



University of Pennsylvania  
**ScholarlyCommons**

---

Publicly Accessible Penn Dissertations

---

2020

## Essays On Decentralized Financial Markets

Ruslan Sverchkov  
*University of Pennsylvania*

Follow this and additional works at: <https://repository.upenn.edu/edissertations>



Part of the [Finance and Financial Management Commons](#)

---

### Recommended Citation

Sverchkov, Ruslan, "Essays On Decentralized Financial Markets" (2020). *Publicly Accessible Penn Dissertations*. 3805.

<https://repository.upenn.edu/edissertations/3805>

This paper is posted at ScholarlyCommons. <https://repository.upenn.edu/edissertations/3805>  
For more information, please contact [repository@pobox.upenn.edu](mailto:repository@pobox.upenn.edu).

---

## Essays On Decentralized Financial Markets

### Abstract

In the first chapter, *To Pool or Not to Pool? Security Design in OTC Markets* with Vincent Glode and Christian C. Opp, we study security issuers' decision whether to pool assets when facing counterparties endowed with market power, as is common in over-the-counter markets. Unlike in competitive markets, pooling assets may be suboptimal in the presence of market power --- both privately and socially --- in particular, when the potential gains from trade are large. In these cases, pooling assets reduces the elasticity of trade volume in the relevant part of the payoff distribution, exacerbating inefficient rationing associated with the exercise of market power. Our results shed light on recently observed time-variation in the prevalence of pooling in financial markets.

In the second chapter, *Selling to Investor Network: Allocations in the Primary Corporate Bond Market*, I develop a model of the primary market for corporate bonds, in which an issuer optimally chooses an issuance price and allocations to investors based on their trading connections in the secondary over-the-counter market. Expected secondary market liquidity, which depends on the structure of the trading network in this market, determines investors' demands in the primary market and, in turn, the issuer's revenues. I show that trading by less connected investors has a relatively high negative impact on expected secondary market liquidity and disproportionately reduces the demands of all investors in the primary market. As a result, the issuer can increase her profits by restricting allocations of new bonds only to more connected investors. This explains the commonly observed exclusion of small institutional investors from the primary market, which is often coupled with seemingly underpriced bonds.

In the third chapter, *Initial Coin Offerings as a Commitment to Competition* with Itay Goldstein and Deeksha Gupta, we model Initial Coin Offerings (ICOs) of utility tokens, which are increasingly used to finance the development of online platforms where buyers and sellers can meet to exchange services or goods. Utility tokens serve as the sole medium of exchange on a platform and can be traded in a secondary market. We show that such a financing mechanism allows an entrepreneur to give up monopolistic rents associated with the control of the platform and make a credible commitment to long-run competitive prices. The entrepreneur optimally chooses to have an ICO, rather than operate as a monopolist, only if future consumers of the platform participate in financing. ICOs, therefore, endogenously require crowd-funding to be viable.

### Degree Type

Dissertation

### Degree Name

Doctor of Philosophy (PhD)

### Graduate Group

Finance

### First Advisor

Vincent Glode

### Second Advisor

Itay Goldstein

---

**Keywords**

Corporate Finance, Decentralized Markets, Financial Economics Theory, FinTech, Primary Markets

**Subject Categories**

Finance and Financial Management

ESSAYS ON DECENTRALIZED FINANCIAL MARKETS

Ruslan Sverchkov

A DISSERTATION

in

Finance

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the  
Degree of Doctor of Philosophy

2020

Supervisor of Dissertation

---

Itay Goldstein  
Joel S. Ehrenkranz Family Professor  
Professor of Finance  
Professor of Economics

Co-Supervisor of Dissertation

---

Vincent Glode  
Associate Professor of Finance

Graduate Group Chairperson

---

Nancy Zhang, Professor of Statistics

Dissertation Committee:

Chaojun Wang, Assistant Professor of Finance

ESSAYS ON DECENTRALIZED FINANCIAL MARKETS

© COPYRIGHT

2020

Ruslan Sverchkov

This work is licensed under the  
Creative Commons Attribution  
NonCommercial-ShareAlike 3.0  
License

To view a copy of this license, visit

<http://creativecommons.org/licenses/by-nc-sa/3.0/>

*For my parents*

## ACKNOWLEDGMENT

I have benefited a great deal from the help, suggestions, and encouragement of many people. I would like to thank my dissertation committee for their invaluable input: I am deeply indebted to Vincent Glode for the continuous support and attentive guidance throughout my Ph.D. studies; I am immensely grateful to Itay Goldstein for the thoughtful advice that enormously benefited my development as a researcher; and I wish to thank Chaojun Wang who provided the detailed feedback that helped significantly improve this manuscript.

I would also like to give special thanks to Christian C. Opp and Deeksha Gupta for numerous fruitful discussions that contributed towards the growth of the key ideas in the dissertation.

In addition, I wish to acknowledge my classmates and colleagues that I met at the Wharton School for their joyful company that made the Ph.D. journey more pleasant and productive.

Finally, I would like to thank my partner for helping me weather some of the most stressful times during the writing of the final version of this dissertation. I thank my parents for their patience and unwavering support throughout my life. This would not be possible without you.

# ABSTRACT

## ESSAYS ON DECENTRALIZED FINANCIAL MARKETS

Ruslan Sverchkov

Vincent Glode, Itay Goldstein

In the first chapter, *To Pool or Not to Pool? Security Design in OTC Markets* with Vincent Glode and Christian C. Opp, we study security issuers' decision whether to pool assets when facing counterparties endowed with market power, as is common in over-the-counter markets. Unlike in competitive markets, pooling assets may be suboptimal in the presence of market power — both privately and socially — in particular, when the potential gains from trade are large. In these cases, pooling assets reduces the elasticity of trade volume in the relevant part of the payoff distribution, exacerbating inefficient rationing associated with the exercise of market power. Our results shed light on recently observed time-variation in the prevalence of pooling in financial markets.

In the second chapter, *Selling to Investor Network: Allocations in the Primary Corporate Bond Market*, I develop a model of the primary market for corporate bonds, in which an issuer optimally chooses an issuance price and allocations to investors based on their trading connections in the secondary over-the-counter market. Expected secondary market liquidity, which depends on the structure of the trading network in this market, determines investors' demands in the primary market and, in turn, the issuer's revenues. I show that trading by less connected investors has a relatively high negative impact on expected secondary market liquidity and disproportionately reduces the demands of all investors in the primary market. As a result, the issuer can increase her profits by restricting allocations of new bonds only to more connected investors. This explains the commonly observed exclusion of small institutional investors from the primary market, which is often coupled with seemingly underpriced bonds.



In the third chapter, *Initial Coin Offerings as a Commitment to Competition* with Itay Goldstein and Deeksha Gupta, we model Initial Coin Offerings (ICOs) of utility tokens, which are increasingly used to finance the development of online platforms where buyers and sellers can meet to exchange services or goods. Utility tokens serve as the sole medium of exchange on a platform and can be traded in a secondary market. We show that such a financing mechanism allows an entrepreneur to give up monopolistic rents associated with the control of the platform and make a credible commitment to long-run competitive prices. The entrepreneur optimally chooses to have an ICO, rather than operate as a monopolist, only if future consumers of the platform participate in financing. ICOs, therefore, endogenously require crowd-funding to be viable.

## TABLE OF CONTENTS

ACKNOWLEDGMENT . . . . .	iv
ABSTRACT . . . . .	v
LIST OF TABLES . . . . .	ix
LIST OF ILLUSTRATIONS . . . . .	x
CHAPTER 1 : To Pool or Not to Pool? Security Design in OTC Markets . . . . .	1
1.1 Introduction . . . . .	1
1.2 The Environment . . . . .	5
1.3 Main Analysis . . . . .	8
1.4 Robustness . . . . .	24
1.5 Conclusion . . . . .	26
CHAPTER 2 : Selling to Investor Network: Allocations in the Primary Corporate Bond Market . . . . .	27
2.1 Introduction . . . . .	27
2.2 Model . . . . .	34
2.3 Equilibrium Analysis . . . . .	44
2.4 Correlated Liquidity Shocks . . . . .	63
2.5 Discussion . . . . .	67
2.6 Conclusion . . . . .	69
CHAPTER 3 : Initial Coin Offerings as a Commitment to Competition . . . . .	71
3.1 Introduction . . . . .	71
3.2 Model Setup . . . . .	77
3.3 Example with $T = 2$ and $N = 2$ . . . . .	81

3.4	Equilibrium Analysis . . . . .	90
3.5	Endogenous Crowd-Funding . . . . .	96
3.6	Tokenized Platform vs Platform Competition under Network Effects . . . . .	104
3.7	Conclusion . . . . .	108
	APPENDIX . . . . .	110
	BIBLIOGRAPHY . . . . .	139

## LIST OF TABLES

TABLE 2.1 : Example of a network with optimal exclusion . . . . .	45
TABLE 2.2 : Example of comparative statics for the exclusion result . . . . .	62

## LIST OF ILLUSTRATIONS

FIGURE 1.1 : Effect of pooling on the shape of the probability density function	17
FIGURE 1.2 : Effect of pooling on the shape of the elasticity function . . . . .	18
FIGURE 2.1 : Model timeline . . . . .	35
FIGURE 2.2 : Segment of a trading network . . . . .	37
FIGURE 2.3 : Patient and impatient investors within a trading network . . . . .	38
FIGURE 2.4 : Trading by impatient neighbors with the common dealer . . . . .	40
FIGURE 2.5 : Trading of impatient investor and neighbors . . . . .	50
FIGURE 2.6 : Optimal trading of an impatient investor . . . . .	51
FIGURE 2.7 : Total demands with correlated liquidity shocks . . . . .	66
FIGURE 3.1 : Sequence of events during a period . . . . .	78
FIGURE 3.2 : Equilibrium prices of the service . . . . .	86
FIGURE 3.3 : Cumulative welfare over time . . . . .	89
FIGURE 3.4 : Welfare as a function of high-type consumers . . . . .	90
FIGURE 3.5 : Equilibrium prices of a service with financing . . . . .	100
FIGURE A.1 : Effect of pooling on the shape of the probability density function	113
FIGURE A.2 : Effect of pooling on the shape of the elasticity function . . . . .	113
FIGURE A.3 : Effect of pooling on the shape of the probability density function	114
FIGURE A.4 : Effect of pooling on the shape of the elasticity function . . . . .	114

## CHAPTER 1 : To Pool or Not to Pool? Security Design in OTC Markets

### 1.1. Introduction

Structured products are typically originated in over-the-counter (OTC) markets, where asymmetric information and market power have been shown to be prevalent frictions.<sup>1</sup> In these markets, issuers may face prices that are not fully competitive, especially when only few financial institutions are well-positioned to acquire new securities. For example, as most institutions are subject to similar regulatory constraints, holding costs can increase simultaneously for many market participants, leaving only few institutions well positioned to provide liquidity.

Motivated by these observations, we study the security design problem of a privately informed issuer who possesses multiple assets and faces liquidity suppliers, or buyers, that are potentially endowed with market power. Our analysis reveals how the allocation of market power has relevant and robust implications for security design that contrast with the takeaways from models considering only competitive environments. To isolate the effect of market power, we consider both competitive and non-competitive markets.

When buyers act competitively, our results echo the findings of the existing literature (e.g., DeMarzo, 2005) — pooling all assets into one security is optimal for the issuer. As diversification reduces an issuer’s informational advantage, pooling assets helps alleviate adverse selection problems, which is in the interest of the issuer when prices are set competitively, since in this case, the issuer fully internalizes the benefits of improving the efficiency of trade.

In contrast, when an issuer receives non-competitive offers for his securities, pooling assets still has the advantage of reducing adverse selection concerns, but it now also comes at a

---

<sup>1</sup>For evidence that OTC trading often involves heterogeneously informed traders, see Green et al. (2007), Jiang and Sun (2015), and Hollifield et al. (2017). For evidence that OTC trading tends to be concentrated among a small set of players, see Cetorelli et al. (2007), Atkeson et al. (2014), Begenu et al. (2015), Di Maggio et al. (2017a), Hendershott et al. (2020), Li and Schürhoff (2019), and Siriwardane (2019).

cost, namely, a potential reduction in the issuer's information rents. Counter to conventional wisdom, a privately informed issuer may prefer not to pool assets in this case, especially when the potential gains from trade are large relative to the information asymmetry between the issuer and prospective buyers. In fact, any pooling decision that achieves perfect diversification is *never* optimal for an issuer facing market power on the demand side. We provide explicit, sufficient conditions under which the issuer's best option is to simply sell all assets separately. Under these conditions, separate sales are not only privately optimal but also achieve the first-best level of total trade surplus. In contrast, when assets are pooled, both the issuer's private surplus and the total surplus from trade are strictly lower, as diversification invites strategic buyers with market power to choose pricing strategies that lead to inefficient rationing. As pooling affects the *shape* of the distributions characterizing information asymmetries between issuers and buyers, it alters how elastically trade volume responds to prices, which is crucial in settings with market power. In particular, pooling would typically worsen inefficient rationing when selling assets separately leads to little or no exclusion of buyer types. Diversification causes payoff distributions to have thinner tails, which, in turn, leads to less elastic trade volume in the right tail of the distribution and greater rationing in equilibrium.

Our results highlight how, in recent years, liquidity shortages among major institutions actively trading in OTC markets might have been an important driver of the dramatic declines in asset-backed security (ABS) issuances, which occurred concurrently with an increase in the volume of assets sold separately.<sup>2</sup> Our analysis shows that, when liquidity becomes scarce and concentrated among few market participants, the benefits of pooling assets highlighted in the literature can be outweighed by an associated increase in the severity of market power problems. In periods of scarce liquidity, the benefits from unloading the assets are typically large for the issuer, but the few traders with excess liquidity gain

---

<sup>2</sup>In 2015, issuance volume of ABS in the U.S. was 60% lower than it was in 2006, while the issuance volume of CDO was 80% lower. In contrast, the total issuance volume in fixed income markets was 3% higher in 2015 than in 2006. For more data, see the Securities Industry and Financial Markets Association: <http://www.sifma.org/research/statistics.aspx>.

market power. These two conditions, when combined, increase the relative benefits of the separate sale of assets versus the issuance of pooled securities. In that sense, it is during time periods when trade is most valuable but potentially impeded by the presence of market power that our new insights become most relevant. Relatedly, our paper sheds light on the consequences of regulating the liquidity of financial institutions that are often on the buy side of the structured securities market.

Early contributions by Subrahmanyam (1991), Boot and Thakor (1993), and Gorton and Pennacchi have emphasized the diversification benefits of pooling assets when securities are sold in competitive/centralized markets that are subject to asymmetric information problems. Our paper focuses on the impact of market power on the decision to pool assets and derives novel insights that shed light on the securities issued in decentralized markets. The two papers closest to ours are DeMarzo (2005) and Biais and Mariotti (2005). Specifically, our focus on the decision to pool assets relates our analysis to DeMarzo (2005) who builds on the signaling-through-retention framework with price-taking buyers of DeMarzo and Duffie (1999) and shows that the pooling of assets dampens an issuer’s ability to signal individual assets’ quality through retention. However, when the number of assets is large and the issuer can sell debt on the pool of assets, this “information destruction effect” is dominated by the above-mentioned benefits of diversifying the risks associated with the issuer’s private information about each asset’s value. Issuing debt on a large pool of assets reduces residual risks and the information sensitivity of the security being issued.<sup>3</sup> In contrast to DeMarzo (2005) whose setup can be thought of as a centralized market where (price-taking) buyers compete for assets, we consider the case of an issuer who faces buyers endowed with market power, capturing a realistic feature of many over-the-counter markets.

Our focus on the role of market power in an issuer’s security design decision relates our analysis to Biais and Mariotti (2005) who analyze a model where the security design stage

---

<sup>3</sup>See also Hartman-Glaser et al. (2012) who model a moral hazard problem between a principal and a mortgage issuer and show that the optimal contract features pooling of mortgages with independent defaults, as it facilitates effort monitoring.



is followed by a stage where either the issuer or the prospective buyer chooses a trading mechanism (i.e., a price-quantity menu) for selling the designed security. When the buyer can choose the trading mechanism, he effectively screens the issuer, trading off higher volume with lower issuer participation. In contrast, when the issuer can choose the mechanism, the setup becomes equivalent to one with multiple competitive buyers. Biais and Mariotti (2005) show that issuing debt on a risky asset is optimal in both cases, since the debt contract’s low information sensitivity helps avoid market exclusion.<sup>4</sup> However, unlike our paper, Biais and Mariotti (2005) only consider the case of an issuer wishing to sell one asset.

Axelson (2007) studies an uninformed issuer’s decision to design securities that are (centrally) traded in a uniform-price auction with privately informed buyers. Axelson (2007) finds that pooling assets and issuing debt on these assets is always optimal when the number of assets is large, otherwise selling assets separately might be optimal if the signal distribution is discrete and competition is high enough. Since the issuer is uninformed and buyers compete for assets through an auction, Axelson’s (2007) analysis is silent about how security design can be used to prevent being monopolistically screened by liquidity providers, which is a key result of our analysis.<sup>5</sup>

Palfrey (1983) analyzes a firm’s decision to bundle products (or assets) sold in a second-price auction. In his model, customers have private information about their heterogeneous valuations for the products. Selling the products separately is optimal when the sum of the expected second-highest valuation for each product is higher than the expected second-highest valuation for the bundle of all products. This comparison depends on the number of prospective customers and the distribution of their product-specific valuations. Unlike Palfrey (1983), our analysis examines how the degree of competition among buyers with

---

<sup>4</sup>Gorton and Pennacchi (1990), Dang et al. (2015), Farhi and Tirole (2015), and Yang (2020) also study the optimal information sensitivity of securities issued in markets with asymmetric information. These papers highlight the benefits of designing securities that split cash-flows into an information-sensitive part and a risk-less part. These papers are, however, silent about how pooling imperfectly correlated assets affects the issuer’s ability to extract surplus when facing buyers with market power, which is the focus of our paper.

<sup>5</sup>See also DeMarzo et al. (2005) and Inderst and Mueller (2006) who study optimal security design problems with informed buyers and only one asset.

identical valuations affect pooling decisions. The cross-buyer heterogeneity in valuations that is central for Palfrey's (1983) results does not play a role for our findings.

In the next section, we describe our model and provide an illustrative example in which the issuer sells a pool of a continuum of assets. This example highlights that the presence of market power on the demand side greatly affects the issuer's benefits from pooling assets. Section 1.3 presents our main analysis of both a competitive market and one with market power. Section 1.4 discusses the robustness of our results to various alternative specifications of the environment. The last section concludes.

## 1.2. The Environment

Suppose an issuer has  $n \geq 2$  fundamental assets to sell. These assets are indexed by  $i$  and the set of all assets is denoted by  $\Omega \equiv \{1, \dots, n\}$ . Each asset  $i$  produces a random payoff  $X_i$  at the end of the period. The assets' payoffs  $X_i$  are assumed to be identically and independently distributed according to the cumulative distribution function (CDF)  $G(\cdot)$  with a probability density function (PDF)  $g(\cdot)$  that is positive everywhere on its domain  $\chi \equiv [0, \bar{x}]$ .

*Market participants and their liquidity needs.* As is common in the security design literature, agents are risk neutral but can differ in their liquidity (or hedging) needs, which are captured by their discount factors. In the analysis that follows, we will study and compare two (polar) market scenarios to highlight the importance of market power in the decision whether to pool assets.

In the first scenario, we assume that *several* deep-pocketed traders are better equipped to hold claims to future cash-flows than the issuer is (who needs liquidity today). Whereas the issuer applies a discount factor  $\delta \in (0, 1)$  to future cash-flows, these prospective buyers apply a discount factor of 1. Thus, the ex ante private value of each fundamental asset is  $\delta \mathbb{E}(X_i)$  for the issuer and  $\mathbb{E}(X_i)$  for any of these buyers. As a result, there are gains from transferring the issuer's assets to such a buyer in exchange for cash. Since there are

multiple buyers who value assets more than the issuer in this scenario, these buyers make competitive bids for the securities offered by the issuer.

In the second scenario, we assume that only one buyer is better equipped to hold claims to future cash-flows than the issuer is; that is, only one buyer has a discount factor of one. In this case, the one buyer with a superior liquidity position has market power; he is the only one bidding for the issuer's securities.<sup>6</sup> This scenario captures the idea that in some time periods, most potential counterparties in the market face similar regulatory constraints or liquidity needs as the issuer, potentially leading to concentration on the demand side. For both scenarios, we will occasionally refer to the prospective buyers with a discount factor of 1 as "liquidity suppliers" (in line with the literature; see, e.g., Biais and Mariotti, 2005).

*Timing and information structure.* Our specification of the timeline follows the existing literature (see, e.g., DeMarzo and Duffie, 1999; Biais and Mariotti, 2005). First, the issuer designs the securities he plans to sell. Second, the issuer becomes informed about the realizations of each asset payoff  $X_i$ . Third, the buyer(s) make(s) take-it-or-leave-it offers to the issuer. Fourth, the issuer decides whether or not to accept any of these offer(s) in exchange for the securities; if multiple buyers offer an identical price that is accepted by the issuer, the security is randomly allocated among the highest bidders. Finally, all payoffs are realized.

Assuming that the issuer does not have private information at the initial security design stage increases the tractability of the analysis and shares similarities with the shelf registration process commonly used in practice (as also argued by DeMarzo and Duffie, 1999; Biais and Mariotti, 2005). In that process, issuers first specify and register with the Securities and Exchange Commission the securities they intend to issue. Then, potentially after several months, issuers bring these securities to the market. In the meantime, the issuer has typically obtained additional private information about future cash-flows. In Section 1.4, we

---

<sup>6</sup>Going forward, we will refer to this scenario as monopolistic demand or monopolistic liquidity supply. In this context, the buyer can also be referred to as a monopsonist.

discuss the robustness of our main insights to changes in this timeline that would introduce signaling concerns at the security design stage.

*An illustrative example.* Before proceeding with our main analysis, we present a simple, yet generic example that illustrates how the issuer's benefits from pooling assets crucially depend on the allocation of market power. Suppose the issuer owns a continuum of assets of measure one with i.i.d. payoffs  $X_i$  with finite mean and variance. The issuer considers selling the pool of these assets to the prospective buyer(s).

First, we analyze the market scenario in which multiple prospective buyers have abundant liquidity (that is, they have a discount factor equal to one). In this case, they effectively compete in quotes à la Bertrand and offer a price that is equal to the expected security payoff conditional on the issuer accepting the offer. When the issuer offers the assets as one pool, the law of large numbers applies, that is, perfect diversification implies that the pool's payoff is  $\int_0^1 x_i di = \mathbb{E}[X_i]$  almost surely. As a result, adverse selection concerns are completely eliminated, and the competitive buyers offer a price  $\hat{p} = \mathbb{E}[X_i]$  for this pool. The maximum total surplus from trade,  $\mathbb{E}[X_i] \cdot (1 - \delta)$ , is attained and the issuer fully internalizes this surplus. That is, the issuer achieves the optimal expected payoff. The fact that pooling the continuum of assets eliminates information asymmetries is unambiguously beneficial when facing competitive buyers, as the issuer then fully internalizes the resultant improvements in trade efficiency (see also Theorem 5 in DeMarzo, 2005).

In contrast, consider the market scenario in which only one prospective buyer has liquidity to purchase the issuer's assets (i.e., only one buyer has a discount factor of one). Acting as a de-facto monopolist, this buyer can choose the price that maximizes his expected payoff. In this case, this optimally chosen price is the issuer's reservation price for the pool of assets, that is,  $p^* = \mathbb{E}[X_i]\delta$ . As in the scenario with multiple prospective buyers, pooling the continuum of assets yields perfect diversification and eliminates adverse selection concerns. Yet, now that the demand side has market power, fully eliminating these information asymmetries has

*no upside* for the issuer. Facing no informational disadvantage, the monopolistic liquidity supplier then charges a price that leaves the issuer indifferent between trading the security and not trading at all.

This generic result for asset pools that achieve perfect diversification strikingly highlights the relevance of market power for the optimality of pooling assets from the perspective of the issuer. In the presence of such market power, the issuer's only source of surplus are information rents, which require retaining some private information. Thus, any pooling that leads to *perfect* diversification (as was the case in this example) is never optimal for an issuer when facing a prospective buyer with market power. Instead, the issuer prefers to retain some private information, which requires deviating from the pooling of all assets. Being at an informational disadvantage, buyers with market power then strategically choose prices that can jeopardize the realization of gains from trade. When deciding whether to pool assets, the issuer therefore faces an intuitive trade-off: he can only extract rents when retaining some private information, but he still partially internalizes the inefficiencies emerging from adverse selection and the exercise of market power under asymmetric information. As a result, he may only choose to pool a subset of assets in order to achieve partial diversification (but not perfect diversification). Understanding these channels and how they affect the design of optimal securities is the focus of our main analysis below.

### 1.3. Main Analysis

We now formalize our paper's main insights. The issuer decides on the pooling of the  $n$  underlying assets and on the securities that are written on each of the pools. Formally, the issuer chooses a partition of the set  $\Omega$ , that is, he groups the  $n$  assets into  $m \leq n$  disjoint subsets denoted by  $\Omega_j$  with  $j \in \{1, \dots, m\}$ . The corresponding  $m$  pools of assets then have the payoffs:

$$Y_j \equiv \sum_{i \in \Omega_j} X_i, \forall j. \tag{1.1}$$

The CDF  $G_j$  of  $Y_j$  and the associated density  $g_j$  are then defined on the compact interval  $\chi_j \equiv [0, \bar{y}_j]$ , where  $\bar{y}_j \equiv \sum_{i \in \Omega_j} \bar{x}$ . Going forward, we follow the convention of using capitalized letters for random variables and lower-case letters for their realizations. In line with the existing literature (e.g., M:8), we assume that these distributions satisfy a regularity condition that ensures that first-order conditions in the trading game with a monopolistic buyer are sufficient conditions for the optimal pricing decisions.

**Assumption 1.1.** *For any partition of  $\Omega$ , the elasticity functions:*

$$e_j(y) \equiv \frac{g_j(y)}{G_j(y)} \cdot y, \quad \forall j \tag{1.2}$$

*are weakly decreasing on their respective support  $\chi_j$ .*

Throughout our main analysis below, we will discuss examples with distributions satisfying Assumption 1.1 (see also the Appendix 1B for additional illustrations). When interpreting elasticity functions, it is helpful to note that they represent the ratio of the local density  $g_j(y_j)$  to the average density  $G_j(y_j)/y_j$ . These quantities will play an important role in determining a monopolistic buyer's optimal pricing strategy. We also denote by  $e(x_i) \equiv \frac{g(x_i)}{G(x_i)} \cdot x_i$  the elasticity function of each fundamental asset  $i$ .

The issuer chooses for each pooled payoff  $Y_j$  a security that is backed by that payoff. Specifically, the security payoff  $F_j$  is contingent on the realized cash-flow  $Y_j$  according to the function  $\varphi_j : \chi_j \rightarrow \mathbb{R}_+$  such that  $F_j = \varphi_j(Y_j)$ . We impose the standard limited liability condition:

$$(LL) \quad 0 \leq \varphi_j \leq \text{Id}_{\chi_j},$$

where  $\text{Id}_{\chi_j}$  is the identity function on  $\chi_j$ . In addition, as in Harris and Raviv (1989), Nachman and Noe (1994), and Biais and Mariotti (2005), we restrict the set of admissible securities by requiring that both the payoffs to the liquidity supplier and to the issuer be non-decreasing in the underlying cash-flow:

(M1)  $\varphi_j$  is non-decreasing on  $\chi_j$ .

(M2)  $\text{Id}_{\chi_j} - \varphi_j$  is non-decreasing on  $\chi_j$ .

The sets of admissible payoff functions for the securities is therefore given by  $\{\varphi_j : \chi_j \rightarrow \mathbb{R}_+ \mid (\text{LL}), (\text{M1}), \text{ and } (\text{M2}) \text{ hold}\}$ .

### 1.3.1. Competitive Demand

In this subsection, we analyze the (benchmark) scenario in which the issuer faces multiple liquidity suppliers that have a discount factor of one. In this case, the issuer receives competitive ultimatum price quotes, a feature that is common in the literature (see, e.g., Boot and Thakor, 1993; Nachman and Noe, 1994; FHJ) and delivers results that are consistent with DeMarzo’s (2005) seminal analysis of pooling decisions in a competitive environment.<sup>7</sup>

## Optimality of Pooling Assets

Echoing the existing literature, our analysis of this scenario predicts that issuing debt on the pool of all assets is optimal for the issuer.

**Proposition 1.1.** *If  $\mathbb{E}[X_i] \geq \delta\bar{x}$ , the issuer is indifferent between selling assets separately and selling them as a pool. If  $\mathbb{E}[X_i] < \delta\bar{x}$ , the issuer optimally pools all  $n$  assets and issues a debt security on this pool.*

To provide intuition for this result we will discuss the proof of Proposition 1.1 in the main text. At the trading stage, the issuer has perfect knowledge of the realizations  $x_i$  of future cash-flows  $X_i$ . Since the payoff of any security  $F_j$  is only contingent on  $Y_j = \sum_{\Omega_j} X_i$ , the issuer also perfectly knows the realization  $f_j = \varphi_j(y_j)$  of  $F_j$ . Suppose the issuer uses a simple equity security (what DeMarzo and Duffie (1999) refer to as a “passthrough” security). If  $\mathbb{E}[X_i] \geq \delta\bar{x}$ , he can sell the assets separately (as equity), each at price  $p = \mathbb{E}[X_i]$ , since at

---

<sup>7</sup>DeMarzo (2005) considers a setting in which the issuer can post price-quantity menus. In contrast, we follow Biais and Mariotti’s (2005) representation of the competitive market environment. See Section 1.4 for a discussion of how retention would affect our results.

this price, even the highest issuer type  $\bar{x}$  finds it optimal to trade. The issuer obtains the same total payoff when pooling the assets and selling an equity security on the pool. Since the potential gains from trade are large enough ( $\delta$  is sufficiently low), adverse selection does not impede the efficiency of trade even when assets are sold separately. The first-best level of total trade surplus is achieved, and the issuer fully internalizes this surplus.

In contrast, if  $\mathbb{E}[X_i] < \delta\bar{x}$ , the sale of an equity security on a single asset leads to adverse selection, since the highest issuer type  $\bar{x}$  would not accept a price equal to  $\mathbb{E}[X_i]$ . Similarly, the sale of an equity security on a pool of  $\tilde{n}$  assets leads to the exclusion of some issuer types, since the highest issuer type  $\bar{y}_j = \tilde{n}\bar{x}$  would not accept a price equal to  $\mathbb{E}[Y_j] = \tilde{n}\mathbb{E}[X_i]$ . In this case, it is useful to recall the following result from Biais and Mariotti's (2005) analysis of a setting with one underlying asset:

**Lemma 1.1.** *Given an underlying asset with random payoff  $Y$  and  $\mathbb{E}[Y] < \delta\bar{y}$ , the issuer optimally designs a debt security with the highest face value  $d$  such that a buyer just breaks even when purchasing this debt security at a price  $p = \delta d$ .*

*Proof.* See Proposition 4 in Biais and Mariotti (2005). □

Independent of his pooling choice that determines the underlying assets with payoffs  $Y_j$ , the issuer optimally uses a debt security when  $\mathbb{E}[X_i] < \delta\bar{x}$  and equivalently,  $\mathbb{E}[Y_j] < \delta\bar{y}_j$ . To determine the issuer's optimal pooling decision, it is useful to first consider buyers' expected net profits. A buyer purchasing debt with face value  $d$  at a price  $p = \delta d$  obtains the following expected net profit:

$$\int_0^d yg_j(y)dy + [1 - G_j(d)]d - \delta d = (1 - \delta)d - \left( G_j(d)d - \int_0^d yg_j(y)dy \right) \quad (1.3)$$

$$= (1 - \delta)d - \int_0^d G_j(y)dy, \quad (1.4)$$

where the last step follows from integration by parts. Next, we compare buyers' expected net-payoff from the sales of separate debt securities to that from the sale of a debt security



on an underlying pool of assets. Consider first that the issuer sells  $\tilde{n}$  individual debt securities with face value  $d$ . Further, suppose that each debt security is written on a separate underlying asset and the price in each transaction is  $\delta d$ . Then buyers' total expected net-profit (which may be negative)<sup>8</sup> is:

$$\tilde{n} \cdot \left( (1 - \delta)d - \int_0^d G(x)dx \right) = (1 - \delta)\tilde{n}d - \int_0^{\tilde{n}d} G\left(\frac{y}{\tilde{n}}\right) dy, \quad (1.5)$$

where we used a change in variables, with  $y = \tilde{n}x$ . In contrast, consider now that the issuer pools the  $\tilde{n}$  assets and issues one debt security with face value  $d_j = \tilde{n}d$  and buyers purchase this debt at price  $\delta d_j$ . In this case, buyers' total expected net-profit (which again may be negative) is:

$$(1 - \delta)\tilde{n}d - \int_0^{\tilde{n}d} G_j(y)dy. \quad (1.6)$$

The following lemma sheds light on the relative magnitude of the profits in (1.5) and (1.6).

**Lemma 1.2.** *The distribution of the pooled payoff  $Y_j = \sum_{i=1}^{\tilde{n}} X_i$  second-order stochastically dominates the distribution of the payoff  $\tilde{n}X_i$ , that is,*

$$\int_0^s \left[ G\left(\frac{y}{\tilde{n}}\right) - G_j(y) \right] dy \geq 0 \quad (1.7)$$

for any  $s \in [0, \bar{y}_j]$ .

*Proof.* See Appendix 1A. □

Lemma 1.2 implies that buyers' total expected net-profit is *higher* in the scenario with pooling (i.e., (1.6) is greater than (1.5)). Next, recall that, according to Lemma 1.1, the *optimal* face value in each scenario would be set such that buyers break even, that is, the optimal face values would ensure that (1.5) and (1.6) are each equal to zero. The above

---

<sup>8</sup>At this point in the proof, the considered supposition does not impose that the buyers' participation constraint is satisfied. That is, the expected net-profit can be negative.

result implies that if buyers break even at a face value  $d^*$  on separate sales (first scenario), then they make positive profits on the pooled sale if the face value is set equal to  $\tilde{n}d^*$  (second scenario). It follows that the issuer can choose a face value  $d_j^* \geq \tilde{n}d^*$  on the pool while still ensuring that the buyers can break even (as buyers' expected net-profit is a continuous function of  $d_j$ ). Finally, observe that when issuing debt with break-even face values under each of the two scenarios, the issuer's total profits are  $(1-\delta)\delta\tilde{n}d^*$  and  $(1-\delta)\delta d_j^*$ , respectively, and the issuer extracts the full gains from trade in the competitive market. Since  $d_j^* \geq \tilde{n}d^*$ , the issuer obtains a higher expected net-profit when pooling the  $\tilde{n}$  assets and issuing debt with face value  $d_j^*$ .

In sum, the argument for the optimality of pooling in this setting is intuitive. With competitive liquidity suppliers, the issuer extracts all the gains from trade and, thus, fully internalizes any improvements in trade efficiency. As a result, when adverse selection concerns impede trade, the issuer seeks to minimize the information asymmetry between him and his prospective buyers by pooling assets. As pooling leads to diversification, it reduces the information asymmetry and its associated inefficiencies. In other words, the issuer does not face a trade-off when facing competitive buyers — reducing information asymmetry is always weakly beneficial. We will, however, show below that the unambiguous optimality of pooling ceases to hold when the supply of liquidity becomes imperfectly competitive.

### 1.3.2. *Monopolistic Demand*

In this subsection, we derive our paper's main results by considering the scenario in which the issuer faces an imperfectly competitive demand, a feature that is relevant for our understanding of OTC markets in practice. In this setting, only one buyer has a discount factor of one, which imparts him the advantage of being a monopolistic liquidity supplier.<sup>9</sup>

---

<sup>9</sup>While we consider the case in which only one buyer has a discount factor of one, similar outcomes arise when there are multiple buyers with a discount factor of one, but these buyers face position limits (see Section 1.4 for additional details). The central feature of our analysis is the presence of some degree of market power, that is, a buyer can strategically affect the prices of the securities being offered. Biais et al. (2000) show that this type of strategic pricing behavior also arises when multiple risk averse liquidity suppliers compete in mechanisms (see also Vives, 2011).

We start by examining this buyer's optimal pricing decision. Biais and Mariotti (2005) show that for a given security offered, the optimal mechanism for the liquidity supplier with market power can be implemented via a take-it-or-leave-it offer (see also Riley and Zeckhauser, 1983). Specifically, the prospective buyer makes an ultimatum price offer  $p_j$  to maximize his ex-ante profit from purchasing a security with payoff  $F_j$ :

$$\Pr(\delta f_j \leq p_j)(\mathbb{E}[f_j | \delta f_j \leq p_j] - p_j) = \int_0^{p_j/\delta} (\varphi_j(y) - p_j)g_j(y)dy. \quad (1.8)$$

The optimal price  $p_j^m$  set by this buyer identifies a marginal issuer type that is just willing to accept this price:  $f_j^m = p_j^m/\delta$ . Issuer types with security payoffs below the threshold value  $f_j^m$  participate in the trade, whereas issuer types with payoffs above  $f_j^m$  are excluded (i.e., they reject the offer).

### Optimality of Separate Equity Sales

We now establish our first main result, which identifies a sufficient condition for the strict optimality of selling assets separately. This result also provides the necessary and sufficient condition under which selling assets separately yields the first-best level of trade surplus.

**Proposition 1.2.** *Suppose that the following condition holds:*

$$e(\bar{x}) \geq \frac{\delta}{1-\delta}, \quad \text{or equivalently} \quad \delta \leq \bar{\delta}, \quad (1.9)$$

where  $\bar{\delta} \equiv \frac{e(\bar{x})}{1+e(\bar{x})}$ . Then the following results obtain:

(i) *The issuer optimally sells each asset separately to a monopolistic buyer, that is,*

$$\Omega_j = \{j\} \text{ and } \varphi_j(X_j) = X_j \text{ for } j = 1, \dots, n. \quad (1.10)$$

*The first-best level of total surplus from trade,  $n(1-\delta)\mathbb{E}[X_i]$ , is achieved and the issuer collects  $n\delta\bar{x}$ , obtaining a surplus of  $n\delta(\bar{x} - \mathbb{E}[X_i])$ .*

(ii) If the issuer pools any of the assets, the total surplus from trade is strictly below the first-best level  $n(1 - \delta)\mathbb{E}[X_i]$ , and the issuer's surplus is strictly below  $n\delta(\bar{x} - \mathbb{E}[X_i])$ .

