Optimal Exchange Design*

Daniel Chen†

Jun 2022, revised Nov 2024

This draft is preliminary and incomplete and is being updated frequently. Please read with caution.

Abstract

In the context of financial markets, which trading mechanisms are optimal for revenue or efficiency or some combination of the two? Do they resemble institutions we see in practice? How should trade be structured if there is no adverse selection among traders about asset payoffs? If there is adverse selection? If traders are heterogeneous? If there are multiple assets? If there are dynamics? Which mechanisms offer guarantees of efficiency that are robust to adverse selection? Which information structures do robust mechanisms gaurd against? How well do prices aggregate private information when the trading mechanism is endogenous? I investigate these questions for a canonical model setting in finance market microstructure: large traders; quadratic holding costs; private endowments.

*I thank Markus Brunnermeier, Briana Chang, Piotr Dworczak, Darrell Duffie, Marzena Rostek, Robert Wilson, Yuliy Sannikov, Wei Xiong, Motohiro Yogo, Anthony Zhang, and several participants at the Essex-UCL Decentralized Financial Markets Conference for insightful comments and suggestions. I also thank Kayla Jiang and Ethan Wang for excellent research assistance. Part of this research was conducted with the support of the Bradley Graduate Fellowship through a grant to SIEPR.

†dtchen@princeton.edu; Department of Economics, Princeton University

1. Introduction

In the past two decades, trade in financial markets has become increasingly complex and fragmented. This has raised concern among regulators and market practitioners that financial markets may be organized in a way that is socially undesirable. This is especially so because most trading institutions are for profit.

To help further our understanding of these issues, this paper develops a mechanism design analysis of trade for a workhorse model setting in finance market microstructure. In this setting, a finite number of traders are privately endowed with some amounts of the assets in the market and trade to share holding costs that are quadratic in asset positions.[1] I solve for mechanisms that are optimal in that they maximize linear combinations of revenue and allocative efficiency. A key aspect of my analysis is that I am able to do so for a variety of market conditions including when there is adverse selection about fundamental asset payoffs, multiple assets, and heterogeneity among traders, within a unified framework. This allows me to offer insight into how trade should be designed for a range of environments. It also allows me to relate my findings to several results in the literature obtained for a given fixed trading mechanism (or within a parametric class).

I now briefly summarize the main results.

- In general, it is optimal to distort the allocations of traders with extreme types who have the highest marginal value for trade prior to monetary transfers.[2] Moreover, under optimal mechanisms each trader's allocation depends on the private information of *all* traders—it is never optimal to segment the market (in the sense of Malamud & Rostek (2017)).

- Across each of the environments I study, double auctions with transaction fees can often implement *any point* on the revenue-efficiency frontier. In the case with multiple assets, this can be done with double auctions that clear separately. That is, each trader's demand schedule can be contingent on only the price of the security in one exchange but not on the prices of securities in other exchanges.

- The revenue-maximizing incentives of an exchange can lead to less informative prices and harm information aggregation.

[1] Papers with this setup include Vives (2011); Chen & Duffie (2021); Du & Zhu (2017); Rostek & Weretka (2012); Rostek & Yoon (2021); Sannikov & Skrzypacz (2016); Babus & Parlatore (2022); Biais et al. (2005); Glosten (1994); Wittwer (2021); Rostek & Yoon (2023) and many others.

[2] This is in contrast with the typical "no distortion at the top" property of mechanisms in the literature.

- Suppose there is a single asset and that the mechanism designer knows the marginal distributions of traders' posterior expectations of the asset's common value but does not know the correlation among traders values. The worst-case information structure for allocative efficiency consists of independent signals that are single-dimensional and combine additively to yield the best estimate of the fundamental. The optimal mechanism for allocative efficiency under this information structure yields the maximal guarantee of efficiency among all mechanisms (allowing for a budget deficit).

- If the designer can regulate transaction fees and there are renewed endowment shocks over time, it is optimal to continuously run a double auction. This is contrast with Vayanos (1999) or Du & Zhu (2017) who show that, in the absence of transaction fees, an increase in trading frequency often leads to lower allocative efficiency.[3]

The nature of these results suggests two ways to view the contribution of this paper. The first is positive. The model can help us to disentangle the inefficiencies resulting from the primitive environment from inefficiencies arising from the assumed trading mechanisms. It may allows us to understand the extent to which the revenue-maximizing incentives of exchanges are a source of inefficiency. The model can also be used to investigate how existing results in the literature may change when the mechanism itself can change endogenously with the environment. The second is normative. The analysis can offer insight into how trade *should* be designed to target desirable outcomes and the kinds of policies we should consider putting in place as a society. Of course, this second view should be taken with some caution given the stylized nature of the model.

Indeed, the analysis in this paper has some important limitations. Though I anticipate many qualitative insights hold more generally, I rely on the highly stylized quadratic holding cost model for tractability. Also, my analysis of the model when there is adverse selection or multiple assets relies on restrictive distributional assumptions to circumvent the technical challenges associated with multidimensional screening.[4] In particular, I characterize optimal mechanisms for the model with adverse selection only for the information structure that is "worst case" for allocative efficiency (thus following a similar approach taken in the literature on robust mechanism design). This is pedagogically valuable but restrictive. Nevertheless, I am able to identify a set of assumptions where my analysis applies for many environments of interest (e.g., if there is adverse selection, heterogeneity, multiple

[3] I demonstrate this for the case of no private information about fundamental asset payoffs.

[4] The analysis for these settings seems close to the limits of what is achievable with existing tools for multidimensional screening.

assets, etc.). As a result, there is usually a parameter setting where this paper's results can be used as a benchmark of comparison for much existing work. For instance, if one seeks to investigate how a result in the literature may change when the trading mechanism varies endogenously with the environment, there is often some parameter setting where one can apply my results to do so.

