CONTENTS

1

Introduction

Securities are traded in one of two types of market structures: centralized exchanges or decentralized over-the-counter (OTC) markets. Standard models in financial economics, however, are largely designed to study centralized exchanges, and ignore many of the key frictions that affect asset prices, liquidity, and fragility in OTC markets. This oversight is significant because OTC markets are both *large* and *important*: they are at least as large as those markets that are organized as centralized exchanges, and play a crucial role in the functioning of our modern financial system. Moreover, OTC markets are fundamentally *different* than exchanges along a number of key dimensions, including the information available to market participants, the institutions and technologies they use to trade with one another, and the methods they use to determine the terms of trade.

Over the past two decades, in recognition of the dissonance between standard models and the realities of OTC markets, economists have developed new models that are capable of analyzing empirical regularities in the data and providing theoretical predictions for policy proposals and technological innovations in OTC markets. At the time of writing, this large and growing area of research is missing a graduate-level textbook—a synthesis of existing methods and results that can familiarize scholars with the literature and convey the basic technical skills they need to jump-start their own research on the topic. Our objective with this book is to provide such a synthesis. We take the reader on a deep dive into the economics of OTC markets, developing a unified framework that encompasses many of the key theoretical developments in this growing literature and, where possible, relating these developments to recent policy debates and empirical findings that are emerging from the availability of bigger and better data sets.

1.1 OTC Markets Are Large and Important

Many assets trade in OTC markets. Commonly cited examples include fixed income securities, such as corporate and government debt, and derivatives such as asset-backed securities, interest rate swaps, and foreign-exchange forward contracts. Most short-term funding markets also operate as OTC markets, including the markets for fed funds and repurchase agreements or *repos*. In addition, a number of markets for non-financial assets also operate OTC, such as the markets for used capital (e.g., commercial aircraft) and real estate. In fact, even a sizeable fraction of equity trading volume is executed through some form of OTC market.

The total supply of publicly listed securities traded in OTC markets in 2021 was about USD66 trillion, which was approximately 50% more than the supply of securities traded on centralized exchanges. In addition to being large, a number of OTC markets play key roles within the architecture of our current financial system. For example, the fed funds and repo markets are crucial cogs in the process of monetary policy transmission in the United States. Similarly, the secondary market for U.S. Treasuries, which also operates OTC, is central for the provision of liquidity and the allocation of capital across financial institutions. Finally, the U.S. corporate and municipal bond markets are a critical source of funding for large firms, cities, and states.

Given these crucial roles, it is not surprising that OTC markets have been at the center of recent financial crises. For example, the markets for various types of derivatives, such as mortgage-backed securities and credit default swaps, were ground zero for the Global Financial Crisis (GFC) that unfolded in 2007–2008. More recently, the financial system teetered on the edge of a crisis as severe dislocations in the markets for Treasuries and corporate bonds appeared at the onset of the COVID-19 pandemic in March 2020.

1.2 OTC Markets Are Different from Centralized Exchanges

There are many differences between OTC markets and centralized exchanges, and many dimensions of heterogeneity *within* each category. Therefore, it is helpful to highlight a few key differences between the two, often competing, types of market arrangements.

As Dodd [2012] highlights, many exchanges trace their origins to a physical location, such as the New York Stock Exchange or the Chicago Board of Trade, where transactions were executed according to a particular set of rules and traditions. Though these exchanges have largely migrated to virtual platforms, trades are still organized centrally, through a limit-order book that allows for both communication and trade

execution to occur on a multilateral basis. As a result, most exchanges feature pretrade price transparency: prices at which one can sell (the *bid* price) or buy (the *ask* or *offer* price) are continuously updated and communicated to all participants. These posted prices are *firm*, rather than negotiable. Finally, most exchanges feature posttrade price transparency: the prices and quantities of completed transactions are also continuously communicated to market participants.