Suppose, for example, that the payoffs of the fundamental assets follow a uniform distribution,  $X_i \sim U[0, 1]$ , then Proposition 1.2 states that selling each asset separately is strictly optimal for the issuer whenever  $\delta \leq \bar{\delta} = 0.5$ . To provide intuition for the central results provided in Proposition 1.2, we develop the proof here in the main text. First, consider part (i) of the proposition. Suppose that the issuer sells an equity claim on a pool  $j$ , such that,  $\varphi_j(Y_j) = Y_j$ . When designing the optimal security, the issuer anticipates the buyer's optimal pricing response. Using equation (1.8), we can write the buyer's marginal benefit of increasing the threshold type  $f_j^m = y_j^m$  for  $f_j^m \in [0, \bar{y}_j)$  as:

$$(1 - \delta)f_j^m g_j(f_j^m) - \delta G_j(f_j^m). \quad (1.11)$$

This last equation highlights the generic trade-off that a buyer with market power faces when choosing the price he plans to offer. When marginally increasing the threshold type by increasing the price, the buyer benefits from extracting the full gains to trade  $(1 - \delta)f_j^m$  from this type, which has the local density  $g_j(f_j^m)$ . Yet, the associate price increase of magnitude  $\delta$  also comes at the cost of paying more when trading with all infra-marginal types, which have measure  $G_j(f_j^m)$ . In net, the buyer benefits from increasing the marginal buyer type if expression (1.11) takes a strictly positive value (for any  $f_j^m < \bar{y}_j$ ). This condition can be equivalently expressed as a condition applying to the above-defined elasticity function:

$$e_j(f_j^m) > \frac{\delta}{1 - \delta}. \quad (1.12)$$

Now suppose the issuer simply sells all assets separately. Then the condition  $e(\bar{x}) > \frac{\delta}{1 - \delta}$  together with Assumption 1.1 ensures that the buyer's optimal price quote for each asset is  $p_i = \delta\bar{x}$ , allowing the issuer to collect  $n\delta\bar{x}$ . In this case, the marginal issuer type is the highest type on the support  $[0, \bar{x}]$  and trade occurs with probability one, ensuring that the

first-best level of surplus from trade is achieved. The issuer cannot collect a total payment greater than  $n\delta\bar{x}$  from the monopolistic buyer since the best possible payoff that all assets can deliver jointly is  $n\bar{x}$ , and a buyer with market power would never offer a price above  $\delta n\bar{x}$ , even if he believed that this maximum payoff on all assets was attained.

To address part (ii) of the proposition, we show that the issuer's surplus and the total surplus are strictly lower when assets are pooled. First, we introduce the following result:

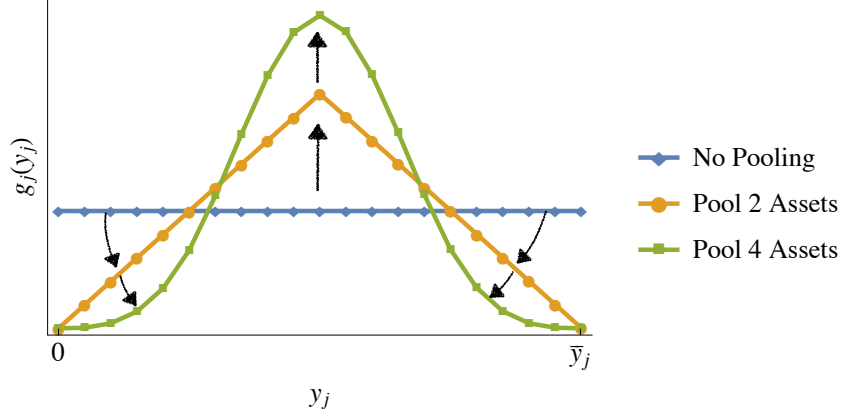
**Lemma 1.3.** *For any set  $\Omega_j$  that contains more than one element (i.e., if there is pooling), the following condition is satisfied:*

$$e_j(\bar{y}_j) = 0 < \frac{\delta}{1 - \delta}. \quad (1.13)$$

*Proof.* See Appendix 1A. □

This lemma states that if the issuer pools assets and issues an equity security on the pool, the elasticity for this security at the upper bound of the support  $\bar{y}_j$  is zero, implying the exclusion of a positive measure of types. The elasticity is zero at the upper bound  $\bar{y}_j$  since the density for the outcome that two assets simultaneously achieve their highest possible value  $\bar{x}$  is zero. The intuitive reason for this result is diversification: the more diversified pool of assets is less likely to generate an extreme outcome than each idiosyncratic asset separately. Figure 1.1 illustrates this result for the case where each separate asset follows a uniform distribution. The figure compares, after rescaling the domains (see caption details), the shapes of the PDFs of a single asset, a pool of two assets, and a pool of four assets. The graph illustrates the familiar notion that diversification leads to a more peaked distribution with thinner tails.

These changes in the shapes of the PDFs map into corresponding changes in the elasticity functions  $e_j(y_j)$ , which govern the pricing behavior in the trading game (see equation (1.12)). Figure 1.2 confirms that as soon as two assets are pooled, the elasticity at the upper bound

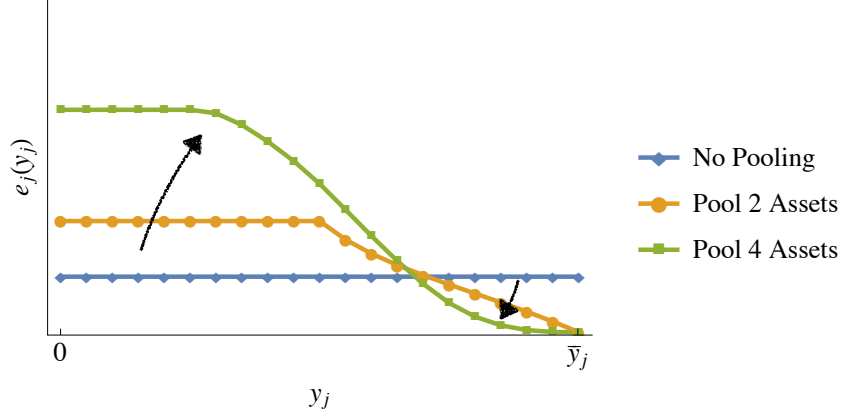


**Figure 1.1: Effect of pooling on the shape of the probability density function.** The graph considers a setting with four assets ( $n = 4$ ), each of which has a payoff  $X_i \sim U[0, 1]$ . The graph plots the PDF of a separate asset, a pool of 2 assets, and a pool of 4 assets. To compare the PDFs' shapes relative to their respective domains ( $[0, 1]$ ,  $[0, 2]$ , and  $[0, 4]$ ), the graph rescales the horizontal axis to represent the interval  $\chi_j = [0, \bar{y}_j]$  for each PDF  $g_j$ .

of the support  $\bar{y}_j$  shrinks to zero. A thinner right tail of the PDF implies a lower elasticity in the right tail of the distribution (recall that the elasticity is the ratio of the local density  $g_j(y_j)$  to the average density  $G_j(y_j)/y_j$ ). Facing a less elastic response from the issuer in that part of the domain, a monopolistic buyer has stronger incentives to offer lower prices, which leads to the exclusion of high issuer types. If  $\tilde{n} \geq 2$  assets are pooled in a set  $\Omega_j$ , then the buyer optimally chooses a marginal issuer type strictly below  $\bar{y}_j = \tilde{n}\bar{x}$ , since  $e_j(\bar{y}_j) = 0 < \frac{\delta}{1-\delta}$ . Correspondingly, the price offered by the buyer is strictly below  $\delta\tilde{n}\bar{x}$  for a pool of  $\tilde{n}$  assets, and the issuer obtains an expected payoff from pooling that is strictly below  $\delta\tilde{n}\bar{x}$ .

To conclude the proof of part (ii) of Proposition 1.2, we address whether the issuer, after pooling assets, could still obtain an equally beneficial payoff as in the case of separate sales by designing an optimal security  $F_j = \varphi_j(Y_j)$  on the pooled payoff  $Y_j$ . The following lemma characterizes the optimal security on a given underlying asset  $Y_j$  when an equity security leads to rationing.

**Lemma 1.4.** *When the trading of an equity security on a payoff  $Y_j$  leads to the exclusion of issuer types (i.e., if  $e(\bar{y}_j) < \delta/(1-\delta)$ ) but sustains trade with positive probability (i.e., if*



**Figure 1.2: Effect of pooling on the shape of the elasticity function.** The graph considers a setting with four assets ( $n = 4$ ), each of which has a payoff  $X_i \sim U[0, 1]$ . The graph plots the elasticity function of a separate asset, a pool of 2 assets, and a pool of 4 assets. To compare the elasticity functions' shapes relative to their respective domains ( $[0, 1]$ ,  $[0, 2]$ , and  $[0, 4]$ ), the graph rescales the horizontal axis to represent the interval  $\chi_j = [0, \bar{y}_j]$  for each elasticity function  $e_j$ .

$e(0) > \delta/(1 - \delta)$ ), the optimal security from the perspective of the issuer is a debt security with face value  $d_j^m$ , i.e.,  $\varphi = \min[Id_{\chi_j}, d_j^m]$ , where  $d_j^m$  is the largest  $d$  such that:

$$\underbrace{\int_0^d f_j g_j(f_j) df_j + [1 - G_j(d)]d - \delta d}_{\text{Net-payoff from offering price } \delta d} - \underbrace{\int_0^{f_j^m} (f_j - \delta f_j^m) g_j(f_j) df_j}_{\text{Net-payoff from offering price } \delta f_j^m < \delta d.} \geq 0, \quad (1.14)$$

and where  $f_j^m$  solves:

$$e_j(f_j^m) = \frac{\delta}{1 - \delta}. \quad (1.15)$$

That is, the optimal debt contract specifies the highest face value such that the buyer weakly prefers offering a price  $\delta d$  for the debt that is always accepted by the issuer over offering a lower price that is only accepted by issuer types below the threshold type  $f_j^m$ .

*Proof.* As each of the pooled payoffs  $Y_j$  satisfy the regularity condition stated in Assumption 1.1, these results follow from Propositions 3, 4, and 5 in Biais and Mariotti's (2005) analysis of a setting with one underlying asset.  $\square$

Since any pooling of  $\tilde{n} \geq 2$  assets in a set  $\Omega_j$  leads to exclusion when an equity security is offered (as  $e_j(\bar{y}_j) < \delta/(1 - \delta)$ ), Lemma 1.4 implies that the best possible security written on that pool is a debt security with face value  $d_j^m$ . Yet, since  $d_j^m < \bar{y}_j = \tilde{n}\bar{x}$ , selling this debt security will also deliver a payoff to the issuer that is strictly below the one he obtains from selling the  $\tilde{n}$  assets separately. Thus, the effects of diversification cannot be undone by designing a security that pays as a function of the pooled (diversified) cash-flow  $Y_j$ . This concludes our proof of Proposition 1.2.

In sum, when separate sales of assets are efficient, pooling assets leads to strictly worse outcomes, both in terms of the issuer's surplus and the total trade surplus. This result emerges as pooling generically leads to a payoff distribution with thinner tails, and equivalently, a less elastic response to price quotes in the right tail of the payoff distribution (see Figure 1.2). A less elastic response causes a liquidity supplier with market power to optimally set prices that lead to inefficient rationing, harming both the issuer and total trade efficiency. Thus, in contrast to the previously analyzed scenario with competitive liquidity suppliers (see Proposition 1.1), pooling assets may hurt the issuer when the demand side has market power.

### Optimality of Separate Debt Sales

Proposition 1.2 provided the condition under which selling assets separately, as equity, is optimal for the issuer and attains the first-best level of trade surplus. We will now show that even when this condition is violated, it may be optimal for the issuer to sell assets separately. However, in those cases, the issuer will opt for separate debt securities rather than equity securities.

**Proposition 1.3.** *Suppose now that each elasticity function  $e_j$  is strictly decreasing on its respective support  $\chi_j$  (recall that Assumption 1.1 only required them to be weakly decreasing). There exists a  $\delta^* \in (\bar{\delta}, 1]$  such that for all  $\delta \in (\bar{\delta}, \delta^*)$ , it is strictly optimal to issue a separate debt security on each asset payoff  $X_i$ .*



To prove this result, it is useful to introduce additional notation. Let  $\Pi(\delta)$  denote the issuer's profit, as a function of the parameter  $\delta$ , from selling one underlying asset separately, and issuing an optimal security on that underlying asset. Further, let  $\Pi_{\tilde{n}}(\delta)$  denote the issuer's profit, also as a function of  $\delta$ , from pooling  $\tilde{n}$  assets and issuing an optimal security on that underlying pool. The basic idea of the proof is to establish that these profits are continuous functions of  $\delta$ , and to use the fact established in Proposition 1.2, which is that for  $\delta = \bar{\delta}$ , selling assets separately yields the issuer a strictly higher expected profit than from pooling assets:

$$\tilde{n}\Pi(\bar{\delta}) > \Pi_{\tilde{n}}(\bar{\delta}). \quad (1.16)$$

First, suppose the issuer issues equity securities. In that case, for all  $\delta \in \left[ \frac{e_j(\bar{x})}{1+e_j(\bar{x})}, \frac{e_j(0)}{1+e_j(0)} \right]$ , the monopolistic buyer would target an *interior* marginal issuer type  $f_j^m$  satisfying:

$$e_j(f_j^m) = \frac{\delta}{1-\delta} \Leftrightarrow f_j^m(\delta) = e_j^{-1} \left( \frac{\delta}{1-\delta} \right), \quad (1.17)$$

where  $e_j$  is an invertible function, since it is assumed to be strictly decreasing on its support. Thus, for all  $\delta \in \left[ \frac{e_j(\bar{x})}{1+e_j(\bar{x})}, \frac{e_j(0)}{1+e_j(0)} \right]$ , this marginal issuer type  $f_j^m$  is a continuous function of the discount factor  $\delta$ . This result is useful, since as shown in Lemma 1.4, the optimal debt security, which will be issued for  $\delta > \frac{e_j(\bar{x})}{1+e_j(\bar{x})}$ , is implicitly characterized as a function of this marginal issuer type obtained when issuing an equity security. Specifically, the optimal security from the perspective of the issuer is a debt security with face value  $d_j^m$ ,  $\varphi = \min[\text{Id}_{\mathcal{X}_j}, d_j^m]$  where  $d_j^m$  is the largest  $d$  such that:

$$\int_0^d f_j g_j(f_j) df_j + [1 - G_j(d)]d - \delta d - \int_0^{f_j^m} (f_j - \delta f_j^m) g_j(f_j) df_j \geq 0, \quad (1.18)$$

where  $f_j^m = e_j^{-1}(\frac{\delta}{1-\delta})$ . Note that this optimal face value  $d_j^m$  is then also a continuous function of  $\delta$ . This continuity result holds for any set  $\Omega_j$ , including the case where  $\Omega_j$  includes only one asset.

Finally, note that if all the optimal face values  $d_j^m$  are continuous functions of  $\delta$ , then the issuer's profit functions  $\Pi(\delta)$  and  $\Pi_{\tilde{n}}(\delta)$  are also continuous functions of  $\delta$  since:

$$\Pi(\delta) = \delta d^m(\delta) - \delta \int_0^{d^m(\delta)} fg(f)df - \delta[1 - G(d^m(\delta))]d^m(\delta) = \delta \int_0^{d^m(\delta)} G(f)df, \quad (1.19)$$

$$\Pi_{\tilde{n}}(\delta) = \delta d_{\tilde{n}}^m(\delta) - \delta \int_0^{d_{\tilde{n}}^m(\delta)} fg(f)df - \delta[1 - G(d_{\tilde{n}}^m(\delta))]d_{\tilde{n}}^m(\delta) = \delta \int_0^{d_{\tilde{n}}^m(\delta)} G_{\tilde{n}}(f)df, \quad (1.20)$$

where we use integration by parts to simplify the expressions.

Given equation (1.16) and the continuity of functions  $\Pi(\delta)$  and  $\Pi_{\tilde{n}}(\delta)$ , we know that there is also a non-empty region  $(\bar{\delta}, \delta^*)$  such that when  $\delta$  lies in that region, we have:

$$\tilde{n}\Pi(\delta) > \Pi_{\tilde{n}}(\delta), \quad (1.21)$$

that is, selling  $\tilde{n} \geq 2$  assets separately (with debt) is strictly better for the issuer than selling debt on a pool of  $\tilde{n}$  assets. The upper bound of the region,  $\delta^*$ , is implicitly defined by the lowest  $\delta$  such that  $\tilde{n}\Pi(\delta) = \Pi_{\tilde{n}}(\delta)$ .

The main insight from Proposition 1.3 is that even when the potential gains to trade are smaller than required by the condition stated in Proposition 1.2, pooling assets may still be suboptimal for the issuer. The main difference relative to the result of Proposition 1.2 is that once separate equity securities do not trade fully efficiently, switching to separate debt securities is optimal. Yet, as the design of these debt securities is still intimately linked to the monopolistic liquidity supplier's incentives to inefficiently screen the issuer (the marginal issuer type from equity sales enters equation (1.18)), the elasticity of trading volume is still an important determinant of the issuer's net-profit. As pooling assets reduces this elasticity in the right tail of the payoff distribution (see Figure 1.2), it is undesirable to do so when the marginal issuer type from separate equity sales is sufficiently high, or equivalently, when the liquidity differences between the issuer and the buyer are sufficiently large (i.e.,  $\delta$  is sufficiently low).

## Optimality of Pooling Assets when Adverse Selection is Severe

Unlike with competitive demand where it is always optimal to pool assets for the issuer, the predictions for the scenario with monopolistic demand are more nuanced and feature a trade-off between the benefits of diversification and the preservation of information rents. Propositions 1.2 and 1.3 have highlighted that the optimality of separate sales emerges when trade is particularly valuable, that is, when the prospective buyer and the issuer differ more in terms of their liquidity. In contrast, when potential gains from trade are smaller, adverse selection concerns and the exercise of market power lead to larger inefficiencies when assets are sold separately. Lower gains from trade (i.e., higher values of  $\delta$ ) cause the liquidity supplier to choose a more aggressive pricing strategy, which leads to the exclusion of a larger range of issuer types when equity securities are issued. In fact, whenever  $\delta > \frac{e(0)}{1+e(0)}$  the trading of separate securities (whether it is equity or debt) fails completely as the elasticity function  $e(x)$  then lies below  $\delta/(1-\delta)$  everywhere on the support — all issuer types are excluded. Yet, as suggested by Figure 1.2, pooling assets increases the elasticity in the *left* tail of the distribution, and thus can allow sustaining trade when separate sales would lead to trade breakdowns. Thus, when adverse selection concerns are severe, relative to the magnitude of the potential gains from trade, the trade-off faced by the issuer is tilted toward favoring the pooling of assets.

**Proposition 1.4.** *Suppose that the issuer has  $n > \frac{\delta}{1-\delta}$  assets. Then at least one of the subsets  $\Omega_j$  will optimally consist of  $n^*$  assets, where  $n^* > \frac{\delta}{1-\delta}$ .*

Proposition 1.4 highlights that for sufficiently high values of the discount factor  $\delta$  the issuer optimally pools multiple assets into a security. This result is directly linked to the previously mentioned fact that trade breaks down completely whenever the elasticity of an underlying asset at the lower bound of the support is lower than  $\delta/(1-\delta)$ . Let  $e_{\tilde{n}}(0)$  denote the elasticity function associated with a pool of  $\tilde{n}$  assets. If  $e_{\tilde{n}}(0) < \frac{\delta}{1-\delta}$ , then trade will break down with probability 1 for any security written on this pool. Yet, as suggested by

Figure 1.2, the elasticity at the lower bound increases when more assets are pooled, a fact that is established in the following lemma.

**Lemma 1.5.** *A pool of  $\tilde{n}$  assets has the elasticity  $e_{\tilde{n}}(0) = \tilde{n}$  at the lower bound of the support.*

*Proof.* See Appendix 1A. □

Since trade breaks down completely whenever  $e_{\tilde{n}}(0) < \frac{\delta}{1-\delta}$ , the issuer can only attain a positive expected surplus when the elasticity of an underlying asset, evaluated at the lower bound, exceeds  $\frac{\delta}{1-\delta}$ . Since, as shown in Lemma 1.5, this elasticity for a pool of  $\tilde{n}$  assets is exactly equal to  $\tilde{n}$ , the issuer will at least pool  $n > \frac{\delta}{1-\delta}$  assets to ensure that he can attain an expected surplus greater than zero. At the same time, we know from our earlier analysis that pooling an infinite number of assets is also suboptimal for the issuer, as perfect diversification leads him to obtain zero surplus. Thus, even when the issuer has a continuum of assets, he prefers to pool only a subset of the assets, or none at all.

Propositions 1.2, 1.3, and 1.4 have highlighted that the trade-offs faced when deciding whether to pool assets are intimately linked to the magnitude of the potential gains from trade. When they are sufficiently large (i.e.,  $\delta$  is sufficiently low) it is optimal to sell assets separately. In this case, the liquidity supplier is less worried about being adversely selected by the issuer and is more cautious in exercising his market power. Moreover, we have shown that when the issuer sells assets separately, the elasticity with which he responds to price changes is larger in the right tail of the distribution than when he is pooling assets. This elasticity in the right tail is relevant when the potential gains from trade are sufficiently large, causing the marginal issuer type to reside in that part of the distribution. Yet, when the potential gains from trade are sufficiently small, adverse selection concerns and the exercise of market power lead to complete market breakdowns when assets are sold separately. In this case, the issuer has to reduce the amount of asymmetric information to

ensure that trade can occur at all. He thus pools assets. In particular, Lemma 1.5 reveals that the elasticity in the left tail of the support rises with the number of assets that are pooled, allowing trade to occur once sufficiently many assets have been pooled.

#### 1.4. Robustness

In this section, we discuss the robustness of our main insights to various changes in the environment.

*Risk aversion.* In line with the existing literature, we have assumed that agents are risk neutral. It is worth noting that, even if we allowed for risk aversion, pooling assets would not by itself lead to better risk sharing among traders. This is because the issuer offers to sell all assets to the buyer(s) independent of whether he pools the assets or not. With risk-averse agents, the main impediment to risk sharing would be the fact that the issuer's private information may result in socially inefficient trade breakdowns, which is already a force at play in our baseline model.

*Correlated asset payoffs.* In our setup, the fundamental payoffs  $X_i$  are identically and independently distributed. The highlighted trade-off between information rents and diversification that is associated with pooling assets would, however, also apply if assets' payoffs exhibited some correlation. Pooling imperfectly correlated payoffs would still lead to qualitatively similar effects on the shape of the payoff distribution — a pool's payoff distribution would still feature thinner tails. As a result, the elasticity function of a pool's payoff would decrease near the upper bound of the support, increasing a monopolistic buyer's incentives to inefficiently screen the issuer. Just like in our baseline model, this downside of pooling assets could then also dominate the diversification benefits highlighted in the existing literature, rendering it optimal for the issuer to sell assets separately.

*Multiple constrained buyers.* The main result of our paper, that is, pooling assets might be suboptimal when liquidity suppliers have market power, is derived in an environment

in which only one buyer has a discount factor of one, but is deep-pocketed. Similar results obtain in the presence of multiple buyers, provided that these buyers face position limits, wealth constraints, or risk aversion. Consider a simple extension of our baseline model in which the aggregate position limit across all prospective buyers (measured in units of underlying assets) is marginally smaller than the total quantity of assets up for sale. In this case, each buyer’s price setting strategy is identical to the one derived in our baseline model — as the total supply always exceeds the total demand, a buyer faces a residual supply curve that is unaffected by other buyers’ pricing strategies.<sup>10</sup> As a result, the issuer still faces the trade-offs featured in our baseline model.

*Signaling through retention.* In the scenario with competing liquidity suppliers, allowing the issuer to signal asset quality through partial retention, as in DeMarzo (2005), would yield results that are (unsurprisingly) consistent with DeMarzo (2005) — issuers with assets of higher quality would retain a higher fraction of the issue.<sup>11</sup> Signaling would then allow the high issuer types to separate themselves from the low types and would resolve the lemons problem for high values of  $\delta$ . In contrast, when facing a liquidity supplier with market power, the issuer can be worse off by signaling asset quality. Since the liquidity supplier makes a take-it-or-leave-it offer, he is able to extract all the surplus from trade when he is able to infer the issuer’s type. In this case, the issuer’s profit from implementing fully revealing retention policies is therefore weakly lower than his profit without any signaling through retention (see also Glode et al., 2018, for related arguments). Moreover, as mentioned earlier, Biais and Mariotti (2005) show that for a given security offered by the issuer, the monopolistic buyer’s optimal mechanism is a take-it-or-leave-it offer for the total supply of the security, rather than a menu of price-quantity offers that could result in the issuer using retention to signal asset quality.

---

<sup>10</sup>The result that capacity constraints can hamper competition is well known in the literature, see, for example, Green (2007a).

<sup>11</sup>See also Williams (2019) who studies the optimality and efficiency of security retention in the presence of search frictions.

## 1.5. Conclusion

This paper studies the optimality of pooling assets when security issuers face a market in which liquidity is scarce and buyers endowed with such liquidity may have market power. Unlike in competitive environments, we find that selling assets separately may be preferred by issuers, in particular when liquidity differences between the buy side and the sell side of the market are sufficiently large. While our results suggest that the dramatic decline of the ABS market post crisis may represent an efficient response by originators to drastic changes in liquidity and market power in OTC markets, it also highlights the potential welfare implications of liquidity constraints imposed on financial institutions in the new market environment.

In future research, the principles uncovered by our analysis could also be applied to shed light on firms' capital structure decisions, specifically, to firms' choices regarding the maturity structure of their debt. To illustrate the mapping between this problem and our setup, suppose a firm generates cash-flows in different time periods and is privately informed about these future cash-flows. Each cash-flow can be viewed as one of the fundamental assets from our baseline setup. The firm then decides whether to pool all cash-flows across time (e.g., by issuing an equity claim or a perpetual debt claim) or not (e.g., by issuing multiple zero coupon bonds of different maturities). Our analysis suggests that when firms face investors with market power, it is relatively more beneficial for them to issue multiple debt securities with different maturities, a practice that is indeed quite common.

## CHAPTER 2 : Selling to Investor Network: Allocations in the Primary Corporate Bond Market

### 2.1. Introduction

The corporate bond market is one of the main sources of capital for US firms, with about \$9 trillion in debt currently outstanding and annual issuance of more than \$1 trillion in recent years.<sup>1</sup> However, despite its size and importance for firms' cost of capital, some essential features of the market have received little attention in the literature. In particular, the issuance process of new bonds, which involves determination of a primary price and allocations to participating investors, remains relatively unexplored and warrants more consideration because of controversial distributions that take place in practice.

Indeed, the question of what determines initial allocations and whether they are distributed fairly is debated among market participants and periodically attracts attention from the financial press.<sup>2</sup> In the primary market, underwriters who are hired by the issuing firms choose an allotment of new bonds. Routinely, they allocate large fractions of new issues to bigger institutional investors while smaller institutional investors, trying to obtain new bonds, receive zero allocations. Coupled with a small underpricing of new issues,<sup>3</sup> this practice is puzzling: it might seem that the issuers are "leaving money on the table." It also provokes the smaller, excluded institutional investors to claim that underwriters act inequitably and prioritize the interests of larger investors.<sup>4</sup> In contrast, the issuers do not express any concerns over this discriminatory allocation practice.

Furthermore, the extent and the significance of the issue is evidenced by the updated European legislative framework for financial regulation MIFID II.<sup>5</sup> Part of the framework is

---

<sup>1</sup>\$9,200.7 billion corporate bonds outstanding as of the end of 2018 and the issuance of \$1,527.7 billion, \$1,652.4 billion, and \$1,336.7 billion in 2016, 2017, and 2018, respectively. SIFMA (Securities Industry and Financial Markets Association).

<sup>2</sup>Ramakrishnan et al. (2014), Reuters. Possible SEC inquiry raises age-old bonds question.

<sup>3</sup>Nikolova et al. (2018) estimate the average underpricing of 37.17 bps for a sample of investment-grade and non-investment-grade bonds issued between 2004 and 2014.

<sup>4</sup>Alloway et al. (2014), Financial Times. Bond syndication bonanza under scrutiny.

<sup>5</sup>The directive became effective as of January 3, 2018.



designed to address the objections of smaller investors and mandates that underwriters provide explicit allocation policies for all new issues.<sup>6</sup> The provision of allocation policies is also instructed under the latest standards established by self-regulatory organizations of market participants.<sup>7</sup>

Unlike the similar discriminatory practices in allocation of stock IPOs, which are studied in the literature, allotments in the primary market for corporate bonds are harder to explain by potential agency problems between issuers and underwriters.<sup>8</sup> In fact, the majority of firms issuing bonds access the market repeatedly over their lifetime and, presumably, can terminate their relationship with underwriters if they become unsatisfied with their service at any point. Additionally, possible information asymmetries between issuers and investors that have been suggested to explain allocations in stock IPOs<sup>9</sup> are likely to be significantly attenuated for the mature public companies regularly issuing corporate bonds.

In this paper, I propose a novel explanation of the common exclusion of smaller investors from the primary market, which relies on a feedback from liquidity of the secondary market and implies that exclusion might be optimal for issuers even if excluded investors are willing to pay a high price for some bonds. To that end, it is necessary to develop a new model of issuance for corporate bonds, since existing models of stock issuance, which examine a feedback from the secondary market, are not directly applicable due to the critical structural differences of the secondary markets for the two asset classes. Specifically, stocks are traded on exchanges, and their secondary market, to the first approximation, can be characterized as centralized. In contrast, corporate bonds are traded over-the-counter (OTC), through dealers, and their secondary market can be described as decentralized.

The secondary market in my model captures two distinctive features of the OTC market

---

<sup>6</sup>Markets in Financial Instruments Directive 2014/65/EU, Articles 38-43.

<sup>7</sup>See New Issues Process for Fixed Income Markets standard instituted by FMSB (FICC Markets Standards Board) which builds on recommendations of ICMA (International Capital Market Association).

<sup>8</sup>Among others, studying agency problems in stock IPOs are Reuter (2006), Nimalendran et al. (2007), Goldstein et al. (2011), Jenkinson et al. (2018).

<sup>9</sup>The seminal contributions of Benveniste and Spindt (1989), Benveniste and Wilhelm (1990) show that allocations can be used to reward investors for revealing their private information about values of new issues.

for corporate bonds that are documented in the literature and set it apart from centralized markets. First, investors participating in the market form persistent trading relationships with dealers, who are the main providers of liquidity. Investors trade with the same dealers over time and rarely search for new counterparties.<sup>10</sup> Second, investors are highly heterogeneous with respect to the number of dealers they trade with.<sup>11</sup> Accordingly, from the market perspective, investors are heterogeneously connected with each other through dealers, and the map of all investor connections can be seen as a stable trading network.

Modeling the secondary market with a decentralized trading network, I study how secondary market liquidity feeds back into the optimal issuance decisions in the primary market. The baseline setup can be summarized as follows. In the centralized, primary market, an underwriter who acts in the best interests of an issuing firm sells homogenous bond units to multiple investors. Aggregating investors' bond orders, represented by demand schedules, the underwriter maximizes the issuer's revenues by choosing a uniform price and allocations to investors, possibly leaving some of the orders unfilled. Next, if an investor wishes to trade out of her bond holdings in the secondary market, the investor contacts her dealers, who provide liquidity. A dealer can be contacted by multiple client investors and their requests for trades affect each others' liquidity obtained from the same dealer. Lastly, investors' anticipation of liquidity in the secondary market feeds back into their primary market demands, which, in turn, determine the issuer's revenues.

There are two main trading frictions in the secondary market of the model that affect liquidity available to investors. First, investors can trade only with dealers with whom they have an established relationship. Second, investors depress each other's liquidity supply obtained from a common dealer due to dealers' limited inventory capacity. As a result, the investor trading network, which is formed through joint dealer connections, determines secondary market liquidity. In the model, as in practice, investors are heterogeneous with

---

<sup>10</sup>See, for instance, Di Maggio et al. (2017b).

<sup>11</sup>Hendershott et al. (2017) document that many small insurance companies trade with only one dealer whereas the largest insurers work with up to forty dealers.

respect to the number of dealers they trade with and, as I show, have different impacts on liquidity depending on their connectedness.

Specifically, even though investors with a higher number of connections demand more bonds in the primary market, they impose less stress on the liquidity available to other investors. The intuition for this result is as follows. In the primary market, more connected investors are willing to order more bonds because of their higher number of dealers and, therefore, greater access to liquidity. This means that, in the secondary market, these investors will have to sell more bonds if they need liquidity. However, more connected investors are able to choose between a greater number of dealers in the secondary market and, thus, trade more with the dealers who can offer better prices because these dealers are not trading with their other clients. This implies that more connected investors demand less liquidity from the dealers who have demands for liquidity from other clients, which, in turn, increases liquidity available to these other clients. As all investors anticipate this beneficial behavior of more connected investors, adjusting for the size of orders at the issuance, trading by more connected investors have a lower impact on expected liquidity.

Since secondary market liquidity feeds back to the primary market demands and less connected investors impose relatively more stress on liquidity available to other investors, the issuer might optimally choose the allocation policy that excludes them from the primary market, i.e., leaving their orders unfilled. Although the exclusion results in a direct loss of sales to these investors at the issuance, it improves expected liquidity for the remaining more connected investors and makes them willing to buy more bonds in the primary market. When the increased demand from the non-excluded investors makes up for the forgone demand, the exclusion is optimal and ultimately improves the issuer's revenues. I find that this is more likely when investors' trading connections are highly heterogenous, when the mass of less connected investors is small, or when they are more likely to demand liquidity in the secondary market. Furthermore, discriminatory allocations become more profitable when dealers' inventory costs increase.

In addition to explaining exclusion, the model provides an intuition for the observed underpricing of corporate bonds. In the model, the primary market demands of all investors are downward-sloping with respect to the price. Thus, if they could, excluded investors would still be willing to buy some bonds at the offering price or even at a higher price right after the primary market and before the realization of liquidity needs. However, the issuer cannot sell any bonds to these investors since she is committed to the chosen allocation policy, which is critical to generating a higher demand from the non-excluded investors. Consequently, in practice, underpricing might result from the attempts of excluded investors to obtain some bonds shortly after the primary market at a price higher than the issuance price.<sup>12</sup>

Lastly, I explore other potential reasons for the exclusion. The baseline model explains exclusion of some investors from the primary market by their higher impact on secondary market liquidity due to their lower connectedness. However, alternative negative externalities imposed by bond holdings of some investors on others can explain exclusion in a similar way. For instance, in the extension of the model, I consider the case of investors with correlated liquidity shocks. Since their presence significantly decreases beliefs of all investors about secondary market liquidity, the issuer might exclude a fraction of investors whose shocks are correlated from the primary market.

More broadly, my paper contributes to the understanding of how illiquidity in the secondary market due to different trading frictions feeds back to the primary market for new issues. The prior literature on this question primarily focuses on centralized markets. For instance, Ellul and Pagano (2006) show how asymmetric information in the secondary market generates underpricing in the primary market for stocks. Similarly, in a comprehensive framework, Vayanos and Wang (2011) summarize how different trading frictions in the secondary market, such as trading costs, asymmetric information and search, affect several liquidity measures and, ultimately, the primary market price.

---

<sup>12</sup>Flanagan et al. (2019) find that only a small fraction of the total amount of bonds is traded shortly after the primary offering and the trades tend to be in smaller sizes.

As a consequence, the literature has already identified several factors that an issuer can control to alter liquidity in the secondary market and, thus, through the feedback, her proceeds in the primary market. Specifically, the issuer can choose information sensitivity of assets, information disclosure policy in the secondary market, or asset maturity. My paper analyzes another significant factor that determines secondary market liquidity in decentralized markets and that the issuer can modify to improve her revenues in the primary market — initial allocations.

**Related literature.** To the best of my knowledge, my paper is the first to study how primary market corporate bond issuance, including allocations, is affected by investors' trading connections in the secondary market. To that end, I model persistent connections as a trading network,<sup>13</sup> which provides an alternative approach to the search and bargaining framework (Duffie et al., 2005, 2007) employed by the majority of papers in the literature that studies trade in decentralized markets.

The paper contributes to the recent theoretical literature on asset issuance in decentralized markets that links secondary market liquidity to primary market outcomes. In the search and bargaining framework, Arseneau et al. (2017) show how the level of bond issuance is determined by investors' portfolio allocation between long-term bonds, traded in the OTC market, and liquid assets; Bethune et al. (2019) study how asset issuance is affected by the division of surplus between buyers and sellers during bargaining in both secondary and primary markets; He and Milbradt (2014) and Bruche and Segura (2017) demonstrate how bond maturity affects secondary market liquidity and total issuance. Finally, Green (2007b) shows that underpricing can result from the lack of competition between underwriters due to their limited retail distribution capacity in the secondary market. My paper complements these studies, since I focus on how investor connections in the secondary market affect bond issuance in the primary market.

---

<sup>13</sup>Among other recent papers that model trading networks in decentralized markets are Gofman (2014), Condorelli et al. (2016), Wang (2016), Babus and Hu (2017), Malamud and Rostek (2017), Babus and Kondor (2018), Manea (2018).

The theoretical literature has also studied the issuance of bonds in multi-unit uniform-price auction setup (Wilson, 1979; Back and Zender, 1993; Wang and Zender, 2002; Back and Zender, 2001). These papers do not model interactions in the secondary market and assume a common value of new bonds for all bidders. Underpricing is explained by the existence of equilibria where buyers strategically bid steep demand curves so that bonds are sold below their value.<sup>14</sup> In contrast, in my paper, the buyers' bond values are derived from investors' access to liquidity in the secondary market, which depends on their connectedness. More importantly, since liquidity demanded by investors from common dealers reduces liquidity available to each other, allocating bonds to buyers in the primary market imposes negative externalities on the bond valuations of other buyers.

The empirical literature on corporate bond issuance focuses mainly on (under)pricing (Cai et al., 2007; Helwege and Wang, 2017; Brugler et al., 2016) due to limited availability of data on allocations. Notably, Goldstein et al. (2019) identify a significant impact of expected secondary market liquidity on primary market prices of new issues. Understanding the link between the two is the main focus of my paper.

Two recent empirical papers attempt to overcome the limitations of the data and consider bond allocations in the primary market: Nagler and Ottonello (2018) infer allocations from the quarterly institutional holdings data; and Nikolova et al. (2018) analyze precise allocations at the issuance date for insurance companies. Both papers find that the volume of secondary market trading with underwriters is a significant determinant of primary market allocations. Since investors with a higher number of dealer connections are more likely to be connected with underwriters and, thus, should be more likely to trade with them, this evidence supports the main prediction of my paper — that issuers should favor allocations to highly connected investors.