Section 2 reviews the related literature and situates the paper at the intersection of mechanism design and finance market microstructure. Section 3 presents the basic model which is static and features symmetric traders with no private information about fundamental asset payoffs. Section 4 characterizes optimal mechanisms for the basic model. Subsequent sections systematically extend the analysis to richer settings. Section 5 introduces the model with adverse selection among traders, solves for worst-case information structures for allocative efficiency, and derives optimal mechanisms for these information structures. Section 6 analyzes a version of the model with multiple assets and Section 7 does the same for a version with heterogeneous traders. Section 8 explains how the results extend to a dynamic setup. In each of these sections, among other results, I demonstrate that double auctions with transaction fees can often implement the outcomes on the revenue-efficiency frontier.

2. Related Literature

This paper is at the intersection of finance market microstructure and mechanism design. To my knowledge, this paper is among the first to consider mechanism design of trade for objectives other than allocative efficiency in a setting with 1) concave utility, 2) a finite number of traders, and 3) traders who may choose to be either buyers or sellers endogenously based on the terms of trade.[5]

I characterize mechanisms that maximize linear combinations of revenue and allocative efficiency for a model setting where traders have private endowments and incur holding costs that are quadratic in their asset positions. This setting has become a workhorse in finance market microstructure in the past fifteen years or so. Papers that analyze variants of this environment include Vives (2011); Chen & Duffie (2021); Du & Zhu (2017); Rostek & Weretka (2012); Rostek & Yoon (2021); Sannikov & Skrzypacz (2016); Babus & Parlatore (2022); Biais et al. (2005); Glosten (1994); Wittwer (2021); Rostek & Yoon (2023) and many others.[6] Existing work typically fixes a trading mechanism or considers a set

[5]There is one paper Lu & Robert (2001) (reviewd below) that I am aware of that precedes this paper, which assumes linear utility up until a commonly known satiation point.

[6]Variants include single and multiperiod models, single and multiple asset models, models with and with-

of mechanisms in a parameterized class. The most commonly assumed mechanism is the (uniform-price) double auction. In contrast, this paper endogenizes the trading mechanism as being chosen by a planner who places weights on revenue and allocative efficiency. As discussed in the Introduction, my paper may be useful to investigate how results may change when the mechanism is endogenous (e.g., chosen by a revenue-maximizing exchange) and may vary with the environment and to see which inefficiencies arise because of the particular form of assumed mechanisms.

This paper is also related to work in finance market microstructure that studies the impact of exchange trading fees (Malinova & Park, 2015; Foucault et al., 2013; Colliard & Foucault, 2012; Jantschgi et al., 2022). I show that trading fees can in fact be a part of optimal mechanism designs. In particular, Jantschgi et al. (2022) develops a theoretical analysis of transaction costs fixing a uniform-price double auction. Relative to Jantschgi et al. (2022), I do not make parametric assumptions on the form of transaction costs that the market designer may have access to and I allow traders to endogenously to choose to be buyers or sellers. I also differ in that I consider environments with adverse selection and multiple assets. Also related is Chen & Zhang (2020), a permanent working paper on which I am a coauthor. Chen & Zhang (2020) studies design of quadratic transaction costs in a uniform-price double auction and does not consider settings with adverse selection or multiple assets. In contrast, the present paper does not restrict the space of mechanisms and analyzes settings with adverse selection and multiple assets.

Analysis of optimal mechanisms for trade in microeconomic theory dates far back and includes classic papers such as Myerson (1981), Myerson & Satterthwaite (1983), Vickrey (1961), Clarke (1971), Groves (1973). Perhaps surprisingly, there have been few papers that study design of optimal *exchange* mechanisms when designer may have revenue as a motive. Further most of these papers consider environments where traders are exogenously designated buyers or sellers and have linear utility. Most of these papers also do not allow for multiple dimensions of private information and interdependent values.

The most closely-related paper is Biais et al. (2000) who study a single asset model with competing mechanisms in a finance setting with a single trader who has CARA utility and may choose to be a buyer or seller endogenously. Because Biais et al. (2000) study the case of a single trader they do not consider exchange mechanisms where the designer does not absorb or supply any net quantity of the asset. Nonetheless, the optimal mechanisms in both of our papers share many properties such as a interior region of binding participation

out private information about asset payoffs, models with symmetric as well as heterogeneous traders.

constraints and kinked transfer rules. One important distinction is that the allocations can be distorted for extreme types in my paper but not in theirs. However, the focuses of our papers are quite different and as a result, many of our results are not directly comparable.

Also related is Lu & Robert (2001) which studies optimal exchange mechanisms in a setting with a finite number of traders who have have linear utility in allocation up to a commonly known satiation point. They allow traders to be either buyers or sellers endogenously. Optimal mechanisms are shown to depend on delicately constructed tie-breaking rules and involve randomization. This is in contrast with the mechanisms in this paper which do not have these properties. Also, Lu & Robert (2001) find there are no distortions at the extremes whereas I find the opposite may be true. The paper Andreyanov & Sadzik (2017) studies optimal exchange mechanisms in a closely-related model setting but focus on allocative efficiency with a robust objective and most results study the large market limit and restrict attention to quadratic subsidies in a double auction. The contemporaneous paper Andreyanov et al. (2023) allows for revenue maximization but imposes ex-post participation constraints which makes the analysis a bit less tractable and so they do not obtain as explicit a characterization of optimal mechanisms.

A contribution of this paper is to apply recent perspectives from information design and robust mechanism design developed, for example, by Brooks & Du (2021), Bergemann & Morris (2013), Bergemann et al. (2015) and others to a finance market microstructure setting. I use these methods to study worst-case information structures (joint distribution of traders' signals and fundamental asset payoffs) and robust mechanisms for allocative efficiency. These methods are especially useful for the model with adverse selection where traders have multiple dimensions of private information and interdependent values so that a traditional analysis for general information structures is intractable or leads to results that are unduly sensitive to common knowledge assumptions. I solve for a strong minimax equilibrium, a solution concept introduced by Brooks & Du (2021) consisting of a worst-case information structure and a guarantee-maximizing mechanism for, in my case, allocative efficiency. I also solve for worst-case information structures for the uniform price double auction, an exercise in the spirit of Bergemann & Morris (2013), Bergemann et al. (2015), and Lambert et al. (2014). These results shed light on the kinds of information structures market designers should seek to gaurd against and yield bounds on efficiency (and price impact) that hold for a variety of information structures.

