k]$ becomes:

$$\frac{Cov(p_\ell, p_k|q_0^i)}{Var(p_k|q_0^i)} = \frac{b_k \cdot b_\ell}{b_k \cdot b_k} \qquad \forall k, \ell \neq k. \tag{135}$$

From the conditional distribution of prices (134)-(135), we get the expected trade (131) for assets $\ell \neq k$ as a linear function of q_0^i and p_k:

$$E[q_\ell^i|q_0^i, p_k] = -\frac{b_k \cdot b_\ell}{b_k \cdot b_k}a_k - (1-\sigma_0^2)(b_\ell - \frac{b_k \cdot b_\ell}{b_k \cdot b_k}b_k) \cdot E[\bar{q}_0] - ((1-\sigma_0^2)b_\ell + \sigma_0^2 \frac{b_k \cdot b_\ell}{b_k \cdot b_k}b_k) \cdot q_0^i - s\frac{b_k \cdot b_\ell}{b_k \cdot b_k}p_k. \tag{136}$$

Substituting for $E[q_\ell^i|q_0^i, p_k]$ from (136) into the first-order condition (126)-(127), or equivalently, into the equation

$$0 = \theta_k - \alpha(q_{k,0}^i + \rho q_{\ell,0}^i) - (\alpha + \lambda)q_k^i - \rho\alpha E[q_\ell^i|q_0^i, p_k] - p_k),$$

we have that this equality holds for all $q_0^i \in \mathbb{R}^2$ and all $p_k \in \mathbb{R}$ by pointwise optimization. Thus, equalizing the coefficients of q_0^i and p_k to zero defines the fixed point problem for demand coefficients $\{a_k, b_k, s_k\}_k$, given the price impact λ.

For simplicity of notation, we write the fixed point problem in matrix form. Let $A \equiv (a_1, a_2)' \in \mathbb{R}^2$ be the vector of intercepts, and let $B \equiv (b_1, b_2)' \in \mathbb{R}^{2\times 2}$ and $S \equiv diag(s_1, s_2) \in \mathbb{R}^{2\times 2}$ be the matrices of coefficients for q_0^i and p in demand (128). Given the price impact matrix $\Lambda \equiv diag(\lambda, \lambda)$, A and B are characterized in closed form:

$$B = ((1-\sigma_0^2)(\alpha\Sigma + \Lambda) + \sigma_0^2 S^{-1})^{-1}\alpha\Sigma, \tag{137}$$

$$A = S\theta - S((\alpha\Sigma + \Lambda) - S^{-1})B(1-\sigma_0^2)E[\bar{q}_0], \tag{138}$$

and an implicit equation for $S = diag(s, s)$:

$$\alpha + \lambda - s^{-1} + \alpha\rho\frac{b_k \cdot b_\ell}{b_k \cdot b_k} = 0. \tag{139}$$

Lastly, from Equation (133), price impact is $\lambda = \frac{1}{I-1}s^{-1}$, which simplifies Equation (139) to

$$\alpha - (I-2)\lambda + \alpha\rho\frac{2xy}{x^2 + y^2} = 0, \tag{140}$$

where $x \equiv (1-\sigma_0^2)(1-\rho^2)\alpha + (1+(I-2)\sigma_0^2)\lambda$ and $y \equiv \rho(1+(I-2)\sigma_0^2)\lambda$ characteirze $b_1 = (x, y)$ and $b_2 = (y, x)$. The price impact λ characterizes a unique equilibrium from Equation (140).

□

Price impacts and incomplete demand conditioning. Consider Example 6.

- *Price impacts are the same in contingent and uncontingent markets if and only if asset payoffs are independent.*

When asset payoffs are independent (i.e., $\rho = 0$), the system of Equation (140) gives $\lambda = \frac{\alpha}{I-2}$. Then, the price impact is the same as with contingent demands:

$$\Lambda = \begin{bmatrix} \frac{\alpha}{I-2} & 0 \\ 0 & \frac{\alpha}{I-2} \end{bmatrix}.$$

From the equivalence between equilibrium in demand schedules and a fixed point in price impacts in Propositions 1 and 6, equilibrium coincides with contingent and uncontingent demands.