Finally, from the modeling perspective the paper is related to the literature that studies

---

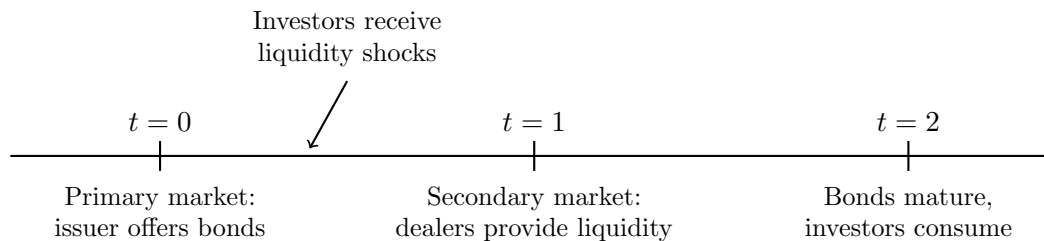
<sup>14</sup>In an equilibrium, an individual bidder does not have incentives to deviate and to submit higher bids because of the steep residual supply curve resulting from the strategies of other bidders. The equilibria in these papers and in Green (2007b) are reminiscent of tacit collusion (see, e.g., Ivaldi et al., 2003).

trade in networks. The majority of this literature studies how a single asset unit can move through networks (e.g., Gofman, 2014; Choi et al., 2017; Siedlarek, 2015). However, there are several recent papers (Manea, 2016, provides an excellent survey) that analyze bargaining in networks over multiple homogenous asset units by several players simultaneously. In their setting, a set of buyers and a set of sellers, all connected in a network, bilaterally bargain over prices of the units. The papers show how the distribution of units in the network affects the bargaining outcomes. This setting is a rough counterpart of the secondary market in my model. The novelty of my paper is that I step one period back and introduce the primary market stage where the issuer effectively determines the set of buyers and the set of sellers in the secondary market. Hence, the problem of the issuer is to find the two optimal sets that maximize issuance revenues given how outcomes in the secondary market feed back to the primary market.

The balance of the paper is organized in the following way. In the next section, I introduce the formal model and describe an investor trading network. Section 2.3 finds the equilibrium of the model and shows that the issuer might prefer to exclude some investors from the primary market based on their connections. Section 2.4 presents an extension to the main model where the issuer can exploit the knowledge of a correlation structure of investors' liquidity needs. Section 2.5 provides discussion of the results and potential generalizations. The last section concludes.

## 2.2. Model

I model the issuance of financial assets, such as corporate bonds, that will later be traded in the secondary over-the-counter market. For specificity, in the remaining sections, I refer to asset units as bond units or bonds. The setup of the model begins with a timeline and then proceeds to describe the agents and their payoffs.



**Figure 2.1: Model timeline.**

### 2.2.1. Timeline

The timeline of the model consists of the three periods  $t \in \{0, 1, 2\}$  and is summarized in Figure 2.1. The primary market for bonds is held at  $t = 0$ . Next, the secondary market opens at  $t = 1$ . Finally, at the last period,  $t = 2$ , bonds mature.

### 2.2.2. Agents

There are three types of risk-neutral agents in the model. First, an issuer who sells bond units in a centralized primary market, at  $t = 0$ . Second, investors who purchase the bond units from the issuer in the primary market. Third, dealers who provide liquidity to their client investors in the secondary market, at  $t = 1$ . The rest of this subsection describes each agent type in more detail.

#### Issuer

In practice, new bonds are sold and allocated on behalf of issuers by underwriters — usually, established investment banks. However, as argued above, underwriters are incentivized to act in the best interests of issuers. Indeed, given the recurrent nature of bond issuance, it is plausible to assume that issuers can terminate their contracts with underwriters if they become unsatisfied with their service at any point. Therefore, underwriters and the issuer are treated as the same agent in my model.

I assume that the issuer produces bond units at a constant cost  $\delta$  per unit. Each bond unit pays one unit of consumption good at maturity, in the final period  $t = 2$ . In the primary



market, the issuer chooses a uniform premium  $\pi$  per unit and offers bonds to investors for a total price  $c = \delta + \pi$  per unit.<sup>15</sup> The condition that bonds are sold for the uniform price is required by law. In particular, issuers can not discriminate investors charging them different prices in the primary market. At the same time, I assume that the issuer can discriminate bond allocations based on investor connections and can potentially exclude some investors from the primary market. I describe the allocation procedure below once investor connections are introduced formally.

### Investors

There is a unit measure of investors. Investors are heterogeneous with respect to a number of dealers that they can trade with in the secondary market. In the baseline model, I assume that each dealer has two client investors.<sup>16</sup> Through joint dealer connections investors are interconnected with each other and form a trading network. Figure 2.2 provides an example of a segment of a trading network. In the figure, investor  $i$  is connected to the four dealers  $d_1, \dots, d_4$  and shares the same common dealer  $d_1$  with investor  $j_1$ . Similarly, investor  $i$  is indirectly connected to investors  $j_2, \dots, j_4$ . Such investors  $j_1, \dots, j_4$  are called *neighbors* of investor  $i$ .

The number of dealers of investor  $i$  is denoted by  $n_i$  and can take values in the set  $N = \{\underline{n}, \dots, \bar{n}\}$ . This captures the empirical fact that investors in the corporate bond market are heterogeneous with respect to the number of dealers that they trade with.<sup>17</sup> The distribution of investor connections in the population is given by the density function  $f : N \rightarrow [0, 1]$ , i.e., the mass of investors with  $n$  dealers is  $f_n$ . This distribution is commonly known.

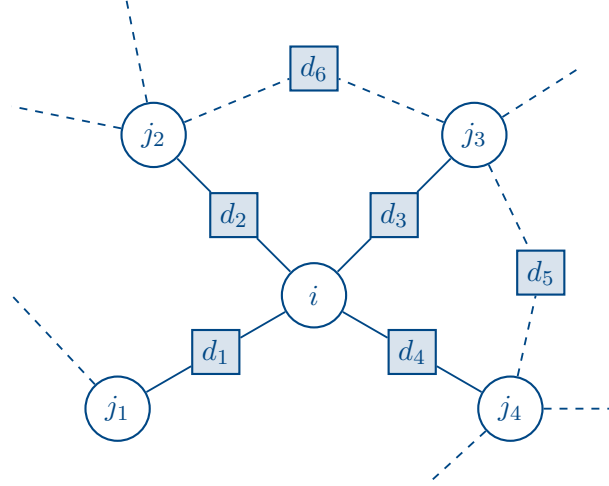
Crucially, I assume that the issuer observes the whole trading network while investors do

---

<sup>15</sup>The choice of the premium  $\pi$  is equivalent to the choice of the issuance quantity  $Q$ . Thus, the model subsumes the case where the bond supply is fixed to some  $\bar{Q}$ .

<sup>16</sup>The model can be generalized to the case where each dealer has more than two clients. See Section 2.5 for the discussion.

<sup>17</sup>The model takes a trading network as given and does not consider the question of network formation.



**Figure 2.2: Segment of a trading network.** An investor  $i$  with  $n_i = 4$  dealers and her investor neighbors  $j_1, \dots, j_4$ . Circles represent investors and squares represent dealers. Connections known (unknown) to investor  $i$  are represented by solid (dashed) lines.

not know the number of their neighbors' dealers. This assumption is motivated by the fact that, in practice, the secondary market is highly opaque and a client of a given dealer can only guess the connectedness of other investors trading with the same dealer. In contrast, underwriters know investor connections because they also act as dealers in the secondary market. Since underwriters and the issuer are the same agent in the model, the knowledge of investor connections allows the issuer to discriminate bond allocations in the primary market.

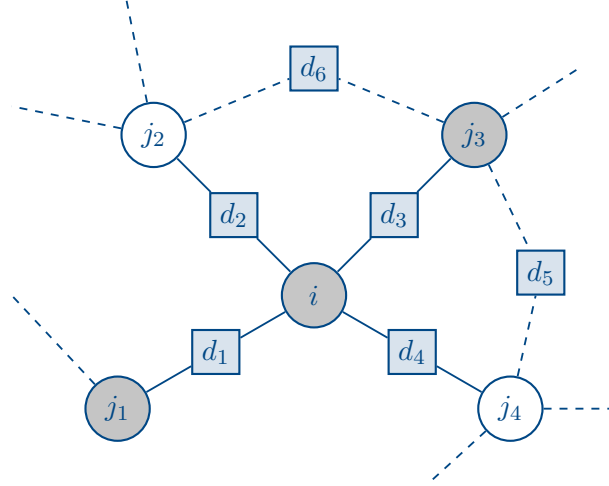
Finally, to formalize investors' beliefs about connectedness of their neighbors, it is useful to introduce the following.

**Assumption 2.1.** *Connections to dealers are formed under a uniform distribution. As a result, the distribution of neighbors' connections in the population is given by:*

$$\tilde{f}_n \equiv \frac{nf_n}{\sum_k kf_k}, \quad (2.1)$$

which is the probability that a given investor's neighbor has  $n$  dealers.<sup>18</sup>

<sup>18</sup>The probability adjusts for the fact that more connected investors have a higher number of dealers and, therefore, through a given dealer, an investor is more likely to be connected to a more connected investor.



**Figure 2.3: Patient and impatient investors within a trading network.** An impatient investor  $i$  with  $k_i = 2$  patient and  $n_i - k_i = 2$  impatient investor neighbors. Transparent and shaded circles represent patient and impatient investors, respectively, and squares represent dealers.

### Investors' liquidity shocks

After the primary market closes at  $t = 0$  and before the secondary market opens at  $t = 1$ , some investors receive an exogenous need for liquidity, which forces them to liquidate their bond positions acquired at issuance.

Formally, each investor with  $n$  dealers receives a liquidity shock with probability  $1 - p_n$ .<sup>19</sup> I call such investors *impatient*. With complimentary probability  $p_n$ , an investor does not need liquidity. Such investors are called *patient*. In the baseline model, I assume that the liquidity shocks are independent across investors.<sup>20</sup> Figure 2.3 illustrates a realization of the liquidity shocks where  $k_i$  denotes the number of patient neighbors of the investor  $i$ .

Patient and impatient investors differ in their value of consumption in the last period,  $t = 2$ . In particular, the utilities of patient and impatient investors are

$$u^{ns} = c_1 + c_2, \quad (2.2)$$

$$u^s = c_1, \quad (2.3)$$

<sup>19</sup>If  $p_n = p$  for all  $n$ , all investors are ex-ante identical except the number of dealer connections.

<sup>20</sup>Section 2.4 considers the case of correlated liquidity shocks and shows how the issuer can exploit the knowledge of a correlation structure of the shocks.

where the superscripts  $\{s, ns\}$  stand for *shocked* (impatient) and *non-shocked* (patient) investors, respectively, while  $c_t$  is consumption in period  $t$ . The above equations indicate that impatient investors, who are in need of liquidity, do not value consumption at  $t = 2$  and are better off selling their bonds in the secondary market for a positive price. In contrast, patient agents do not discount consumption between periods  $t = 1$  and  $t = 2$ .

In addition, it is assumed that investors do not discount consumption between periods  $t = 0$  and  $t = 1$ . As a result, the total surplus from transferring bonds from the issuer to the patient investors is equal to  $1 - \delta$ , which is the difference between the amount of consumption units that a single bond pays at maturity and the marginal cost of a bond unit to the issuer.

An investor  $i$ 's demand for bonds in the primary market,  $q_i$ , is determined by the price per bond unit  $c$  charged by the issuer and the investor's expectations about the secondary market liquidity, which is provided by the dealers and affected by the trading network.

## Dealers

Dealers provide liquidity to their clients non-strategically. The dealer  $d_{ij}$  connecting its two client investors  $i$  and  $j$  provides liquidity to them by absorbing higher quantities  $q$  at lower prices<sup>21</sup> through a downward-sloping asset demand schedule:

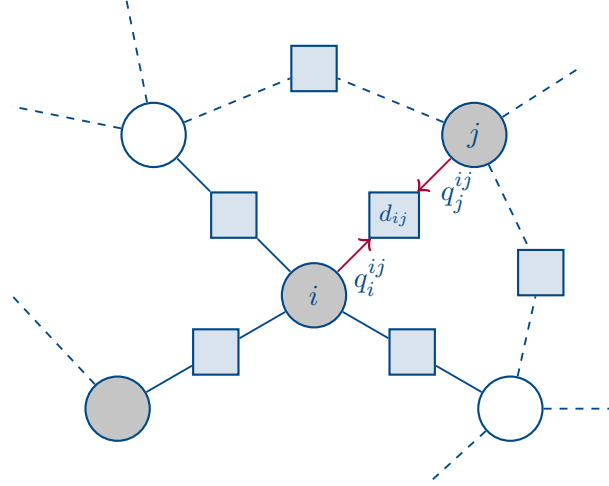
$$P_{ij}(q) = 1 - \lambda \cdot q^\theta, \quad (2.4)$$

where the parameter  $\lambda > 0$  governs the magnitude of a trade's price impact and  $\theta \geq 1$  controls its curvature. Specifically, if investor  $i$  sells  $q_i^{ij}$  and investor  $j$  sells  $q_j^{ij}$  to their common dealer  $d_{ij}$  (Figure 2.4), the price per unit that both of them obtain is given by:

$$P_{ij}(q_i^{ij} + q_j^{ij}) = 1 - \lambda(q_i^{ij} + q_j^{ij})^\theta. \quad (2.5)$$

---

<sup>21</sup>This assumption is consistent with empirical findings that dealer inventory is a significant determinant of bid-ask spreads (Feldhütter and Poulsen, 2018).



**Figure 2.4: Trading by impatient neighbors with the common dealer.** Impatient investors  $i$  and  $j$  trading with their common dealer  $d_{ij}$  in the secondary market. Arrows represent trading activity. Transparent and shaded circles represent patient and impatient investors, respectively, and squares represent dealers.

The specific nature of the liquidity supply can result from inventory costs incurred by dealers, with higher  $\lambda$  implying higher costs, and an increase in the time it takes a dealer to unwind larger positions.<sup>22,23</sup> It is assumed that the bonds obtained by dealers are resold to their patient clients through the inter-dealer market after  $t = 1$ .<sup>24</sup> Since liquidity shocks are realized by that time, the price is equal to each bond's payment at maturity and, therefore, bonds purchased by investors after the secondary market do not affect their primary market demands.

Due to the form of liquidity supply, only impatient investors trade their bond holdings in the secondary market. In contrast, patient investors strictly prefer to hold bonds until maturity. I assume that impatient investors trading in the secondary market submit their orders to their dealers simultaneously and all orders are cleared at once. In addition, it is assumed that, before the secondary market opens at  $t = 2$ , investors learn who among their neighbors received a liquidity shock and became impatient. In practice, investors can learn

<sup>22</sup> Appendix 2B provides a microfoundation for the dealer liquidity supply.

<sup>23</sup> Similarly, adverse selection concerns may result into a downward-sloping liquidity supply.

<sup>24</sup> The assumption that dealers are connected into inter-dealer market allows dealers providing liquidity to both their clients to unwind the accumulated inventory before bond's maturity. Importantly, the access to this market is not frictionless and dealers have to incur inventory costs before they can trade with others.

this through dealers who adjust their price quotes based on the total incoming order flow. Consequently, investors learn which of their dealers will have to provide liquidity to both their clients in the secondary market. This allows more connected investors to direct more of their trades towards dealers offering better prices and reflects natural flexibility of these investors to adjust their trading in the secondary market.

Overall, the setup captures the idea that it is cheaper to trade with the dealers connecting investors to patient rather than impatient neighbors because of the common knowledge that patient neighbors do not trade in the secondary market, while impatient neighbors have to tap dealers' liquidity.<sup>25</sup>

### *2.2.3. Payoffs*

In this subsection, I formally define the issuer's and investors' payoffs, starting with the latter. Since some investors have to trade bonds due to liquidity shocks, I first introduce their payoffs in the secondary market and then characterize their payoffs in the primary market.

#### **Investor's payoff in the secondary market**

As argued above, due to the nature of the liquidity supply, patient investors do not participate in the secondary market and obtain one unit of consumption for each bond unit when the bonds mature. Therefore,  $u_i^{ns}(q_i) = q_i$ .

On the other hand, impatient investors have to liquidate their bond holdings  $q_i$  in the secondary market by trading with their dealers. Thus, the profit from trade of an impatient investor  $i$  depends on her connectedness as well as the number of her patient neighbors and

---

<sup>25</sup>Note that the model implies that the same bonds are sold for different prices in different parts of the trading network, which is consistent with the empirical literature that documents a significant price dispersion in the corporate bond market, e.g., Goldstein and Hotchkiss (2012); Feldhütter (2012); O'Hara et al. (2018).

is given by:

$$u_i^s(q_i, n_i, k_i) = \sum_{j \in N_i} \tilde{\mathbb{E}}_{n_j} \mathbb{E}_{k_j} \left[ P_{ij} (q_i^{ij} + q_j^{ij}) q_i^{ij} \right] \quad s.t. \quad \sum_{j \in N_i} q_i^{ij} = q_i, \quad (2.6)$$

where  $N_i$  is the set of investor  $i$ 's dealers or, equivalently, her neighbors.

The payoff is a sum of revenues from trades with different dealers of investor  $i$ . The constraint indicates that investor  $i$  can split the sale of her total holdings  $q_i$  among her dealers. The quantities  $q_j^{ij}$ , traded by impatient neighbors  $j$  through the common dealers, are similarly determined by their access to liquidity in the secondary market. However, in the secondary market, investor  $i$  knows only which of her neighbors are impatient but not their connectedness. Thus, the expectation is taken with respect to the total number of investor  $j$ 's neighbors  $n_j$  and the number of investor  $j$ 's patient neighbors  $k_j$ .<sup>26</sup>

### Investor's payoff in the primary market

Stepping one period back, the payoff of an investor  $i$  who buys  $q_i$  bonds in the primary market is defined by:

$$U_i(q_i, n_i, c) = p_{n_i} \cdot u_i^{ns}(q_i) + (1 - p_{n_i}) \mathbb{E}_{k_i} u_i^s(q_i, n_i, k_i) - c \cdot q_i, \quad (2.7)$$

where  $c$  is the price charged by the issuer per bond unit.

In the primary market, the realization of each liquidity shock is still unknown: the investor may become patient with probability  $p_{n_i}$  and keep  $q_i$  bonds until maturity; whereas, with probability  $1 - p_{n_i}$ , the investor may become impatient and have to trade in the secondary market for the payoff  $u_i^s(q_i, n_i, k_i)$ . Since realizations of neighbors' liquidity shocks are also unknown at the issuance, the expectation is taken with respect to the number of investor  $i$ 's patient neighbors,  $k_i$ .

---

<sup>26</sup>The expectation with respect to  $n_j$  is taken under the distribution  $\tilde{f}$  defined above.

Maximizing her payoff at the issuance (2.7), investor  $i$  chooses her primary market demand  $q_i = q_i(n_i, c)$ , which depends on her connectedness and submits it to the issuer.

### Issuer's payoff in the primary market

Finally, I define the issuer's payoff in the primary market. The issuer first aggregates all individual investors' demands to determine the total demand for the bonds. If the issuer does not discriminate allocations by the number of investor connections the total demand for the bonds is equal to:

$$Q(c) = \sum_{n \in N} q(n, c) \cdot f_n, \quad (2.8)$$

which is the total quantity demanded by investors with different numbers of dealers weighted by their masses in the population. In this case, the profit of the issuer, who only sets a uniform price  $c = \delta + \pi$  for all investors, is

$$V = c \cdot Q(c) - \delta \cdot Q(c) = \pi \cdot Q(c), \quad (2.9)$$

which is the premium  $\pi$  charged by the issuer times the total quantity of bonds sold to investors.

If, however, the issuer discriminates based on the number of investor connections, she can restrict bond allocations to some investors in the primary market. In this case, I assume that the issuer chooses *fill rates*  $0 \leq \alpha_n \leq 1$  for all levels of investor connectedness  $n$ . The fill rate  $\alpha_n$  represents the fraction of investors with  $n$  connections that obtain their desired demand  $q(n, c)$  in the primary market. For instance, if  $\alpha_k = 0$  for some  $k$ , investors with  $k$  dealers are excluded from the primary market. Similarly, if  $\alpha_n = 1$  for all  $n$ , the issuer does not discriminate allocations based on investor connectedness and all investors receive their desired demands.

Importantly, I assume that the fill rates are known to investors before they submit their demands and the issuer is committed to the chosen allocation policy. Since investors'



expectations about future liquidity depend on the fill rates, which determine who is expected to trade in the secondary market, this assumption guarantees that the issuer’s policy agrees with investors’ beliefs. In practice, issuers’ past allocation policies are known to investors and can be used to form expectations about future allocations of new issues while the commitment can be supported by reputation concerns of issuers and their underwriters.

When the issuer discriminates allocations by the number of investor connections the total demand for bonds adjusts to:

$$Q_d(c, \alpha) = \sum_{n \in N} \alpha_n \cdot q(n, c) \cdot f_n, \quad (2.10)$$

where  $\alpha = (\alpha_n, \dots, \alpha_{\bar{n}})$ , while the issuer’s profit becomes:

$$V_d(\alpha) = \pi \cdot Q_d(c, \alpha). \quad (2.11)$$

### 2.3. Equilibrium Analysis

Before introducing the details of the equilibrium analysis, I present an example of a trading network, in which exclusion of less connected investors from the primary market is optimal for the issuer.

**Example.** I determine the issuer’s optimal actions for a parameterized trading network with two levels of investor connectedness. The parameters of the network are chosen so that the average number of investor’s dealers, and the connectedness of the most and least connected investors closely match those values observed for a broad sample of US insurance companies. Specifically, I assume that a less connected investor has  $n_1 = 1$  dealers, while a more connected investor has  $n_2 = 25$  dealers. The masses of the two investor groups are  $f_1 = 0.75$  and  $f_2 = 0.25$ , respectively. Thus, an average investor trades with 7 dealers in the network, which closely matches the reported average number of dealers of a US insurance

	$n$	$f_n$	$p_n$
Less Connected	1	0.75	0.25
More Connected	25	0.25	0.95

**Table 2.1: Example of a network with optimal exclusion.** The parameters, characterizing a trading network and liquidity needs of investors, under which exclusion of less connected investors is optimal.

company.<sup>27</sup> I also assume that a less connected investor remains patient with probability  $p_1 = 0.25$ , while a more connected investor remains patient with probability  $p_2 = 0.95$ .

Table 2.1 provides the summary of the parameters.

Under these parameters, the exclusion of less connected investors from the primary market is optimal for the issuer, i.e., she optimally sets the allocation policy  $\alpha = (\alpha_1, \alpha_2) = (0, 1)$ . The optimal premium is  $\pi = \frac{1-\delta}{2}$ , which leads to the issuance price of  $c = 0.5$ , assuming  $\delta = 0$ . The exclusion of less connected investors is profitable because it significantly improves expectations of more connected investors about liquidity in the secondary market. As a consequence, more connected are willing to buy more bonds in the primary market,  $q_d(n_2, c) > q(n_2, c)$ , and their increased demand makes up for the forgone sales to less connected,  $Q_d(c, \alpha) > Q(c)$ , i.e., the total demand increases. Ultimately, restricting allocations to only more connected investors raises the total primary market demand at the issuance price by 0.02 percent and, therefore, allows the issuer to increase her profits.

Importantly, less connected investors are still willing to buy bonds at the issuance price  $c = 0.5$  or even at a higher price because the demands of all investors in the primary market are downward-sloping. In particular, the demand of less connected at the issuance price is  $q_d(n_1, c) = 0.03$  and these investors are ready to buy, for instance,  $q_d(n_1, 0.8) = 0.01$  at a price of 0.8. However, the issuer cannot sell any bonds to less connected investors since she is committed to the chosen allocation policy, which is crucial to generating a higher demand from the non-excluded more connected investors. The following provides the intuition for the improvement in expected liquidity in the secondary market, which raises the demands of more connected investors at issuance, while the details of the feedback to the primary

<sup>27</sup>See Table 1 in Hendershott et al. (2017).

market are covered in the general analysis.

The reason why expectations of more connected investors improve significantly is that less connected investors have a relatively high trading impact on the secondary market liquidity. There are two factors that generate the higher impact. First, there is a difference between the ways less and more connected investors trade in the secondary market. If a less connected investor is hit by a liquidity shock, she has to liquidate all her holdings through her sole dealer, even if the investor knows that the dealer is connected to an impatient neighbor. In this case, it is very costly for both the investor herself and her impatient neighbor. In contrast, if a more connected investor is hit by a liquidity shock, she moves her trades to the dealers who are connected to patient neighbors. This is not only beneficial for the investor herself but also for all her impatient neighbors. Consequently, trading by more connected investors has a lower impact on the expected liquidity available to other investors in the secondary market.

The second factor, which amplifies the effect of the first, is that, from any investor's perspective, her impatient neighbor is more likely to be less connected, even though, as a group, less connected investors have many fewer links. Indeed, although less connected investors are more numerous,  $\frac{f_1}{f_2} = 3$ , through any given dealer, an investor is less likely to be connected to a less connected investor because more connected investors trade with many more dealers in total. Specifically, the ratio of the number of links leading to less connected investors to the number of links leading to more connected investors is  $\frac{\tilde{f}_1}{\tilde{f}_2} = \frac{n_1 f_1}{n_2 f_2} = 0.12$ . However, after liquidity shocks are realized, it is more likely that an impatient neighbor is a less connected investor: the ratio of the number of links leading to less connected impatient investors to the number of links leading to more connected impatient investors is  $\frac{(1-p_1)\tilde{f}_1}{(1-p_2)\tilde{f}_2} = \frac{(1-p_1)n_1 f_1}{(1-p_2)n_2 f_2} = 1.8$ . Consequently, the way less connected impatient investors trade in the secondary market impacts its expected liquidity even more. The combination of the two factors results in an impact on the secondary market liquidity, the magnitude of which is high enough, relative to that of more connected investors, to make exclusion of

less connected investors from the primary market profitable.

I next turn to the general analysis. To determine the issuer's optimal decision on the fill rates and, therefore, to find out if she chooses to exclude some investors from the primary market, it is necessary first to derive their equilibrium demand functions at the issuance. Since primary market demands depend on the secondary market liquidity available to investors, I start the analysis by solving for investors' optimal actions and payoffs in the secondary market, which in turn govern its liquidity. To derive closed-form expressions, I set  $\theta = 1$  in this section.

### 2.3.1. Investor's problem in the secondary market

This subsection solves for the optimal actions of an impatient investor in the secondary market, which depend on liquidity demanded by her neighbors from common dealers and, thus, neighbors' bond holdings. The next subsection will show then that there is a unique equilibrium in the primary market that determines these holdings.

In the secondary market, an impatient investor  $i$  optimally trades with her dealers to maximize the payoff (2.6). Before solving the optimization problem, it is instructive to rearrange the trade profit in three steps. First, the payoff can be broken into two parts:

$$u_i^s(q_i, n_i, k_i) = \sum_{j \in N_i^{ns}} P_{ij}(q_i^{ij})q_i^{ij} + \sum_{j \in N_i^s} \left[ \tilde{\mathbb{E}}_{n_j} \mathbb{E}_{k_j} P_{ij}(q_i^{ij} + q_j^{ij})q_i^{ij} \right], \quad (2.12)$$

where  $N_i^{ns}$  and  $N_i^s$  are the sets of investor  $i$ 's patient and impatient neighbors.

Second, it is convenient to introduce illiquidity discounts  $d_{ij}(q) = 1 - P_{ij}(q)$  that investor  $i$  obtains by trading with her dealers. A discount  $d_{ij}(q)$  increases in liquidity  $q$  demanded from the dealer  $ij$  by its client investors  $i$  and  $j$ . The trade profit can then be rewritten as:

$$u_i^s(q_i, n_i, k_i) = q_i - \sum_{j \in N_i^{ns}} d_{ij}(q_i^{ij})q_i^{ij} - \sum_{j \in N_i^s} \left[ \tilde{\mathbb{E}}_{n_j} \mathbb{E}_{k_j} d_{ij}(q_i^{ij} + q_j^{ij})q_i^{ij} \right], \quad (2.13)$$

where the first term emerges because of the constraint  $\sum_{j \in N_i} q_i^{ij} = q_i$  that impatient investor  $i$  sells all her bonds in the secondary market. Hence, the payoff (2.13) of an impatient investor from liquidating her bonds is equal to their payment at maturity  $q_i$  minus the sum of discounts incurred from trading with the investor's dealers. The discounts obtained from the dealers who are connected to patient neighbors increase only because of investor's own trades while the discounts obtained from the dealers who are connected to impatient neighbors also grow due to neighbors' trades. Specifically, high demand for liquidity by impatient neighbors from the common dealers reduces liquidity available to investor  $i$ .

Third, since the functional form of discounts in the baseline model is linear, the expectation in the third term of (2.13) can be brought in the argument of the discount function:

$$u_i^s(q_i, n_i, k_i) = q_i - \sum_{j \in N_i^{ns}} d_{ij}(q_i^{ij})q_i^{ij} - \sum_{j \in N_i^s} d_{ij}(q_i^{ij} + \tilde{\mathbb{E}}_{n_j} \mathbb{E}_{k_j} q_j^{ij})q_i^{ij}. \quad (2.14)$$

Thus, the expected neighbors' demand for liquidity from the common dealers is summarized by the term  $\tilde{\mathbb{E}}_{n_j} \mathbb{E}_{k_j} q_j^{ij}$ , which depends on investor  $i$ 's expectation about her neighbors' connectedness and the number of their impatient neighbors.

When an investor  $i$  considers the number of impatient neighbors of an investor  $j$ ,  $k_j$ , she knows only that her neighbor  $j$  has at least one impatient neighbor — investor  $i$  herself. Consequently, the number of investor  $j$ 's patient neighbors must be no greater than  $n_j - 1$ , i.e.,  $k_j \leq n_j - 1$ . Beyond this knowledge, investor  $i$  does not have any other information about the neighbors of investor  $j$ .

The expected neighbors' demand for liquidity, therefore, is

$$\bar{q}^s \equiv \tilde{\mathbb{E}}_{n_j} \mathbb{E}_{k_j} q_j^{ij} = \tilde{\mathbb{E}}_{n_j} [\mathbb{E}_{k_j} [q_j^{ij} | k_j \leq n_j - 1]], \quad (2.15)$$

which is the average quantity sold by an impatient investor, who has at least one impatient neighbor, on a link with an impatient neighbor. This quantity can be viewed as a measure

of the secondary market illiquidity for connections with dealers who trade with both of their clients — I call such dealers *stressed*. The higher  $\bar{q}^s$ , the larger is the discount that is incurred on the links with stressed dealers. Then, the total trade discount for an investor  $i$  is

$$d_i(q_i, n_i, k_i, \bar{q}^s) \equiv \sum_{j \in N_i^{n_i}} d_{ij}(q_i^{ij})q_i^{ij} + \sum_{j \in N_i^s} d_{ij}(q_i^{ij} + \bar{q}^s)q_i^{ij}, \quad (2.16)$$

which shows explicitly that an impatient investor faces differential liquidity on links with patient and impatient neighbors, with expected liquidity being “worse” on the latter by  $\bar{q}^s$ .

Finally, combining the above derivations, the investor’s profit maximization problem in the secondary market reduces to the minimization of the total discount. An investor solves the problem given the expected level of the secondary market illiquidity  $\bar{q}^s$  on the links with stressed dealers. This level is then determined in equilibrium by how her impatient neighbors trade with the common dealers by solving the same problem. Formally, the problem is

$$\min_{\{q_i^{ij}\}_{j \in N_i}} d_i(q_i, n_i, k_i, \bar{q}^s) \quad s.t. \quad \sum_{j \in N_i} q_i^{ij} = q_i. \quad (2.17)$$

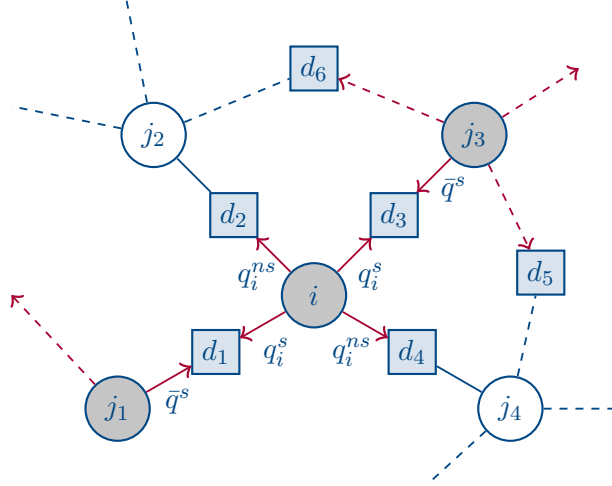
In particular, an investor  $i$  minimizes the total discount obtained from the secondary market by splitting the sale of her total bond holdings  $q_i$  between her dealers. Figure 2.5 illustrates the trading activity in the secondary market from the standpoint of an investor  $i$  and the optimal trading quantities are given by the following.

**Proposition 2.1.** *For given bond holdings  $q_i$  and expected liquidity demanded by impatient neighbors  $\bar{q}^s$ , an impatient investor  $i$  with  $n_i$  dealers and  $k_i$  patient neighbors trades a larger quantity of bonds on the links with patient neighbors than on the links with impatient neighbors. The latter quantity decreases with  $k_i$ .*

*Proof.* The solution to the impatient investor’s problem is presented in the Appendix 2A.<sup>28</sup>

---

<sup>28</sup>Other omitted proofs and derivations are presented in the Appendix 2A.



**Figure 2.5: Trading of impatient investor and neighbors.** An impatient investor  $i$  trading with her dealers in the secondary market. Solid arrows represent trading activity for the optimal quantities traded by the investor  $i$  and expected quantities traded by her neighbors (given under the arrows). Transparent and shaded circles represent patient and impatient investors, respectively, and squares represent dealers.

From the solution, the quantity traded on all the links with patient neighbors is

$$q_i^{ns}(q_i, n_i, k_i, \bar{q}^s) = \frac{q_i}{n_i} + \frac{(n_i - k_i) \bar{q}^s}{n_i} \frac{1}{2}, \quad (2.18)$$

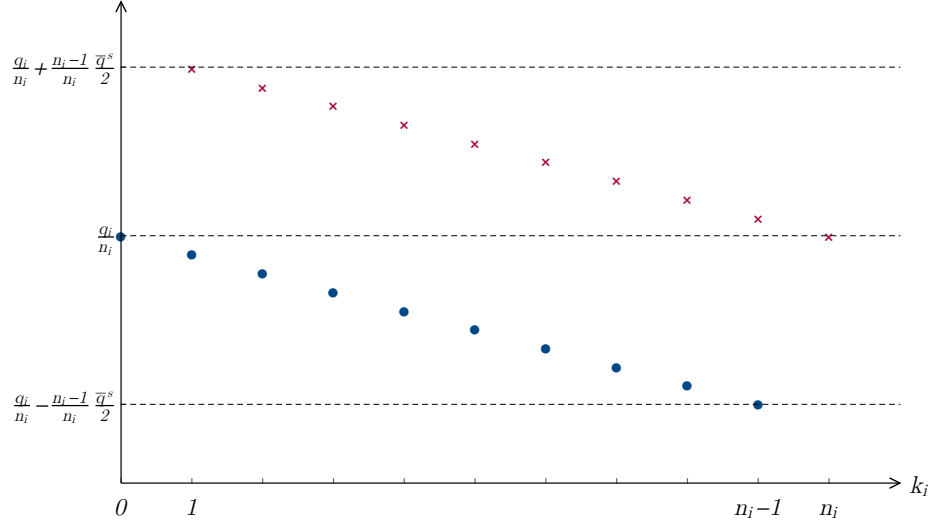
while the quantity sold on all the links with impatient neighbors is

$$q_i^s(q_i, n_i, k_i, \bar{q}^s) = \frac{q_i}{n_i} - \frac{k_i \bar{q}^s}{n_i} \frac{1}{2}. \quad (2.19)$$

Clearly,  $q_i^{ns} > q_i^s$  as the price impact on the links with impatient neighbors is higher due to their demands for liquidity. This induces an impatient investor to tilt her trades away from stressed dealers. Figure 2.6 plots the optimal quantities as functions of the number of patient neighbors  $k_i$ . □

Before studying expected liquidity demanded by impatient neighbors it is natural to make the following.

**Assumption 2.2.** *Investors learn which of their neighbors, if any, have not received allo-*



**Figure 2.6: Optimal trading of an impatient investor.** The optimal trading quantities on the links with patient neighbors  $q_i^{ns}$  (crosses) and on the links with impatient neighbors  $q_i^s$  (dots) as functions of the number of patient neighbors  $k_i$ .

*cation, at the time of a realization of liquidity shocks.*

Since excluded investors do not demand liquidity from their dealers in the secondary market, from the impatient investors' perspective, they are equivalent to patient investors. This implies that, in the primary market, the probability of an investor's neighbor being patient equals to:

$$p(\alpha) = \tilde{\mathbb{E}}p_n\alpha_n + \tilde{\mathbb{E}}(1 - \alpha_n). \quad (2.20)$$

In this equation, the first term covers all non-excluded neighbors.<sup>29</sup> Specifically, the fraction  $\alpha_n$  of neighbors with  $n$  dealers receives allocations in the primary market and has a chance to stay patient with probability  $p_n$ . The second term deals with the excluded fraction  $1 - \alpha_n$  of neighbors, who are patient with probability 1, according to Assumption 2.2. Therefore, if some investors are excluded from the primary market, the probability of an investor's neighbor being patient becomes larger. Correspondingly, the probability of an investor's

<sup>29</sup>The expectation is taken with respect to  $n$  under the distribution  $\tilde{f}$ . In the following, the subscript  $n$  is omitted to reduce notations.



neighbor being impatient:

$$1 - p(\alpha) = \tilde{\mathbb{E}}(1 - p_n)\alpha_n \quad (2.21)$$

becomes smaller. Note that, since an investor does not have any information about her neighbors and their neighbors, in the secondary market, the probability of a neighbor's neighbor being patient is also given by  $p(\alpha)$ .

Having established the optimal trading quantities of impatient investors as a function of their connectedness, I next study how these trades, through neighbors' activity, impact the expected liquidity provided by stressed dealers. The following proposition shows that impatient investors with a higher number of dealers contribute less to the equilibrium level of  $\bar{q}^s$ , which proxies for the degree of illiquidity in the secondary market.

**Proposition 2.2.** *If an impatient neighbor  $j$  needs to trade the same quantity of bonds per dealer in the secondary market for different levels of her connectedness, the expected liquidity demanded by this neighbor from the dealers connected to impatient investors decreases with the number of neighbor's dealers  $n_j$ .*

*Proof.* From standpoint of an impatient investor  $i$ , consider the expected quantity of bonds traded by her impatient neighbor  $j$  through their common dealer as a function of the neighbor's connectedness  $n_j$ . As argued above, it is equal to the conditional expectation, with respect to  $k_j$ , of the neighbor's optimal quantity  $q_j^s(q_j, n_j, k_j, \bar{q}^s)$  traded on links with stressed dealers:

$$\mathbb{E}_{k_j}[q_j^s(q_j, n_j, k_j, \bar{q}^s)|k_j \leq n_j - 1] = \frac{q_j}{n_j} - \frac{p(\alpha)(n_j - 1)\bar{q}^s}{n_j} \frac{1}{2} = \frac{q_j}{n_j} - g_n \frac{p(\alpha)}{2}\bar{q}^s, \quad (2.22)$$

where  $g_n \equiv \frac{n-1}{n}$  and  $p(\alpha)$  is the probability of a neighbor  $j$ 's neighbor being patient.