3. Basic Model

3.1. Environment

In what follows, all random variables are defined on $(\Omega, \mathcal{F}, \mathbb{P})$. There is a single asset with a payoff π that is a random variable with mean μ_π. There are $N \geq 2$ traders indexed by $i \in \{1, 2, ..., N\}$. Trader i is endowed with some privately known quantity e_i of the asset prior to trade. The endowment e_i is a finite variance random variable with cumulative distribution function (CDF) F defined on the interval $\mathbb{I} \subseteq \mathbb{R}$. F has a continuous density $f > 0$. Traders' endowments $\{e_i\}$ and the asset's payoff π are jointly independent.

If trader i purchases X_i units of the asset at a total price T_i then her utility is

$$U(e_i + X_i, T_i) = \pi(e_i + X_i) - T_i - \frac{1}{2\kappa}(e_i + X_i)^2 \tag{1}$$

for some constant $\kappa > 0$. The quadratic cost represents the disutility from exposure to risk or any other costs associated with holding a net position in the asset. I therefore refer to κ as a trader's holding capacity.

As reviewed in Section 2, variants of this setup are popular in the literature on financial market microstructure. Preferences of the form in (1) have appeared in Vives (2011); Chen & Duffie (2021); Du & Zhu (2017); Rostek & Weretka (2012); Rostek & Yoon (2021); Sannikov & Skrzypacz (2016); Babus & Parlatore (2022); Wittwer (2021) and many others. I have adopted a baseline model that is the simplest "common denominator" among these papers. Subsequent sections will systematically introduce further modeling ingredients found in the literature.

3.2. Trading Mechanisms

I seek to conduct a mechanism design analysis of the above environment. To proceed, I first define the concept of a trading mechanism.

Definition 3.1. A trading mechanism $(\{M_i\}, \{X_i\}, \{T_i\})$ consists of

- a message space M_i for trader i,
- an allocation rule $X_i : \Pi_{j=1}^N M_j \to \mathbb{R}$ mapping profiles of reported messages to the quantity purchased by trader i,

- and a transfer rule $T_i : \Pi_{j=1}^N M_j \to \mathbb{R}$ mapping profiles of reported messages to the price paid to the mechanism by trader i,

for each trader i.[7]

The most commonly assumed trading mechanism in the literature for the environments I study in this paper is the (uniform-price) double auction.

Example 1. *In a double auction, trader i's message space M_i is the set of measurable functions $q_i : \mathbb{R} \to \mathbb{R}$ specifying how much trader i purchases for each realization of the asset's price. Given a vector $\vec{q} = (q_1, ..., q_N)$ of reported demand functions, allocation rules are such that,*

$$X_i(\vec{q}) = q_i(p^*)$$

and transfer rules are such that

$$T_i(\vec{q}) = p^* q_i(p^*)$$

for each i whenever there is a unique market clearing price p^* defined by

$$\sum_{i=1}^N q_i(p^*) = 0.$$

If no such p^* exists, then trade shuts down in that $X_i(\vec{q}) = T_i(\vec{q}) = 0$, for each i.

The double auction will serve as a useful benchmark of comparison in the analysis to come. An important property of the double auction is that the mechanism does not absorb or supply any net quantity of the asset: $\sum_{i=1}^N X_i = 0$. I call any trading mechanism with this property an exchange mechanism. The rest of this paper analyzes exchange mechanisms that are optimal in a sense that I now specify.[8]

[7]Formally, the mechanism also specifies a σ-algebra $\mathcal{F}_i$ along with each message space M_i and I require that $X_i : (\Pi_{j=1}^N M_j, \otimes_{j=1}^N \mathcal{F}_j) \to (\mathbb{R}, \mathcal{L})$ and $T_i : (\Pi_{j=1}^N M_j, \otimes_{j=1}^N \mathcal{F}_j) \to (\mathbb{R}, \mathcal{L})$ where $\mathcal{L}$ is the Lebesgue σ-algebra.

[8]In some cases, it may be of interest to allow the mechanism to absorb some amount of the asset. I shut this channel down for this paper, though it is easy to extend the analysis to allow the designer to absorb some amount of the asset. For example, if the designer also has a quadratic holding cost, then the derivation of optimal mechanisms is largely the same.

3.3. Objective

I seek to derive exchange mechanisms that maximize the expectation of convex combinations of revenue generated by the mechanism and allocative efficiency (measured by the negative of the sum of traders' holding costs) that arise in some Bayes Nash equilibrium of the mechanism.[9] By the revelation principle, it is without loss of generality to restrict attention to direct mechanisms where each trader's message space M_i is the type space $\mathbb{I}$ and such that 1) it is individually rational for each trader to participate in the mechanism, 2) incentive compatible for each trader to report her endowment truthfully, and 3) the mechanism is an exchange mechanism.

Formally, given $\alpha \in [0,1]$, the objective is to solve

$$\max_{\{X_i\},\{T_i\}} \mathbb{E}\left[\alpha \underbrace{\sum_{i=1}^{N} T_i(\vec{e})}_{\text{Revenue}} + (1-\alpha) \underbrace{\sum_{i=1}^{N} -\frac{1}{2\kappa}\left(e_i + X_i(\vec{e})\right)^2}_{\text{Allocative Efficiency}}\right] \tag{2}$$

such that,

$$\begin{aligned}
&(\mathrm{P}): \quad \mathbb{E}\left[U\left(e_i + X_i(\vec{e}), T_i(\vec{e})\right) \middle| e_i\right] \geq U(e_i, 0), \quad \forall e_i, \forall i \\
&(\mathrm{IC}): \quad e_i \in \arg\max_{m_i} \mathbb{E}\left[U\left(e_i + X_i(m_i, \vec{e}_{-i}), T_i(m_i, \vec{e}_{-i})\right) \middle| e_i\right], \quad \forall e_i, \forall i \\
&(\mathrm{E}): \quad \sum_{i=1}^{N} X_i(\vec{e}) = 0, \quad \forall \vec{e}
\end{aligned}$$

where $\vec{e} = \{e_i\}_{i=1}^{N}$ is the vector of traders' endowments.