- *Equilibrium in uncontingent markets is equivalent to that in contingent markets if and only if all assets are either perfectly correlated or independent.*

When $\rho = \pm 1$, the unique solution to the system of Equation (140) is:

$$\Lambda = \begin{bmatrix} \frac{2\alpha}{I-2} & 0 \\ 0 & \frac{2\alpha}{I-2} \end{bmatrix}.$$

The price impact *per unit* of the perfectly correlated asset is the same as in the contingent market (i.e., $\frac{\alpha}{I-2}$). It follows from Propositions 1 and 6 that the equilibrium utilities are the same in the uncontingent and the contingent markets. In particular, inference is perfect across exchanges, i.e., $E[p_\ell | q_0^i, p_k] = p_\ell$, because asset prices are perfectly correlated.[141]

- *With imperfectly correlated assets, the price impact λ increases as the price correlation $\frac{Cov(p_\ell, p_k | q_0^i)}{Var(p_k | q_0^i)} = \frac{b_k \cdot b_\ell}{b_k \cdot b_k}$ increases.*

When assets are correlated, the price impact Λ in uncontingent markets is a diagonal matrix, and hence is not proportional to Σ in contrast to the price impact in the contingent market. From Equation (139), λ is increasing and convex in $\frac{b_k \cdot b_\ell}{b_k \cdot b_k}$. This effect captures adverse selection across assets.

[141] The perfect inference is shown in Equation (140): Substituting $\rho = 1$ into x and y gives $b_1 = b_2 = (1 + (I-2)\sigma_0^2)\lambda \mathbf{1}_2$. The proportionality of b_1 and b_2 to vector $\mathbf{1}_2 \equiv (1, 1)$ implies, by the price Equation (130), that both prices p_1 and p_2 are perfectly correlated with $\bar{q}_{1,0} + \bar{q}_{2,0}$. Similarly, when $\rho = -1$, both b_1 and b_2 are proportional to $(1, -1)$, and prices p_1 and p_2 are perfectly correlated with $\bar{q}_{1,0} - \bar{q}_{2,0}$. Inference is perfect $E[p_\ell | q_0^i, p_k] = p_\ell$ for all k and $\ell \neq k$.

- *Incomplete demand conditioning can improve welfare relative to the contingent market.*

The (expected) welfare is strictly larger than that in the contingent market if the expected trading needs of all traders i are such that $sign((E[\bar{q}_{k,0}]-E[q^i_{k,0}])(E[\bar{q}_{\ell,0}]-E[q^i_{\ell,0}])) = sign(\sigma_{k\ell})$ for all asset k and $\ell \neq k$.

When the market is competitive (i.e., $I \to 0$), $\Lambda^i \to 0$ for all i. The *ex ante* welfare is unambiguously lower in the uncontingent market relative to the contingent market due to the information error $E[p_\ell|q^i_0, p_k] - p_\ell$ for $\ell \neq k$ and k. Taking the limit of the demand coefficient B (Equation (137)) as $I \to \infty$, we have:

$$B \to ((1-\sigma_0^2)\alpha\Sigma + \sigma_0^2 S^{-1})^{-1}\alpha\Sigma \quad \neq \quad Id,$$

because the diagonal matrix $S = diag(s_1, s_2) > 0$ is not proportional to Σ^{-1} except when asset payoffs are independent. This implies that equilibrium is not *ex post* in the competitive market even when traders' private information q^i_0 is independent across traders i and assets k.

Prices and allocations. When the matrix coefficients $S^i \equiv diag(s^i_k) \in \mathbb{R}^{K\times K}$ and $B^i \equiv (b^i_1, ..., b^i_K)' \in \mathbb{R}^{K\times K}$ are defined as in Example 6,[142] equilibrium prices and allocations are:

$$p = \overline{v} - Q, \tag{141}$$

$$q^i + q^i_0 = S^i(v^i - \overline{v}) + B^i\ ((S^i)^{-1}B^i)^{-1}Q + (Id - B^i)q^i_0, \tag{142}$$

where the aggregate valuation $\overline{v}$ and the aggregate risk Q are:

$$\overline{v} \equiv \theta - E[Q^c] + E[Q]; \qquad Q \equiv (\sum_j S^j)^{-1} \sum_j B^j q^j_0. \tag{143}$$