Next, if investors need to liquidate the same number of bonds per dealer in the secondary market, the fraction  $\frac{q_j}{n_j}$  is the same for all  $n_j$  in the previous equation. Thus, since  $g_n$  is an increasing function of  $n$ , the quantity given by the equation decreases in  $n_j$ . Specifically,

investors with a higher number of dealer connections demand less liquidity from stressed dealers.  $\square$

Intuitively, an impatient investor expects that her more connected impatient neighbor will be able to tilt away her trades from their common dealer to a greater extent compared to her less connected impatient neighbor. At one extreme, an impatient neighbor with only one dealer is expected to sell everything through the common dealer — no matter whether the dealer is stressed or not. In contrast, at the other extreme, an impatient neighbor with many connections has greater flexibility and predictably tilts trades away from the common stressed dealer. The greater ability of more connected investors to move their trades away from stressed dealers is reflected in the increasing function  $g_n$ .

The previous equation (2.22) is further used to determine the equilibrium level of liquidity demanded by impatient neighbors  $\bar{q}^s$ . It is equal to the expected quantity of bonds traded by neighbors with different levels of connectedness  $n$  conditional on being impatient. Importantly, the pool of impatient investors is adjusted on the fill rates chosen by the issuer, considering that only non-excluded investors can become impatient:

$$\bar{q}^s = \frac{\tilde{\mathbb{E}} \left( \frac{q_n}{n} - g_n \frac{p(\alpha)}{2} \bar{q}^s \right) (1 - p_n) \alpha_n}{\tilde{\mathbb{E}}(1 - p_n) \alpha_n}. \quad (2.23)$$

Solving for  $\bar{q}^s$  yields:

$$\bar{q}^s = \frac{\frac{\tilde{\mathbb{E}} \frac{q_n}{n} (1 - p_n) \alpha_n}{\tilde{\mathbb{E}}(1 - p_n) \alpha_n}}{1 + \frac{p(\alpha)}{2} \frac{\tilde{\mathbb{E}} g_n (1 - p_n) \alpha_n}{\tilde{\mathbb{E}}(1 - p_n) \alpha_n}}. \quad (2.24)$$

In equilibrium, the quantity  $\bar{q}^s$ , which proxies for illiquidity on the links with impatient neighbors, is positively affected through the nominator by the weighted average of investors' per-dealer bond holdings obtained in the primary market. Critically, it also depends negatively on the weighted average “connectedness” of impatient bondholders, manifested in  $g_n$ , through the denominator. When the average connectedness of impatient investors is high, investors expect their impatient neighbors to demand less liquidity from common dealers.

To complete the description of the impatient investor's problem in the secondary market, I derive the expected total discount from trade by plugging the optimal trading quantities  $q_i^{ns}$  and  $q_i^s$  back into equation (2.17):

$$d_i(q_i, n_i, k_i, \bar{q}^s) = \lambda \left( \frac{q_i^2}{n_i} + \frac{(n_i - k_i)q_i\bar{q}^s}{n_i} - \frac{k_i(n_i - k_i)(\bar{q}^s/2)^2}{n_i} \right). \quad (2.25)$$

The illiquidity discount increases with the price impact parameter  $\lambda$  and with the total bond holdings  $q_i$  that an investor  $i$  has to liquidate in the secondary market. Importantly, the expected discount decreases with the number of her patient neighbors  $k_i$  and increases with expected liquidity demanded by neighbors  $\bar{q}^s$ .

Having established the optimal payoff of an impatient investor in the secondary market as a function of her bond holdings obtained in the primary market, I step one period back and solve for investors' equilibrium demands at issuance.

### 2.3.2. Investor's problem in the primary market

This subsection derives investors' primary market demands and analyzes how bonds purchased at issuance affect expected liquidity demanded by neighbors in the equilibrium.

**Proposition 2.3.** *For any price  $\delta \leq c \leq 1$  and fill rates  $\alpha$  chosen by the issuer, there exists a unique Bayes-Nash equilibrium in the primary market demands  $q(n, c)$ .*

*Proof.* The following derives the optimal demands while the proof of existence and uniqueness is relegated to the Appendix 2A. In the primary market, the number of patient neighbors  $k_i$  of an investor  $i$  is unknown. Therefore, taking the expectation of the optimal discount (2.25) with respect to this number and considering that some neighbors might be excluded, the investor  $i$ 's expected payoff from trade in the secondary market is

$$\mathbb{E}_{k_i} u_i^s(q_i, n_i, k_i) = q_i - \lambda \left( \frac{(q_i)^2}{n_i} + (1 - p(\alpha))q_i\bar{q}^s - p(\alpha)(1 - p(\alpha))\frac{(\bar{q}^s)^2}{4}(n_i - 1) \right). \quad (2.26)$$

The payoff is affected by the number of investor's dealers  $n_i$  through the second and the last term and increases with  $n_i$ . The last term emerges due to a better ability of a more connected investor to tilt her trades away from dealers who are connected to impatient neighbors. It increases when the illiquidity of the secondary market, proxied by  $\bar{q}^s$ , is high. However, this term does not depend on the total bond holdings  $q_i$  obtained in the primary market and, thus, the benefit of better dealer diversification does not affect the optimal demand at the issuance.

Combining the expected payoff from trade in the secondary market (2.26) with the equation for the total expected payoff in the primary market (2.7), an investor  $i$  solves the following problem in the period  $t = 0$ :

$$U_i(q_i, n_i, c) = \max_{q_i} (1 - c)q_i - \lambda(1 - p_{n_i}) \left( \frac{(q_i)^2}{n_i} + (1 - p(\alpha))q_i\bar{q}^s \right). \quad (2.27)$$

The solution to this problem determines the investor  $i$ 's optimal demand for the bonds at the issuance and is given by:

$$q_i(n_i, c) = n_i \left( \frac{(1 - c)}{2\lambda(1 - p_{n_i})} - \frac{(1 - p(\alpha))\bar{q}^s}{2} \right). \quad (2.28)$$

□

Notably, the primary demand scales with the number of investor's dealers. The per-dealer optimal quantity decreases with the price  $c$  charged by the issuer. It is larger when the investor is more likely to remain patient, i.e., when  $p_{n_i}$  is higher. More importantly, it is negatively affected by the illiquidity of the secondary market through the price impact parameter  $\lambda$  and through the expected quantity of bonds traded by neighbors with common dealers  $(1 - p(\alpha))\bar{q}^s$ . The latter is composed of two factors: i) the probability that an investor's neighbor becomes impatient  $1 - p(\alpha)$ ; and ii) the expected liquidity demanded by impatient neighbors  $\bar{q}^s$ .

What is the impact of the investors who are more likely to remain patient, i.e., who have higher  $p_n$ , on the expected quantity of bonds traded by neighbors  $(1 - p(\alpha))\bar{q}^s$ ? There are two opposing effects acting through the above two factors. On the one hand, when they are more likely to remain patient, the per-dealer quantity of bonds demanded by the investors is high. Therefore, if the investors become impatient, the amount of bonds that needs to be liquidated in the secondary market is high, which leads to the larger  $\bar{q}^s$ . On the other hand, the probability of investors becoming impatient and having to trade in the secondary market is low, which leads to the smaller  $1 - p(\alpha)$ .

To see how these two effects interact, I derive the equilibrium values for  $\bar{q}^s$  and  $(1 - p(\alpha))\bar{q}^s$ . Combining the investors' demand in the primary market (2.28) with the equation for  $\bar{q}^s$  (2.24), the expected liquidity demanded by impatient neighbors is

$$\bar{q}^s = \frac{1 - c}{2\lambda} \frac{\frac{\tilde{\mathbb{E}}\alpha_n}{\tilde{\mathbb{E}}(1-p_n)\alpha_n}}{1 + \frac{p(\alpha)}{2} \frac{\tilde{\mathbb{E}}g_n(1-p_n)\alpha_n}{\tilde{\mathbb{E}}(1-p_n)\alpha_n} + \frac{(1-p(\alpha))}{2}}. \quad (2.29)$$

Consequently, the equilibrium level of expected liquidity demanded by neighbors is

$$(1 - p(\alpha))\bar{q}^s = \frac{1 - c}{2\lambda} \frac{\tilde{\mathbb{E}}\alpha_n}{1 + \frac{p(\alpha)}{2} \frac{\tilde{\mathbb{E}}g_n(1-p_n)\alpha_n}{\tilde{\mathbb{E}}(1-p_n)\alpha_n} + \frac{(1-p(\alpha))}{2}}. \quad (2.30)$$

It can be seen that per-dealer holdings of the investors who are more likely to remain patient indeed lead to the larger  $\bar{q}^s$  through its numerator. However, this effect is mitigated in  $(1 - p(\alpha))\bar{q}^s$  since its numerator does not depend on  $p_n$ . Even though relatively more patient investors buy more bonds in the primary market, their contribution to the equilibrium level of liquidity demanded by neighbors has the similar magnitude as that of less patient investors. In the special case, when  $p_n = p$  for all  $n$ , all investors demand the same quantity of bonds per dealer in the primary market, but the numerator of the equation for  $(1 - p(\alpha))\bar{q}^s$  is exactly the same.

The relative patience of investors appears only in the denominator of  $(1 - p(\alpha))\bar{q}^s$  through

the probability of an investor's neighbor being patient  $p(\alpha)$  and the average connectedness of impatient investors, reflected in  $g_n$ . In particular, the connectedness of relatively more impatient investors has a higher weight in the average. This effect plays a key role for the result of the next subsection.

To complete the description of the investor's problem in the primary market, the individual per-dealer demand as a function of the primitives is

$$q(n, c) = n \frac{1 - c}{2\lambda} \left( \frac{1}{1 - p_n} - \frac{\frac{1}{2} \tilde{\mathbb{E}} \alpha_n}{1 + \frac{p(\alpha)}{2} \frac{\tilde{\mathbb{E}} g_n (1 - p_n) \alpha_n}{\tilde{\mathbb{E}} (1 - p_n) \alpha_n} + \frac{(1 - p(\alpha))}{2}} \right). \quad (2.31)$$

### 2.3.3. Issuer's problem in the primary market

Having established individual equilibrium demands for the bonds in the primary market, as a function of the issuance price and investor connectedness in the secondary market, it is now possible to solve the issuer's problem.

## Homogenous investor connections

Before analyzing the general case of investors with heterogeneous connections to dealers it is instructive to first consider the case of homogenous investor connections. It illustrates that, in the equilibrium, more connected investors impose less stress on the secondary market liquidity by moving their trades away from stressed dealers and highlights how that feeds back to their primary market demands. Specifically, suppose all investors have the same number of dealers  $n$  and let  $p_n = p$ . Then, the following holds.

**Proposition 2.4.** *Holding the number of dealers in the secondary market fixed, the issuance premium  $\pi$  from a sale of a fixed amount of bonds  $\bar{Q}$  to a homogenous investor network increases with the number of investors' dealers  $n$ . The increase is higher when dealers' inventory costs  $\lambda$  are larger.*

*Proof.* In the case of a homogenous investor network, the issuer does not find it optimal to exclude any investors from the primary market and the individual demand at the issuance is simplified to:

$$q_n(c) = n \frac{(1-c)}{\lambda} \frac{1 + \frac{p}{2}g_n}{2(1-p)(1 + \frac{p}{2}g_n) + (1-p)^2}. \quad (2.32)$$

Thus, the total demand in the primary market is

$$Q_n(c) = q_n(c)f_n = \frac{(1-c)}{\lambda} \beta(n)nf_n, \quad (2.33)$$

where

$$\beta(n) \equiv \frac{1 + \frac{p}{2}g_n}{2(1-p)(1 + \frac{p}{2}g_n) + (1-p)^2} \quad (2.34)$$

is an increasing function of  $n$  since  $g_n$  is increasing.

Therefore, when the issuer offers  $\bar{Q}$  bonds<sup>30</sup> in the primary market the offering premium is

$$\pi_n = 1 - \delta - \bar{Q} \frac{\lambda}{\beta(n)nf_n}. \quad (2.35)$$

This equation illustrates explicitly how investor connectedness affects the issuance and the primary market price. In particular, consider the primary market outcomes in the two scenarios of investor connectedness in the secondary market: *i*) where all investors have  $n_l$  dealers and *ii*) where all investors have a higher number of dealers  $n_h > n_l$ .

To make a fair comparison of the outcomes in the two scenarios, it is important to condition on the total amount of liquidity available from dealers in the secondary market, because it feeds back to the primary market demands. Thus, it is necessary that the total number of dealers in the secondary markets of the two scenarios be the same. Since every dealer has two client investors, this number is equal to  $nf_n/2$  — the half of the total number of connections in the population. Consequently, assume that  $n_l f_{n_l}/2 = n_h f_{n_h}/2$ , which requires the mass of investors in the second scenario to be smaller,  $f_{n_h} < f_{n_l}$ , because they

---

<sup>30</sup>Assuming the total quantity  $\bar{Q}$  is low enough for  $\pi$  to be positive.

have more connections,  $n_h > n_l$ .

Under this assumption, the total demands in the primary market (2.33) in the two scenarios differ only in the term  $\beta(n)$ . Since the function  $\beta(n)$  increases in  $n$ , i.e.,  $\beta(n_h) > \beta(n_l)$ , the total demand for bonds is higher in the second scenario, when all investors are more connected. This allows the issuer to sell the same quantity of bonds at a higher price. Moreover, the effect is stronger when the price impact parameter  $\lambda$  is higher, i.e., when dealers' inventory costs are higher.  $\square$

Even though the total number of dealers in the secondary market is the same, the greater ability of investors, when they are all more connected, to move their trades away from stressed dealers improves its liquidity. The improvement in the liquidity of the secondary market feeds back to the primary market, increasing investors' demands. This is beneficial for the issuer because she can sell the same quantity of bonds at a higher price or, equivalently, more bonds at a given price.

### Heterogeneous investor connections

I now analyze the general case. In the primary market, the issuer aggregates all investor demands (2.31) into the total demand and solves the following maximization problem:

$$V_d(\alpha) = \pi \cdot Q_d(c, \alpha) = \max_{\pi, \{\alpha_n\}} \pi \sum_{n \in N} \alpha_n \cdot q(n, c) \cdot f_n \quad s.t. \quad q(n, c) = \arg \max U(q, n, c). \quad (2.36)$$

By choosing different order fill rates  $\alpha_n$  for investors with different levels of connectedness  $n$ , the issuer is able to alter secondary market liquidity through the expected quantity of bonds traded by neighbors  $(1 - p(\alpha))\bar{q}^s$ . This choice feeds back into the primary market demands of participating investors and, in turn, affects the issuer's profits.

When investors are heterogeneous with respect to the number of dealers that they can trade with in the secondary market, the following holds.



**Proposition 2.5.** *For a heterogeneous investor network, if the probability  $p_n$  that an investor with  $n$  dealers remains patient strictly increases with  $n$ , the issuer's profit is maximized when*

$$\begin{aligned}\alpha_n &= 0 & \text{if } n < n', \\ \alpha_n &= 1 & \text{if } n \geq n'\end{aligned}$$

for some  $\underline{n} \leq n' < \bar{n}$  and  $\pi = \frac{1-\delta}{2}$ .

The proposition asserts that the issuer might find it optimal to exclude less connected investors from the primary market. This result is due to the fact that bonds allocated to less connected investors and subsequently traded by them have relatively high impact on liquidity provided by dealers in the secondary market, compared to bonds allocated to more connected investors.

Specifically, as it can be seen from the equation for the issuer's profit (2.36), reducing  $\alpha_n$  for a given  $n$  involves a trade-off. On the one hand, since the orders from investors with  $n$  dealers are not fully filled, the total demand for the bonds in the primary market decreases. This lowers the number of bonds sold  $Q_d$  and the issuer's profit. On the other hand, expectations of non-excluded investors about liquidity available from dealers in the secondary market improve, through  $(1-p(\alpha))\bar{q}^s$ . This raises their primary market demands and, as a result, the issued quantity of bonds  $Q_d$  as well as the issuer's profit increase. For less connected investors, i.e., low  $n$ , the second effect can dominate since their trading has a relatively high impact on the secondary market liquidity. When this is the case, the issuer's payoff increases overall as a consequence of their exclusion from the primary market.

To see why the result requires less connected investors to be relatively more impatient, consider the total demand at the issuance:

$$Q_d(c, \alpha) = \frac{1-c}{2\lambda} \left( \sum_{n \in N} \left( \frac{\alpha_n n f_n}{1-p_n} \right) - \frac{\frac{1}{2}(\tilde{\mathbb{E}}\alpha_n) \sum_{n \in N} \alpha_n n f_n}{1 + \frac{p(\alpha)}{2} \frac{\tilde{\mathbb{E}}g_n(1-p_n)\alpha_n}{\tilde{\mathbb{E}}(1-p_n)\alpha_n} + \frac{(1-p(\alpha))}{2}} \right). \quad (2.37)$$

The first term in the brackets is positive and due to the purchases of investors with different

number of dealers  $n$ . When investors are more impatient they have smaller primary market demand and contribute less to this term. The second term is the sum of negative effects due to the expected quantity of bonds traded by neighbors  $(1 - p(\alpha))\bar{q}^s$ . As it is shown in the previous subsection, more impatient investors primarily affect  $(1 - p(\alpha))\bar{q}^s$  through the weighted average connectedness of impatient investors:

$$\frac{\tilde{\mathbb{E}}g_n(1 - p_n)\alpha_n}{\tilde{\mathbb{E}}(1 - p_n)\alpha_n}, \quad (2.38)$$

which gauges how well impatient investors move their trades away from stressed dealers. When less connected investors are relatively more impatient, their weight in the average is higher and, hence, the average itself is smaller, since  $g_n$  is increasing. Therefore, the exclusion of less connected investors from the primary market significantly increases the weighted average connectedness of impatient investors in the secondary market and the second term gets smaller.<sup>31</sup>

Summing up, the exclusion of less connected investors leads to a small decrease in the first (positive) term and a bigger decrease in the second (negative) term, yielding an overall net positive effect. In other words, the increase in the total demand of the remaining, more connected investors makes up for the forgone sales to less connected, which ultimately improves the issuer's revenues.

The relative magnitude of the two effects, caused by moving a fill rate, determines the likelihood of the exclusion.

**Proposition 2.6.** *The less connected investors are more likely to be excluded from the primary market when:*

- i) the difference  $\bar{n} - \underline{n}$  is high, i.e., investors are highly heterogeneous with respect to dealer connections;*
- ii) the difference  $p_{\bar{n}} - p_{\underline{n}}$  is high, i.e., less connected investors are significantly more impatient*

---

<sup>31</sup>Since  $g_n$  is concave the effect is stronger when less connected investors have a low number of dealers in absolute value.

	Exclusion	No Exclusion
$n_1$	1	3
$p_1$	0.05	0.25
$f_1$	0.30	0.81
$n_2$	25	3
$p_2$	0.90	0.80
$f_2$	0.70	0.19

**Table 2.2: Example of comparative statics for the exclusion result.** Exclusion of less connected investors is optimal under parameters in the left column. Changing the value of one of the parameters to the value in the right column while holding other parameters fixed makes the exclusion suboptimal.

*compared to more connected;*

*iii) the distribution  $f_n$  is skewed to the right, i.e., more connected investors have a higher mass.*

**Example.** To illustrate the above, I compute the issuer’s profits for a parameterized trading network with two levels of investor connectedness. In the network, a less connected investor has  $n_1 = 1$  dealers and remains patient with probability  $p_1 = 0.05$ , while a more connected investor has  $n_2 = 25$  dealers and remains patient with probability  $p_2 = 0.90$ . The masses of the two investor groups are  $f_1 = 0.30$  and  $f_2 = 0.70$ , respectively. Under these parameters, the exclusion of less connected investors from the primary market is optimal for the issuer. Changing one of the parameters, while holding others fixed, it can be established when the exclusion becomes suboptimal. It happens either when:  $n_2$  declines to 3;  $n_1$  rises to 3;  $p_2$  declines below 0.80;  $p_1$  rises above 0.25; or  $f_1$  rises above 0.81. Table 2.2 provides the summary of these values.

In addition, the model sheds some light on the observed underpricing of corporate bonds. In practice, underpricing might indicate the attempts of excluded investors to obtain some bonds shortly after the primary market at a price higher than the issuance price. This corresponds to the model in the following way. In the model, the primary market demands of all investors are downward-sloping with respect to the price. Thus, if they could, the excluded investors would still be willing to buy some bonds at the issuance price or even at

a higher price right after the primary market and before the realization of liquidity shocks. Indeed, their demand at this time is given by  $\sum_{\{n \in N: n < n'\}} q(n, c) f_n$ , which is positive even if the price  $c$  is close to 1 — the bond’s payoff at maturity. Note that the issuer cannot sell any bonds to these investors since she is committed to the chosen allocation policy, which is critical to generating a higher demand from the non-excluded investors.

Proposition 2.5 derives the result in the baseline model with the linear dealer liquidity supply (2.4), when  $\theta = 1$ . However, the effects should be amplified for the general liquidity supply with  $\theta > 1$ .<sup>32</sup> Indeed, when the supply function is concave, investors dislike variation in liquidity demanded by impatient neighbors from stressed dealers and, therefore, exclusion of less connected investors, whose demand for liquidity from these dealers is more variable, should be more likely. Section 2.4 illustrates this intuition for the case of investors with correlated liquidity shocks and homogenous dealer connections.

## 2.4. Correlated Liquidity Shocks

In this section, I explore the case of correlated liquidity shocks. It is shown that the issuer can exploit the knowledge of a correlation structure and might choose to exclude neighbor investors whose liquidity shocks are correlated from the primary market.

I extend the simplified version of the baseline model to allow for correlated liquidity shocks. Specifically, I assume that investor connections are homogenous and each investor has just one dealer and, therefore, one neighbor investor. It is assumed that, for the fraction  $\rho$  of neighbor pairs, the liquidity shocks are perfectly correlated while, for the remaining fraction  $1 - \rho$  of neighbor pairs, the liquidity shocks are perfectly negatively correlated.<sup>33</sup> Crucially, investors know the general correlation structure but do not know whether their neighbor has correlated liquidity shock or not.<sup>34</sup> In contrast, the issuer is informed about which investor neighbors have correlated shocks. Additionally, I use the general form of the

---

<sup>32</sup>A closed form solution of the general model is not available for this specification.

<sup>33</sup>The symmetry requires that  $p = \frac{1}{2}$  in this case.

<sup>34</sup>Similarly to the baseline model, this assumption can be motivated by the opacity of the OTC secondary market.

dealers' liquidity supply function:

$$P_{ij}(q) = 1 - \lambda \cdot q^\theta, \quad (2.39)$$

which adds curvature to its version in the baseline model when  $\theta > 1$ .

In this simplified setup, the investor's problem in the secondary market becomes trivial since each investor has just one dealer. If an investor becomes impatient, which happens with probability  $1 - p$ , she trades all her bond holdings to her single dealer. Moreover, since liquidity shocks are perfectly correlated for the fraction  $\rho$  of investor neighbors and perfectly negatively correlated for the remaining pairs, an impatient investor expects that with probability  $\rho$  her neighbor is also trading with their common dealer in the secondary market. Therefore, an impatient investor's payoff from the secondary market is

$$u_i^s(q_i) = q_i - \lambda \left( (1 - \rho)(q_i)^\theta + \rho(q_i + q^s)^\theta \right) q_i \quad (2.40)$$

where  $q^s$  is a quantity traded by an impatient neighbor through the common dealer.

Thus, combining it with the equation for the total investor's expected payoff in the primary market (2.7), an investor solves the following problem at the issuance:

$$U_i(c) = \max_{q_i} (1 - c)q_i - (1 - p)\lambda \left( (1 - \rho)(q_i)^\theta + \rho(q_i + q^s)^\theta \right) q_i. \quad (2.41)$$

The solution to this problem determines the investor's demand for the bonds at the issuance:

$$q_i(c) = \left( \frac{(1 - c)}{\lambda(1 - p)} \frac{1}{(1 - \rho)(\theta + 1) + \rho(2 + \theta)2^{\theta-1}} \right)^{\frac{1}{\theta}}. \quad (2.42)$$

Accordingly, if the issuer does not discriminate allocations and the whole unit measure of investors participates in the primary market, the total demand for bonds is

$$Q(c) = \left( \frac{(1 - c)}{\lambda(1 - p)} \frac{1}{(1 - \rho)(\theta + 1) + \rho(2 + \theta)2^{\theta-1}} \right)^{\frac{1}{\theta}}. \quad (2.43)$$

The total demand decreases with  $\rho$ , the probability that investor neighbors' liquidity shocks are correlated. Crucially, the decline for small  $\rho$  happens faster when  $\theta$  is larger, when variation in liquidity supplied by dealers in the secondary market is more costly.

Alternatively, suppose the issuer discriminates allocations in the primary market. Specifically, the issuer excludes one of the two neighbor investors whose liquidity shocks are correlated. Since the fraction  $\rho$  of investor neighbors has perfectly correlated liquidity shocks the total mass of excluded investors is  $\frac{\rho}{2}$ .

This exclusion improves investors' expectations about liquidity in the secondary market. Specifically, each investor knows that when she needs to trade with her dealer in the secondary market the investor's neighbor will not trade with their common dealer. Thus, the investor's problem in the primary market can be obtained by letting  $\rho = 0$  in (2.41):

$$U_i(c) = \max_{q_i} (1 - c)q_i - (1 - p)\lambda(q_i)^{\theta+1}. \quad (2.44)$$

Consequently, the investor's demand for the bonds at the issuance is

$$q_i(c) = \left( \frac{(1 - c)}{\lambda(1 - p)} \frac{1}{(\theta + 1)} \right)^{\frac{1}{\theta}} \quad (2.45)$$

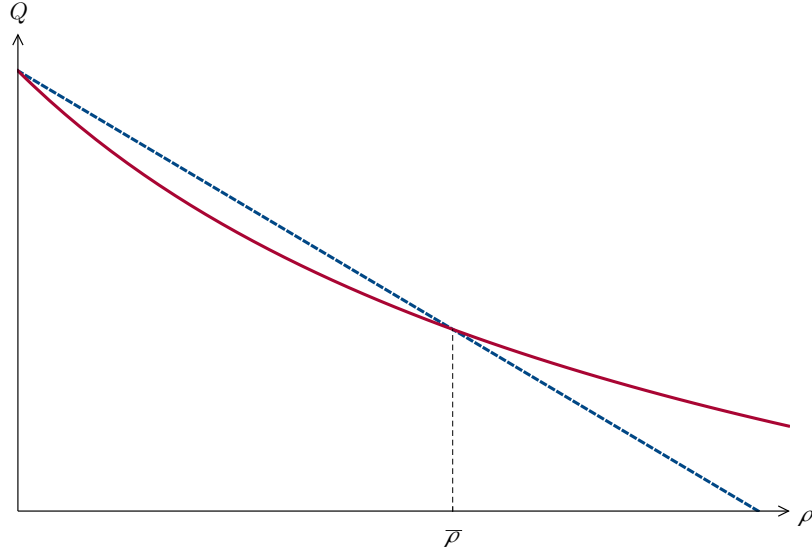
while the total demand in the primary market is

$$Q(c) = \left(1 - \frac{\rho}{2}\right) \left( \frac{(1 - c)}{\lambda(1 - p)} \frac{1}{(\theta + 1)} \right)^{\frac{1}{\theta}}, \quad (2.46)$$

which accounts for the fact that only the mass  $1 - \frac{\rho}{2}$  of investors receive allocations. Figure 2.7 plots the total demands for the two cases.

Comparing the total demands in the two cases: when allocations are discriminated (2.46) and when they are not (2.43), the following can be established.

**Proposition 2.7.** *The issuer is better off excluding investors with correlated liquidity shocks from the primary market when  $\rho \leq \bar{\rho}(\theta)$ , i.e., the mass of excluded investors is sufficiently*



**Figure 2.7: Total demands with correlated liquidity shocks.** The total demands without exclusion (solid) and when investors with correlated liquidity shocks are excluded (dashed). The parameters used are  $c = 1/2$ ,  $p = 1/2$ ,  $\lambda = 1$  and  $\theta = 2$ .

*small. The threshold  $\bar{\rho}(\theta)$  is increasing in  $\theta$ , i.e., the exclusion is more beneficial when variation in liquidity of the secondary market is more important.*

Similarly to the baseline model, the exclusion of investors from the primary market involves a trade-off. On the one hand, since the orders from some investors are not filled, the total demand for the bonds in the primary market decreases. This lowers the number of bonds sold  $Q$  and the issuer's profit. On the other hand, expectations of other investors about liquidity available from dealers in the secondary market improve. This raises their primary market demands and, as a result, the issued quantity of bonds  $Q$  and the issuer's profit increase. For the lower values of  $\rho$ , when the mass of investors with correlated liquidity shocks is small and variation in liquidity is high, the second effect dominates. Therefore, the issuer's payoff increases overall as a consequence of their exclusion from the primary market.

## 2.5. Discussion

### *2.5.1. Implications*

#### **Empirical predictions**

The results of Propositions 2.1 and 2.2 indicate that highly connected investors have a lower impact on liquidity provided by dealers in the secondary market because these investors are able to move their trades to dealers offering better prices. In practice, this implies that highly connected investors should trade at lower bid-ask spreads on average (Di Maggio et al., 2017b; Hendershott et al., 2017; Kondor and Pinter, 2019). Similarly, bonds that are traded in densely connected trading networks should have lower bid-ask spreads on average.

The exclusion results suggest that issuers might choose to discriminate allocations in the primary market in order to improve liquidity of the secondary market for the participating investors. In practice, issuers and their underwriters take into consideration future secondary market liquidity when they allocate new bonds to investors. In particular, secondary market price volatility is one of the factors mentioned in the underwriters' allocation policies, which are now mandatorily disclosed under MIFID II in Europe. For instance, Rabobank Bond Syndicate Allocation Policy<sup>35</sup> states: "The basic objective of allocation is to produce an appropriate spread of investors with a view to achieving an orderly aftermarket with sufficient liquidity and reasonable price stability."

Additionally, it is established that issuers are more likely to discriminate allocations of new bonds in the primary market if investors are highly heterogeneous with respect to the number of dealers they trade with or if liquidity shocks of investors are correlated. Moreover, from Proposition 4, exclusion is more profitable if the price impact of trades in the secondary market is high, which corresponds to higher inventory costs incurred by the dealers. This implies that if there is some additional cost to implement the exclusion, allocations to more connected investors should increase if the secondary market is expected

---

<sup>35</sup>Rabobank (2017).



to be more illiquid (Nagler and Ottonello, 2018).

### **Underwriters and direct offerings**

The result in Proposition 2.5 shows that the discriminatory allocation of bonds to investors in the primary market increases the issuer’s profits. Since exclusion of the “right” investors from the primary market is possible only when the issuer knows investor connections in the secondary market, the increase in the profits can be seen as a value of the knowledge of investor connectedness.

It is not a coincidence then, that in practice, new bonds are sold and allocated on behalf of issuers by underwriters, usually established investment banks, who are much better informed about investor connections in the secondary market compared to issuing firms. Underwriters’ knowledge of the trading network enables them to provide valuable service to issuers by coordinating allocations in the primary market.

Finally, this reasoning also suggests why proposed fintech platforms for direct offerings of bonds to investors, bypassing underwriters, might not get traction in this decentralized market. Indeed, since participation on such platforms is supposed to be open to everyone, issuers might lose their valuable ability to coordinate allocations, with the help of underwriters, among heterogeneously connected investors.

#### *2.5.2. Alternative specifications*

### **Number of client investors per dealer**

The baseline model of the paper assumes that each dealer is connected to two client investors while, in practice, dealers usually have more clients. To better accommodate this observation, it is straightforward to extend the baseline setup to the case where each dealer has  $m > 2$  clients. The extended model would be cumbersome notation-wise and harder to follow, since, for a given impatient investor  $i$  connected to a dealer  $d$ , the number of impatient neighbors connected to the same dealer  $d$  would vary between 0 and  $m - 1$ .

This extension, however, does not affect the main intuition behind the result of the paper: that an investor increases her primary market demand for bonds when she knows that she is connected to dealers who have highly connected clients. A high number of connections allows them to rebalance their trading in the secondary market more efficiently and guarantees better liquidity for the investor.

### **General form of liquidity supply**

Appendix 2C considers a more general form of the dealers' liquidity supply function  $d(q)$ , which is given by (2.4) in the baseline model. It shows that, in general, more connected investors might demand more bonds per dealer in the primary market compared to less connected investors even if all investors have the same probability of becoming impatient. However, more connected investors still trade less with stressed dealers and have a lower impact on expected liquidity of the secondary market. Therefore, the intuition behind the results should be preserved under more general assumptions on the dealers' liquidity supply function.

## 2.6. Conclusion

In this paper, I develop a model to study primary market allocations of corporate bonds and show how they are determined by investors' trading connections in the secondary market. I model the secondary market with a trading network to capture two distinctive features of the OTC market for corporate bonds. First, investors participating in the market have persistent trading relationships with dealers, who are the main providers of liquidity. Second, investors are highly heterogeneous with respect to the number of dealers they trade with. In the model, secondary market liquidity, which depends on the structure of the trading network, feeds back into investors' primary market demands and, therefore, determines the issuer's revenues.

I show that trading by less connected investors has a disproportionately high negative impact on secondary market liquidity. As a result, the issuer can increase her profits by excluding

less connected investors from the primary market and allocating new bonds only to more connected investors. Moreover, the exclusion is optimal even if less connected investors are willing to buy bonds at a price, which is higher than the issuance price. This result provides a rationale for the commonly observed exclusion of small institutional investors from the primary market, the fairness of which is debated by market participants since most of new bond issues appear to be underpriced.

In the paper, I focus on the heterogeneity of investor connections in the secondary market and abstract away from any potential asymmetric information frictions between the issuer and investors. It is interesting to explore the interaction between the two frictions in more detail. Another intriguing direction for future work is to study the question of network formation in light of the intuition that investors prefer to be indirectly linked, through dealers, to highly connected investors although costs of building a relationship with a dealer are private.

## CHAPTER 3 : Initial Coin Offerings as a Commitment to Competition

### 3.1. Introduction

Over the past couple of years, the market for Initial Coin Offerings (ICOs) has grown rapidly. In 2016, 52 ICOs collectively raised about \$283 million in this nascent market. Only two years later, in 2018, over 3,800 ICOs raised close to \$29.7 billion, which is almost 90% of the size of the IPO market that year.<sup>1</sup> The recent proliferation of ICOs raises several important questions that need to be answered to better understand the phenomenon. Does the ICO mechanism represent a novel way to raise funds for early ventures? Did ICOs attract such a large amount of funds because they are simply a means of regulatory arbitrage or do they offer new, attractive features that are not available from other established forms of financing? If yes, what are these features and what ventures are the most suitable for this mechanism of fundraising?

We attempt to shed light on some of these questions and highlight a novel feature of ICOs that distinguish them from other forms of financing. In particular, we focus on ICOs in which entrepreneurs obtain financing to develop a platform by selling digital assets, commonly referred to as "utility" tokens. These tokens can be later exchanged for services on the completed platform and are typically traded in a secondary market. In this paper, we show that such a token-based mechanism allows entrepreneurs to credibly commit to the long-run competitive pricing of services. During an ICO, consumers are able to finance the platform and commit some of their future consumer surplus to the entrepreneur. This generates incentives for the entrepreneur to run a token-based platform rather than operate as a monopolist. We show that ICOs can improve trade and generate welfare gains in markets which are prone to rent-seeking.

A sale of utility tokens via an ICO is typically used by entrepreneurs to develop a decentralized, online platform that facilitates trade by matching potential sellers and buyers

---

<sup>1</sup>See Davydiuk et al. (2018) and Ritter (2014).

of a good or service. The token sold is the sole currency that is used on the platform. Buyers pay sellers in tokens, which sellers can later monetize by selling these tokens in a secondary market on an exchange. For example, a prominent ICO called Filecoin, which raised \$257 million in 2017, is developing a platform in which users can buy and sell online data storage.<sup>2</sup> The company is developing a blockchain-based interface which allows users who need additional storage to rent this space using Filecoin tokens from users who have excess storage on their devices. A user in need of storage purchases Filecoin tokens on an exchange,<sup>3</sup> and uses it to buy storage on the platform. A user with excess storage sells it on the platform for Filecoin tokens, which they can then sell on the exchange. Such a common market where customers can buy tokens, in order to spend them later on the platform, and service providers can sell tokens, which they received from customers on the platform is a crucial aspect of the ICO mechanism.

Our analysis is motivated by the observation that marketplaces that match buyers and sellers are growing increasingly common and naturally lead to rent-seeking by developers due to associated network effects. Consider ride-sharing applications such as Uber or Lyft, apartment-rental services such as AirBnB, or the above example of Filecoin. These platforms require a large number of users looking for and providing rides, apartments and storage, giving rise to natural monopolies and oligopolies. For example, a user looking for a taxi, will be unlikely to download a hundred different ride-sharing applications and compare prices and wait times across them. Moreover, if an application does not have a critical number of drivers to match riders with the closest driver and optimize wait times, users looking for rides may not value these services over traditional taxi companies. There are, therefore, obvious efficiency gains from all users being on a single platform. However, rent-seeking can erode many of the welfare gains from these marketplaces. The ICO mechanism can help improve welfare by limiting the rent-seeking of entrepreneurs building a platform.

In the model, a penniless entrepreneur seeks investment to develop a platform, on which

---

<sup>2</sup>See <https://icobench.com/ico/filecoin>.

<sup>3</sup>Examples of cryptocurrency exchanges include Coinbase, Coinmama and CEX.io.

competitive service providers are matched with consumers. Consumers are heterogeneous in their valuations of the service, with some consumers valuing it more than others. The entrepreneur issues tokens and each token can be exchanged for one unit of the service. The entrepreneur chooses how many tokens to issue and whether to allow a common marketplace, in which tokens can be traded by everyone, or, alternatively, to retain sole rights to sell and redeem tokens.

If the entrepreneur does not allow the common marketplace for tokens, the situation is equivalent to operating as a monopolist. Without commitment to a common market for tokens, the entrepreneur, as the sole seller of tokens, has the power to charge any price for a token, which is the cost of a unit of service. As the sole redeemer of tokens, the entrepreneur also has full discretion over how much to pay a service provider. In equilibrium, the entrepreneur will operate exactly as a monopolist — charging customers more than the marginal cost of service production for each token even if providers are fully competitive. The entrepreneur will reimburse service providers at a price per token that is equal to their marginal cost of service provision, just enough to reimburse them for their costs. Thus, the entrepreneur can earn a spread from each service exchange and fully controls the quantity and pricing of the service. The entrepreneur will, therefore, optimally set an equilibrium price and quantity resembling that of a monopolistic service provider, even though service providers are perfectly competitive.