In the constraint (IC), with a transparent abuse of notation, I write trader i's report m_i as the first argument in X_i and I denote the vector of the other traders' endowments by $\vec{e}_{-i}$. I will continue to do so whenever convenient.

By solving (2) for each $\alpha \in [0,1]$ I am able to identify the revenue-efficiency frontier. The revenue-efficiency frontier consists of the combinations of expected revenue and expected allocative efficiency such that there does not exist a mechanism that simultaneously achieves a higher level of both.

I focus on the objective (2) for several reasons. When α is zero, the objective reflects

[9] That is, I allow the designer to select the equilibrium (partial implementation). However, this will turn out to be without loss because the optimal direct mechanism will have a unique equilibrium.

the preferences of a social planner (such as the government) who cares only about allocative efficiency, incurs no fixed costs of running the exchange or else is willing to subsidize trade in illiquid markets. On the other hand, the case when $\alpha > 0$ may represent the problem of a planner who has a binding budget constraint. The planner can determine which mechanism to run by calibrating α to the lowest level that satisfies her constraint. Another reason $\alpha > 0$ is useful to analyze is because the planner may want to generate revenue, because she can use it for some other productive purpose for society. In that case, $\alpha > 0$ is determined by the marginal social value of the revenue. When $\alpha = 1$, the planner may represent a profit-maximizing exchange (most exchanges in practice are privately run with profit-maximizing motives). In this case, the analysis sheds light on how such an exchange may structure itself to extract revenue and the associated inefficiencies.[10]

3.4. Model Discussion

Despite the prevalence of the model setting in the literature, there has been surprisingly little analysis using mechanism or information design tools as discussed in Section 2 on related literature. Below, I briefly discuss model assumptions.

- To ease the exposition, I restrict attention to deterministic mechanisms where there is no residual uncertainty in allocations or transfers conditional on traders' messages. I formally show in Lemma 2 in Appendix A that this is without loss: one can not achieve a higher value for the objective using stochastic mechanisms.

- The quadratic holding cost utility in (1) and CARA utility are equivalent (up to a monotone transformation) when the allocation and transfer that a trader receives is known to her conditional on her message in a direct mechanism. These preferences are generally not equivalent. Typically, in an exchange mechanism, the allocation of a trader is not known to her given her report because she must absorb some amount of the other traders' endowments. Quadratic utility is more tractable than CARA utility in this case and is why I am able to solve the model with multiple traders and also why I do not need to restrict attention to deterministic transfers. However, it will turn out that the allocation rules I identify as optimal with quadratic holding costs are implementable when traders instead have CARA utility though I do not know whether they are necessarily optimal in that case.

[10]Indeed, classical auction theory focuses on revenue maximization, yet there has been far less analysis of revenue-maximizing designs of exchange mechanisms.

- The model assumes that endowments are independent which is an assumption that appears often in the literature (as, for example, in Chen & Duffie (2021)). Independence allows me to avoid delicate mechanisms of the type in Crémer & McLean (1988). There is broad consensus that such mechanisms are impractical because of their sensitivity to common knowledge assumptions.

- Other assumptions will be relaxed in subsequent sections. I allow for private information about the fundamental π in Section 5, multiple assets in Section 6, heterogeneity among traders in Section 7, and dynamics in Section 8. However, I obtain the sharpest characterizations with the least conditions for the basic model of the present section.

4. Optimal Mechanisms

I now give an informal sketch of the derivation of optimal mechanisms. In terms of technique, I build on methods developed for abstract settings with type-dependent participation constraints by Jullien (2000) when there is a single agent. The analysis is also similar to that of Biais et al. (2005) though they study the CARA setting with a single trader.[11] Relative to these papers, the setting I consider involves several traders. Because of the constraint that mechanisms are exchange mechanisms, the problem of deriving an optimal mechanism can not be solved trader-by-trader. Nevertheless, we will see that the problem remains highly tractable.

4.1. Sketch of the Derivation

Step 0: Notation. To start, I define some notation to ease the exposition. I denote the expected utility, expected trade quantity, and expected transfer by

$$\hat{U}_i(e_i) \equiv \mathbb{E}\left[U\left(e_i + X_i(\vec{e}), T_i(\vec{e})\right)\middle| e_i\right],$$

$$\hat{X}_i(e_i) \equiv \mathbb{E}\left[X_i(\vec{e})|e_i\right],$$

and

$$\hat{T}_i(e_i) \equiv \mathbb{E}\left[T_i(\vec{e})|e_i\right]$$

[11] One other difference is that Biais et al. (2005) work with traders' rents as the control variables and restrict attention to deterministic mechanisms.

respectively.

Also, whenever $b > a$, I let $\int_b^a := -\int_a^b$.

Step 1: Incentive Compatibility. Trader i solves

$$\max_{m_i} \mathbb{E}\left[U\Big(e_i + X(m_i, \vec{e}_{-i}), T_i(m_i, \vec{e}_{-i})\Big)|e_i\right].$$

By standard arguments, local incentive compatability of truth telling is equivalent to

$$\hat{U}_i(e_i) - U\Big(e_i, 0\Big) = -\frac{1}{\kappa}\int_{\mu_e}^{e_i} \hat{X}_i(s)\,\mathrm{d}s + \hat{U}_i(\mu_e) - U_i\Big(\mu_e, 0\Big), \tag{3}$$

and

$$\hat{T}(e_i) = \mathbb{E}\left[-\frac{1}{2\kappa}X_i(\vec{e})^2\Big|e_i\right] - \left(\frac{1}{\kappa}e_i - \pi\right)\hat{X}_i(e_i) + \frac{1}{\kappa}\int_{\mu_e}^{e_i} \hat{X}_i(s)\,\mathrm{d}s$$
$$- \left[\hat{U}(\mu_e) - U\Big(\mu_e, 0\Big)\right]. \tag{4}$$

Moreover, a mechanism is globally incentive compatible if and only if it is locally incentive compatible and each $\hat{X}_i$ is weakly decreasing.