OTC markets, in contrast, are decentralized. They are typically not located in a particular location, but rather represent what Dodd [2012] describes as a network of trading relationships, often (but not always) organized through a set of intermediaries or *broker-dealers*. Without a centralized limit-order book, most OTC markets do not feature pretrade price transparency. In particular, in the absence of publicly available bid and ask prices, investors have to *search* for quotes to buy or sell a particular quantity of an asset. Traditionally, search occurred on a bilateral basis, in the sense that investors contacted intermediaries one at a time. Though some OTC markets have introduced electronic platforms that allow investors to search on a multilateral basis—running auctions with a small number of dealers—trading in OTC markets remains fragmented. As a result, the terms of trade continue to vary across transactions and often depend on the identities of the counterparties. OTC markets traditionally did not feature posttrade price transparency either, though recent regulations—such as the Trade Reporting and Compliance Engine (TRACE) in the United States and the Markets in Financial Instruments Directive (MiFID) regulations in Europe—have been enacted to collect and share the terms of trade in completed transactions, albeit with some delay.

1.3 A Paradigm to Study OTC Markets

The unique structure, size, and central role of OTC markets in the financial system have prompted policymakers to confront a number of important issues. For example, in the early 2000s, concerns over a lack of transparency and competition in OTC markets prompted new regulations in the markets for corporate and municipal bonds, whereby certain market participants (mainly dealers) were required to report the price and quantity of their transactions soon after they occurred [Bessembinder and Maxwell, 2008]. With the collapse of several OTC markets during the GFC of 2007–2008, policymakers were faced with a new set of questions: How should they respond when trade collapses in an important OTC market? Should the government engage in ex post interventions, such as large-scale asset purchases? What regulations should be put in place ex ante to prevent such market failures in the future? Should OTC markets be forced to migrate to a centralized exchange?

More recently, economists have debated whether post-GFC regulations put in place to promote financial stability—such as stricter capital and liquidity requirements—have had the unintended consequence of reducing liquidity in certain OTC markets. For example, as Duffie [2012a] argues, intermediaries in OTC markets (i.e., dealers) have to use their balance sheets to provide liquidity to investors: they buy assets, hold them for a period of time as inventory, and resell them later, hopefully at a higher price. Therefore, policies that make it more expensive to hold assets reduce the dealers' incentives to provide liquidity, thereby increasing transaction costs and reducing market depth. These concerns materialized dramatically at the onset of the COVID-19 crisis, when various fixed income markets were subject to unprecedented selling pressures that dealers were unable or unwilling to absorb [Kargar, Lester, Lindsay, Liu, Weill, and Zúñiga, 2021]. In the wake of this crisis, regulators have been tasked with understanding the trade-off between promoting financial stability and maintaining liquidity in OTC markets.

Traditional models of financial markets are not appropriate to address these policy issues and, more generally, to analyze the functioning of OTC markets. Indeed, the vast majority of the asset pricing literature is based on the perfectly competitive Walrasian paradigm, in which investors submit their demand and supply curves to a fictitious auctioneer who sets a market-clearing price for multilateral trade. While this model of equilibrium prices and allocations is quite similar to centralized exchanges in practice, it bears little resemblance to OTC markets.

In particular, the Walrasian paradigm recognizes neither the decentralized nature of trade in OTC markets nor the crucial role that intermediaries play in facilitating trade and providing liquidity. Indeed, the fictitious auctioneer in the background of the Walrasian model leaves little room for understanding a variety of arrangements and institutions that have evolved to overcome the spatial and informational frictions inherent to over-the-counter trading. The Walrasian paradigm also ignores a key dimension of market quality in OTC markets, namely the *time* it takes investors to find a counterparty and trade. Last, but not least, models that ignore the decentralized (and, often, bilateral) nature of interactions among market participants are ill-suited to understand the sources of private information in OTC markets and the implications for prices and allocations.

To address these shortcomings, scholars interested in OTC markets have stepped away from the Walrasian paradigm and built new theoretical market settings to capture the key features of OTC markets. While various approaches have emerged, our objective in this book is to provide a unified treatment of one important strand of this literature that is now sufficiently mature to warrant its own textbook: the search-theoretic approach.

We start, in chapter 2, by discussing a number of empirical regularities that are common across many OTC markets, using data from the corporate bond market as a leading example. This introductory chapter is intended to establish a set of stylized facts and motivate the theoretical frameworks that we develop throughout the book.