In contrast, when the entrepreneur allows a common market place for tokens, the model setting resembles the mechanism typical for many ICOs. We show that, in this case, when agents can trade tokens directly with each other, the entrepreneur is able to commit to give up pricing power. With a common marketplace, providers, who receive tokens in exchange for services, can resell tokens directly to consumers of the service instead of redeeming them with the entrepreneur. Therefore, each time the entrepreneur releases additional tokens, she is increasing the number of tokens that are sold in the future in the common marketplace, thereby generating competition for herself. Intuitively, we can think of the entrepreneur,

in this case, as having a limited stock of market power. Every time she wishes to monetize the platform, she necessarily creates future competition for herself, and uses up some of her market power.

We show that, in the presence of a common marketplace, the entrepreneur optimally chooses to release tokens over time rather than all at once, gradually increasing the number of consumers who purchase the service. Eventually, enough tokens are released so that all consumers who value the service above its marginal cost of production are able to access the service and the total surplus equals that in a competitive market. Over time, the equilibrium price of the service falls and the quantity increases to those that would occur in a competitive equilibrium. The long-run surplus is, therefore, always higher under the ICO mechanism. However, the surplus in the short-run may be higher under a monopolist, since all tokens are not released immediately.

We also consider the entrepreneur's choice between having an ICO and operating as monopolist when raising money to finance the development of the platform. The entrepreneur always generates a higher profit in the latter case, making it seem that the entrepreneur will never optimally choose to have an ICO. We show that if the entrepreneur is raising money from outside investors, who do not derive any value from consuming the product and only benefit from the return on their investment, the entrepreneur indeed always prefers to operate as a monopolist. However, if the entrepreneur is raising money from investors who also get utility from consuming the service, she may prefer to have an ICO and be better off committing to long-run competition. Consumers of the platform get higher surplus when prices are lower. They will, therefore, take into account their future surplus from consuming the product as well as the return on their investment when they are funding the entrepreneur. Consumers of the platform effectively subsidize the entrepreneur during an ICO, allowing her to keep a larger share of her profit. If this subsidy is large enough, the entrepreneur may prefer to have an ICO rather than to operate as a monopolist. The ICO mechanism, therefore, gives rise to *endogenous crowd-funding*, in which future consumers

of the platform are the only investors who can successfully fund its creation.

Our analysis predicts that token prices during an ICO, when tokens are initially released, can be much higher than token prices in secondary markets. This is because during financing, consumers are rationally subsidizing the entrepreneur up to an amount equal to their future consumer surplus and "over-paying" for the initial service the token can buy them. After the ICO has taken place, tokens will trade at a price that reflects the value of the service they can be exchanged for.

We further show that if the entrepreneur can break up financing into multiple rounds than the ICO mechanism is preferable to operating as a monopolist. The entrepreneur will choose to have financing rounds equal to the number of heterogeneous valuations by consumers, and monetize all future consumer surplus. In this case, all the tokens will be released over the multiple financing rounds, and the platform will operate at the competitive price and quantity from its first time of operation.

**Related literature.** Our paper contributes to the nascent but rapidly growing literature that studies various aspects of the ICO mechanism: Catalini and Gans (2018), Chod and Lyandres (2018), Cong et al. (2019b), Bakos and Halaburda (2019).<sup>4</sup>

The closest papers to ours are Li and Mann (2018), Sockin and Xiong (2018), Lee and Parlour (2019), Canidio (2018). Both Li and Mann (2018) and Sockin and Xiong (2018) show that the ICO mechanism allows entrepreneurs to resolve the coordination failure problem between consumers and providers who decide whether to participate in a new platform developed by entrepreneurs. We abstract from the coordination problem and study how token issuance affects pricing of a service exchanged between providers and consumers on a platform.

---

<sup>4</sup>Empirical literature on ICOs is also rapidly expanding, e.g., see Davydiuk et al. (2018); Adhami et al. (2018); Amsden and Schweizer (2018); Boreiko and Sahdev (2018); Bourveau et al. (2018); Deng et al. (2018); Fisch (2019); Jong et al. (2018); Howell et al. (2019); Lyandres et al. (2018); and Benedetti and Kostovetsky (2018).



Lee and Parlour (2019) show how crowdfunding mechanism allows consumers to finance socially efficient service provision that might be forgone by traditional profit-maximizing intermediaries in light of potential competition. In contrast, we show that entrepreneurs can commit to competitive platform pricing via the ICO mechanism.

Canidio (2018) considers how entrepreneurs dynamically sell tokens in the post-ICO period, which creates incentives and generates financial resources for further development of the platform. We, instead, focus on how the ICO mechanism allows an entrepreneur to commit to letting a platform run in a truly decentralized way and, thereby, supports commitment to competitive pricing.

Other papers study the economics of blockchains including benefits and limitations of adopting cryptocurrencies, such as Bitcoin, as a means of payment: Yermack (2013), Harvey (2014), Chiu and Koepl (2018), Abadi and Brunnermeier (2018), Budish (2018), Pagnotta (2018), Hinzen et al. (2019b), Biais et al. (2019), Chiu and Koepl (2019), Cong et al. (2019a), Saleh (2019), Easley et al. (2019), and Huberman et al. (2019).<sup>5</sup> Our paper is related to the work by Cong and He (2019) who develop a model in which blockchain technology can increase competition by allowing entrants to commit to delivering goods. This helps to overcome barriers to entry arising from information asymmetry problems, which give rise to a lemons problem in traditional markets. In contrast, we focus on a case in which natural monopolies can commit to competitive pricing.

Our paper also relates to the literature on monopolist selling durable goods. Coase (1972) shows that when a monopolist sells durable goods, in the continuous time limit, the monopolist immediately saturates the market. Stokey (1981) shows that in a discrete time version, the speed of market saturation depends on the interval of time between periods. The monopolist reaps larger profit when the period of times between successive intervals lengthens. Bulow (1982) shows that durable-good monopolists have an incentive to produce

---

<sup>5</sup>See Chen et al. (2019) for an overview of the recent research into blockchain economics. See also Hu et al. (2018), Liu and Tsyvinski (2018), Hinzen et al. (2019a), and Li et al. (2019) for empirical analysis of cryptocurrencies.

less durable products. In our paper, the re-tradability of tokens generates durability even though the service provided is not durable. This durability allows the entrepreneur to commit to acting like a monopolist selling durable goods, rather than one selling a non-durable good.

The rest of this paper is organized in the following way. In the next section, we introduce the formal setup of the model. Section 3.3 illustrates the main intuition behind our results in an example. Section 3.4 analyzes equilibrium in the general model. Section 3.5 discusses when an entrepreneur optimally chooses to have an ICO. Section 3.6 shows that, under network effects, welfare on the tokenized platform is greater than welfare delivered by two competing standard platforms. The last section concludes.

### 3.2. Model Setup

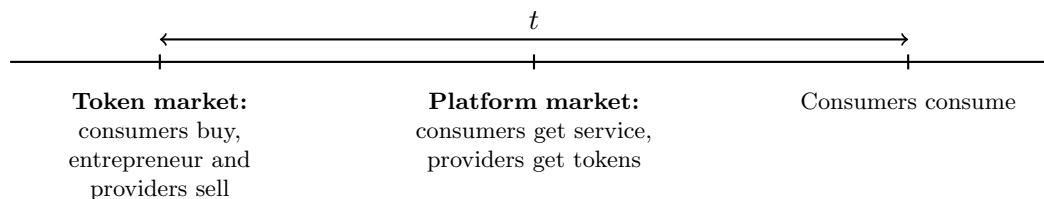
The model comprises of  $T$  periods. There are three types of agents: a long-lived entrepreneur who develops a platform and issues tokens, long-lived service providers who produce a service and can sell it on the platform, and long-lived consumers who value the service and can buy it on the platform. All agents are risk-neutral and have a common discount factor  $\delta \leq 1$ .

#### **Platform and tokens**

The platform is initiated at  $t = 1$  by the entrepreneur and allows consumers to obtain services from service providers by matching them in all periods  $t \geq 1$ . We assume that the service can only be purchased through the platform and there is no another way for service providers to match with consumers looking for the service.<sup>6</sup> Tokens are the only means of payment on the platform. Thus, in order to acquire the service, consumers have to get tokens first — this requirement generates a non-zero value for tokens. We assume that each token can be exchanged for 1 unit of the service. In addition to the platform exchange, each

---

<sup>6</sup>Matching, in this case, can be more involved than consumers and service providers simply being able to meet. Matching can involve using the platform's technology to facilitate provision of a service. For example, on a platform that connects users looking for taxi rides, matching involves mapping technology and optimization to connect each user with the closest driver. We also assume away the problem of platform leakage.



**Figure 3.1:** Sequence of events during a period  $t$ .

period  $t$ , there is also a market for tokens, in which service providers and the entrepreneur can sell any tokens they have to consumers for a price  $p_t$  that is determined in equilibrium. The sequence of events during a period  $t$  is summarized in Figure 3.1.

### Entrepreneur

Before the entrepreneur can start the platform at  $t = 1$ , she first raises investment  $I$  from investors, which is required to develop the platform. In our analysis, investors can be either complete outsiders or consumers that are described in more detail below. In the case of outsiders, the investment is raised with an equity contract. Alternatively, the entrepreneur can sell tokens to raise the required investment from consumers before the platform becomes operational. When the platform is initiated at  $t = 1$ , the entrepreneur chooses the total number of tokens to create,  $F \in [0, \infty)$ . Its value is common knowledge. Each period  $t$ , the entrepreneur can sell  $q_t \geq 0$  tokens to consumers in the token market. As described above, consumers value tokens since they can be exchanged for the service on the platform. We define  $F_t$  as the total number of tokens that the entrepreneur owns at the start of each period  $t$ . Since the entrepreneur initially controls all tokens  $F_1 = F$ . It also follows that the entrepreneur's budget constraint is  $0 \leq q_t \leq F_t$ , i.e. the entrepreneur cannot sell more tokens than she owns.

### Service providers

A large mass of service providers can access the platform and sell their service in exchange for tokens. Their marginal cost of producing a unit of the service is  $c$ . Since service providers

can participate in the market for tokens, they accept tokens at  $t$  as payment for the service knowing that they can sell tokens in the next period  $t + 1$  in this market. We assume that, at the end of last period, service providers can redeem their tokens with the issuer for a price  $c$ . This assumption is necessary for tokens to be a credible medium of exchange on the platform in a finite horizon model. The assumption can be relaxed in an infinite horizon model.

### Consumers

There is a unit mass  $[0, 1]$  of consumers who are long-lived. Each period, each consumer values only one unit of the service. There are  $N \leq T$  types of consumers. Every period  $t$ , a consumer of type  $i$  values a unit of service at  $v_i \in [\underline{v}, \bar{v}]$  where  $\underline{v} \geq c$ . Any subsequent units of the service in the same period are valued at 0. Without loss of generality,  $v_i$  is decreasing in  $i$  with  $v_1 = \bar{v}$  and  $v_N = \underline{v}$ . The mass of type  $i$  consumers is equal to  $\alpha_i$  and, therefore,  $\sum_{i=1}^N \alpha_i = 1$ . As mentioned above, consumers purchase tokens in the token market in order to exchange them for the service on the platform. We assume consumers are deep pocketed and, therefore, unconstrained in their ability to buy tokens.

### Price of tokens

Finally, we define how the price of tokens is set in the token resale market. We assume that if the mass of buyers is larger than the mass of sellers, then the token price is given by the value of the marginal buyer. In the opposite case, when the mass of sellers is larger than the mass of buyers, the token price is given by the value of the marginal seller.

### Key assumptions

There are three key features that distinguish tokens in our model and are crucial for our results:

1. The entrepreneur can commit to issue a fixed stock of tokens  $F$ .

2. Tokens are the sole medium of exchange on the platform that allows transfer of the service between providers and consumers.
3. There is an active token resale market, in which tokens can be sold and bought by providers, consumers and the entrepreneur.

In practice, these commitments are important features of utility tokens in a typical ICO and are implemented through smart contracts. We discuss the importance of each of these features in the equilibrium analysis in more detail.

### 3.2.1. Entrepreneur's Problem

The entrepreneur decides how many tokens  $F$  to create initially, at  $t = 1$ . Subsequently, she decides how many tokens to sell each period in the token market. Importantly, the entrepreneur understands that the total number of tokens she creates and the amount she decides to release each period will affect the current as well as future token prices. The entrepreneur solves the following problem:

$$\max_{F, \{q_t\}_{t=1}^T} \sum_{t=1}^T \delta^{t-1} q_t \cdot p_t(F, q_1, \dots, q_T) - \delta^{T-1} \sum_{t=1}^T q_t \cdot c \quad s.t. \quad \sum_{t=1}^T q_t \leq F. \quad (3.1)$$

The first term in the entrepreneur's maximization problem is the discounted sum of revenues from tokens sold each period while the second term is the amount she is committed to pay to service providers in the final period when they redeem their tokens.

### 3.2.2. Equilibrium Definition

An equilibrium of this model is given by:

1. The amount of tokens  $F$  created by the entrepreneur.
2. The number of tokens  $q_t \leq F_t$  that the issuer sells at each period  $t$ .
3. A price  $p_t$  that is determined in the market for tokens at each period  $t$ .

In the next two sections, we start our equilibrium analysis assuming the platform is already financed and fully operational, which is equivalent to the model with  $I = 0$ . We compare equilibrium quantities, prices, profits and welfare to the case, in which the entrepreneur operates as a monopolist. Next, in Section 3.5, we step back and consider the financing stage when  $I > 0$  and the entrepreneur can choose between raising funds with an ICO or raising funds with an equity contract and operating as a monopolist.

### 3.3. Example with $T = 2$ and $N = 2$

To illustrate the intuition behind our results, we start the analysis with an example, in which we set  $T = 2$  and  $N = 2$ . Additionally, in the most of our analysis, we assume that  $\delta = 1$  and only briefly discuss how results change when  $\delta < 1$ .<sup>7</sup> Thus, the platform operates two periods and there are two types of consumers. In this section, we refer to these two types as a high-type ( $H$ ) and a low-type ( $L$ ) and their respective utilities are  $v_H$  and  $v_L$ , where  $v_H > v_L \geq c$ . In the analysis, we further assume that the total number of tokens created is  $F = 1$  which is equal to the total mass of consumers. In Lemma 3.1, when we analyze the general model, we prove that this is the optimal quantity chosen by the entrepreneur.

#### 3.3.1. *Monopolist Entrepreneur*

To contrast our results for the entrepreneur who issues tokens and allows a token resale market, we first consider a situation where the token resale market is shut down. In this case, the entrepreneur acts as a monopolistic service provider.

When the exchange of the service on the platform is implemented with tokens but there is no resale market, the entrepreneur can sell tokens to consumers for one price and redeem them from service providers for another price. By setting an appropriate spread between the two prices, the entrepreneur can act as a monopolist who produces the service herself at constant marginal cost  $c$  and charges a unit price  $p_t$  to consumers. Specifically, this is

---

<sup>7</sup>See Appendix 3B for a complete analysis of the example with  $\delta < 1$ .

the case when the entrepreneur charges consumers the monopolistic price  $p_t$  for tokens and redeems them from service providers for  $c$  — the providers’ marginal production cost. Note that this is also optimal for the entrepreneur since no rents are given to service providers and the maximum monopolistic rents are extracted from consumers.

To differentiate between the two cases in the rest of the paper we call the monopolistic entrepreneur, who does not allow an active token resale market, the *monopolist* while we call the entrepreneur, who allows the token resale market, simply the *entrepreneur*. Below, we derive the optimal token price and release schedule for the monopolist.

Without an active token resale market, in each of the two periods the monopolist holds all the tokens, i.e.,  $F_t = 1$ , and chooses the number of tokens  $q_t \in [0, 1]$  to sell to consumers. Therefore, the entrepreneur’s multi-period problem (3.1) separates into two identical one-period problems, in which the monopolist trades off rents extracted from high-type consumers versus rents collected by serving a larger mass of consumers.

The monopolist will find it optimal to sell  $q_1 = q_2 = \alpha_H$  fraction of tokens, serving only high-type consumers, for a price  $p_1 = p_2 = v_H$  if

$$\alpha_H(v_H - c) \geq v_L - c. \tag{3.2}$$

In this case, extracting the maximum rents from high consumer types is more profitable than selling to both high- and low-type consumers. The monopolist’s total profit over the two periods is

$$2\alpha_H(v_H - c). \tag{3.3}$$

If (3.2) does not hold, the monopolist will optimally sell  $q_1 = q_2 = 1$  fraction of tokens, serving both types of consumers, for a price  $p_1 = p_2 = v_L$ . In this case, forgoing some rents from high-type consumers and instead serving all consumers is more profitable. The

monopolist's total profit over the two periods is

$$2(v_L - c). \tag{3.4}$$

### 3.3.2. Entrepreneur with Active Token Resale Market

When the token resale market is active, tokens that service providers received from consumers in exchange for the service at  $t = 1$  are sold by providers in the token market at  $t = 2$  to cover their production costs. Therefore, if the entrepreneur sells  $q_1$  tokens at  $t = 1$  to consumers, the consumers exchange these tokens for the service at  $t = 1$ , and service providers sell  $q_1$  tokens directly to consumers in the token resale market at  $t = 2$ . Service providers have an incentive to provide the service at  $t = 1$  as long as the token price is higher than their production cost,  $p_2 \geq c$ . This implies that consumers purchase tokens at  $t = 2$  both from providers and from the entrepreneur.

Since the platform operates for two periods the entrepreneur is committed to redeem all tokens owned by service providers for  $c$  at the end of  $t = 2$ . Absent such a commitment, in a finite horizon model, tokens have no value after  $t = 2$  and, thus, cannot act as a credible medium of exchange. The absence of the commitment will cause service providers to refuse the provision of service at  $t = 2$ . This will cause the market for tokens to break down at the start of the period as consumers will not want to purchase tokens they cannot exchange for the service. This will further cause the market to break down at  $t = 1$  as service providers will know that tokens will be worthless at  $t = 2$ .

We define  $Q_t = \sum_{s=1}^t q_s$  as the number of tokens released by the entrepreneur up to the date  $t$ . We show in the Appendix that the total supply of tokens sold in the market is always  $Q_t$ , i.e., service providers and consumers have no incentives to hoard tokens in order to instead sell or redeem them in the future.<sup>8</sup> Of this supply,  $q_t$  are tokens that are newly

---

<sup>8</sup>In the equilibrium, token prices decrease over time. Service providers, therefore, want to sell tokens as soon as possible. Similarly, since the price of tokens falls over time, consumers have no incentive to hoard tokens as this means they are paying more than they have to for a service they will receive in the future.



sold by the issuer, and the remaining  $Q_t - q_t$  are tokens that are sold by service providers and have been in circulation before.

When  $F = 1$ , depending on the entrepreneur's token release, there are three candidate equilibrium token quantity and price schedules:

1. If  $q_1 > \alpha_H$ , the equilibrium price in both periods is low,  $p_1 = p_2 = v_L$ .
2. If  $q_1 \leq \alpha_H$  and  $q_2 \leq \alpha_H - q_1$ , the equilibrium price in both periods is high,  $p_1 = p_2 = v_H$ .
3. If  $q_1 \leq \alpha_H$  and  $q_2 > \alpha_H - q_1$ , the equilibrium price is high in the first period and is low in the second period,  $p_1 = v_H$  and  $p_2 = v_L$ .

Note that the third case never occurs in the equilibrium with the monopolist since her problem is the same in each period and, therefore, the price will either be always high  $v_H$  or always low  $v_L$ . We can show that, with an active resale market, the equilibrium quantity and price schedule will always be according to the third case above, in which the token price is high  $v_H$  at  $t = 1$  and then falls to  $v_L$  at  $t = 2$ .

Consider the first candidate pricing schedule, in which there is a high price for tokens in both periods. The entrepreneur issues the maximum amount of tokens possible that can be sold at this price, i.e.  $q_1 + q_2 = \alpha_H$ . Therefore, the entrepreneur's total profit over the two periods is

$$\alpha_H(v_H - c). \tag{3.5}$$

The entrepreneur can always do strictly better by selling  $\alpha_H$  tokens at  $t = 1$  and  $(1 - \alpha_H)$  tokens at  $t = 2$  since such a token release schedule would yield the total profit of:

$$\alpha_H v_H + (1 - \alpha_H)v_L - c > \alpha_H(v_H - c). \tag{3.6}$$

Now, consider the second candidate pricing schedule, in which there is a low price for tokens

in both periods. Again, the entrepreneur will issue the maximum amount of tokens possible that can be sold at this price, i.e.  $q_1 + q_2 = 1$ . Therefore, the entrepreneur's total profit over the two periods is

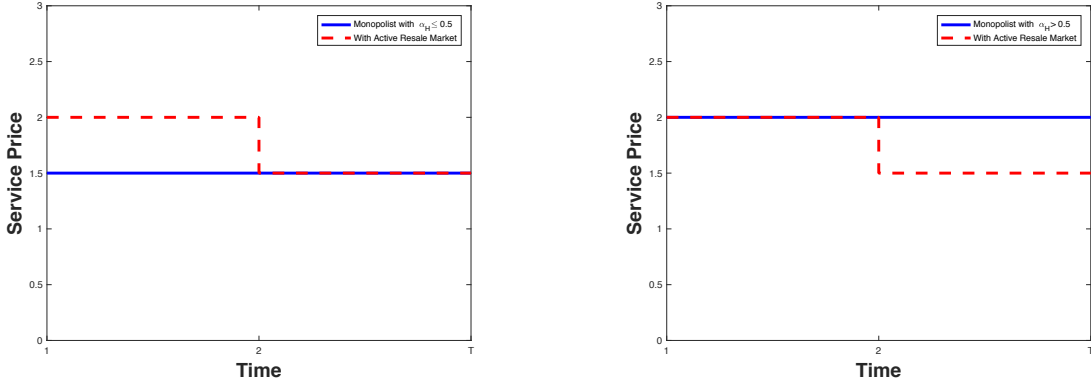
$$v_L - c. \tag{3.7}$$

Similarly to the previous case, the entrepreneur can do strictly better by selling  $\alpha_H$  tokens at  $t = 1$  and  $(1 - \alpha_H)$  tokens at  $t = 2$ . Such a token release schedule allows the entrepreneur to make  $v_H$  instead of  $v_L$  on the first  $\alpha_H$  proportion of tokens sold.

There are two key mechanisms at play here. First, each time the entrepreneur wants to monetize the platform and sell additional tokens, she increases competition for herself with service providers in subsequent periods due to the token resale market. Indeed, any tokens the entrepreneur releases will be subsequently sold by service providers. Over time, as the total quantity of tokens in circulation grows, competition in the resale market increases, reducing the price of tokens.

Second, the entrepreneur can only profit from each token once, since any released tokens will be subsequently resold each period by competitive service providers. The token issuer will, therefore, price discriminate to get the maximum surplus from each token. Intuitively, we can think of the entrepreneur as having a limited stock of market power, which eventually runs out.

As a result, in the equilibrium, not every consumer is served at first but, eventually, everyone who values the service more than its marginal cost will be able to obtain the service. Absent time discounting, when  $\delta = 1$ , the entrepreneur will practice perfect price discrimination over time. As a result, only high value consumers obtain the service at  $t = 1$  while low value consumers obtain it at  $t = 2$ . If the entrepreneur releases tokens to both consumer types simultaneously at  $t = 1$ , the token price will drop to  $v_L$  at  $t = 1$ . This will force the entrepreneur to forego making  $v_H - v_L$  additional revenue from each high-type consumer. Consequently, in exactly 2 periods, competitive pricing of the service is reached, in which



**Figure 3.2: Equilibrium prices of the service.** The equilibrium price of the service under the monopolist and the one under the entrepreneur with an active token resale market. The parametrization is as follows:  $\delta = 1$ ,  $v_H = 2$ ,  $v_L = 1.5$ ,  $c = 1$ ,  $N = 2$ ,  $T = 2$ .

all consumers who value the service above its marginal cost are able to obtain it. As we show in the analysis of the general model with  $N$  types this competitive pricing is reached in exactly  $N$  periods.

Figure 3.2 plots the equilibrium price of the service or, equivalently, the token price under the monopolist and under the entrepreneur with an active token resale market. The price under the entrepreneur is independent of  $\alpha_H$  as she always finds it optimal to price discriminate over time. In contrast, the monopolist prefers to serve only high-type consumers and excludes low-type consumers from the market for high values of  $\alpha_H$ . In this case, an active token resale market can help commit to competitive pricing over time.

When the entrepreneur discounts revenues from future periods,  $\delta < 1$ , she might choose to sell tokens to multiple types at once. In our example, the entrepreneur prefers to release all tokens at  $t = 1$  if

$$v_L > \alpha_H v_H + \delta(1 - \alpha_H)v_L, \quad (3.8)$$

equivalently,

$$\delta < \frac{v_L - \alpha_H v_H}{(1 - \alpha_H)v_L}. \quad (3.9)$$

Time-discounting can, therefore, speed up the process of getting to the competitive price if

the entrepreneur has a high discount factor for future payoffs. With time-discounting, the price, at which all consumers who value the service above its marginal cost are willing to buy it, is reached in *at most*  $N$  periods.

With an active token resale market, a competitive pricing is reached over time. Specifically, all consumers who value the service above its marginal cost to service providers obtain the service over time. Therefore, the long-run quantity of tokens released will always be weakly higher than under a monopolist and the long-run price will always be weakly lower than the monopolist's price. Due to the perfect price discrimination by the entrepreneur, the short-run price and quantity supplied may be less than those under the monopolist. As we showed above, under some conditions, the monopolist releases all tokens in both periods charging  $v_L$ . In contrast, with an active resale market, the entrepreneur will always price discriminate.

It is clear from the analysis that an active resale market for tokens is the key factor that allows a commitment to long-run competition. There are two other features of the token market that are required for our mechanism to work. First, the entrepreneur can commit to a fixed supply of tokens  $F$ . To see why this is important, consider the opposite case, in which the entrepreneur cannot commit to the total supply of tokens. Then, the entrepreneur has an incentive to create more tokens at  $t = 2$  since she is always willing to produce an extra token and sell it at a price greater than or equal to  $c$ . The entrepreneur will not sell tokens for any price lower than  $c$  because she has promised to redeem tokens in the future at  $c$ . The over-supply of tokens at  $t = 2$  will cause the price to fall to  $c$ , which means that service providers expect to earn strictly less than  $c$  for the tokens they sell at  $t = 2$ , since with a positive probability some tokens will remain unsold. This causes unraveling of the service exchange at  $t = 1$  because service providers are unwilling to accept tokens as payment for their service since they do not expect to recoup their marginal cost of providing the service.

The second feature that is required for our mechanism, is that the token is the sole medium of exchange used to purchase the service on the platform. Imagine that the platform simply

allowed consumers to connect with service providers and that, besides tokens, consumers could pay with cash for the service. Then, since the service providers' side of the market is competitive, the price of the service would fall to  $c$  in the first period, forcing the token price to also fall to  $c$ . The entrepreneur will, therefore, make no profit from the platform. If the entrepreneur faces any costs of platform development, i.e.  $I > 0$ , she has no way of recovering those and, therefore, will not design the platform.

### 3.3.3. Profits

The monopolist always earns higher profit than the entrepreneur. There are two cases for the monopolist. When  $\alpha_H(v_H - c) \geq v_L - c$ , the monopolist's profit is  $2\alpha_H(v_H - c)$ . This is larger than the profit of the entrepreneur if

$$2\alpha_H(v_H - c) > \alpha_H(v_H - c) + (1 - \alpha_H)(v_L - c) \quad (3.10)$$

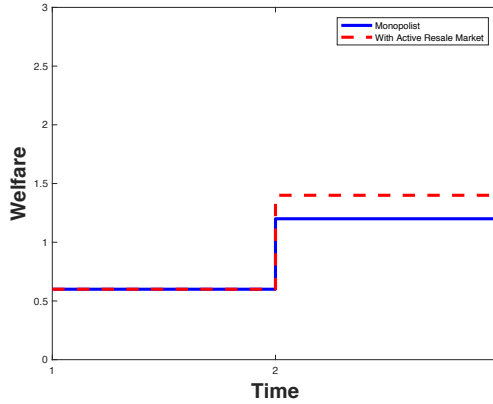
which is true since  $\alpha_H(v_H - c) \geq v_L - c$ . Therefore, in this case, the monopolist earns a higher profit than an entrepreneur.

When  $\alpha_H(v_H - c) < v_L - c$ , the monopolist's profit is  $2(v_L - c)$ . This is larger than the profit of the entrepreneur if

$$2(v_L - c) > \alpha_H(v_H - c) + (1 - \alpha_H)(v_L - c). \quad (3.11)$$

The above is also true since  $\alpha_H(v_H - c) < v_L - c$  in this case.

Therefore, the monopolist always earns a higher profit than the entrepreneur with an active token resale market. Intuitively, since the monopolist has greater market power, she can always choose to replicate the cash flow that is optimal with the active resale market. For example, the monopolist can always choose to sell  $\alpha_H$  tokens in the first period, charging  $v_H$  for them, and then sell  $\alpha_L = 1 - \alpha_H$  tokens in the second period and charge  $v_L$  for them. However, the monopolist's greater market power makes it profitable to deviate and either



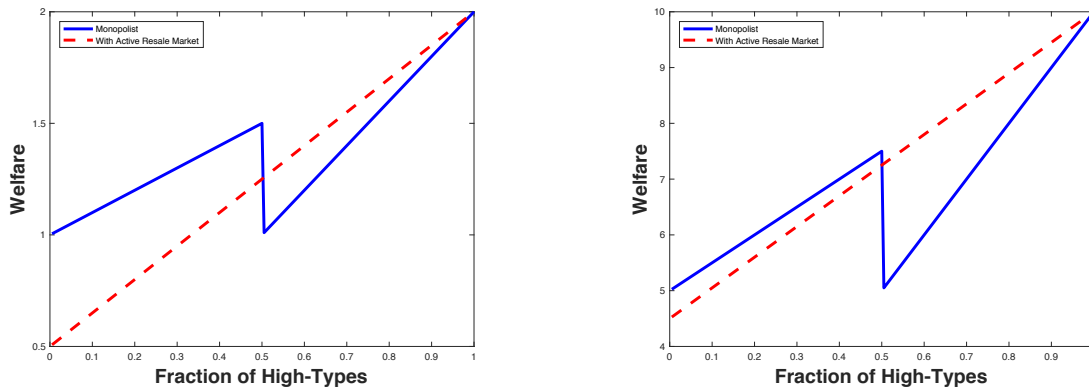
**Figure 3.3: Cumulative welfare over time.** The welfare under the monopolist and the one under entrepreneur with an active token resale market over time. The parametrization is as follows:  $\delta = 1$ ,  $v_H = 2$ ,  $v_L = 1.5$ ,  $c = 1$ ,  $N = 2$ ,  $\alpha_H = 0.6$ .

restrict or increase the supply of tokens. The monopolist can, therefore, always choose to replicate the total producer surplus, the sum of what the entrepreneur and the service providers get, which is generated in a market with active resale. Therefore, any alternative equilibrium strategy chosen by the monopolist must be more profitable.

#### 3.3.4. Welfare

There are two forces affecting the ranking of welfare, which we define as the total surplus, under the monopolist versus under the entrepreneur. On the one hand, with the active token resale market, in equilibrium, we always achieve the competitive outcome which maximizes total surplus. On the other hand, due to price discrimination by the entrepreneur, this outcome is reached after some time. If the monopolist makes enough profit by providing a large mass of consumers with the service, the total welfare under the monopolist may be higher because the initial token price for consumers is lower.

In our example, when  $\alpha_H(v_H - c) < v_L - c$ , the total welfare under the monopolist is higher than that under the entrepreneur with an active token resale market. Otherwise, the total welfare under the entrepreneur is always higher than the total welfare under the monopolist. Figure 3.3 illustrates this case, in which  $\alpha_H(v_H - c) \geq v_L - c$ . Here, the first period welfare



**Figure 3.4: Welfare as a function of high-type consumers.** The total welfare under the monopolist and the one under entrepreneur with an active token resale market as functions of  $\alpha_H$ . The parametrization is as follows:  $\delta = 1$ ,  $v_H = 2$ ,  $v_L = 1.5$ ,  $c = 1$ ,  $N = 2$ . The left panel shows the welfare when  $T = 2$  while the right panel illustrates the welfare when  $T = 10$ .

is identical under the monopolist and under the entrepreneur as only high-type consumers are served in the first period. However, the second period welfare is higher with an active token resale market since low-type consumers are also able to obtain the service.

Figure 3.4 plots the total welfare under the monopolist and that under the entrepreneur as a function of  $\alpha_H$ . For low values of  $\alpha_H$ , the monopolist finds it optimal to serve both high- and low-type consumers. Therefore, the welfare under the monopolist is always higher than that under the entrepreneur since in the latter case the entrepreneur price discriminates in the first period. In contrast, for high values of  $\alpha_H$ , the monopolist only services high-types. Therefore, the welfare is higher with an active token resale market. As the number of periods  $T$  increases (the right panel), the welfare gain from having an active resale market increases for high values of  $\alpha_H$  and the welfare loss from low values of  $\alpha_H$  decreases.

We next turn to the analysis of the equilibrium in the general model with  $N$  types and  $T$  periods. The insights from the example of this section all carry over.

### 3.4. Equilibrium Analysis

We now consider the main model setting, which lasts  $T$  periods and has  $N$  consumer types. As in the previous section, we compare the two cases: i) the entrepreneur who operates

the platform with an active token resale market; and ii) the monopolist who operates the platform without the market for tokens which is equivalent to the monopolistic service provider.

#### 3.4.1. Entrepreneur with Active Token Resale Market

We start by showing that the entrepreneur always chooses to create the number of tokens equal to the total measure of consumers,  $F = 1$ .

**Lemma 3.1.** *At date  $t = 1$ , the entrepreneur creates a number of tokens equal to the total mass of consumers each period, i.e.  $F = 1$ .*

*Proof.* On the one hand, the entrepreneur never wants to create less than a measure 1 of tokens because she is leaving money on the table by doing so. Specifically, since the entrepreneur can price discriminate consumer types by delaying the release of tokens, she always wants to sell tokens to all consumer types. On the other hand, if the entrepreneur tries to create a greater number of tokens than the per-period mass of consumers, once all the tokens are released, the expected token price falls below  $c$ .<sup>9</sup> Indeed, since the supply of tokens is greater than the demand for tokens in this case, each service provider expects to make less than  $c$  from selling a token. At this expected token value, service providers refuse to accept tokens in exchange for their service. Therefore, over-supplied tokens can not be a credible medium of exchange for services on the platform.  $\square$

With the active token resale market, the entrepreneur faces a similar problem to the one in the example. When  $\delta = 1$ , the entrepreneur wants to release  $q_t = \alpha_i$  tokens in period  $t = i$  where  $\alpha_i$  is the measure of consumers who have the highest value for the service among consumers who have not yet obtained the service. Specifically, at  $t = 1$ , the entrepreneur releases  $\alpha_1$  measure of tokens, which is equal to the measure of consumers who have the highest value for the service, and the token price is  $\bar{v}$ . At  $t = 2$ , providers sell these tokens,

<sup>9</sup>The entrepreneur knows that in the final period tokens are redeemable for  $c$ , so the lowest price she will be willing to sell a token for is  $\delta^{T-t}c$ .



received as a payment for their service in the previous period, and the entrepreneur releases an additional  $\alpha_2$  tokens. The price of tokens falls to the new level  $v_2$ . This continues until the last period  $N$ , in which the entrepreneur sells tokens to the group of consumers who value the service the least and the token price falls to  $\underline{v}$ . By using this delayed token release schedule, the entrepreneur is able to price-discriminate perfectly and maximizes her profit.

When  $\delta < 1$ , the entrepreneur discounts future revenues and may choose to sell new tokens to a larger measure of consumers at a lower price. As in our example, a high  $\delta$  can speed up the time it takes to get to competitive pricing. In the general model, we can establish the following proposition.

**Proposition 3.1.** *With the active token resale market, there is a unique equilibrium, in which the total quantity of tokens released increases over time and the price of tokens decreases over time. With  $N$  different consumer types, the competitive outcome is achieved in exactly  $N$  periods when  $\delta = 1$ . When  $\delta < 1$ , the competitive outcome is achieved in at most  $N$  periods.*

*Proof.* See Appendix 3A. □

The key insight is that the entrepreneur can profit from each token only once and is limited in how many tokens she can create for the platform to be feasible. Whenever the entrepreneur wants to monetize the platform by selling tokens, she is also necessarily creating competition in the future token resale markets as more tokens will be sold by service providers. We can think of the entrepreneur as having a limited stock of market power. The more new tokens the entrepreneur sells, the less market power she has in the token market going forward. As time passes, we eventually reach a competitive outcome in the market, in which all consumers who value the service above its marginal cost are able to consume it. The ICO structure, therefore, allows commitment to long-run competitive price of the service to consumers.

As in our example, we can compare this equilibrium to the one that would exist if the entrepreneur operated as a monopolistic service provider. As argued above, this is equivalent to the case of the entrepreneur who operates the platform without token resale market and buys back tokens at  $c$  from service providers at the end of every period.

### 3.4.2. Monopolist Platform Developer

In the case of the platform without an active token resale market, the monopolistic entrepreneur holds all the tokens each period, i.e.  $F_t = 1$ , and chooses the number of tokens  $q_t \in [0, 1]$  to sell for all  $t$ . Thus, the entrepreneur's multi-period problem (3.1) separates into  $T$  identical one-period problems, in which the monopolist trades off rents extracted from serving higher consumer types versus rents collected by serving to a larger measure of consumer types at a lower price.

Irrespective of time discounting, the monopolist solves the following problem each period:

$$i_m \in \arg \max_i \sum_{j=1}^i \alpha_j (v_i - c), \quad (3.12)$$

which determines the marginal consumer type that is served. Consequently, the monopolist sells  $q_t = \sum_{j=1}^{i_m} \alpha_j$  tokens for the price  $p_t = v_{i_m}$  to consumers and redeems them from the service providers for  $c$ . Only the mass  $\sum_{j=1}^{i_m} \alpha_j$  of consumers obtains tokens, and, thus, are able to buy a service every period while the rest of consumers are not able to get the service. With a monopolist entrepreneur, there are, therefore, gains from trade between consumers and service providers that are not realized and some consumers who value the service above its marginal cost are not able to purchase it.

We now turn to comparing equilibrium profits and welfare under the entrepreneur with the active token resale market and the monopolist. For the rest of the analysis, we assume  $\delta = 1$ , as this simplifies equilibrium expressions. However, all our results are qualitatively similar when  $\delta < 1$ .