By inspection, (3) implies that the types e for whom (P) binds expect to trade zero $\hat{X}_i(e) = 0$ and these types must fall in some closed interval $[e_a, e_b]$.

Step 2: Lagrangian Relaxation. The next step is to form the Lagrangian for the relaxed version of (2) that ignores global incentive constraints.

Let Ω_i be the Lagrange multiplier on trader i's participation constraint (P). Ω_i is a bounded variation measure on $\mathbb{I}$. With abuse of notation, I let Ω_i also denote the distribution function of this measure.

Using equations (3) and (4), I form the Lagrangian, integrate by parts, rearrange, and drop some constant terms to arrive at

$$\max_{\{X_i\},\{\hat{U}_i(\mu_e)\}} \left\{\sum_{i=1}^{N} \mathbb{E}\left[-\frac{1}{2\kappa}\left(e_i - \alpha\frac{\Omega_i(e_i) - F(e_i)}{f(e_i)} + X_i(\vec{e})\right)^2\right] - \alpha\Big(1 - \lim_{e\to\sup\mathbb{I}} \Omega_i(e)\Big)\hat{U}_i(\mu_e)\right\}$$

such that (E): $\sum_{i=1}^{N} X_i = 0$.

Step 3: Candidate Optimal Mechanism. It follows from Jensen's inequality that the solution for the optimal allocation rule, ignoring global incentive compatibility, is

$$X_i(\vec{e}) = -\left(e_i - \alpha\frac{\Omega_i(e_i) - F(e_i)}{f(e_i)}\right) + \frac{1}{N}\sum_{j=1}^{N} e_j - \alpha\frac{\Omega_i(e_j) - F(e_j)}{f(e_j)}$$

Next, we observe that for there to be an interior solution for $\hat{U}_i(\mu_e)$, it must be that

$$\lim_{e_i \to \sup \mathbb{I}} \Omega_i(e_i) = 1.$$

Thus, by complementary slackness (which implies that Ω_i is flat outside $[e_a, e_b]$), we have

$$\Omega_i(e) = 0, \quad e < e_a$$
$$\Omega_i(e) = 1, \quad e > e_b.$$

The $\{\Omega_i\}_{i=1}^{N}$ are constructed in $[e_a, e_b]$ to ensure that $\hat{X}_i = 0$ on $[e_a, e_b]$ for each i. In the Appendix A, I show that the Ω_i are the same for each i and are also continuous.

Step 4: Verification. The last step is to verify that strong duality holds (which I do in the Appendix) and to establish conditions for when the allocation rule from Step 3 is weakly decreasing so that global incentive constraints are not binding. The following condition is clearly sufficient for this and is analogous to the regularity condition in Myerson (1981).

Condition 1. *It holds that $e - \alpha\frac{1-F(e)}{f(e)}$ and $e + \alpha\frac{F(e)}{f(e)}$ are weakly increasing in e.*

Condition 1 holds whenever F is log-concave, a stronger condition than is necessary. Many common distributions satisfy Condition 1 including the uniform and Gaussian distributions.

4.2. Characterization

The following Theorem 1 summarizes the results of the derivation and provides a sharp characterization of optimal mechanisms.

Theorem 1. *Suppose that Condition 1 holds. Then the unique solution of the optimiza-*

tion problem (2) *for* $\{X_i\}$ *sets*

$$X_i(\vec{e}) = -v(e_i) + \frac{1}{N}\sum_{j=1}^{N} v(e_j)$$

for each trader i *where*

$$\begin{aligned} v(e) &= e + \alpha\frac{F(e)}{f(e)}, \quad e < e_a \\ v(e) &= e - \alpha\frac{1-F(e)}{f(e)}, \quad e > e_b \\ v(e) &= \mathbb{E}\left[v(e_j)\right], \quad e \in [e_a, e_b] \end{aligned}$$

for e_a *and* e_b *such that* v *is continuous:* $\lim_{e\to e_b} v(e) = v(e_a) = v(e_b)$.

In all instances, computing the equilibrium entails solving a system of three equations in the three unknowns $(e_a, e_b, \mathbb{E}[v(e_i)])$. There are many examples for which there are nearly closed-form solutions. For instance, this is so whenever the pdf f is symmetric (as for example in the Gaussian case).

Corollary 1.1. *Suppose that the conditions of Theorem 1 are satisfied and that* f *is symmetric about the mean* μ_e *of endowments. Then*

$$\mathbb{E}\left[v(e_j)\right] = \mu_e,$$

and

$$e_a = \mu_e - (e_b - \mu_e),$$

where e_b *solves*

$$e_b - \alpha\frac{1-F(e_b)}{f(e_b)} = \mu_e.$$

Thus, the optimal mechanism is characterized in closed form given a single unknown e_b*. If, in addition F is uniform, then the optimal mechanism, including* e_b *is computable in closed form.*

There are severable notable implications of Theorem 1. One is that optimal mechanisms, perhaps surprisingly, depend on the number N of traders in only a very limited way.

Corollary 1.2. *The region* $[e_a, e_b]$ *of endowments for which the participation constraint* (P) *binds does not depend on the number* N *of traders. Thus, the expected trade quantity* $\hat{X}_i$*, for any given trader, is invariant to the number* N *of traders.*

A second is that optimal mechanisms typical feature distortions in allocations even at the extremes.

Remark. *An important property of the optimal mechanisms in Theorem 1 that distinguishes it from optimal mechanisms in prior work is that the ex-post allocation and the interim allocation of a trader are, in general, both distorted from first best even for traders with extreme types. That is, though there is no distortion in the virtual types of traders at extremes in that* $\lim_{e \to \sup \mathbb{I}} e - v(e) = 0$ *and* $\lim_{e \to \inf \mathbb{I}} e - v(e) = 0$*, there is distortion in allocations at the extremes. This is in contrast with Biais et al. (2005), Lu & Robert (2001), and to my knowledge, almost all other models of optimal trade. One can show that there is no distortion in interim allocations when* f *is symmetric about its mean, but, as shown in Example 3 in Appendix A, this is not generally so when* f *is asymmetric.*

A third notable implication is that it is in fact possible to implement the efficient allocation with ex-ante budget balance.