Valuation adjustment $S^i(v^i - \overline{v})$ in allocation $q^i + q^i_0$ is present due to the imperfect inference as in Section 4.3. The valuation adjustment is zero for all i if and only if traders are symmetric in $\alpha^i = \alpha$ and $E[q^i_0] = E[q_0]$ for all i, or when all assets are independent so cross-asset inference does not affect equilibrium.[143]

(Non-)Redundant financial innovation. Consider an uncontingent market with two assets and suppose that a new security is introduced to be traded in a separate exchange. Its return is a linear combination of the returns of the existing assets: $r_d = \omega_1 r_1 + \omega_2 r_2$, $\omega = (\omega_1, \omega_2)' \in \mathbb{R}^2$. Traders hold no additional endowment of the new asset. Denote equilibrium

[142] The matrices S^i and B^i are counterparts of the coefficients $(\alpha^i\Sigma + \Lambda^i)^{-1}$ and $(\alpha^i\Sigma + \Lambda^i)^{-1}\alpha^i\Sigma$, respectively, for the contingent demand schedule in Equation (84). This implies that $q^{i,*} \equiv ((S^i)^{-1}B^i)^{-1}Q$ is the counterpart of $(\alpha^i\Sigma)^{-1}Q^c$, which is the aggregate risk portfolio in the centralized market.

[143] $S^i(v^i - \overline{v}) = E[(B^i q^i_0 - S^i Q) - (\alpha^i\Sigma + \Lambda^i)^{-1}(\alpha^i\Sigma q^i_0 - Q^c)]$, where Q^c is the aggregate risk in the contingent market in Equation (26).

price impact in the new market by $\Lambda \equiv diag(\Lambda_a, \Lambda_d) \in \mathbb{R}^{3\times 3}$, where $\Lambda_a \in \mathbb{R}^{2\times 2}$ corresponds to the price impact for two existing assets and $\Lambda_d \in \mathbb{R}$ is the price impact for the derivative asset. Similarly, $q_a^i \in \mathbb{R}^2$ and $q_d^i \in \mathbb{R}$ denote equilibrium trades for the underlying assets and the derivative.

Definition 10 (Projected Price Impact) *The* projected price impact $\widetilde{\Lambda} \in \mathbb{R}^{2\times 2}$ *is the unique matrix such that*

$$q_a^i + q_d^i \omega = (\alpha\Sigma + \widetilde{\Lambda})^{-1}\alpha\Sigma(\bar{q}_0 - q_0^i) \qquad \forall i, \tag{144}$$

for any $\{q_0^i\}_i$.

In Equation (144), equilibrium demand for the existing assets is proportional to $(\bar{q}_0 - q_0^i)$ with a proportionality coefficient $(\alpha\Sigma + \widetilde{\Lambda})^{-1}\alpha\Sigma$.

By substituting equilibrium allocation (142) into Equation (144), the projected price impact $\widetilde{\Lambda}$ is characterized as:

$$\widetilde{\Lambda} = (\Lambda_a^{-1} + \omega\Lambda_d^{-1}\omega')^{-1} = \Lambda_a - \Lambda_a\omega(\Lambda_d + \omega'\Lambda_a\omega)^{-1}\omega'\Lambda_a. \tag{145}$$

By Proposition 6, in an imperfectly competitive market, a new asset is neutral for equilibrium utilities if and only if its introduction does not change the projected price impact: i.e., $\widetilde{\Lambda} = \Lambda_a$.

D Relation to Other Models and Results

D.1 Shapley and Shubik model

The Shapley and Shubik (1977) model is often considered the closest to a demand submission game. The key difference is how the models define the counterfactual that determines a trader's price impact. In the Shapley-Shubik model, traders commit to the money they bid without knowing the equilibrium price, and price is determined as the ratio of the money bid to the total supply. If the price changes (following a deviation by one of the players), the quantity allocated to each trader would still be determined by the same amount of money bid, which is *fixed* off equilibrium. That is, traders are not allowed to adjust optimally the amount they bid for different prices, and their off-equilibrium allocation is suboptimal. This critically affects the properties of the counterfactual (price impact).