### 3.4.3. Profits

As in the example, we can calculate the monopolist's and the entrepreneur's profits. In particular, the monopolist is always more profitable than the entrepreneur.

**Proposition 3.2.** *The monopolistic entrepreneur always earns a higher profit than the entrepreneur who allows the active token resale market.*

*Proof.* The total profit in the monopolist case is the lifetime sum of one-period profits:

$$T \sum_{j=1}^{i_m} \alpha_j (v_{i_m} - c). \quad (3.13)$$

At the same time, the total profit of the entrepreneur who releases all the tokens with delay and allows the active resale market is

$$\sum_{j=1}^N \alpha_j (v_j - c) = \sum_{j=1}^N \alpha_j v_j - c \quad (3.14)$$

The monopolist always earns a higher profit than the entrepreneur. First, note that the monopolist earns a positive profit after period  $t = N$  while the entrepreneur sells all her tokens by that time and earns zero in subsequent periods. Second, even if the number of periods  $T$  is small the monopolist has a greater market power and she can always choose to replicate the cash flow that is optimal for the entrepreneur with the active token resale market. In particular, the monopolist achieves this by selling  $\alpha_t$  tokens for  $v_t$  at every period  $1 \leq t \leq N$ . Therefore, any alternative equilibrium strategy chosen by the monopolist must be more profitable.  $\square$

Recall that, with the active token resale market, the entrepreneur can profit from each token only once. The monopolist, on the other hand, has the ability to earn continued profits as her market power gets reset every period. The monopolist can choose equilibrium quantities and prices to mimic those of the entrepreneur and replicate the entrepreneur's

payoff. However, in equilibrium, she would always find it more profitable to sell to the same measure  $\sum_{j=1}^{i_m} \alpha_j$  of consumers each period.

#### 3.4.4. Welfare

The intuition on the ranking of the total welfare under the monopolist and under the entrepreneur with the token resale market, which is developed in the example, carries through in the general setup.

With the active token resale market, competitive pricing, which maximizes the total per-period surplus, is always achieved in the equilibrium. However, due to price discrimination by the entrepreneur, this outcome is reached only after some time. If the monopolist makes enough profit by providing a large mass of consumers with the service, the total welfare under the monopolist may be higher because the token markets are more competitive from the beginning. Specifically, if  $i_m$  is the marginal consumer type served by the monopolist, the per-period surplus will be lower under the entrepreneur relative to that under the monopolist for the first  $m$  periods. Formally, we can establish the following proposition.

**Proposition 3.3.** *The total welfare under the entrepreneur is higher than the total welfare under the monopolist when the number of periods  $T$  is sufficiently high. The opposite is true if  $T$  is small and  $i_m$  is high, i.e. when the monopolist serves a sufficiently large mass of consumers.*

*Proof.* See Appendix 3A. □

In the above analysis, we focus on the case when  $\delta = 1$ . When  $\delta < 1$ , the qualitative results are similar but there are two additional forces. On the one hand, the competitive outcome in the service market is reached sooner and, therefore, the total welfare is more likely to be higher under the entrepreneur who allows an active resale market. On the other hand, the discounted surplus from the future periods contributes less to the total surplus and early price discrimination reduces the welfare under the entrepreneur relative to that under the

monopolist.

Note that our competitive token price is  $\underline{v}$  and not the marginal cost  $c$ . The total surplus is maximized for any price between  $c$  and  $\underline{v}$ . Only the division of the surplus between service providers and consumers is affected. The token price does not necessarily fall to  $c$  because consumers have discrete valuations for the service, rather than continuous valuations around the marginal cost. For the next section it is at times useful to think of the case when  $\underline{v}$  is close to the marginal cost  $c$ , which tilts the division of the surplus  $(\underline{v} - c)$  once the competitive price is reached towards consumers. Intuitively, this means that a non-zero measure of consumers value the service at close to its marginal cost. In this case, the competitive price with an active resale market will be approximately  $c$ .

### 3.5. Endogenous Crowd-Funding

Our analysis so far has focused on contrasting equilibrium quantities and prices on the platform in the case of a token-based system versus a monopoly. We can now extend the model to allow for a fundraising stage at  $t = 1$ , in which an entrepreneur needs to raise an amount  $I$  to develop the platform and can choose to either operate as a monopolist or an entrepreneur who allow token resale market. If the project gets funded, service providers and consumers can access the platform at  $t = 1$  and the economy continues as in the previous section. We assume that consumers are deep-pocketed at  $t = 1$  and can fund the creation of the platform. We also normalize their outside investment option to 1. For simplicity, we focus on the case when  $\delta = 1$ .

We start analyzing the optimal fund-raising by continuing our example with  $N = 2$  and  $T = 2$ . Recall that, in the case, when  $\alpha_H(v_H - c) > v_L - c$ , the total surplus under the entrepreneur is higher than that under the monopolist. However, the profit of the monopolist is always higher.

It is straightforward to show that if the entrepreneur was raising funds from investors who do not value the service available on the platform, i.e., the sole way to attract investors

was through profit-sharing contracts, the entrepreneur would always choose to operate as a monopolist rather than an entrepreneur who allows token resale market. To get funded from such investors, a monopolist would have to offer a fraction of the platform  $s_m$  for sale such that

$$s_m = \frac{I}{2\alpha_H(v_H - c)}. \quad (3.15)$$

Analogously, with an active token resale market, the entrepreneur would have to offer a fraction of the platform  $s_e$  for sale such that

$$s_e = \frac{I}{\alpha_H v_H + (1 - \alpha_H)v_L - c}. \quad (3.16)$$

It follows then that the monopolist's total payoff is always higher than the entrepreneur's total payoff

$$(1 - s_m)2\alpha_H(v_H - c) > (1 - s_e)(\alpha_H v_H + (1 - \alpha_H)v_L - c). \quad (3.17)$$

Since the platform controlled by the monopolist generates higher profit she has to offer the smaller share of the firm to investors to generate the necessary return, i.e.,  $s_m > s_e$ . Further, we have previously shown that the monopolist's profit,  $\alpha_H(v_H - c)$ , is higher than the profit of the entrepreneur with an active token resale market,  $(\alpha_H v_H + (1 - \alpha_H)v_L - c)$ . Therefore, when trying to raise money from investors who do not value consuming the service, the entrepreneur would prefer to operate as a monopolist.

In contrast, when the entrepreneur can raise money from future consumers of the service, she may prefer to hold an ICO and operate as an entrepreneur who allows token resale market rather than to operate as a monopolist. Consider a candidate equilibrium, in which each high-type consumer is investing an amount  $\frac{I}{\alpha_H}$  to fund platform's creation. Then a high-type consumer who values the service at  $v_H$  and whose investment is pivotal for the platform's success will invest in the ICO if

$$\frac{s_e}{\alpha_H}(\alpha_H v_H + (1 - \alpha_H)v_L - c) - \frac{I}{\alpha_H} + (v_H - v_L) \geq 1. \quad (3.18)$$

The entrepreneur will choose  $s_e$  so that the above condition binds

$$s_e = \frac{I - \alpha_H(v_H - v_L)}{\alpha_H v_H + (1 - \alpha_H)v_L - c}. \quad (3.19)$$

In this case, the share of the firm the entrepreneur needs to sell is lower than the share she needs to sell to investors who do not value the service. The high-type consumers know that prices for the service will be lower on a token-based platform at  $t = 2$  and expect to get consumer surplus in this period. Therefore, these consumers are willing to accept a lower share of the platform's profit.

The entrepreneur will prefer an ICO over being a monopolist if

$$(1 - s_e)(\alpha_H v_H + (1 - \alpha_H)v_L - c) \geq (1 - s_m)2\alpha_H(v_H - c), \quad (3.20)$$

which is equivalent to

$$(v_L - c)(1 - 2\alpha_H) \geq 0. \quad (3.21)$$

Alternatively, the above can be represented as the entrepreneur selling tokens instead of equity. Specifically, suppose the entrepreneur sells  $q_1 = \alpha_H$  tokens to raise required investment amount  $I$ . In this case, if a high type, whose participation is pivotal for the development of the platform, is willing to pay up to  $v_H + (v_H - v_L)$  for a token. The first term  $v_H$  is the utility the consumer gets from exchanging the token for the service in the first period. The second term  $(v_H - v_L)$  is the future consumer surplus she expects to obtain in the second period. In this case, the entrepreneur prefers to do an ICO over being a monopolist if

$$\alpha_H(2v_H - v_L) - I + (1 - \alpha_H)v_L - c \geq (1 - s_m)2\alpha_H(v_H - c), \quad (3.22)$$

which simplifies to the same condition (3.21).

Since  $v_L > c$  by assumption, the entrepreneur prefers to have an ICO whenever  $\alpha_H < \frac{1}{2}$ . The ICO allows the entrepreneur to extract the consumer surplus that high-types expect to

enjoy because of low prices in the future periods. At the same time, the ICO also continues to be able to price discriminate over time and earn rents of  $v_L - c$  from low-type consumers. Since price does not fall to  $c$  at  $t = 2$ , the ICO only allows the entrepreneur to extract  $v_H - v_L$  of  $t = 2$  surplus from high-types rather than get rents of  $v_L - c$  from high-types as the monopolist does. However, because of price discrimination over time, the ICO does not get  $v_L - c$  from high-types but does get it from low-types. Therefore if the fraction of low-types is higher than the fraction of high-types, the entrepreneur would prefer to have an ICO. The ICO, essentially, gives the entrepreneur a unique form of price discrimination that the monopolist is not able to do since the monopolist cannot commit to charging lower prices in the future.

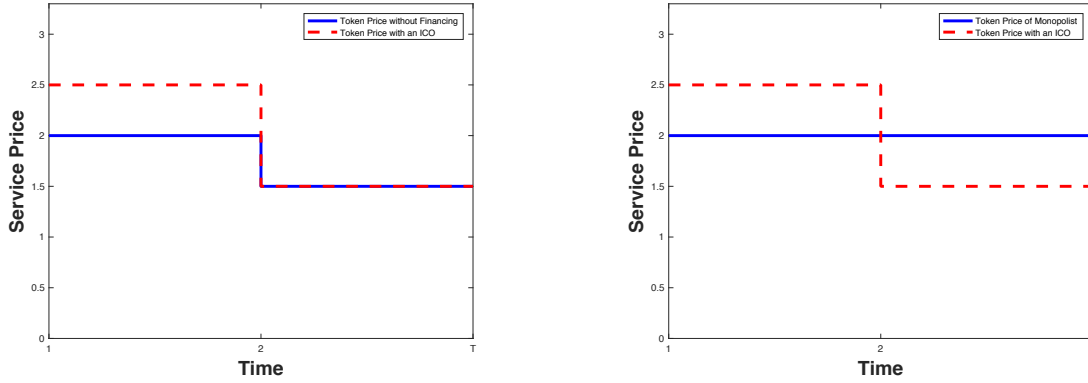
We can show that if (3.21) holds, then the candidate equilibrium described above is the preferred equilibrium of the entrepreneur and can be implemented by her.<sup>10</sup> When  $\alpha_H(v_H - c) < v_L - c$ , the total surplus under a monopolist is higher than under an ICO and it will never be optimal for the entrepreneur to have an ICO.

Our analysis has a few important implications. If investors are only accounting for the share of a profit they have access to, an entrepreneur will always prefer to operate as a monopolist. However, if agents who get utility from consuming the service in the future participate in fundraising, the entrepreneur at times prefers an ICO. We, therefore, endogenously require that ICOs are financed through crowd-funding. Furthermore, conditional on the type of consumer who is participating in crowd-funding, the entrepreneur wants to design the ICO in a way that the maximum measure of that type of consumer participate. In an ICO, the platform will be funded by the consumers who value the service the most as they get the

---

<sup>10</sup>In the above equilibrium, every high-type needs to believe that they are pivotal to the platform being successfully funded. If a high-type consumer believes that the platform would operate without his purchase of tokens, then he would prefer to not participate in the ICO and get his consumer surplus in the subsequent period. Other equilibria of this game exist, in which some high type consumers would invest more than others. The entrepreneur prefers that all high-type consumers invest since this maximizes the subsidy she receives from future consumer surplus. To get high-type consumers to co-ordinate on this equilibrium, the entrepreneur can simply set a minimum amount that needs to be raised for the project to proceed, otherwise the money would be returned to investors. In practice, many ICOs implement such limits known as soft caps. Typically, if the soft cap is not reached, the entrepreneurs return any money raised to investors. Smart contracts implement these terms so the entrepreneur can commit to returning the money ex-ante.





**Figure 3.5: Equilibrium prices of a service with financing.** The left panel plots the equilibrium price with an ICO and the equilibrium price with an active resale market without financing. The right panel plots the equilibrium price with an ICO and the one with a monopolist. The parametrization is as follows:  $\delta = 1$ ,  $v_H = 2$ ,  $v_L = 1.5$ ,  $c = 1$ ,  $N = 2$ ,  $T = 2$ ,  $\alpha_H = 0.6$ .

highest surplus from the operational platform and are, therefore, willing to pay the highest token price in an ICO. Since, high-type consumers are willing to pay more than their one-period value of the service, the ICO price can be much higher than the subsequent secondary market price for tokens.

The left panel of Figure 3.5 shows that if, in the financing round, high-type consumers believe they are pivotal to the platform’s success the initial token price increases. This higher price is due to the subsidy that high-type consumers provide to the entrepreneur since they account for their future consumer surplus. The right panel compares the equilibrium price with an ICO to that of the monopolist. Since tokens have a higher value at  $t = 1$ , the entrepreneur may benefit more from an ICO than from operating as a monopolist.

We now turn to the the full model with  $N$  types and  $T$  periods. We can establish the following lemma, stating that if investors only care about their returns from profit-sharing and do not value consuming services on the platform, the entrepreneur always prefers to operate as a monopolist.

**Lemma 3.2.** *If investors in the platform do not obtain any benefit from consuming services on the platform, the entrepreneur always prefers to operate as a monopolist and offers*

investors an equity contract.

*Proof.* See Appendix 3A. □

In contrast, when investors of the platform also value services, which become available if the platform is operational, they are willing to subsidize the entrepreneur if she operates as a token issuer since they expect to benefit from lower prices for the service in the future periods. If consumers participate in financing, the entrepreneur may prefer to hold an ICO. We can establish the following proposition.

**Proposition 3.4.** *If consumers are able to participate in financing of the platform the entrepreneur prefers to issue tokens and have an active resale market rather than operate as a monopolist when*

$$\begin{aligned}
& T \left( \sum_{i=1}^{i_e} \alpha_i v_{i_e} - \sum_{j=1}^{i_m} \alpha_j v_{i_m} \right) + \sum_{j=1}^{N-i_e} v_{j+i_e} \left( \alpha_{j+i_e} - \sum_{i=1}^{i_e} \alpha_i \right) \\
& - (T - N + i_e - 1) \underline{v} \sum_{i=1}^{i_e} \alpha_i - c \left( 1 - T \sum_{j=1}^{i_m} \alpha_j \right) \geq 0
\end{aligned} \tag{3.23}$$

where

$$i_m \in \arg \max_i \sum_{j=1}^i \alpha_j (v_i - c), \tag{3.24}$$

$$i_e \in \arg \max_i \sum_{j=1}^i \alpha_j \left( v_i + \sum_{j=1}^{N-i} (v_i - v_{j+i}) + (T - N + i - 1)(v_i - \underline{v}) \right). \tag{3.25}$$

To help interpret the above expression, it is helpful to consider the case, in which the entrepreneur chooses  $i_e = i_m$  (as was the case in our example) and when  $\underline{v}$  is close to  $c$ , i.e., a positive measure of consumers values the service at about its marginal cost of production.

In this case, the entrepreneur prefers an ICO if

$$\sum_{j=1}^{N-i_m} (v_{j+i_m} - c) \left( \alpha_{j+i_m} - \sum_{i=1}^{i_m} \alpha_i \right) \geq 0. \quad (3.26)$$

By allowing high-type consumers to commit their consumer surplus during the ICO and then releasing tokens to low-type consumers over time, the entrepreneur is able to generate a unique form of price-discrimination. She is able to get some rents from high-types who commit their future surplus during the ICO which she combines with rents from low-type consumers who she sells to over time. When the measure of low-type consumers not served by the monopolist is high relative to the measure of high-type consumers, the entrepreneur's relative preference for issuing tokens increases. This generally happens when there is a small measure of consumers who value the service very high. In this case, it is optimal for the monopolist to forego selling the service to lower types. Issuing tokens allows the entrepreneur to benefit from the high valuations of the small measure of consumers — they will fund the offering as they benefit the most from the fall in prices. At the same time, the entrepreneur sells tokens to a large measure of consumers who value the token less, generating a profit that the monopolist is not able to.

In other words, when operating as a monopolist leads to a low amount of surplus relative to the competitive equilibrium, an entrepreneur may prefer to have an ICO rather than raise funds through a traditional equity offering. The full expression when  $i_e$  can differ from  $i_m$  accounts for the fact that the entrepreneur may optimally prefer to raise money from a different collection of consumers than the ones the monopolist serves. Additionally, if  $\underline{v} > c$  the future consumer surplus that is committed to the ICO, once the competitive pricing is reached, is  $(\underline{v} - c)$  less per consumer than what the entrepreneur can get if she operates as a monopolist on the  $\alpha_{i_m}$  consumers that are served by the monopolist.

Therefore, an entrepreneur will optimally choose to issue tokens rather than be a monopolist only if future consumers participate in financing of the platform. Our model, therefore,

endogenously generates crowd-funding as imperative for an ICO's success. Since consumers account for their future consumer surplus when paying for tokens during the ICO, this also means that the price of a token during an ICO can be much higher than any future price of the token. Therefore, the model implies that token prices during ICOs are higher than secondary market prices after the ICO has taken place.

### *3.5.1. First Best Through Multiple Financing Rounds*

In our model, we can achieve the first best level of welfare by allowing the entrepreneur to have multiple financing rounds before the platform is operational. Additional financing rounds are always welfare-improving and make the entrepreneur more likely to benefit from an ICO. An extra financing round, in which the entrepreneur can charge a different price for the token than during the ICO allows the entrepreneur to price-discriminate and obtain different amounts of consumer surplus from different consumer types during financing. Consumers who value the service more, are willing to give a larger subsidy to the entrepreneur. This increases the entrepreneur's profit during financing, improving her relative payoff from having an ICO.

If there is no limit on the number of financing rounds an entrepreneur can have, they will choose to have  $N$  rounds, and release all tokens over the different financing rounds. In the case, in which a positive measure of consumers value the service at close to its marginal cost, this allows the entrepreneur to get a subsidy equal to the total consumer surplus that is generated under a competitive equilibrium during financing.<sup>11</sup> Since a monopolist's profit is always less than consumer surplus under competition, the entrepreneur would always prefer to have an ICO. Once the platform is operational, the equilibrium quantity and pricing on the platform immediately reach the competitive level, which maximizes the welfare. Formally, we can establish the following proposition if we allow the entrepreneur to break

---

<sup>11</sup>As before, each type will only commit future surplus if they believe they are pivotal. The entrepreneur can simply use soft caps, i.e. set a minimum fundraising amount, to implement this. If the entrepreneur does not hit this amount, the money is returned to investors. Smart contracts allow the implementation of such a procedure, and soft caps are commonly used by ICOs in practice.

up financing into multiple rounds.

**Proposition 3.5.** *If  $\underline{v}$  is close to  $c$ , i.e. a non-zero measure of consumers value the service at close to its marginal cost, the entrepreneur will always choose to have  $N$  financing rounds and have an active resale market. In each financing round  $i$ , the entrepreneur will sell  $\alpha_i$  tokens at a price of  $v_i + (T - 1)(v_i - c)$ .*

*Proof.* See Appendix 3A. □

Multiple rounds of financing also, in theory, benefit the entrepreneur if she operates as a monopolist and raises money from consumers. Any consumer who values the service more than  $v_{i_m}$  will be willing to subsidize the monopolist and multiple financing rounds will help the monopolist price-discriminate and get this surplus. However, since the monopolist cannot commit to charging a price any lower than  $v_{i_m}$ , the monopolist is not able to generate any surplus for consumers who value the service less than this amount. Therefore, if the entrepreneur can have multiple financing rounds, she always prefers to commit to future competitive pricing.

In practice, many ICOs have pre-sales, in which some tokens are released before the scheduled ICO. However, pre-sales are often used to attract early investors and generate interest in the ICO, so prices of tokens are typically lower in early financing than later rounds. Our analysis demonstrates that the pre-sale mechanism could in theory be used to get entrepreneurs who would otherwise be monopolists and seek rents to have an ICO instead and improve welfare.

### 3.6. Tokenized Platform vs Platform Competition under Network Effects

In this section, we compare outcomes for a platform operating with tokens and for two competing platforms. We show that if network effects are present and high enough, a tokenized platform delivers higher total welfare than two competing standard platforms. Intuitively, if network effects are relatively high it is efficient to exchange the service only

on one platform, even if competition between platforms is possible. Therefore, given our prior analysis, a tokenized platform delivers the highest total welfare in the long run.

### 3.6.1. Platform Exchange with Network Effects

To demonstrate this idea, we model network effects in the following way. We assume that a higher mass of consumers on a platform leads to a lower marginal cost for service providers, i.e.

$$c = c(\alpha), \tag{3.27}$$

with  $c'(\alpha) \leq 0$  and where  $\alpha \in [0, 1]$  is a mass of consumers on a given platform. In this case, we can prove the following.

**Proposition 3.6.** *If a platform exchange exhibits network effects, i.e.  $c'(\alpha) \leq 0$ , it follows that: i) the long-run welfare under a tokenized platform is always higher than that under two competing platforms; ii) if the magnitude of network effects is high enough, i.e.  $|c'(\alpha)| > C$  for some  $C > 0$ , the welfare under a monopolistic platform is higher than that under two competing platforms.*

*Proof.* See Appendix 3A. □

The proposition asserts that if network effects are high enough, it is inefficient to split consumers between several platforms, i.e. the efficiency gains due to network effects are higher than the efficiency gains due to competition between standard platforms. Therefore, the *long-run* welfare under a tokenized platform is always higher than that under two competing platforms. That is, under network effects, competition within a token market of a single platform is more efficient than competition between several standard platforms.

### 3.6.2. Example

As an illustration of the above result for platform exchange with network effects, we consider an example with fully specified marginal cost function and service values for consumers.

Particularly, the marginal cost for service providers as a mass of consumers  $\alpha$  on a given platform is assumed to be:

$$c(\alpha) = \frac{1}{4} - c_n \alpha, \quad (3.28)$$

where  $c_n$  is a parameter gauging the size of network effects. Thus, higher  $c_n$  means stronger network effects. Additionally, consumers' values have a linear form,<sup>12</sup> i.e. the consumer's inverse demand function is

$$v(\alpha) = \frac{1}{2} - \frac{1}{2} \alpha, \quad (3.29)$$

where the total mass of consumers is 1, as in the baseline model, while  $v(0) = \bar{v} = \frac{1}{2}$  and  $v(1) = \underline{v} = 0$ .

We next present the equilibrium outcomes in the three different scenarios: with a monopolist, with an entrepreneur, and with competing platforms.

### Monopolist

Solving the problem of a monopolist, it can be shown that the optimal mass of consumers on the platform is  $\alpha_m = \frac{1}{4(1-2c_n)}$ . The monopolist serves more consumers when network effects are stronger, i.e. when  $c_n$  is higher. The optimal price charged by the monopolist is

$$p_m = \frac{3 - 8c_n}{8(1 - 2c_n)}, \quad (3.30)$$

while the total surplus in this scenario is given by:

$$TS_m = \frac{3 - 4c_n}{4} \left( \frac{1}{4(1 - 2c_n)} \right)^2. \quad (3.31)$$

### Tokenized platform

In this scenario, the *long-run* mass of consumers served by the platform is  $\alpha_e = \frac{1}{2(1-2c_n)}$

---

<sup>12</sup>This specification diverges slightly from our inverse demand in the baseline model where we use a step function. However, it simplifies the exposition here.

and the *long-run* competitive token price in the token resale market is

$$p_e = \frac{1 - 4c_n}{4(1 - 2c_n)}. \quad (3.32)$$

The total surplus is given by:

$$TS_e = \left( \frac{1}{4(1 - 2c_n)} \right)^2. \quad (3.33)$$

### Two competing platforms

Here, the mass of consumers served by each platform is  $\alpha_c = \frac{1}{4(1-c_n)}$  while the competitive price charged by each is

$$p_c = \frac{1 - 2c_n}{4(1 - c_n)}. \quad (3.34)$$

The total surplus in this scenario is given by:

$$TS_c = \left( \frac{1}{4(1 - c_n)} \right)^2. \quad (3.35)$$

Having established the total welfare under different scenarios, we can make several observations. First, consistent with our baseline analysis, it can be seen that the *long-run* welfare under the tokenized platform is always higher than that under the monopolistic platform,  $TS_e > TS_m$ . Second, confirming the results in the Proposition 3.6, it can be shown that, when network effects are high enough,  $c_n > 0.21$ , the welfare under the monopolistic platform is higher than the welfare under the two competing platforms,  $TS_m > TS_c$ . In this case, the efficiency gains due to network effects are higher than the efficiency gains due to competition between standard platforms. Third, and by the same reason, the *long-run* welfare under the tokenized platform is always higher than that under the two competing platforms,  $TS_e > TS_c$ .

In addition, assuming specific values for  $c_n$  allows to make predictions about *short-run* com-



petition. For example, suppose  $c_n = \frac{1}{4}$ , then

$$p_m = \frac{1}{4}, TS_m = \frac{1}{8}, \alpha_m = \frac{1}{2}, c_m = \frac{1}{8}; \quad (3.36)$$

$$p_e = 0, TS_e = \frac{1}{4}, \alpha_e = 1, c_e = 0; \quad (3.37)$$

$$p_c = \frac{1}{6}, TS_c = \frac{1}{9}, \alpha_c = \frac{1}{3}, c_c = \frac{1}{6}. \quad (3.38)$$

It can be seen that, in this case, the price set by the two competing platforms is lower than the monopolistic price and, therefore, an entry of a second platform is a threat to the monopolistic platform when consumers can switch platforms with relatively low costs. If this threat is credible the monopolist cannot set the service price above  $p_c$ . Thus, the entry threat enhances welfare in this scenario, since the monopolist have to lower the price while all consumers remain on the single platform.

At the same time, the price set by two competing platforms is higher than the *long-run* token price on the tokenized platform and, therefore, an entry of the second platform is not a threat for the tokenized platform in the *long run*. However, if this threat is credible in the *short run*, the tokenized platform have to sell enough tokens in the first period so that the token price is no greater than  $p_c$ . In the subsequent periods, the token price declines to  $p_e$  as in the baseline model.

### 3.7. Conclusion

This paper shows that the ICO mechanism allows an entrepreneur to give up control of an online service exchange platform and can help her to commit to competitive pricing of the exchanged service. Due to network effects, many online service exchange platforms, which require a critical number of users to be operational, are natural monopolies and give rise to inefficient rent-seeking by their developers. Our theory can help rationalize the emergence of ICOs, many of which seek funding for such platforms. Our model demonstrates that many features of ICOs — a fixed supply of tokens, secondary market for tokens, and pre-sales — help developers to commit to competition and can greatly improve efficiency.

We show that a traditional form of financing, such as a venture capital market, in which investors benefit only from profit-sharing, cannot support the ICO mechanism as investors would prefer that developers retain control of a platform and operate as monopolists. In contrast, ICOs are only feasible when consumers themselves invest. Crowd-funding is, therefore, an integral part of the ICO market.

## APPENDIX

### Appendix 1A. Proofs Omitted from the Text

**Proof of Lemma 1.2:** Without loss of generality, suppose  $i = 1$ . We can express  $\tilde{n}X_1$  as follows:

$$\tilde{n}X_1 = \sum_{i=1}^{\tilde{n}} X_i + \left[ (\tilde{n} - 1)X_1 - \sum_{k=2}^{\tilde{n}} X_k \right], \quad (1A.1)$$

where  $\left[ (\tilde{n} - 1)X_1 - \sum_{k=2}^{\tilde{n}} X_k \right]$  has a conditional expected value of zero:

$$\mathbb{E} \left[ (\tilde{n} - 1)X_1 - \sum_{k=2}^{\tilde{n}} X_k \left| \sum_{i=1}^{\tilde{n}} X_i \right. \right] = (\tilde{n} - 1) \mathbb{E} \left[ X_1 \left| \sum_{i=1}^{\tilde{n}} X_i \right. \right] - \sum_{k=2}^{\tilde{n}} \mathbb{E} \left[ X_k \left| \sum_{i=1}^{\tilde{n}} X_i \right. \right] \stackrel{a.s.}{=} 0. \quad (1A.2)$$

It directly follows that  $\tilde{n}X_1$  is a mean-preserving spread of  $Y_j$ , and the distribution of  $Y_j$  thus second-order stochastically dominates the distribution of  $\tilde{n}X_1$ .

**Proof of Lemma 1.3:** Consider the convolution of  $Y_{\tilde{n}} = \sum_{i=1}^{\tilde{n}} X_i$  and  $X_k$  where  $k > \tilde{n}$ , that is,  $Y_{\tilde{n}+1} \equiv Y_{\tilde{n}} + X_k$ . Since these  $Y_{\tilde{n}}$  and  $X_k$  are independent, we can write:

$$g_{\tilde{n}+1}(y_{\tilde{n}+1}) = \int_0^{\bar{x}} g_{\tilde{n}}(y_{\tilde{n}+1} - x)g(x)dx. \quad (1A.3)$$

Now evaluate  $g_{\tilde{n}+1}$  at the upper bound of the support  $\bar{y}_{\tilde{n}+1} = (\tilde{n} + 1)\bar{x}$ :

$$g_{\tilde{n}+1}((\tilde{n} + 1)\bar{x}) = \int_0^{\bar{x}} g_{\tilde{n}}((\tilde{n} + 1)\bar{x} - x)g(x)dx = 0, \quad (1A.4)$$

since the density  $g_{\tilde{n}}$  is equal to zero for any outcome above  $\tilde{n}\bar{x}$ . As a result, the elasticity  $e_{\tilde{n}+1}(\bar{y}_{\tilde{n}+1}) = g_{\tilde{n}+1}(\bar{y}_{\tilde{n}+1})\bar{y}_{\tilde{n}+1}/G(\bar{y}_{\tilde{n}+1})$  is also zero for all  $\tilde{n} \geq 1$ , that is, as soon as at least two assets are pooled, such that  $\tilde{n} + 1 \geq 2$ , the elasticity of the pool will be zero at the upper bound  $\bar{y}_{\tilde{n}+1}$ .

**Proof of Lemma 1.5:** First, suppose that  $g(0) > 0$  and  $g'(0)$  is finite. By L'Hôpital's rule, the elasticity is:

$$\lim_{y \rightarrow 0} \frac{g(y)y}{G(y)} = \lim_{y \rightarrow 0} \frac{g'(y)y + g(y)}{g(y)} = \frac{g'(0)0 + g(0)}{g(0)} = 1. \quad (1A.5)$$

Next, suppose that  $g(0) = 0$ ,  $g'(0) > 0$ , and  $g''(0)$  is finite. Then the elasticity is:

$$\lim_{y \rightarrow 0} \frac{g(y)y}{G(y)} = \lim_{y \rightarrow 0} \frac{g'(y)y + g(y)}{g(y)} = \lim_{y \rightarrow 0} \frac{g''(y)y + 2g'(y)}{g'(y)} = 2. \quad (1A.6)$$

Then, suppose that  $g(0) = 0$ ,  $g'(0) = 0$ ,  $g''(0) > 0$ , and  $g'''(0)$  is finite. The elasticity is:

$$\lim_{y \rightarrow 0} \frac{g(y)y}{G(y)} = \lim_{y \rightarrow 0} \frac{g'(y)y + g(y)}{g(y)} = \lim_{y \rightarrow 0} \frac{g''(y)y + 2g'(y)}{g'(y)} = \lim_{y \rightarrow 0} \frac{g'''(y)y + 3g''(y)}{g''(y)} = 3. \quad (1A.7)$$

More generally, if the  $n$ -th derivative of the density function  $g$  is the first derivative to be positive and finite, then the elasticity is  $(n + 1)$ .

It remains to be shown that if the density function of one underlying asset is positive at the lower bound (i.e.,  $g(0) > 0$ ), then if we construct a pool of  $\tilde{n}$  assets, the first derivative of the density function of this pool that is positive (and non-zero) is the  $(\tilde{n} - 1)$ -th derivative.

Consider the convolution of  $Y_{\tilde{n}} = \sum_{i=1}^{\tilde{n}} X_i$  and  $X_k$  where  $k > \tilde{n}$ , that is,  $Y_{\tilde{n}+1} \equiv Y_{\tilde{n}} + X_k$ .

Since these  $Y_{\tilde{n}}$  and  $X_k$  are independent, we can write:

$$g_{\tilde{n}+1}(y_{\tilde{n}+1}) = \int_0^{\bar{x}} g_{\tilde{n}}(y_{\tilde{n}+1} - x)g(x)dx, \quad (1A.8)$$

and for  $0 \leq y_{\tilde{n}+1} \leq \bar{x}$  we can write:

$$g_{\tilde{n}+1}(y_{\tilde{n}+1}) = \int_0^{y_{\tilde{n}+1}} g_{\tilde{n}}(y_{\tilde{n}+1} - x)g(x)dx. \quad (1A.9)$$

Thus, the derivatives become:

$$g'_{\tilde{n}+1}(y_{\tilde{n}+1}) = g_{\tilde{n}}(0)g(y_{\tilde{n}+1}) + \int_0^{y_{\tilde{n}+1}} g'_{\tilde{n}}(y_{\tilde{n}+1} - x)g(x)dx, \quad (1A.10)$$

$$g''_{\tilde{n}+1}(y_{\tilde{n}+1}) = g_{\tilde{n}}(0)g'(y_{\tilde{n}+1}) + g'_{\tilde{n}}(0)g(y_{\tilde{n}+1}) + \int_0^{y_{\tilde{n}+1}} g''_{\tilde{n}}(y_{\tilde{n}+1} - x)g(x)dx, \quad (1A.11)$$

$$g'''_{\tilde{n}+1}(y_{\tilde{n}+1}) = g_{\tilde{n}}(0)g''(y_{\tilde{n}+1}) + g'_n(0)g'(y_{\tilde{n}+1}) + g''_{\tilde{n}}(0)g(y_{\tilde{n}+1}) + \int_0^{y_{\tilde{n}+1}} g'''_{\tilde{n}}(y_{\tilde{n}+1} - x)g(x)dx. \quad (1A.12)$$

Hence, when evaluated at  $y_{\tilde{n}+1} = 0$ , we obtain the following derivatives:

$$g'_{\tilde{n}+1}(0) = g_{\tilde{n}}(0)g(0), \quad (1A.13)$$

$$g''_{\tilde{n}+1}(0) = g_{\tilde{n}}(0)g'(0) + g'_{\tilde{n}}(0)g(0), \quad (1A.14)$$

$$g'''_{\tilde{n}+1}(0) = g_{\tilde{n}}(0)g''(0) + g'_n(0)g'(0) + g''_{\tilde{n}}(0)g(0). \quad (1A.15)$$

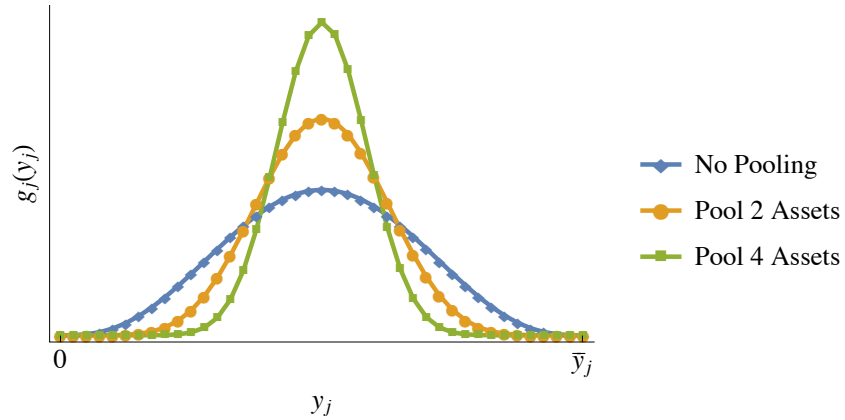
Next consider the following iteration:

- Suppose we have  $\tilde{n} = 1$ . Then  $g_1(0) = g(0) > 0$  and adding an asset yields  $g_2(0) = 0$  (see above integral), and  $g'_2(0) = g_1(0)g(0) = g(0)^2 > 0$ .
- Suppose we have  $\tilde{n} = 2$ . Then, as just shown,  $g_2(0) = 0$  and  $g'_2(0) > 0$ . Now if we add an asset, then it yields  $g_3(0) = 0$  (integral equation), and  $g'_3(0) = g_2(0)g(0) = 0$ . Now consider  $g''_3(0) = g_2(0)g'(0) + g'_2(0)g(0) = g'_2(0)g(0) > 0$ .
- Suppose we have  $\tilde{n} = 3$ . Then, as just shown,  $g_3(0) = 0$ ,  $g'_3(0) = 0$ , and  $g''_3(0) > 0$ . Now if we add an asset, then it yields  $g_4(0) = 0$  (integral equation),  $g'_4(0) = g_3(0)g(0) = 0$ , and  $g''_4(0) = g_3(0)g'(0) + g'_3(0)g(0) = 0$ . Now consider  $g'''_4(0) = g_3(0)g''(0) + g'_3(0)g'(0) + g''_3(0)g(0) = g''_3(0)g(0) > 0$ .
- ...

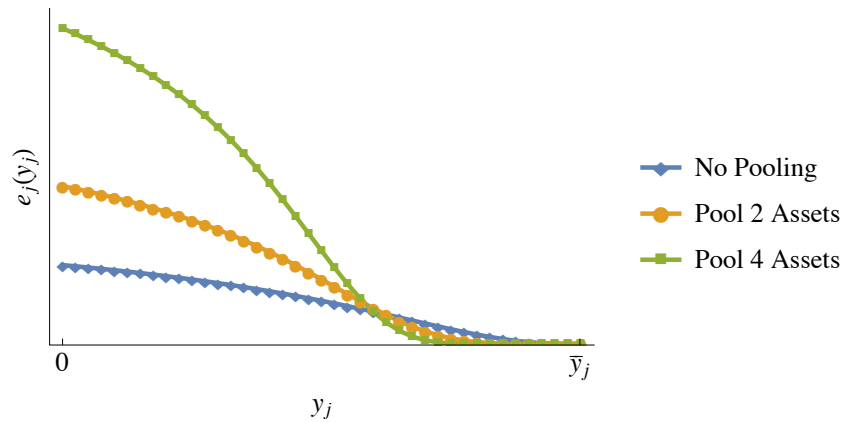
More generally, every time we add an asset to the pool, the next-higher derivative of the density function turns to zero, while leaving the derivatives thereafter positive.