Then, in chapter 3, we introduce the dynamic problem of an investor with time-varying preferences who searches for an acceptable price to trade, taking as given that price offers are drawn at random arrival dates from an exogenous distribution. This analysis is a crucial building block of the equilibrium models that follow. It also provides an opportunity to introduce and develop a number of technical tools and results, such as dynamic programming and Hamilton-Jacobi-Bellman (HJB) equations, that are used throughout the text. More generally, we aim for the book to be as self-contained as possible, minimizing the number of results and techniques that the reader needs to find elsewhere.

Understanding the basic decision problem of an individual investor sets the stage for chapters 4 and 5, where we present what we view as the two workhorse search-theoretic models of OTC markets. The first, which we refer to as the model of *semicentralized* OTC markets, considers an environment where investors with heterogeneous preferences and asset holdings periodically get the opportunity to trade with a dealer, who has access to a frictionless interdealer market. This model is meant to capture dealer-intermediated OTC markets, such as the U.S. corporate bond market, where most (if not all) trades are executed through intermediaries, who are able to reallocate assets through a relatively efficient interdealer market. The second model, which we call the *pure decentralized* market, considers an environment without dealers, where investors periodically meet and trade directly with other investors. This model is best applied to OTC markets such as the U.S. Fed Funds market, where financial institutions trade directly with other financial institutions, without any dedicated intermediaries or dealers. The analysis of these two workhorse models, and the various extensions and generalizations described below, derive in large part from Hugonnier, Lester, and Weill [2022], which itself extends the seminal contributions of Duffie, Gârleanu, and Pedersen [2005, 2007].

The remainder of the book builds off of these two workhorse models to create richer frameworks, which can be used to address a variety of important issues and questions. The first extension addresses the fact that both workhorse models described above abstract from the intensive margin of trade—that is, trade size—by assuming that assets are indivisible. In chapter 6, we relax this assumption and characterize the equilibrium in a semicentralized market setting with divisible asset holdings, as in Gârleanu [2009] and Lagos and Rocheteau [2009]. This extension allows us to study, among other things, the relationship between transaction costs and trade size.

Another important simplifying assumption in the model of semicentralized markets is that dealers trade among themselves in a frictionless, competitive market. While this is a good approximation of some dealer-intermediated OTC markets, it abstracts from the process of reallocation within the dealer sector, which is known to be frictional in some important OTC markets. For example, in their study of the U.S. municipal bond market, Li and Schürhoff [2019] document that trades between customers often involve long chains of trades within the dealer sector. Moreover, they show that the length of these *intermediation chains*, and the characteristics of the dealers who participate in them, have systematic relationships with the size of the transaction costs that investors pay. To study the importance of frictions *within* the dealer sector, chapter 7 develops a model in which investors trade with dealers, who themselves trade in a frictional interdealer model, as in Hugonnier, Lester, and Weill [2020].

Next, in chapter 8, we take a deep dive into the price-setting mechanisms that are used to determine the terms of trade between investors and dealers. We first explore the microfoundations of the bargaining game that is employed in the workhorse models described above, and then explore two alternative price-setting mechanisms. First, we consider a game based on the model of Burdett and Judd [1983], which is meant to capture the multilateral "request for quote" (RFQ) system that prevails on many of the increasingly popular electronic OTC trading platforms. Second, we analyze a framework where investors *direct* their search according to the prices posted by dealers, as in Moen [1997] and Burdett, Shi, and Wright [2001]. The latter extension is important for understanding, e.g., the trade-off that has been documented between trading cost and the speed of execution.

Given the central role of OTC markets in recent financial crises, it is also important to study how OTC markets perform in times of stress, which requires going beyond the analysis of steady-state equilibria. We tackle this problem in chapter 9, where we study the nonstationary dynamics of our semicentralized benchmark in response to a shock that generates a sudden surge in selling pressure. We characterize the optimal inventory accumulation of dealers as they attempt to "lean against the wind," and derive the subsequent implications for asset prices, trading volume, and overall market resilience.