Proposition 1. *The efficient allocation is implementable with ex-ante budget balance in that expected revenue is zero.*

I note that McAfee (1991) also derive conditions for when ex post efficiency is achievable with ex-ante budget balance in a closely related setting. Proposition 1 can also be proven by checking those conditions. However, subsequent sections will investigate analogs of Proposition 1 under different model conditions (adverse selection, multiple assets, heterogeneity, dynamics etc.) which all fall outside the analysis of McAfee (1991).

Before, turning to an illustrative example, I present some comparative statics results which show that, even accounting for the designer's revenue maximization motives, both revenue and efficiency improve when the market is thicker and that traders are worse off when the designer places mroe weight on revenue in her objective.

Proposition 2. *Suppose that Condition 1 holds. Then under the optimal mechanism,*

1. *As a function of the number of traders* N*, the expected holding cost relative to autarky* $\frac{1}{2\kappa}\mathbb{E}\left[(e_i + X_i(\vec{e}))^2 - e_i^2\right]$*, expected transfer* $\mathbb{E}\left[T_i(\vec{e})\right]$*, and the expected utility gain* $\hat{U}_i(\cdot) - U_i(0)$*, are all proportional to* $(N-1)/N$ *whenever* $\alpha > 0$*. Thus expected*

allocative efficiency, expected revenue, and the welfare of each trader increases as N increases.

2. *If the weight α on revenue increases, then expected utility of traders $\mathbb{E}\left[\sum_{i=1}^{N} \hat{U}_i(e_i)\right]$ decreases. If, in addition F is symmetric about its mean, then an increase in α also leads the total expected trade volume $\mathbb{E}\left[\sum_{i=1}^{N} |X_i(\vec{e})|\right]$ to decrease.*

Part 1 of Proposition 2 implies that the revenue-efficiency frontier shifts out as the number of traders increases. Intuitively, with more traders, traders should more easily be able to share their holding costs. As a result, the expected holding cost decreases for each trader. When $\alpha > 0$ the designer can also extract some fraction of this surplus gain and so the expected transfer of each trader also rises.

Part 2 of the proposition states that, even when accounting for the revenue-maximization incentives of the designer, traders nonetheless benefit overall from a thicker market.

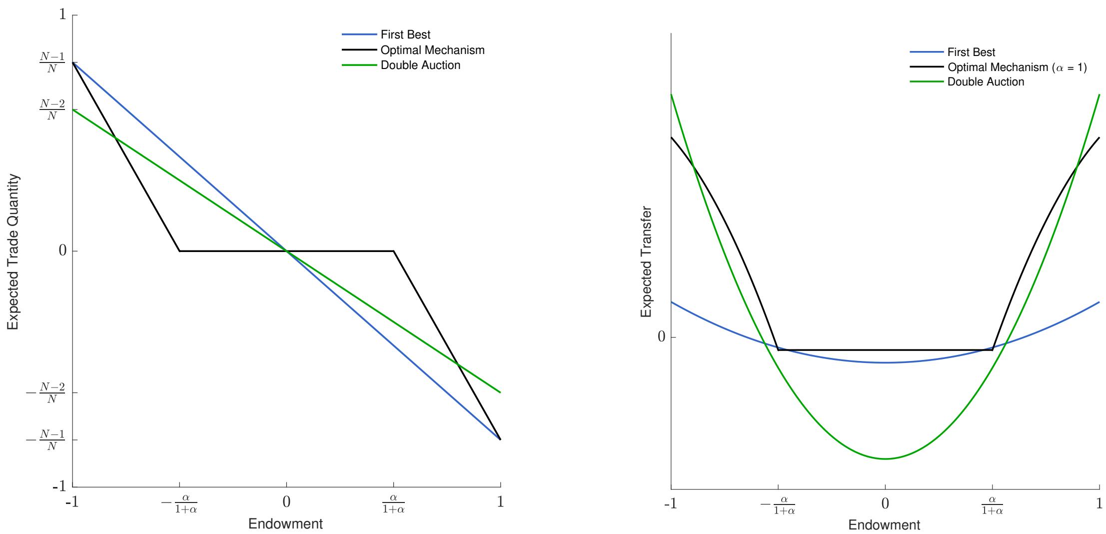
Part 3 implies that as the revenue-maximizing motive increases for the designer, total trade volume decreases and each trader is worse off which is intuitive.

4.3. Illustrative Example

To further highlight some properties of the optimal mechanism and to compare it with the double auction, I present an illustrative example for the case when $F = \mathrm{U}[-1, 1]$ (which, as shown in Corollary 1.1, is a case when there is a closed-form solution for optimal mechanisms).

Panel (a) in Figure 1 depicts the expected trade quantity for the first best efficient allocation (in blue), under the optimal mechanism (in black) for arbitrary α, and for the (symmetric-linear) equilibrium of the double auction (in green). As N increases, the double auction converges to the first best—this is to be expected because price impact disappears which is the fundamental source of inefficiency. In constrast, the optimal mechanism stays bounded away for any given $\alpha > 0$. When α is higher, the region of binding participation constraints expands. Compared with the double auction, the efficient trade outcome is achieved at the extremes. Interestingly, this is only efficiency in expectation and is *not a generic property* whenever F is not symmetric (see the Remark in the previous susection). It is, however, a generic property of the virtual endowment $v(e)$.

Panel (b) plots the expected transfer. Because the expected trade quantity is a linear transformation of the endowments the price schedule qualitatively the same. Note that the expected transfer is more convex under the double auction than under the first best. This



(a) (b)

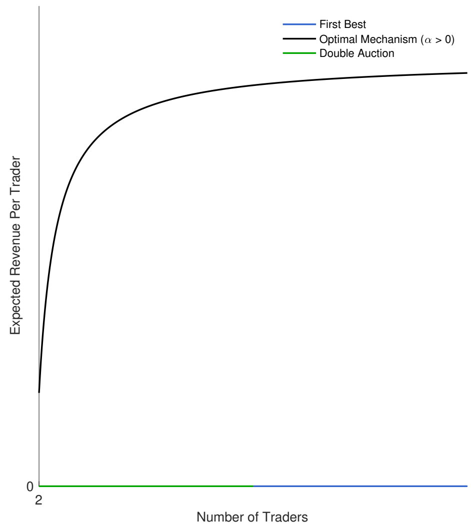
Figure 1: Interim allocation rule and interim transfer rule

Notes: Comparison of interim allocation rules and interim transfer rules when $F = \mathrm{U}[-1, 1]$.