## Appendix 1B. Additional Examples of Distributions

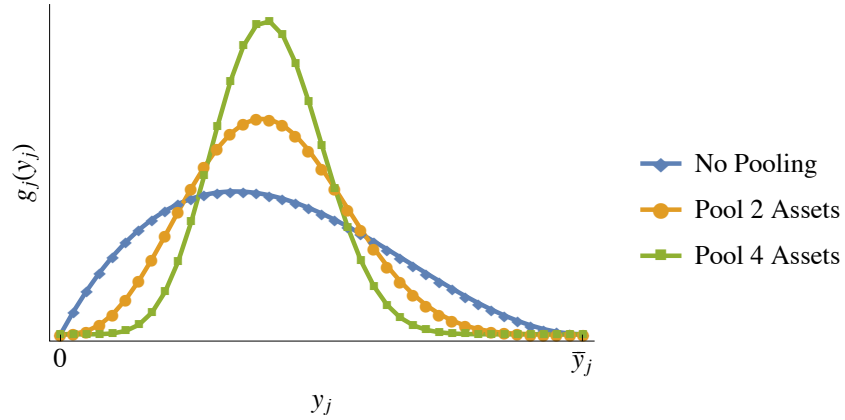
In this Appendix, we provide additional examples of distributions satisfying the assumptions of our setup (including Assumption 1.1). The figures below show the effects of pooling on the shapes of the PDF and the elasticity function.



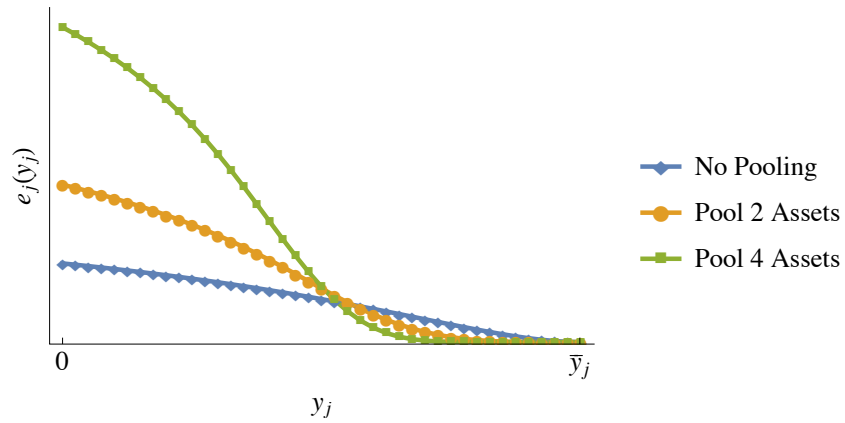
**Figure A.1: Effect of pooling on the shape of the probability density function.** The graph considers a setting with four assets ( $n = 4$ ), each of which has a payoff  $X_i$  that follows a beta distribution, with shape parameters  $\alpha = 4$  and  $\beta = 4$ , that is truncated on the interval  $[0.001, 0.999]$ . The graph plots the PDF of a separate asset, a pool of 2 assets, and a pool of 4 assets. To compare the PDFs' shapes relative to their respective domains, the graph rescales the horizontal axis to represent the interval  $\chi_j = [0, \bar{y}_j]$  for each PDF  $g_j$ .



**Figure A.2: Effect of pooling on the shape of the elasticity function.** The graph considers a setting with four assets ( $n = 4$ ), each of which has a payoff  $X_i$  that follows a beta distribution, with shape parameters  $\alpha = 4$  and  $\beta = 4$ , that is truncated on the interval  $[0.001, 0.999]$ . The graph plots the elasticity function of a separate asset, a pool of 2 assets, and a pool of 4 assets. To compare the elasticity functions' shapes relative to their respective domains, the graph rescales the horizontal axis to represent the interval  $\chi_j = [0, \bar{y}_j]$  for each elasticity function  $e_j$ .



**Figure A.3: Effect of pooling on the shape of the probability density function.** The graph considers a setting with four assets ( $n = 4$ ), each of which has a payoff  $X_i$  that follows a beta distribution, with shape parameters  $\alpha = 2$  and  $\beta = 3$ , that is truncated on the interval  $[0.001, 0.999]$ . The graph plots the PDF of a separate asset, a pool of 2 assets, and a pool of 4 assets. To compare the PDFs' shapes relative to their respective domains, the graph rescales the horizontal axis to represent the interval  $\chi_j = [0, \bar{y}_j]$  for each PDF  $g_j$ .



**Figure A.4: Effect of pooling on the shape of the elasticity function.** The graph considers a setting with four assets ( $n = 4$ ), each of which has a payoff  $X_i$  that follows a beta distribution, with shape parameters  $\alpha = 2$  and  $\beta = 3$ , that is truncated on the interval  $[0.001, 0.999]$ . The graph plots the elasticity function of a separate asset, a pool of 2 assets, and a pool of 4 assets. To compare the elasticity functions' shapes relative to their respective domains, the graph rescales the horizontal axis to represent the interval  $\chi_j = [0, \bar{y}_j]$  for each elasticity function  $e_j$ .

## Appendix 2A.

### 2A.1. Investor's problem in the secondary market

In the secondary market, the minimization problem of an impatient investor  $i$  with  $n_i$  dealers and  $k_i$  patient neighbors, which is given by the equation (2.17), can be reduced by symmetry to:

$$d_i(q_i, n_i, k_i, \bar{q}^s)/\lambda \equiv \min_{\{q_i^{ns}, q_i^s\}} k_i(q_i^{ns})^2 + (n_i - k_i)(q_i^s + \bar{q}^s)q_i^s \quad s.t. \quad k_i q_i^{ns} + (n_i - k_i)q_i^s = q_i. \quad (2A.1)$$

The FOCs to this problem are

$$k_i 2q_i^{ns} = k_i \theta \quad (2A.2)$$

$$(n_i - k_i)(2q_i^s + \bar{q}^s) = (n_i - k_i) \theta \quad (2A.3)$$

$$k_i q_i^{ns} + (n_i - k_i)q_i^s = q_i \quad (2A.4)$$

where  $\theta$  is Lagrange multiplier. Eliminating the multiplier yields:

$$2q_i^{ns} = (2q_i^s + \bar{q}^s) \quad (2A.5)$$

$$k_i(q_i^s + \bar{q}^s/2) + (n_i - k_i)q_i^s = q_i. \quad (2A.6)$$

Finally, the optimal quantities sold on the links with unstressed and stressed dealers are

$$q_i^{ns} = \frac{q_i}{n_i} + \frac{(n_i - k_i) \bar{q}^s}{n_i} \frac{1}{2} \quad (2A.7)$$

$$q_i^s = \frac{q_i}{n_i} - \frac{k_i \bar{q}^s}{n_i} \frac{1}{2}. \quad (2A.8)$$

Plugging back into the investor's objective function, the total expected discount from the secondary market of an impatient investors  $i$  with  $n_i$  dealers and  $k_i$  patient investors is

$$d_i(q_i, n_i, k_i, \bar{q}^s)/\lambda$$



$$\begin{aligned}
&= k_i \left( \frac{q_i + (n_i - k_i)\bar{q}^s/2}{n_i} \right)^2 + (n_i - k_i) \left( \frac{q_i - k_i\bar{q}^s/2}{n_i} \right) \left( \frac{q_i - k_i\bar{q}^s/2 + n_i\bar{q}^s}{n_i} \right) \\
&= \frac{1}{n_i^2} (k_i[q_i^2 + 2(n_i - k_i)q_i\bar{q}^s/2 + (n_i - k_i)^2(\bar{q}^s/2)^2] + (n_i - k_i)[q_i^2 - 2k_iq_i\bar{q}^s/2 \\
&+ k_i^2(\bar{q}^s/2)^2 + q_in_i\bar{q}^s - k_i\bar{q}^s/2n_i\bar{q}^s]) = \frac{1}{n_i^2} (q_i^2[k_i + n_i - k_i] + q_i[k_i2(n_i - k_i)\bar{q}^s/2 \\
&- (n_i - k_i)2k_iq_i\bar{q}^s/2 + (n_i - k_i)n_i\bar{q}^s] + [k_i(n_i - k_i)^2(\bar{q}^s/2)^2 + k_i^2(n_i - k_i)(\bar{q}^s/2)^2 \\
&- (n_i - k_i)k_in_i2(\bar{q}^s/2)^2]) = \frac{1}{n_i^2} (n_iq_i^2 + n_i(n_i - k_i)q_i\bar{q}^s/2 - n_ik_i(n_i - k_i)(\bar{q}^s/2)^2) \\
&= \frac{q_i^2}{n_i} + \frac{(n_i - k_i)q_i\bar{q}^s}{n_i} - \frac{k_i(n_i - k_i)(\bar{q}^s/2)^2}{n_i}. \quad (2A.9)
\end{aligned}$$

### 2A.2. Investor's problem in the primary market

The expected total discount from the secondary market of an impatient investors  $i$  with  $n_i$  dealers before realization of liquidity shocks is obtained by taking expectation of the equation (2A.9) with respect to  $k_i$ :

$$\mathbb{E}_{k_i} d_i(q_i, n_i, k_i, \bar{q}^s)/\lambda = \frac{(q_i)^2}{n_i} + (1 - p(\alpha))q_i\bar{q}^s - p(\alpha)(1 - p(\alpha))\frac{(\bar{q}^s)^2}{4}(n_i - 1). \quad (2A.10)$$

The problem of an investor  $i$  with  $n_i$  dealers is equivalent to:

$$\max_{q_i} (1 - c)q_i - \lambda(1 - p_{n_i}) \left( \frac{(q_i)^2}{n_i} + (1 - p(\alpha))q_i\bar{q}^s \right). \quad (2A.11)$$

The FOC for this problem is

$$(1 - c) - \lambda(1 - p_{n_i}) \left( 2\frac{q_i}{n_i} + (1 - p(\alpha))\bar{q}^s \right) = 0. \quad (2A.12)$$

Therefore, the demand of an investor  $i$  with  $n_i$  dealers in the primary market for the bonds is given by:

$$q_i(n_i, c) = n_i \left( \frac{(1 - c)}{2\lambda(1 - p_{n_i})} - \frac{(1 - p(\alpha))}{2}\bar{q}^s \right). \quad (2A.13)$$

2A.3. *Equilibrium demands without exclusion and  $p_n = p$*

In this case, the equilibrium average quantity traded by impatient investors on the links with impatient neighbors is simplified to

$$\bar{q}^s = \frac{\tilde{\mathbb{E}} \frac{q_n}{n}}{1 + \frac{p}{2} \tilde{\mathbb{E}} g_n}. \quad (2A.14)$$

where the expectation is taken with respect to the distribution of links in the population

$$\tilde{f}_n \equiv \frac{n f_n}{\sum_k k f(k)}.$$

Therefore, writing out the expectations explicitly, yields:

$$\bar{q}^s = \frac{\sum_n \frac{q_n}{n} n f_n}{\sum_n n f_n + \frac{p}{2} \sum_n g_n n f_n}. \quad (2A.15)$$

Additionally, it is established that the primary market demands as a function of  $\bar{q}^s$  are

$$q(n, c) = n \frac{(1-c)/\lambda - (1-p)^2 \bar{q}^s}{2(1-p)}. \quad (2A.16)$$

Thus, plugging it into the equation for  $\bar{q}^s$ , obtain:

$$\bar{q}^s = \frac{\sum_n \frac{(1-c)/\lambda - (1-p)^2 \bar{q}^s}{2(1-p)} n f_n}{\sum_n n f_n + \frac{p}{2} \sum_n g_n n f_n}. \quad (2A.17)$$

Collecting terms:

$$\bar{q}^s = \frac{\frac{(1-c)}{2(1-p)\lambda} \sum_n n f_n}{\sum_n n f_n + \frac{p}{2} \sum_n g_n n f_n + \frac{(1-p)^2}{2(1-p)} \sum_n n f_n} \quad (2A.18)$$

and simplifying:

$$\bar{q}^s = \frac{(1-c)/\lambda \cdot \sum_n n f_n}{2(1-p)(\sum_n n f_n + \frac{p}{2} \sum_n g_n n f_n) + (1-p)^2 \sum_n n f_n}. \quad (2A.19)$$

Finally, plugging back to the primary market demand equation obtain its final expression:

$$\begin{aligned}
q(n, c) &= n \frac{(1-c)}{2(1-p)\lambda} \frac{2(1-p)(\sum_n n f_n + \frac{p}{2} \sum_n g_n n f_n)}{2(1-p)(\sum_n n f_n + \frac{p}{2} \sum_n g_n n f_n) + (1-p)^2 \sum_n n f_n} \\
&= n \frac{(1-c)}{\lambda} \frac{\sum_n n f_n + \frac{p}{2} \sum_n g_n n f_n}{2(1-p)(\sum_n n f_n + \frac{p}{2} \sum_n g_n n f_n) + (1-p)^2 \sum_n n f_n}. \quad (2A.20)
\end{aligned}$$

### Homogenous investors.

When investors are homogenous they all have the same number of dealers  $n$ . In this case, the individual primary market demand from the above is simplified to:

$$q(n, c) = n \frac{(1-c)}{\lambda} \frac{n f_n + \frac{p}{2} g_n n f_n}{2(1-p)(n f_n + \frac{p}{2} g_n n f_n) + (1-p)^2 n f_n} \quad (2A.21)$$

or

$$q(n, c) = n \frac{(1-c)}{\lambda} \frac{1 + \frac{p}{2} g_n}{2(1-p)(1 + \frac{p}{2} g_n) + (1-p)^2}. \quad (2A.22)$$

Thus, the total demand is

$$Q(c) = q(n, c) f_n = \frac{(1-c)}{\lambda} \frac{1 + \frac{p}{2} g_n}{2(1-p)(1 + \frac{p}{2} g_n) + (1-p)^2} n f_n. \quad (2A.23)$$

#### 2A.4. Equilibrium demands with exclusion

If the issuer chooses fill rates  $\alpha_n$  for investors with  $n$  dealers then the average quantity sold by impatient investors on the links with impatient neighbors is

$$\bar{q}^s = \frac{\frac{\tilde{\mathbb{E}}_n g_n (1-p_n) \alpha_n}{\tilde{\mathbb{E}}(1-p_n) \alpha_n}}{1 + \frac{p(\alpha)}{2} \frac{\tilde{\mathbb{E}}_n g_n (1-p_n) \alpha_n}{\tilde{\mathbb{E}}(1-p_n) \alpha_n}}. \quad (2A.24)$$

where the expectation is taken with respect to the distribution of links in the population  $\tilde{f}_n \equiv \frac{n f_n}{\sum_k k f(k)}$  and the equation accounts for the fact that only a fraction  $\alpha_n$  of investors with  $n$  dealers receive their desired allocation in the primary market.

Therefore, writing out the expectations explicitly, yields:

$$\bar{q}^s = \frac{\frac{\sum_{n \in N} \frac{q_n}{n} (1-p_n) \alpha_n n f_n}{\sum_{n \in N} (1-p_n) \alpha_n n f_n}}{1 + \frac{p(\alpha)}{2} \frac{\sum_{n \in N} g_n (1-p_n) \alpha_n n f_n}{\sum_{n \in N} (1-p_n) \alpha_n n f_n}}. \quad (2A.25)$$

The expression for the primary market demand as a function of  $\bar{q}^s$  is given by:

$$\frac{q(n, c)}{n} = \frac{(1-c)}{2\lambda(1-p_n)} - \frac{(1-p(\alpha))}{2} \bar{q}^s. \quad (2A.26)$$

Define  $a_n \equiv \frac{(1-c)}{2\lambda(1-p_n)}$ . Plugging the primary demands into the equation for  $\bar{q}^s$ , obtain:

$$\bar{q}^s = \frac{\frac{\sum_{n \in N} a_n (1-p_n) \alpha_n n f_n}{\sum_{n \in N} (1-p_n) \alpha_n n f_n} - \frac{(1-p(\alpha))}{2} \bar{q}^s}{1 + \frac{p(\alpha)}{2} \frac{\sum_{n \in N} g_n (1-p_n) \alpha_n n f_n}{\sum_{n \in N} (1-p_n) \alpha_n n f_n}}. \quad (2A.27)$$

Collecting terms:

$$\bar{q}^s = \frac{\frac{\sum_{n \in N} a_n (1-p_n) \alpha_n n f_n}{\sum_{n \in N} (1-p_n) \alpha_n n f_n}}{1 + \frac{p(\alpha)}{2} \frac{\sum_{n \in N} g_n (1-p_n) \alpha_n n f_n}{\sum_{n \in N} (1-p_n) \alpha_n n f_n} + \frac{(1-p(\alpha))}{2}}. \quad (2A.28)$$

Plugging back to the equation for the primary market demand obtain:

$$\frac{q(n, c)}{n} = a_n - \frac{(1-p(\alpha))}{2} \frac{\frac{\sum_{n \in N} a_n (1-p_n) \alpha_n n f_n}{\sum_{n \in N} (1-p_n) \alpha_n n f_n}}{1 + \frac{p(\alpha)}{2} \frac{\sum_{n \in N} g_n (1-p_n) \alpha_n n f_n}{\sum_{n \in N} (1-p_n) \alpha_n n f_n} + \frac{(1-p(\alpha))}{2}}. \quad (2A.29)$$

Since the expanded expression for  $1-p(\alpha)$  is

$$1-p(\alpha) = \frac{\sum_n (1-p_n) n f_n \alpha_n}{\sum_n n f_n}, \quad (2A.30)$$

the above becomes

$$\frac{q(n, c)}{n} = a_n - \frac{\frac{1}{2} \frac{\sum_{n \in N} a_n (1-p_n) \alpha_n n f_n}{\sum_{n \in N} n f_n}}{1 + \frac{p(\alpha)}{2} \frac{\sum_{n \in N} g_n (1-p_n) \alpha_n n f_n}{\sum_{n \in N} (1-p_n) \alpha_n n f_n} + \frac{(1-p(\alpha))}{2}}. \quad (2A.31)$$

Finally, simplifying, the primary market demand is

$$\frac{q(n, c)}{n} = \frac{(1-c)}{2\lambda} \left( \frac{1}{(1-p_n)} - \frac{\frac{1}{2} \frac{\sum_{n \in N} \alpha_n n f_n}{\sum_{n \in N} n f_n}}{1 + \frac{p(\alpha)}{2} \frac{\sum_{n \in N} g_n (1-p_n) \alpha_n n f_n}{\sum_{n \in N} (1-p_n) \alpha_n n f_n} + \frac{(1-p(\alpha))}{2}} \right). \quad (2A.32)$$

Thus, the total primary market demand is

$$Q_d(c, \alpha) = \frac{(1-c)}{2\lambda} \left( \sum_{n \in N} \frac{\alpha_n n f_n}{(1-p_n)} - \frac{\frac{1}{2} \frac{\sum_{n \in N} \alpha_n n f_n}{\sum_{n \in N} n f_n} \sum_{n \in N} \alpha_n n f_n}{1 + \frac{p(\alpha)}{2} \frac{\sum_{n \in N} g_n (1-p_n) \alpha_n n f_n}{\sum_{n \in N} (1-p_n) \alpha_n n f_n} + \frac{(1-p(\alpha))}{2}} \right). \quad (2A.33)$$

#### 2A.5. Correlated liquidity shocks

Since any investor does not know if she has a neighbor with correlated liquidity shock or not all investors solve the same problem in the primary market. The investor's maximization problem is

$$\max_{q_i} (1-c)q_i - (1-p)\lambda \left( (1-\rho)(q_i)^\theta + \rho(q_i + q^s)^\theta \right) q_i. \quad (2A.34)$$

The FOC for the problem is

$$(1-c) - (1-p)\lambda \left( (1-\rho)(\theta+1)(q_i)^\theta + \rho((q_i + q^s)^\theta + \theta(q_i + q^s)^{\theta-1} q_i) \right) = 0. \quad (2A.35)$$

Simplifying:

$$(1-c) - (1-p)\lambda \left( (1-\rho)(\theta+1)(q_i)^\theta + \rho((1+\theta)q_i + q^s)(q_i + q^s)^{\theta-1} \right) = 0. \quad (2A.36)$$

Due to the symmetry, the quantity sold by an impatient neighbor  $q^s = q_i$  in the equilibrium.

Therefore, an investor's demand in the primary market is

$$q_i(c) = \left( \frac{(1-c)}{\lambda(1-p)} \frac{1}{(1-\rho)(\theta+1) + \rho(2+\theta)2^{\theta-1}} \right)^{\frac{1}{\theta}}. \quad (2A.37)$$

## 2A.6. Proofs

*Proof of Proposition 2.3.* The results of Glaeser and Scheinkman (2002) can be used to prove existence and uniqueness of the equilibrium. For this, rewrite the investor's utility in the primary market in per-dealer terms:

$$U_i^n(q_i^n, \bar{q}^s, c) \equiv U_i(q_i, n_i, c)/n_i = (1-c)q_i^n - \lambda(1-p_{n_i})(q_i^n)^2 - \lambda(1-p_{n_i})^2 q_i^n \bar{q}^s \quad (2A.38)$$

where  $q_i^n \equiv \frac{q_i}{n_i}$ . Thus, the investor's original problem is equivalent to the maximization of  $U_i^n(q_i^n, \bar{q}^s, c)$  with respect to  $q_i^n$ . Note that  $\bar{q}^s$  represents the scaled weighted quantity demanded by the reference group.

From Proposition 1 of Glaeser and Scheinkman (2002), there exists an equilibrium if  $\forall \lambda, p, c$   $\exists \bar{q} \geq 0$  such that  $\forall \bar{q}^s \in [-\bar{q}, \bar{q}]$  the following two conditions hold

$$\frac{\partial U_i^n}{\partial q_i^n}(-\bar{q}, \bar{q}^s, c) \geq 0, \quad (2A.39)$$

$$\frac{\partial U_i^n}{\partial q_i^n}(\bar{q}, \bar{q}^s, c) \leq 0. \quad (2A.40)$$

Since the partial derivative decreases in  $\bar{q}^s$  it follows  $\frac{\partial U_i^n}{\partial q_i^n}(-\bar{q}, \bar{q}^s, c) \geq \frac{\partial U_i^n}{\partial q_i^n}(-\bar{q}, \bar{q}, c)$  and  $\frac{\partial U_i^n}{\partial q_i^n}(\bar{q}, \bar{q}^s, c) \leq \frac{\partial U_i^n}{\partial q_i^n}(\bar{q}, -\bar{q}, c)$ . Therefore, we need to verify that

$$\frac{\partial U_i^n}{\partial q_i^n}(-\bar{q}, \bar{q}, c) = 1 - c + \lambda(1-p)[2 - (1-p)]\bar{q} \geq 0, \quad (2A.41)$$

$$\frac{\partial U_i^n}{\partial q_i^n}(\bar{q}, -\bar{q}, c) = 1 - c - \lambda(1-p)[2 - (1-p)]\bar{q} \leq 0 \quad (2A.42)$$

holds  $\forall \lambda, p, c$  for some  $\bar{q}$  which is the case since  $\lambda(1-p)[2 - (1-p)] > 0$ .

Additionally, from Proposition 3 of Glaeser and Scheinkman (2002), the equilibrium is unique if

$$\left| \frac{\partial^2 U_i^n}{\partial q_i^n \partial \bar{q}^s}(q_i^n, \bar{q}^s, c) \right| / \left| \frac{\partial^2 U_i^n}{(\partial q_i^n)^2}(q_i^n, \bar{q}^s, c) \right| < 1 \quad (2A.43)$$

which is the case since the term in the brackets is equal to  $\frac{1-p}{2}$ .  $\square$

*Proof of Proposition 2.5.* From the above, the total primary market demand

$Q_d(c, \alpha) = \sum_{n \in N} n \frac{q(n, c)}{n} \cdot \alpha_n \cdot f_n$  is given by:

$$Q_d(c, \alpha) = \frac{(1-c)}{2\lambda} \left( \sum_{n \in N} \frac{\alpha_n n f_n}{(1-p_n)} - \frac{\frac{1}{2} \frac{\sum_{n \in N} \alpha_n n f_n}{\sum_{n \in N} n f_n} \sum_{n \in N} \alpha_n n f_n}{1 + \frac{p(\alpha)}{2} \frac{\sum_{n \in N} g_n (1-p_n) \alpha_n n f_n}{\sum_{n \in N} (1-p_n) \alpha_n n f_n} + \frac{(1-p(\alpha))}{2}} \right). \quad (2A.44)$$

Since the price  $c$  enters investors' optimal demand multiplicatively the issuer's decisions about the optimal premium  $\pi = c - \delta$  and potential exclusion of any investors separate. In particular, the issuer's profit  $V_d = \pi \cdot Q_d(c, \alpha)$  is maximized at the premium  $\pi = \frac{1-\delta}{2}$  and, thus, the optimal issuance price is  $c = \frac{1+\delta}{2}$ .

To find out if there is any exclusion of investors with  $n$  dealers, take a derivative with respect to  $\alpha_n$  at a general profile of  $\alpha = (\alpha_n, \dots, \alpha_{\bar{n}})$ :

$$(Q_d(c, \alpha))'_{\alpha_n} = \frac{q(n, c)}{n} \cdot n f_n + \left( \frac{q(n, c)}{n} \right)'_{\alpha_n} \sum_{k \in N} \alpha_k \cdot k f_k \quad (2A.45)$$

The first term is positive and is due to the additional allocation to investors with  $n$  dealers while the second term is negative and is due to the decrease in demands of all other investors participating in the primary market. In the following, I show that the total derivative can be negative, if investors are heterogeneous with respect to the number of dealer connections and the assumption of increasing relative patience is satisfied. In that case, the issuer excludes investors from the primary market.

To avoid cumbersome derivations, I present the total derivative for an investor network with two levels of connectedness which is analogous to the derivative in the general case and allows to illustrate the trade-off. In particular, consider an investor network such that a mass  $f_1$  of investors have  $n_1$  dealers each and a mass  $f_2$  of investors have  $n_2$  dealers each. Investors from the two groups remain patient in the secondary market with probabilities  $p_1$

and  $p_2 < p_1$ , respectively. Then,

$$p(\alpha) = \frac{p_1\alpha_1 n_1 f_1 + p_2\alpha_2 n_2 f_2 + (1 - \alpha_1)n_1 f_1 + (1 - \alpha_2)n_2 f_2}{n_1 f_1 + n_2 f_2} \quad (2A.46)$$

and

$$gp(\alpha) \equiv \frac{g_1(1 - p_1)\alpha_1 n_1 f_1 + g_2(1 - p_2)\alpha_2 n_2 f_2}{(1 - p_1)\alpha_1 n_1 f_1 + (1 - p_2)\alpha_2 n_2 f_2}. \quad (2A.47)$$

Thus, the total demand is

$$Q_d(c, \alpha) = \frac{(1 - c)}{2\lambda} \frac{1}{\sum_{n \in N} n f_n} \left( (n_1 f_1 + n_2 f_2) \sum_{n \in N} \frac{\alpha_n n f_n}{(1 - p_n)} - \frac{(\alpha_1 n_1 f_1 + \alpha_2 n_2 f_2)^2}{2 + p(\alpha)gp(\alpha) + 1 - p(\alpha)} \right). \quad (2A.48)$$

It can be verified that the derivative of the total demand with respect to  $\alpha_2$  is positive for any  $\alpha = (0, \alpha_2)$  with  $0 \leq \alpha_2 \leq 1$ . Therefore, it is not optimal to exclude any more connected investors. Next, take the derivative of the total demand with respect to  $\alpha_1$  at the profile  $\alpha = (0, 1)$  to determine if it is optimal to include any less connected investors. The derivatives of  $p(\alpha)$  and  $gp(\alpha)$  at  $\alpha = (0, 1)$  are

$$p'_{\alpha_1}(\alpha) = -\frac{(1 - p_1)n_1 f_1}{n_1 f_1 + n_2 f_2} \quad (2A.49)$$

and

$$gp'_{\alpha_1}(\alpha) = \frac{(g_1 - g_2)(1 - p_1)n_1 f_1}{(1 - p_2)n_2 f_2}. \quad (2A.50)$$

Plugging to the derivative of the total demand, the term which determines its sign is

$$\begin{aligned} & (n_1 f_1 + n_2 f_2) \frac{n_1 f_1}{(1 - p_1)} (2 + p(\alpha)gp(\alpha) + 1 - p(\alpha))^2 \\ & - 2(n_1 f_1)(n_2 f_2) (2 + p(\alpha)gp(\alpha) + 1 - p(\alpha)) + (n_2 f_2)^2 (p'_{\alpha_1} gp(\alpha) + p(\alpha)gp'_{\alpha_1}(\alpha) - p'_{\alpha_1}(\alpha)). \end{aligned} \quad (2A.51)$$

The first line comes from taking the derivative of the first term in (2A.48), while the second line is from taking the derivative of the second term in (2A.48). Define  $z = \frac{n_1 f_1}{n_2 f_2}$ .



Substituting the equations for  $p(\alpha)$ ,  $gp(\alpha)$  and their derivatives and extracting common factors yields:

$$\begin{aligned} & \frac{(n_1 f_1)(n_2 f_2)^2}{(n_1 f_1 + n_2 f_2)} \left( \frac{1}{(1-p_1)} \left( 2(z+1) + (p_2+z)g_2 + 1 - p_2 \right) \right)^2 \\ & - 2 \left( 2(z+1) + (p_2+z)g_2 + 1 - p_2 \right) + (1-p_1)(1-g_2) + (p_2+z) \frac{(1-p_1)(g_1-g_2)}{(1-p_2)}. \end{aligned} \tag{2A.52}$$

Since  $n_1 < n_2$  and, thus,  $g_1 < g_2$  the last term is negative. Additionally, the term is high in absolute value when  $p_2$  is high and  $p_1$  is low. Therefore, since all other terms are bounded the sign of the total derivative can be negative. In this case, it is optimal for the issuer to exclude less connected investors, who have  $n_1$  dealers, from the primary market.

There are two more notes to make here. First, the magnitude of the first term, which is positive, is primarily determined by  $z = \frac{n_1 f_1}{n_2 f_2}$  and increases with it. Therefore, the exclusion is more likely when  $z$  is small, i.e., when either  $n_1$  is small or  $f_1$  small. Second, the exclusion result requires both assumptions: i) that  $n_1 < n_2$  and ii) that  $p_1 < p_2$  because both them are needed to generate the large negative term in the derivative.

In the general case, because of increasing  $g_n$  and  $p_n$ , if the derivative with respect to  $\alpha_n$  is positive for the profile  $\alpha = (0, \dots, 0, \alpha_n, \dots, \alpha_{\bar{n}})$ , then the derivative with respect to  $\alpha_k$  at  $\alpha$  is positive for all  $k > n$ . At the same time, if the derivative with respect to  $\alpha_n$  is negative for a profile  $\alpha = (0, \dots, 0, \alpha_n, \dots, \alpha_{\bar{n}})$ , the derivative with respect to  $\alpha_{n+1}$  is higher for a profile  $\alpha = (0, \dots, 0, \alpha_{n+1}, \dots, \alpha_{\bar{n}})$ .

Finally, from the above, the derivative with respect to  $\alpha_{\bar{n}}$  is always positive for a profile  $\alpha = (0, \dots, 0, \alpha_{\bar{n}})$ . At the same time, the derivative with respect to  $\alpha_n$  can be negative for low  $n$ , starting from  $\underline{n}$ , for a profile  $\alpha = (0, \dots, 0, \alpha'_n, \dots, \alpha_{\bar{n}})$ . Therefore, the issuer might find it optimal to set a profile of  $\alpha = (0, \dots, 0, \alpha'_n, \dots, \alpha_{\bar{n}})$ , i.e., excluding some investors from the primary market.  $\square$

*Proof of Proposition 2.6.* From the proof of Proposition 2.5, the main term that affects the exclusion decision is

$$gp(\alpha) \equiv \frac{\sum_{n \in N} g_n(1-p_n)\alpha_n n f_n}{\sum_{n \in N} (1-p_n)\alpha_n n f_n}. \quad (2A.53)$$

When less connected investors participate in the primary market this term decreases. If the magnitude of the decline is sufficiently high the total demand decreases and the issuer prefers to exclude these investors.

The derivative of the term with respect to  $\alpha_n$  at the profile  $\alpha = (\alpha_n, \dots, \alpha_{\bar{n}})$  is

$$gp'_{\alpha_n}(\alpha) = -\frac{(1-p_n)n f_n \sum_{k \in N} (g_k - g_n)(1-p_k)\alpha_k k f_k}{(\sum_{k \in N} (1-p_k)\alpha_k k f_k)^2}. \quad (2A.54)$$

Since  $g_n$  is increasing the derivative is negative for small  $n$ . It can be large in magnitude if  $p_l$  and  $g_l$  are small for  $l < n'$  while  $p_h$  and  $g_h$  are high for  $h \geq n'$  which is the case when  $p_{\bar{n}} - p_n$  is high and  $\bar{n} - n$  is high.

Additionally, from the proof of Proposition 2.5 the positive part of the derivative of the total demand for small  $n$  is smaller when the distribution  $\tilde{f}_n$  is skewed to the right which is the case when either  $\bar{n} - n$  is high or the distribution  $f_n$  is skewed to the right. □

*Proof of Proposition 2.7.* The comparison of the total demand functions in the two cases reduces to the comparison of:

$$\left(1 - \frac{\rho}{2}\right) \left(\frac{1}{(\theta+1)}\right)^{\frac{1}{\theta}} \quad \vee \quad \left(\frac{1}{(1-\rho)(\theta+1) + \rho(2+\theta)2^{\theta-1}}\right)^{\frac{1}{\theta}} \quad (2A.55)$$

which is equivalent to:

$$\left(1 - \frac{\rho}{2}\right)^{\theta} \left((1-\rho)(\theta+1) + \rho(2+\theta)2^{\theta-1}\right) \quad \vee \quad (\theta+1). \quad (2A.56)$$

The two sides are equal when  $\rho = 0$  while the LHS is smaller than the RHS when  $\rho = 1$ . The derivative of the LHS with respect to  $\rho$  is

$$\left(1 - \frac{\rho}{2}\right)^{\theta-1} \left(-\frac{\theta}{2} \left[(1 - \rho)(\theta + 1) + \rho(2 + \theta)2^{\theta-1}\right] + \left(1 - \frac{\rho}{2}\right) \left[-(\theta + 1) + (2 + \theta)2^{\theta-1}\right]\right). \quad (2A.57)$$

At  $\rho = 0$ , it is monotone and equal to:

$$\frac{1}{2}(2 + \theta)(2^\theta - 1 - \theta) \quad (2A.58)$$

which is positive for  $\theta > 1$ . Thus, since the derivative of the RHS is zero for all  $\rho$ , the LHS is larger than the RHS for small  $\rho$  and  $\theta > 1$ .

Additionally, since the derivative of LHS changes its sign only once the inequality holds for all  $\rho \leq \bar{\rho}(\theta)$ . Finally, applying the implicit function theorem it can be shown that  $\bar{\rho}(\theta)$  is increasing in  $\theta$ .

□

## Appendix 2B.

### 2B.1. Dealers' liquidity supply

Assume that a dealer can resell the amount of bonds  $q$  that it bought from its client investors in the secondary market to some patient investors or on the inter-dealer market after  $t = 1$ . Since all liquidity shocks are realized by that moment the price that patient investors are willing to pay for a bond unit is 1 which is equal to a bond's payout at maturity. Assume also that to locate these patient investors a dealer incurs a holding cost  $h(q)$  which is increasing convex function of  $q$  — the total amount that it has to resell. Then, the total dealer's profit is

$$u_d(q) = q - P(q)q - h(q). \quad (2B.1)$$

Therefore, to breakeven, a dealer have to charge the price:

$$P(q) = \frac{q - h(q)}{q} = 1 - \frac{h(q)}{q} \quad (2B.2)$$

that is downward-sloping since  $h(q)$  is increasing convex function.

Finally, if it is further assumed that  $h(q) = \lambda q^{\theta+1}$ , dealers charge the price  $P(q) = 1 - \lambda q^{\theta}$ .

**Intuition.**

Dealers incur holding costs between the secondary market and inter-dealer market. They pass these costs to their client investors in the secondary market. At the same time, they charge the price which is equal to the bond payoff after the liquidity shocks are realized. Therefore, with some probability, the primary market investors have to incur the liquidation cost in the secondary market which is convex function of their holdings. The presence of liquidation cost generates a downward-sloping investor demand in the primary market. Additionally, the liquidation cost also depends on how many other investors around them in the trading network liquidate their holdings.

Appendix 2C.

*2C.1. General form of secondary market discount*

Assume that a discount offered by dealers takes a general form and is given by an increasing convex function  $d(q_i)$ , i.e.,  $d'(\cdot) > 0$  and  $d''(\cdot) > 0$ .

An ex-ante total discount from the secondary market of an impatient investor  $i$  with one dealer,  $n = 1$ :

$$pq_i d(q_i) + (1 - p)q_i d(q_i + \bar{q}^s) \quad (2C.1)$$

where  $\bar{q}^s$  is a certainty equivalent of a quantity of bonds traded by impatient neighbors. The investor faces uncertainty about whether she will have to trade with stressed or unstressed dealer.

The derivative of the discount with respect to  $q_i$ , which determines ex-ante willingness of the investor to buy bonds, is given by:

$$p[d(q_i) + q_i d'(q_i)] + (1 - p)[d(q_i + \bar{q}^s) + q_i d'(q_i + \bar{q}^s)]. \quad (2C.2)$$

In contrast, an impatient investor  $i$  with  $n \rightarrow \infty$  dealers is almost sure that the fraction  $1 - p$  of her neighbors becomes impatient. Thus, her ex-ante discount from the secondary market is given by the following problem:

$$\min_{\{q_i^{ns}, q_i^s\}} p q_i^{ns} d(q_i^{ns}) + (1 - p) q_i^s d(q_i^s + \bar{q}^s) \quad s.t. \quad p q_i^{ns} + (1 - p) q_i^s = q_i \quad (2C.3)$$

where  $q_i$  is a quantity of bonds per dealer. The discount is smaller when  $n \rightarrow \infty$ , compared to the case of  $n = 1$ , due to a better ability of the investor to move her trades between a greater number of dealers.

The optimal  $q_i^{ns}$  and  $q_i^s$  are given by the constraint in the previous equation and the following FOC:

$$d(q_i^{ns}) + q_i^{ns} d'(q_i^{ns}) = d(q_i^s + \bar{q}^s) + q_i^s d'(q_i^s + \bar{q}^s). \quad (2C.4)$$

Naturally,  $q_i^s < q_i < q_i^{ns}$ , i.e., the investor with  $n \rightarrow \infty$  sells lower quantities to stressed dealers compared to those traded with non-stressed dealers.