As in the Shapley-Shubik model, traders in the uniform-price double auction submit their demands without knowing the equilibrium price, but the demand schedules allow them to make quantity choices contingent on price realizations. The traders can, thus, adjust their quantity demanded optimally for different price realizations. Price impact is then determined

by optimization and market clearing both on and off equilibrium.[144]

Another challenge with market games is that, for traded assets, arbitrage pricing does not hold for traded assets. Indeed, in the Shapley-Shubik model, derivatives cannot be priced from the prices of fundamental assets, even if such derivatives are traded in the market. Peck and Shell (1989) and Koutsougeras (2003) showed that state prices need not be well defined at the equilibria of the strategic market games of Shapley and Shubik (1977).

D.2 Multiplicity of Nash Equilibria

The equilibrium multiplicity was observed in demand submission games with and without uncertainty (e.g., exogenous supply or private information). Fix a profile of strategies for all players but i (Grossman (1981)). In a game with no uncertainty, Nash Equilibrium requires that a player best responds (i.e., equalizes his marginal utility and marginal revenue) *at* the equilibrium price p^* ($p \in \mathbb{R}^K$ or, for simplicity, assume $K = 1$):

$$\frac{\partial u^i(q_0^i + q^i)}{\partial q^i} = p^* + \Lambda^i q^i, \tag{146}$$

but it does not restrict quantities at other prices. However, there are many best response schedules that satisfy condition (146). Since i's best response defines the residual supply for other players, the indeterminacy of best responses induces the multiplicity of equilibria. More precisely, the quantities optimal for players $j \neq i$ at p^* are functions of the slope of i's demand (and, hence, of trades at prices different than equilibrium price). The corresponding multiplicity of equilibrium outcomes is not 'local' around some noncompetitive outcome. In fact, in a double-auction game with $I > 2$ players, the Nash equilibrium outcomes vary between competitive and autarky ones.[145] This result can be seen in the first-order condition (146): a player's best-response behavior depends on what he anticipates to be the consequences of trading a bit more — as captured precisely by the counterfactual that defines a trader's price impact (Λ^i). In a model with no randomness, this counterfactual is not specified. If a player assumes that the price will not change following his unilateral deviation (i.e., $\Lambda^i = 0$ for all i), this counterfactual yields the competitive outcome; if all traders submit quantity schedules that are not contingent on price (i.e., vertical (net) demands), the corresponding counterfactual represents an illiquid market ($\Lambda^i \to \infty$ for all i) and leads to no trade. Hence, Nash equilibrium has no predictive power: by studying Nash equilibrium behavior, one does not learn anything about outcomes that is not already implied by the primitives of the market.

[144] In a demand-function version of the Shapley-Shubik model, the demand has a particular functional form, regardless of the primitive utility functions (e.g., Peck and Shell (1989)).

[145] This result was shown by Grossman (1981) in a complete information game. Bergemann, Heumann, and Morris (2018) make a related point in an incomplete information game, allowing the distributions to vary. Similarly, in Negishi (1961), demands are restricted only *at* the equilibrium price; Gary-Bobo (1986) gave the corresponding result about equilibrium multiplicity.

To summarize, in a game with no randomness, the equilibrium multiplicity is due to the fact that the off-equilibrium counterfactual is not specified. With randomness, the main challenge in characterizing equilibrium is the determination of price impacts. With quadratic payoffs and jointly Gaussian random variables, price impacts are constant.

Subsequent literature introduced uncertainty. The source of uncertainty can be exogenous to all players (e.g., due to exogenous supply in Klemperer and Meyer (1989) or trembles in the symmetric information model in Vayanos (1999, 2001)) or private information (e.g., asymmetric information about the asset value or traders' endowments in Kyle (1989), Vayanos (1999), Vives (2011), and Rostek and Weretka (2012)). With uncertainty about the price to be realized, traders respond optimally for all prices in the support rather than just the realized equilibrium price; equivalently, they respond to each realization of the residual supply whose intercept depends on the realization of the random variable (see, e.g., Equation (77)). All of these papers assume quadratic payoffs (or an equivalent) and establish the uniqueness of the linear equilibrium. (In a model with more general utilities, Klemperer and Meyer (1989) also show that the set of equilibria is connected.)