Up until this point in the book, we maintained the assumption that trade occurs under perfect information. In particular, this means that both parties to a transaction observe their counterparty's *private value* for the asset, and that they are symmetrically informed about the quality—the *common value* component—of the asset. However, OTC markets are often considered to be "opaque": trade is decentralized with a lack of public signals or benchmarks that aggregate information. It is thus important to study the effects of asymmetric information over both private and common values. We

tackle these two cases in chapters 10 and 11, respectively. We study how information and search frictions can combine to reduce trading volume and even cause markets to freeze; how the deleterious effects of private information can be exacerbated by search frictions; and how policymakers can intervene to help correct the distortions that arise from asymmetric information.

Finally, in chapter 12, we develop a theoretical model to make progress on several fundamental questions regarding the microstructure of financial markets. In particular, economists have long struggled to understand *why* some asset markets are organized through centralized exchanges while others trade OTC, and whether one type of structure generates higher welfare than the other [see, e.g., Rust and Hall, 2003]. Though there are many dimensions to this question, and many possible theoretical frameworks to study it, we contribute to the discussion by developing a model in which investors *choose* whether to trade through a centralized exchange or through an OTC market. We identify conditions under which the OTC market is too large or too small, and analyze how the optimal market structure depends on the primitives of the economy.

1.4 Some Final Notes

Before jumping in, we think it is worthwhile to discuss the way we have chosen to write the book and its relationship to other books in the literature. To start, we should emphasize that there are several different approaches to studying OTC markets, each with its own advantages and disadvantages. For example, a growing literature using network theory has evolved alongside—and, in some rare instances, made contact with—the literature using search theory to study OTC markets [see, e.g., Malamud and Rostek, 2017; Atkeson, Eisfeldt, and Weill, 2015]. While one would obviously like to cover all approaches to studying OTC markets, we have chosen to be narrow yet deep, as opposed to broad and, by necessity, more shallow. For the same reasons, we do not cover the small, but important literature that studies OTC markets in a broader macroeconomic context, by analyzing the *interaction* between prices and allocations in OTC markets and the decisions and/or outcomes in other areas of the economy.

These choices allow us to develop and extend a single theoretical framework throughout the book. We believe this approach has several advantages. First, from a pedagogical point of view, we are able to provide a unified discussion of existing work with a consistent set of assumptions and notation. Second, by incorporating a number of earlier studies into the context of a unified framework, we uncover several deep theoretical connections between seemingly different modeling approaches and assumptions within the literature. Finally, by showing how to modify our workhorse models with various extensions, our analysis equips the reader with a versatile framework

that can be further extended and manipulated to conduct empirical, quantitative, and theoretical research.

Next, it is worth pointing out that we have striven to make this book self-contained by including nearly all the technical results *and the corresponding proofs* that one needs to understand our theoretical analysis. This makes the book well suited for graduate students and other researchers who are interested in the details of solving this class of models, along with other related search models that are used in labor economics, monetary economics, and industrial organization. However, the book is also written for those with more applied interests who desire an "off-the-shelf" model of OTC markets; wherever possible, we present or propose exercises on simplified special cases of our general framework, and also indicate the theoretical deep dives that can be skipped by the reader with more applied interests.

Finally, let us end this introduction with a discussion of related textbooks. Major references in the field of market microstructure include O'Hara [1995], Amihud, Mendelson, and Pedersen [2005], Hasbrouck [2007], and Foucault, Pagano, and Roell [2013]. We add to these important contributions by focusing on the recent literature addressing market liquidity from the point of view of search theory. The only earlier book on the search-theoretic approach to OTC markets is the Princeton Lectures of Duffie [2012b]. The present book has a different objective: we aim to provide a more systematic overview of the literature, including new methodological developments, applications, and empirical evidence. As such, it is a self-contained reference for a semester class on the topic. In addition, our work is related to several other important books on search theory, including Pissarides [2000], Shimer [2010], and Petrosky-Nadeau and Wasmer [2017]—who use search theory to study labor economics—and Nosal and Rocheteau [2011], who present the search-theoretic approach to monetary economics. Our treatment should be viewed as complementary to these books, with a new focus on financial market applications and a significant number of new technical results to support our exploration.

INDEX

References in *Italics* refer to figures.