Finally, by the envelope theorem, the derivative of the discount with respect to  $q_i$ , which determines ex-ante willingness of the investor to buy bonds, is

$$d(q_i^{ns}) + q_i^{ns} d'(q_i^{ns}) = p[d(q_i^{ns}) + q_i^{ns} d'(q_i^{ns})] + (1 - p)[d(q_i^s + \bar{q}^s) + q_i^s d'(q_i^s + \bar{q}^s)]. \quad (2C.5)$$

Thus, the derivative of the discount for  $n \rightarrow \infty$  lies weakly below the derivative of the discount for  $n = 1$ . As a result, the discount does not grow as fast with  $q_i$  in the former case and an investor with  $n \rightarrow \infty$  is willing to buy weakly more  $q_i$  in the primary market.

However, it can also be seen that  $q_i^s$  for  $n \rightarrow \infty$  cannot be larger than  $q_i$  for  $n = 1$  since it is necessary that  $q_i^s < q_i^{ns}$  and respective derivatives should be zero at  $q_i(n \rightarrow \infty)$  and  $q_i(n = 1)$ . Therefore, an investor with  $n \rightarrow \infty$  sells less to stressed dealers even if she potentially buys more in the primary market.

### 2C.2. Exponential secondary market discount (limit cases)

Assume the specific functional form of the discount offered by a dealer:

$$d(q_i) = (q_i)^\theta \quad (2C.6)$$

with  $\theta \geq 1$ .

Substituting it into the expressions from the previous section, the ex-ante discount from the secondary market of an impatient investor  $i$  with one dealer,  $n = 1$ :

$$p(q_i)^{\theta+1} + (1-p)q_i(q_i + \bar{q}^s)^\theta. \quad (2C.7)$$

while the derivative of the discount with respect to  $q_i$  is

$$p(\theta + 1)(q_i)^\theta + (1-p)[(\theta + 1)q_i + \bar{q}^s](q_i + \bar{q}^s)^{\theta-1}. \quad (2C.8)$$

Thus, when  $\theta = 1$  the discount of an investor with  $n = 1$  is

$$(q_i)^2 + (1-p)q_i\bar{q}^s. \quad (2C.9)$$

Similarly, substituting into the above, the ex-ante discount of an impatient investor  $i$  with  $n \rightarrow \infty$  dealers is given by the following problem:

$$\min_{\{q_i^{ns}, q_i^s\}} p(q_i^{ns})^{\theta+1} + (1-p)q_i^s(q_i^s + \bar{q}^s)^\theta \quad s.t. \quad pq_i^{ns} + (1-p)q_i^s = q_i. \quad (2C.10)$$

while the derivative of the discount with respect to  $q_i$  is:

$$(\theta + 1)(q_i^{ns})^\theta = p(\theta + 1)(q_i^{ns})^\theta + (1 - p)[(\theta + 1)q_i^s + \bar{q}^s](q_i^s + \bar{q}^s)^{\theta-1}. \quad (2C.11)$$

Therefore, when  $\theta = 1$  and  $n \rightarrow \infty$  the optimal quantities are

$$q_i^{ns} = q_i + (1 - p)\frac{\bar{q}^s}{2} \quad (2C.12)$$

$$q_i^s = q_i - p\frac{\bar{q}^s}{2} \quad (2C.13)$$

and, the ex-ante discount of an investor is

$$(q_i)^2 + (1 - p)q_i\bar{q}^s - p(1 - p)\frac{(\bar{q}^s)}{4}. \quad (2C.14)$$

The last term explicitly shows the decrease in the expected secondary market discount when  $n \rightarrow \infty$  due to a better ability of the investor to move her trades between dealers compared to the case of  $n = 1$ .

However, it can be also seen that since the derivatives of the discounts with respect to  $q_i$  in the two cases coincide, the ex-ante choice of  $q_i$  is identical for the two investors with  $n = 1$  and with  $n \rightarrow \infty$ . Thus, the benefit from diversification of sales is private and does not affect initial choice of  $q_i$ . Therefore, the stressed investor with  $n = 1$  sells  $q_i$  on links with impatient neighbors while the stressed investor with  $n \rightarrow \infty$  sells a smaller amount on such links, as  $q_i^s = q_i - p\frac{\bar{q}^s}{2} < q_i$ .

### Appendix 3A. Proofs Omitted from the Text

*Proof of Proposition 3.1.* We formally prove the proposition by backward induction. In Section 3.3, we have already shown that the statement of the proposition holds when there are 2 consumer types. Therefore, we need to show that if the entrepreneur optimally releases tokens to  $N - 1$  consumer types in  $N - 1$  periods, then she finds it optimal to release tokens

to  $N$  consumer types in  $N$  periods.

Without loss of generality, suppose the additional  $N$ -th consumer type is the one that has the highest value for the service. Define also the entrepreneur's optimal payoff that she obtains when she releases tokens to  $N-1$  lower consumer types in  $N-1$  periods as  $V_{N-1}^*(\alpha_2, \dots, \alpha_N)$ . Given this definition, if the entrepreneur serves all consumers of the highest type in the first period, we reach the induction step and the entrepreneur optimally releases her remaining tokens in the remaining  $N-1$  periods for the payoff  $V_{N-1}^*(\alpha_2, \dots, \alpha_N)$ . Consequently, we need to show that the entrepreneur does not have incentives to speed up the release of tokens by serving two or more consumer types in the first period.

Specifically, consider the two possibilities. If the entrepreneur releases  $q_1 = \alpha_1$  tokens in the first period then the token price is  $p_1 = v_1 = \bar{v}$  and her continuation payoff is  $V_{N-1}^*(\alpha_2, \dots, \alpha_N)$ . Clearly, there is no incentive to release  $q_1 < \alpha_1$  tokens since the token price is the same when  $q_1 = \alpha_1$ . If, however, the entrepreneur releases slightly more tokens  $q_1 = \alpha_1 + \epsilon$  then their price falls below the value of the highest consumer type,  $p_1 = v_2 < v_1 = \bar{v}$ , and the entrepreneur's continuation payoff also decreases,  $V_{N-1}^*(\alpha_2 - \epsilon, \dots, \alpha_N) < V_{N-1}^*(\alpha_2, \dots, \alpha_N)$ , because her remaining stock of tokens gets smaller. Finally, since

$$\alpha_1 v_1 + V_{N-1}^*(\alpha_2, \dots, \alpha_N) > (\alpha_1 + \epsilon) v_2 + V_{N-1}^*(\alpha_2 - \epsilon, \dots, \alpha_N) \quad (3A.1)$$

the former release schedule yields a higher total payoff. Therefore, it is suboptimal for entrepreneur to speed up the release of tokens in the first period and she finds it optimal to release tokens to  $N$  consumer types in  $N$  periods.  $\square$

*Proof of Proposition 3.3.* The total welfare in the case of the monopolist is the lifetime sum of her one-period profits and one-period surpluses of consumers who are able to obtain the service:

$$T \sum_{j=1}^{i_m} \alpha_j (v_{i_m} - c) + T \sum_{j=1}^{i_m} \alpha_j (v_j - v_{i_m}) = T \sum_{j=1}^{i_m} \alpha_j (v_j - c). \quad (3A.2)$$



Since the monopolist charges the same price  $v_{i_m}$  for tokens in every period, each term in the sum is a one period surplus of the respective agent type multiplied by the total number of periods  $T$ .

When  $T \geq N$ , the total welfare in the case of the entrepreneur who releases all the tokens eventually but with an initial delay is the sum of entrepreneur's, consumers', and providers' surpluses:

$$\begin{aligned} \sum_{j=1}^N \alpha_j (v_j - c) + \sum_{j=1}^N \sum_{i=1}^{j-1} \alpha_i (v_i - v_j) + \sum_{j=1}^N \sum_{i=1}^{j-1} \alpha_i (v_j - c) + (T - N) \sum_{i=1}^N \alpha_i (v_i - c) \\ = \sum_{j=1}^N \sum_{i=1}^j \alpha_i (v_i - c) + (T - N) \sum_{i=1}^N \alpha_i (v_i - c). \end{aligned} \quad (3A.3)$$

The sum of the first three terms represents the total surplus in the first  $N$  periods when the entrepreneur gradually releases tokens to consumers. Specifically, in period  $j$ , the entrepreneur releases  $\alpha_j$  tokens, in addition to the current outstanding stock of tokens  $\sum_{i=1}^{j-1} \alpha_i$ , and the token price is  $v_j$ . In this period, the total surplus generated by consumers of type  $i < j$  is split between consumers and service providers while the surplus generated by consumers of type  $j$  is entirely captured by the entrepreneur.

Additionally, the last term in the sum (3A.3) is the total surplus from periods  $t > N$  when the service market reaches the competitive outcome, in which all  $N$  consumer types are served. When this happens, the per-period surplus is maximized and is strictly higher than the per-period surplus under the monopolist who does not serve all consumers, which is the case when  $i_m < N$ .

Therefore, if  $T$  is sufficiently large and  $i_m < N$ , the total surplus under the entrepreneur is higher than that under the monopolist since (3A.2) is smaller than the last term in (3A.3). In the opposite case, if  $T$  is small and  $i_m$  is sufficiently close to  $N$ , the total surplus under the monopolist can be higher since (3A.2) can be larger than (3A.3).  $\square$

*Proof of Lemma 3.2.* If investors get no utility from consuming the service, the entrepreneur can only offer them profit-sharing contracts in order to provide return on their investment. Let  $V_m$  represent the profit that the entrepreneur obtains as a monopolist and  $V_e$  represent the profit that she makes with an active token resale market. We know from Proposition 3.2 that  $V_m > V_e$ .

Next, if the entrepreneur operates as a monopolist she has to offer investors  $s_m$  such that

$$s_m = \frac{I}{V_m} \tag{3A.4}$$

while if the entrepreneur allows for an active token resale market she has to offer investors  $s_e$  such that

$$s_e = \frac{I}{V_e}. \tag{3A.5}$$

Since  $V_m > V_e$  it follows that  $s_e > s_m$ .

Finally, the entrepreneur's payoff when operating as a monopolist is

$$(1 - s_m)V_m = V_m - I \tag{3A.6}$$

while the entrepreneur's payoff when she allows for an active token resale market is

$$(1 - s_e)V_e = V_e - I. \tag{3A.7}$$

Therefore, the entrepreneur is always better off by operating as a monopolist.

□

*Proof of Proposition 3.4.* As before, we can calculate the proportion of tokens, the entrepreneur would sell in an ICO. Consider a candidate equilibrium, in which a proportion  $\sum_{i=1}^{i_e} \alpha_i$  of consumers participate in the ICO. Then a consumer of type  $i_e$  who believes they

are pivotal will pay up to

$$v_{i_e} + \sum_{j=1}^{N-i_e} (v_{i_e} - v_{j+i_e}) + (T - N + i_e - 1)(v_{i_e} - \underline{v}). \quad (3A.8)$$

Note that if the price is equal to the above, everyone with type  $i < i_e$  will always want to buy a token. If the entrepreneur has an ICO it is optimal to choose  $i_e$  such that

$$i_e \in \arg \max_i \sum_{j=1}^i \alpha_j \left( v_i + \sum_{j=1}^{N-i} (v_i - v_{j+i}) + (T - N + i - 1)(v_i - \underline{v}) \right) \quad (3A.9)$$

Each type  $i_e$  needs to believe they are pivotal for them to finance the platform. This can be implemented by setting a minimum fund raising amount of

$$\sum_{i=1}^{i_e} \alpha_i \left( v_{i_e} + \sum_{j=1}^{N-i_e} (v_{i_e} - v_{j+i_e}) + (T - N + i_e - 1)(v_{i_e} - \underline{v}) \right) \quad (3A.10)$$

If this amount is not met, the money raised is returned to investors. In this case, all types  $i \leq i_e$  will purchase a token during financing at a price given by (3A.8) with type  $i_e$  being just indifferent between financing and not-financing the platform.

An entrepreneur would prefer to have an ICO rather than operate as a monopolist if

$$\begin{aligned} \sum_{i=1}^{i_e} \alpha_i \left( v_{i_e} + \sum_{j=1}^{N-i_e} (v_{i_e} - v_{j+i_e}) + (T - N + i_e - 1)(v_{i_e} - \underline{v}) \right) \\ - I + \sum_{j=1}^{N-i_e} \alpha_{j+i_e} (v_{j+i_e}) - c \geq T \sum_{j=1}^{i_m} \alpha_j (v_{i_m} - c) - I. \end{aligned} \quad (3A.11)$$

This can be simplified to

$$T \left( \sum_{i=1}^{i_e} \alpha_i v_{i_e} - \sum_{j=1}^{i_m} \alpha_j v_{i_m} \right) + \sum_{j=1}^{N-i_e} v_{j+i_e} \left( \alpha_{j+i_e} - \sum_{i=1}^{i_e} \alpha_i \right)$$

$$-(T - N + i_e - 1)v \sum_{i=1}^{i_e} \alpha_i - c \left( 1 - T \sum_{j=1}^{i_m} \alpha_j \right) \geq 0. \quad (3A.12)$$

□

*Proof of Proposition 3.5.* If the entrepreneur has  $N$  financing rounds where, in each round  $i$ , she sells  $\alpha_i$  tokens at a price of  $v_i + (T - 1)(v_i - c)$  then the entrepreneur gets a total profit of

$$(T - 1) \sum_{j=1}^N \alpha_j (v_j - c) + \sum_{j=1}^N \alpha_j v_j - c. \quad (3A.13)$$

This is always greater than the monopolist's profit  $T \sum_{j=1}^{i_m} \alpha_j (v_{i_m} - c)$ .

Moreover, the profit (3A.13) for  $N$  financing rounds is also greater than having less than  $N$  financing rounds and releasing some tokens later. Specifically, assume the entrepreneur decides to have  $K < N$  financing rounds. In this case, the entrepreneur's profit is

$$T \sum_{j=1}^K \alpha_j (v_j - c) + \sum_{j=K+1}^N \alpha_j (v_j - c) \quad (3A.14)$$

It straightforward to see that the above is maximized when  $K = N$ . Therefore, it is optimal for the entrepreneur to have  $N$  financing rounds. This can be implemented by setting a minimum fundraising amount equal to (3A.13). In this case, the investment by every consumer is pivotal and all consumers will be just indifferent between investing and not investing in the ICO.

□

*Proof of Proposition 3.6.* To prove the proposition, we first derive the equilibrium outcomes in different scenarios under network effects: with a monopolist, with an entrepreneur, and with competing platforms. Next, we compare welfare in these scenarios.

*Monopolist.* The monopolist who controls a standard platform solves the following problem:

$$\max_{\alpha} (v(\alpha) - c(\alpha))\alpha. \quad (3A.15)$$

Compared to the baseline model where the marginal cost  $c$  was constant, the monopolist has incentives to serve more consumers since  $c(\alpha)$  is decreasing with higher  $\alpha$ . The price is  $p_m = v(\alpha_m)$  and the total welfare in this scenario is

$$TS_m = \int_0^{\alpha_m} (v(\alpha) - p_m)d\alpha + (p_m - c(\alpha_m))\alpha_m = \int_0^{\alpha_m} (v(\alpha) - c(\alpha_m))d\alpha. \quad (3A.16)$$

*Platform with tokens.* As we noted in the analysis of the baseline model, *in the long run* a tokenized platform operates at full capacity and the price in the token market is set competitively such that

$$v(\alpha_e) = c(\alpha_e). \quad (3A.17)$$

This price is  $p_e = v(\alpha_e)$  and the total welfare in this scenario is

$$TS_e = \int_0^{\alpha_e} (v(\alpha) - c(\alpha_e))d\alpha. \quad (3A.18)$$

*Two competing platforms.* Finally, consider two standard platforms that compete à la Bertrand by setting price of the service to consumers. In a symmetric equilibrium, prices on the platforms are the same and consumers are split equally between the two. Therefore, each platform faces a modified inverse demand function  $v(2\alpha)$ , which is twice steeper than that faced by a monopolistic platform or by a tokenized platform. Given perfect competition, the mass of consumers  $\alpha_c$  served by each platform is such that

$$v(2\alpha_c) = c(\alpha_c). \quad (3A.19)$$

The price on each platform is  $p_c = v(\alpha_c)$  and the total welfare in this scenario is

$$TS_c = 2 \int_0^{\alpha_c} (v(2\alpha) - c(\alpha_c)) d\alpha. \quad (3A.20)$$

*Proof of i).* Since the demand faced by two competing platforms is steeper than that faced by a single platform, it is clear that  $\alpha_e > \alpha_c$ . Therefore, since  $c(\alpha_e) \leq c(\alpha_c)$  it follows from (3A.17) and (3A.19) that  $\alpha_e \geq 2\alpha_c$ . Finally, the welfare under competing platforms can be modified to:

$$TS_c = \int_0^{\alpha_c} (v(2\alpha) - c(\alpha_c)) d(2\alpha) = \int_0^{2\alpha_c} (v(u) - c(\alpha_c)) du. \quad (3A.21)$$

Thus,

$$TS_e = \int_0^{\alpha_e} (v(u) - c(\alpha_e)) du \geq \int_0^{2\alpha_c} (v(u) - c(\alpha_e)) du \geq \int_0^{2\alpha_c} (v(u) - c(\alpha_c)) du = TS_c, \quad (3A.22)$$

where the first inequality is due to  $\alpha_e \geq 2\alpha_c$  and the second is due to  $c(\alpha_e) \leq c(\alpha_c)$ .

*Proof of ii).* Since the demand faced by two competing platforms is steeper than that faced by a single platform, it follows that  $\alpha_m > \alpha_c$ . If  $\alpha_m \geq 2\alpha_c$  then the proof is the same as in *i)*. Alternatively, if  $\alpha_m < 2\alpha_c$  then

$$TS_m - TS_c = \int_0^{\alpha_m} (c(\alpha_c) - c(\alpha_m)) d\alpha - \int_{\alpha_m}^{2\alpha_c} (v(u) - c(\alpha_c)) du, \quad (3A.23)$$

which is positive if  $c(\alpha_c) - c(\alpha_m)$  is sufficiently high, i.e.  $c(\cdot)$  decreases fast enough. Thus,  $TS_m > TS_c$  if  $|c'(\alpha)| > C$  for some  $C > 0$ .

□

Appendix 3B. Baseline Example ( $T = 2$  and  $N = 2$ ) with  $\delta < 1$

In this Appendix, we solve the model of the baseline example with positive discounting, i.e.  $\delta < 1$ . We explicitly show that the results derived for  $\delta = 1$  carry over.

**Monopolist's profit.** If  $\alpha_H(v_H - c) \geq v_L - c$ , the monopolist's profit is

$$(1 + \delta)\alpha_H(v_H - c). \quad (3B.1)$$

Alternatively, if  $\alpha_H(v_H - c) < v_L - c$ , the monopolist's profit is

$$(1 + \delta)(v_L - c). \quad (3B.2)$$

**Entrepreneur's profit.** The entrepreneur promises to buy back tokens from service providers in the end of the last period for  $c$ . With discounting, the cost of this buyback is  $\delta c$ . Since providers offer the service upfront and get paid only in the next period when they sell tokens, it is necessary that

$$\delta v_L \geq c \quad (3B.3)$$

for the platform to be operational. This condition guarantees that, in the second period, when they participate in the token market, service providers recoup their costs incurred in the first period, i.e. the discounted token price from the second period  $v_L$  is higher than the cost  $c$ .

If  $\alpha_H v_H + \delta(1 - \alpha_H)v_L \geq v_L$ , the entrepreneur releases tokens gradually, in 2 periods, and her profit is

$$\alpha_H v_H + \delta(1 - \alpha_H)v_L - \delta c. \quad (3B.4)$$

Alternatively, if  $\alpha_H v_H + \delta(1 - \alpha_H)v_L < v_L$ , the entrepreneur releases all tokens at once, in the first period, and her profit is

$$v_L - \delta c. \quad (3B.5)$$

It can be verified that the monopolist's profit is higher than the entrepreneur's profit under all possible scenarios as long as condition (3B.3) is satisfied.

**Welfare under monopolist.** If  $\alpha_H(v_H - c) \geq v_L - c$ , the welfare under the monopolist is

$$(1 + \delta)\alpha_H(v_H - c). \quad (3B.6)$$

Alternatively, if  $\alpha_H(v_H - c) < v_L - c$ , the welfare under the monopolist is

$$(1 + \delta)\alpha_H(v_H - c) + (1 + \delta)(1 - \alpha_H)(v_L - c). \quad (3B.7)$$

**Welfare under entrepreneur.** If  $\alpha_H v_H + \delta(1 - \alpha_H)v_L \geq v_L$ , the welfare under the entrepreneur is

$$(1 + \delta)\alpha_H(v_H - c) + \delta(1 - \alpha_H)(v_L - c). \quad (3B.8)$$

Alternatively, if  $\alpha_H v_H + \delta(1 - \alpha_H)v_L < v_L$ , the welfare under the entrepreneur is

$$(1 + \delta)\alpha_H(v_H - c) + (1 + \delta)(1 - \alpha_H)(v_L - c). \quad (3B.9)$$

It can be seen that the welfare under the monopolist is lower than the welfare under the entrepreneur if  $\alpha_H(v_H - c) \geq v_L - c$ , i.e. when the monopolist excludes some consumers from the platform.



## BIBLIOGRAPHY

- J. Abadi and M. Brunnermeier. Blockchain economics. Working paper, 2018.
- S. Adhami, G. Giudici, and S. Martinazzi. Why do businesses go crypto? an empirical analysis of initial coin offerings. *Journal of Economics and Business*, 100:64 – 75, 2018.
- T. Alloway, M. Mackenzie, and V. Rodrigues. Bond syndication bonanza under scrutiny, 2014. Available at <https://www.ft.com/content/55406aea-a30a-11e3-ba21-00144feab7de> [Online; accessed 24-March-2019].
- R. Amsden and D. Schweizer. Are blockchain crowdsales the new “gold rush”? success determinants of initial coin offerings. Working paper, 2018.
- D. M. Arseneau, D. E. Rappoport, and A. Vardoulakis. Private and public liquidity provision in over-the-counter markets. 2017.
- A. G. Atkeson, A. L. Eisfeldt, and P.-O. Weill. The market for otc credit derivatives. Working Paper, UCLA, 2014.
- U. Axelson. Security design with investor private information. *The Journal of Finance*, 62(6):2587–2632, 2007.
- A. Babus and T.-W. Hu. Endogenous intermediation in over-the-counter markets. *Journal of Financial Economics*, 125(1):200–215, 2017.
- A. Babus and P. Kondor. Trading and information diffusion in over-the-counter markets. *Econometrica*, 86(5):1727–1769, 2018.
- K. Back and J. F. Zender. Auctions of divisible goods: On the rationale for the treasury experiment. *The Review of Financial Studies*, 6(4):733–764, 1993.
- K. Back and J. F. Zender. Auctions of divisible goods with endogenous supply. *Economics Letters*, 73(1):29–34, 2001.
- Y. Bakos and H. Halaburda. Funding new ventures with digital tokens: Due diligence and token tradability. Available at SSRN 3335650, 2019.
- J. Begenau, M. Piazzesi, and M. Schneider. Banks’ risk exposures. Working Paper, National Bureau of Economic Research, 2015.
- H. Benedetti and L. Kostovetsky. Digital tulips? returns to investors in initial coin offerings. Working paper, 2018.

- L. M. Benveniste and P. A. Spindt. How investment bankers determine the offer price and allocation of new issues. *Journal of financial Economics*, 24(2):343–361, 1989.
- L. M. Benveniste and W. J. Wilhelm. A comparative analysis of ipo proceeds under alternative regulatory environments. *Journal of financial economics*, 28(1-2):173–207, 1990.
- Z. Bethune, B. Sultanum, and N. Trachter. Asset issuance in over-the-counter markets. *Review of Economic Dynamics*, 2019.
- B. Biais and T. Mariotti. Strategic liquidity supply and security design. *The Review of Economic Studies*, 72(3):615–649, 2005.
- B. Biais, D. Martimort, and J.-C. Rochet. Competing mechanisms in a common value environment. *Econometrica*, 68(4):799–837, 2000.
- B. Biais, C. Bisière, M. Bouvard, and C. Casamatta. The Blockchain Folk Theorem. *Review of Financial Studies*, 32(5):1662–1715, 2019.
- A. W. Boot and A. V. Thakor. Security design. *The Journal of Finance*, 48(4):1349–1378, 1993.
- D. Boreiko and N. Sahdev. To ico or not to ico – empirical analysis of initial coin offerings and token sales. Working paper, 2018.
- T. Bourveau, E. T. De George, A. Ellahie, and D. Macciocchi. Initial coin offerings: Early evidence on the role of disclosure in the unregulated crypto market. Working paper, 2018.
- M. Bruche and A. Segura. Debt maturity and the liquidity of secondary debt markets. *Journal of Financial Economics*, 124(3):599–613, 2017.
- J. Brugler, C. Comerton-Forde, and J. S. Martin. Do you see what i see? transparency and bond issuing costs. 2016.
- E. Budish. The Economic Limits of Bitcoin and the Blockchain Bitcoin and the Blockchain. 2018.
- J. I. Bulow. Durable-goods monopolists. *Journal of Political Economy*, 90(2):314–332, 1982.
- N. Cai, J. Helwege, and A. Warga. Underpricing in the corporate bond market. *The Review of Financial Studies*, 20(6):2021–2046, 2007.
- A. Canidio. Financial incentives for open source development: the case of blockchain. 2018.
- C. Catalini and J. S. Gans. Initial coin offerings and the value of crypto tokens. Technical report, National Bureau of Economic Research, 2018.
- N. Cetorelli, B. Hirtle, D. P. Morgan, S. Peristiani, and J. A. Santos. Trends in financial

- market concentration and their implications for market stability. *Economic Policy Review*, 13(1), 2007.
- L. Chen, L. Cong, and Y. Xiao. A Brief Introduction to Blockchain Economics. Working paper, 2019.
- J. Chiu and T. V. Koepl. The Economics of Cryptocurrencies Bitcoin and Beyond. Working paper, 2018.
- J. Chiu and T. V. Koepl. Blockchain-Based Settlement for Asset Trading. 32(5):1716–1753, 2019. ISSN 14657368.
- J. Chod and E. Lyandres. A theory of icos: Diversification, agency, and information asymmetry. *Agency, and Information Asymmetry (July 18, 2018)*, 2018.
- S. Choi, A. Galeotti, and S. Goyal. Trading in networks: Theory and experiments. *Journal of the European Economic Association*, 15(4):784–817, 2017.
- R. Coase. Durability and Monopoly. *Journal of Law and Economics*, 12(1):143–149, 1972.
- D. Condorelli, A. Galeotti, and L. Renou. Bilateral trading in networks. *The Review of Economic Studies*, 84(1):82–105, 2016.
- L. W. Cong and Z. He. Blockchain disruption and smart contracts. *The Review of Financial Studies*, 32(5):1754–1797, 2019.
- L. W. Cong, Y. Li, and N. Wang. Tokenomics and Platform Finance. Working paper, 2019a.
- L. W. Cong, Y. Li, and N. Wang. Tokenomics: Dynamic adoption and valuation. *Becker Friedman Institute for Research in Economics Working Paper*, (2018-49):2018–15, 2019b.
- T. V. Dang, G. Gorton, and B. Holmström. The information sensitivity of a security. Working Paper, MIT, 2015.
- T. Davydiuk, D. Gupta, and S. Rosen. De-crypto-ing Signals in Initial Coin Offerings: Evidence of Rational Token Retention. *SSRN Electronic Journal*, 2019(October 2018), 2018.
- P. DeMarzo and D. Duffie. A liquidity-based model of security design. *Econometrica*, 67(1):65–99, 1999.
- P. M. DeMarzo. The pooling and tranching of securities: A model of informed intermediation. *The Review of Financial Studies*, 18(1):1–35, 2005.
- P. M. DeMarzo, I. Kremer, and A. Skrzypacz. Bidding with securities: Auctions and security design. *American Economic Review*, 95(4):936–959, 2005.

- X. Deng, Y.-T. Lee, and Z. Zhong. Decrypting coin winners: Disclosure quality, governance mechanism, and team networks. Working paper, 2018.
- M. Di Maggio, A. Kermani, and Z. Song. The value of trading relations in turbulent times. *Journal of Financial Economics*, 124(2):266–284, 2017a.
- M. Di Maggio, A. Kermani, and Z. Song. The value of trading relations in turbulent times. *Journal of Financial Economics*, 124(2):266–284, 2017b.
- D. Duffie, N. Gârleanu, and L. H. Pedersen. Over-the-counter markets. *Econometrica*, 73(6):1815–1847, 2005.
- D. Duffie, N. Gârleanu, and L. H. Pedersen. Valuation in over-the-counter markets. *The Review of Financial Studies*, 20(6):1865–1900, 2007.
- D. Easley, M. O’Hara, and S. Basu. From mining to markets: The evolution of bitcoin transaction fees. *Journal of Financial Economics*, 134(1):91–109, 2019. ISSN 0304405X.
- A. Ellul and M. Pagano. Ipo underpricing and after-market liquidity. *The Review of Financial Studies*, 19(2):381–421, 2006.
- E. Farhi and J. Tirole. Liquid bundles. *Journal of Economic Theory*, 158:634–655, 2015.
- P. Feldhütter. The same bond at different prices: identifying search frictions and selling pressures. *The Review of Financial Studies*, 25(4):1155–1206, 2012.
- P. Feldhütter and T. K. Poulsen. What determines bid-ask spreads in over-the-counter markets? *Available at SSRN*, 2018.
- C. Fisch. Initial coin offerings (icos) to finance new ventures. *Journal of Business Venturing*, 34(1):1–22, 2019.
- T. Flanagan, S. Kedia, and X. A. Zhou. Secondary market liquidity and primary market allocations in corporate bonds. *Available at SSRN 3449431*, 2019.
- E. Glaeser and J. A. Scheinkman. Non-market interactions. advances in economics and econometrics: theory and applications. In *Eight World Congress, M. Dewatripont, LP Hansen, and S. Turnovsky (eds.)*, Cambridge University Press, 2002.
- V. Glode, C. C. Opp, and X. Zhang. Voluntary disclosure in bilateral transactions. *Journal of Economic Theory*, 175:652–688, 2018.
- M. Gofman. A network-based analysis of over-the-counter markets. 2014.
- M. A. Goldstein and E. S. Hotchkiss. Dealer behavior and the trading of newly issued corporate bonds. 2012.

- M. A. Goldstein, P. Irvine, and A. Puckett. Purchasing ipos with commissions. *Journal of Financial and Quantitative Analysis*, 46(5):1193–1225, 2011.
- M. A. Goldstein, E. S. Hotchkiss, and D. J. Pedersen. Secondary market liquidity and primary market pricing of corporate bonds. *Journal of Risk and Financial Management*, 12(2):86, 2019.
- G. Gorton and G. Pennacchi. Security baskets and index-linked securities. *Journal of Business*, pages 1–27.
- G. Gorton and G. Pennacchi. Financial intermediaries and liquidity creation. *The Journal of Finance*, 45(1):49–71, 1990.
- R. C. Green. Presidential address: Issuers, underwriter syndicates, and aftermarket transparency. *The Journal of Finance*, 62(4):1529–1550, 2007a.
- R. C. Green. Presidential address: Issuers, underwriter syndicates, and aftermarket transparency. *The Journal of Finance*, 62(4):1529–1550, 2007b.
- R. C. Green, B. Hollifield, and N. Schürhoff. Financial intermediation and the costs of trading in an opaque market. *The Review of Financial Studies*, 20(2):275–314, 2007.
- M. Harris and A. Raviv. The design of securities. *Journal of Financial Economics*, 24(2):255–287, 1989.
- B. Hartman-Glaser, T. Piskorski, and A. Tchisty. Optimal securitization with moral hazard. *Journal of Financial Economics*, 104(1):186–202, 2012.
- C. R. Harvey. Cryptofinance. Working paper, 2014.
- Z. He and K. Milbradt. Endogenous liquidity and defaultable bonds. *Econometrica*, 82(4):1443–1508, 2014.
- J. Helwege and L. Wang. Liquidity and price pressure in the corporate bond market: Evidence from mega-bonds. 2017.
- T. Hendershott, D. Li, D. Livdan, and N. Schürhoff. Relationship trading in otc markets. 2017.
- T. Hendershott, D. Li, D. Livdan, and N. Schürhoff. Relationship trading in over-the-counter markets. *The Journal of Finance*, 75(2):683–734, 2020.
- F. J. Hinzen, F. Irresberger, K. John, and F. Saleh. The Public Blockchain Ecosystem: An Empirical Analysis. Working paper, 2019a.
- F. J. Hinzen, K. John, and F. Saleh. Bitcoin’s Fatal Flaw The Limited Adoption Problem. Working paper, 2019b.

- B. Hollifield, A. Neklyudov, and C. Spatt. Bid-ask spreads, trading networks, and the pricing of securitizations. *The Review of Financial Studies*, 30(9):3048–3085, 2017.
- S. Howell, M. Niessner, and D. Yermack. Initial Coin Offerings: Financing Growth with Cryptocurrency Token Sales. *The Review of Financial Studies*, 10 2019.
- A. Hu, C. A. Parlour, and U. Rajan. Cryptocurrencies: Stylized Facts on a New Investible Instrument. Working paper, 2018.
- G. Huberman, J. D. Leshno, and C. Moallemi. An Economic Analysis of the Bitcoin Payment System. Working paper, 2019.
- R. Inderst and H. M. Mueller. Informed lending and security design. *The Journal of Finance*, 61(5):2137–2162, 2006.
- M. Ivaldi, B. Jullien, P. Rey, P. Seabright, J. Tirole, et al. The economics of tacit collusion. *Final report for DG competition, European Commission*, pages 4–5, 2003.
- T. Jenkinson, H. Jones, and F. Sunthelm. Quid pro quo? what factors influence ipo allocations to investors? *The Journal of Finance*, 73(5):2303–2341, 2018.
- H. Jiang and Z. Sun. News and corporate bond liquidity. Working paper, UC Irvine, 2015.
- A. d. Jong, P. Roosenboom, and T. v. d. Kolk. What determines success in initial coin offerings? Working paper, 2018.
- P. Kondor and G. Pinter. Clients’ connections: Measuring the role of private information in decentralised markets. 2019.
- J. Lee and C. A. Parlour. Consumers as financiers: Crowdfunding, initial coin offerings and consumer surplus. 2019.
- D. Li and N. Schürhoff. Dealer networks. *The Journal of Finance*, 74(1):91–144, 2019.
- J. Li and W. Mann. Initial coin offerings and platform building. 2018.
- T. Li, D. Shin, and B. Wang. Cryptocurrency Pump-and-Dump Schemes. Working paper, 2019.
- Y. Liu and A. Tsyvinski. Risks and Returns of Cryptocurrency. 2018.
- E. Lyandres, B. Palazzo, and D. Rabetti. Are tokens securities? an anatomy of initial coin offerings. Working paper, 2018.
- S. Malamud and M. Rostek. Decentralized exchange. *American Economic Review*, 107(11): 3320–62, 2017.
- M. Manea. Models of bilateral trade in networks. In *The Oxford Handbook of the Economics of Networks*. Citeseer, 2016.

- M. Manea. Intermediation and resale in networks. *Journal of Political Economy*, 126(3): 1250–1301, 2018.
- D. C. Nachman and T. H. Noe. Optimal design of securities under asymmetric information. *The Review of Financial Studies*, 7(1):1–44, 1994.
- F. Nagler and G. Ottonello. Structural changes in corporate bond underpricing. *BAFFI CAREFIN Centre Research Paper*, (2017-48), 2018.
- S. Nikolova, L. Wang, and J. Wu. Institutional allocations in the primary market for corporate bonds. *Available at SSRN 3181983*, 2018.
- M. Nimalendran, J. R. Ritter, and D. Zhang. Do today’s trades affect tomorrow’s ipo allocations? *Journal of Financial Economics*, 84(1):87–109, 2007.
- M. O’Hara, Y. Wang, and X. A. Zhou. The execution quality of corporate bonds. *Journal of Financial Economics*, 130(2):308–326, 2018.
- E. Pagnotta. Bitcoin As Decentralized Money: Prices, Mining Rewards, and Network Security. Working paper, 2018.
- T. R. Palfrey. Bundling decisions by a multiproduct monopolist with incomplete information. *Econometrica.*, pages 463–483, 1983.
- Rabobank. Rabobank bond syndicate allocation policy, 2017. Available at <https://www.rabobank.com/en/images/rabobank-bonds-allocation-policy.pdf> [Online; accessed 24-March-2019].
- S. Ramakrishnan, D. Robinson, and P. Scipio. Possible sec inquiry raises age-old bonds question, 2014. Available at <https://www.reuters.com/article/sec-bonds-allocation/possible-sec-inquiry-raises-age-old-bonds-question-idUSL1N0M019520140303> [Online; accessed 24-March-2019].
- J. Reuter. Are ipo allocations for sale? evidence from mutual funds. *The Journal of Finance*, 61(5):2289–2324, 2006.
- J. Riley and R. Zeckhauser. Optimal selling strategies: When to haggle, when to hold firm. *The Quarterly Journal of Economics*, 98(2):267–289, 1983.
- J. R. Ritter. Initial Public Offerings : Updated Statistics. pages 1980–2013, 2014. ISSN 00221082.
- F. Saleh. Blockchain Without Waste: Proof-of-Stake. 2019.
- J.-P. Siedlarek. Intermediation in networks. 2015.
- E. N. Siriwardane. Limited investment capital and credit spreads. *The Journal of Finance*, 74(5):2303–2347, 2019.

- M. Sockin and W. Xiong. A model of cryptocurrencies. *Unpublished manuscript, Princeton University*, 2018.
- N. L. Stokey. Rational Expectations and Durable Goods. *The Bell Journal of Economics*, 12(1):112–128, 1981.
- A. Subrahmanyam. A theory of trading in stock index futures. *The Review of Financial Studies*, 4(1):17–51, 1991.
- D. Vayanos and J. Wang. Liquidity and asset returns under asymmetric information and imperfect competition. *The Review of Financial Studies*, 25(5):1339–1365, 2011.
- X. Vives. Strategic supply function competition with private information. *Econometrica*, 79(6):1919–1966, 2011.
- C. Wang. Core-periphery trading networks. *Available at SSRN 2747117*, 2016.
- J. J. Wang and J. F. Zender. Auctioning divisible goods. *Economic theory*, 19(4):673–705, 2002.
- B. Williams. Search, liquidity, and retention: Screening multidimensional private information. Working Paper, New York University, 2019.
- R. Wilson. Auctions of shares. *The Quarterly Journal of Economics*, pages 675–689, 1979.
- M. Yang. Optimality of debt under flexible information acquisition. *The Review of Economic Studies*, 87(1):487–536, 2020.
- D. Yermack. Is Bitcoin a Real Currency? Working paper, 2013.