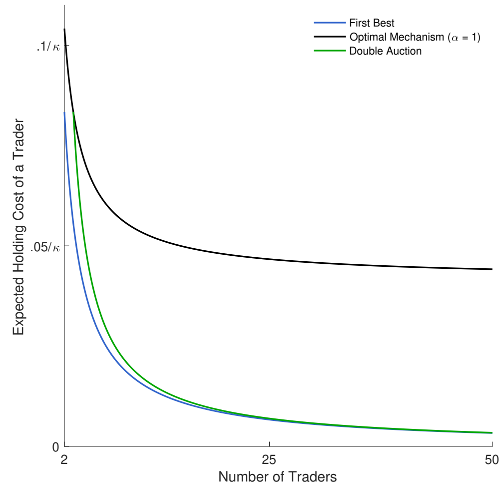
is because of price impact costs (later we will see the reverse can be true when there is adverse selection). Under the optimal mechanism there are kinks corresponding to bid-ask spreads. A similar feature appears in Biais et al. (2005) for the case of a single trader and a market maker who functions as the designer. The concave regions correspond to quantity discounts and are needed to encourage extreme types to trade more aggressively. These regions gradually become convex as α decreases.

Figure 2 plots the expected revenue per trader in panel (a). As you can see it is increasing. Intuively, there is less randomness in the average position of traders and because traders are risk-averse over asset allocations, they like this and are willing to pay more. Thus, in thicker markets, the designer can extract more from each individual trader. Panel (b) plots the expected holding cost of a trader. As seen in the figure, inefficiency disappears for the double auction and first best mechanism but it persists in the long run under the optimal mechanism for revenue. In fact it does so for any $\alpha > 0$. Even under revenue maximization, each trader is better off when there are more traders. Conincidentally, efficiency when there are 3 traders is the same for the double auction as it is for the revenue-maximizing mechanism.

Thus, the double auction does not lie on the revenue-efficiency frontier plotted below



(a)



(b)

Figure 2: Revenue and allocative efficiency

Notes: Revenue per trader and expected holding cost per trader as a function of the number of traders N when $F = \text{U}[-1, 1]$.

in Figure 3. One might wonder whether the mechanisms that comprise the frontier have indirect implementations that might resemble trade in practice financial markets or how one might be able to alter the double auction to bring it closer to the frontier. I investigate this next.

4.4. Implementation: Double Auction with Transaction Fees

It turns out that the double auction can be altered to reach the revenue-efficiency frontier simply by introducing a transaction fee.

Definition 4.1. A double auction with transaction fee $\mathcal{T} : \mathbb{R}^2 \Rightarrow \mathbb{R}$ is an exchange mechanism that operates in three stages:

1. Each trader i submits a demand schedule: $q_i : \mathbb{R} \Rightarrow \mathbb{R}$.
2. The clearing price p is computed: $\sum_{i=1}^{N} q_i(p) = 0$. If there does not exist a unique clearing price, then no trades or transfers are executed.
3. If a unique clearing price p exists then trader i pays $pq_i(p) + \mathcal{T}(q_i(p), p)$ in return for $q_i(p)$ units of the asset.

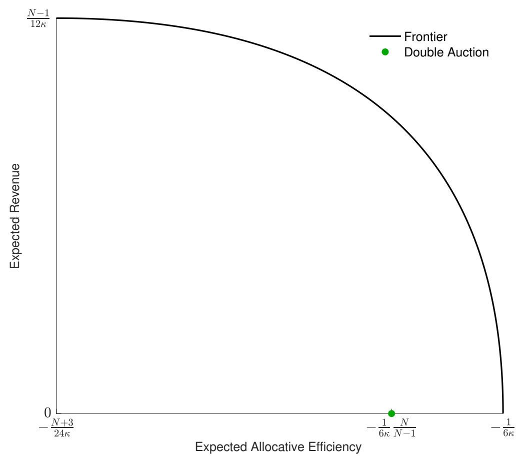


Figure 3: Revenue-efficiency frontier

Notes: The revenue-efficiency frontier when $F = \mathrm{U}[-1, 1]$

Given the flexibility of the transaction fees one might be tempted to say we can achieve practically anything with them. That is not so. Consider for example a fragmented market considered in Malamud & Rostek (2017). Clearly, those allocations are not achievable by running a single double auction with transaction fee. (Other additively separable examples?)

Proposition 3. *For any given $\alpha \in [0, 1]$, the allocation rule $\{X_i\}_{i=1}^{N}$ in Theorem 1 is implementable by a double auction with transaction fee $\mathcal{T} : \mathbb{R}^2 \Rightarrow \mathbb{R}$ defined by*

$$\mathcal{T}\left(p, q_i(p)\right) = \frac{a}{2} q_i(p)^2 - b\left(-q_i(p) - \frac{1}{N} p\right) - c$$

where

$$a = -\frac{N}{N-1},$$

$b : \mathbb{R} \to \mathbb{R}$ satisfies

$$b(e) = \frac{N-1}{N\kappa}\left[\frac{1}{2} e^2 - \int_{\mu_e}^{e} v^{-1}(s)\, \mathrm{d}s\right], \quad e \in \mathbb{I},$$

and c is a constant set sufficiently high so that participation is individually rational for all types of all traders.

Before I offer intuition for Proposition 3 let me first remark on some of the economic implications. Proposition 3 suggests that acheiving desirable market outcomes may not require a major overhaul of existing market infrastructure. In the model, by carefully calibrating transaction fees a planner can target any point on the revenue-efficiency frontier. In practice, exchanges often charge transaction fees such as maker-taker fees which are subject of much policy debate—these fees may not be "just a detail." There is a large literature, both empirical and theoretical, studying these fees (Malinova & Park, 2015; Colliard & Foucault, 2012; Foucault et al., 2013; Jantschgi et al., 2022). Proposition 3 also suggests that depending on the extent of competition among exchanges, allowing exchanges to set transaction fees may either exacerbate or mitigate allocative inefficiency.

To provide intuition I now give a sketch of the proof.

Proof Sketch. Recall that

$$X_i(e_i) = -v(e_i) + \underbrace{\frac{1}{N}\sum_{j=1}^{N} v(e_j)}_{\text{common to all}} .$$

In a double auction $p = \sum_{j=1}^{N} v(e_j)$ if each trader i submits

$$q_i(p) = -v(e_i) - \kappa p. \tag{5}$$

Reverse engineer $\mathcal{T}$ so that this is optimal for each trader i:

$$\mathcal{T}(q,p) = a + \frac{b}{2}q^2 + \tau\left(q_i(p) + \kappa p\right)$$

for some a, b, and $\tau(\cdot)$.

The function τ affects incentives for the intercept of the demand schedule and the coefficient b affects incentives for the slope of the demand schedule. One can calibrate τ and b so that each trader submits the demand schedule (5) and set a so that each trader finds it individually rational to participate. □

5. Adverse Selection

For trade of many securities, adverse selection is a first-order concern (Kyle, 1985, 1989; Du & Zhu, 2017). In this section, I investigate the optimal design of trade in the presence of adverse selection. Loosely speaking, the main results of this section are 1) a characterization of the worst-case information structures for allocative efficiency, 2) a characterization of optimal mechanisms for these worst-case information structures, and 3) a designer's motives for revenue maximization can inhibit information aggregation in prices. Exactly what is meant by "worst-case information structure" will be made precise below.

5.1. Environment

I retain the setup of the basic model except now I assume that traders additionally see private signals $\{\zeta_i\}$ of the asset's payoff π. These signals are random variables that may be correlated with π. I will not take a stance, yet, on the distribution of these signals or on the spaces where they live. I continue to assume that traders' endowments $\{e_i\}$ are independent, both of each other and the asset's payoff π, and moreover that they are independent of the signals $\{\zeta_i\}$. This assumption simplifies the analysis which has as its main focus the role of adverse selection. Also, to ease the exposition, for this section only, I assume that $\mu_\pi = 0$ though all results extend to the case of nonzero μ_π.

Mechanism design in this environment is difficult for two well-known reasons. First, each trader i now has two dimensions of private information e_i and ζ_i. One can not meaningfully simplify the model by removing endowments $\{e_i\}$ because heterogeneous endowments are the source of gains from trade. Thus, the problem of optimal mechanism design is a multidimensional screening problem, which is known to be intractable except for some special cases. Second, it is unclear what particular information structure is appropriate to model. Most of the market microstructure literature assumes signals that are one dimensional and equal to the fundamental π corrupted by additive noise that is independent across traders. However, because these signals are correlated, optimal mechanisms for this information structure would be of the sort in Crémer & McLean (1988) which are very sensitive to the fine details of the environment and not robust. To motivate the approach I take to circumvent these issues, I first take a detour and conduct a worst-case analysis of information structures for the double auction.

5.2. Worst-Case Information Structures: Double Auction

The analysis here is in the spirit of the literature on "robust prediction in games" or "information design" (Bergemann & Morris, 2013; Bergemann et al., 2017; Lambert et al., 2014). I assume that each trader sees a real-valued signal of π. Without loss I take trader i's signal ζ_i to be her posterior expectation of π: $\zeta_i = \mathbb{E}[\pi|\zeta_i]$.

I assume that it is known that

1. $\zeta_i \sim N(\mu_\pi, \sigma_\zeta^2)$
2. $\zeta_1, ..., \zeta_N, \pi$ are jointly Gaussian
3. $\mathrm{corr}(\zeta_i, \zeta_j) \geq 0$ is the same for $\forall(i,j)$
4. the joint distribution of ζ_i and π is the same $\forall i$.

In what follows, let $\mathcal{I} \subset \Delta(\vec{\zeta}, \pi)$ denote the set of information structures satisfying conditions 1-4 above. In what follows, I take the perspective of an outside observer who knows the marginal distributions of traders' signals and that the the correlation between any two signals is weakly positive (which seems to be an empirically plausible assumption (Vives, 2011)). The observer also knows that traders' signals are symmetric in the sense of conditions 3. and 4. above (an assumption I make for tractability). However, the observer does not know how traders' signals combine to give the best estimate of π. That is, she does not know the nature of the interdependency in traders' signals: e.g., she does not know if traders may have formed their expectations using the same or very different news sources.

In what follows, I seek to derive predictions for information structures in $\mathcal{I}$ for perfect Bayesian equilibria in symmetric linear demand schedules where each trader optimizes by submiting a demand function of the form

$$q_i(p) = -ae_i + b\zeta_i - cp \tag{6}$$

for some constants a, b, and c if she anticipates that the other traders will do the same. Note that the restriction to symmetric Gaussian environments with linear equilibria is also made in Bergemann & Morris (2013).

Fixing a double auction, what information structure is worst for allocative efficiency? I say that an information structure $I^* \in \Delta(\vec{\zeta}, \pi)$ is worst case if it minimizes allocative efficiency among equilibria in symmetric linear demand schedules or if there does not exist a symmetric linear equilibrium for I^*. Let SLE(I) denotes the set of symmetric linear equilibria for the information structure I. Formally, I^* is worst case if either SLE(I^*) is empty or

$$I^* \in \underset{I \in \mathcal{I},\, \vec{q} \in \mathrm{SLE}(I)}{\arg\min} \ \mathbb{E}_I \left[-\sum_{i=1}^{N} \left(e_i + q_i(p^*)\right)^2 \right]$$

where p^*, defined formally in Example 1, is the market clearing price induced by $\vec{q} = \{q_i\}$. The subscript on the expectation operator indicates its dependence on I.

The following Theorem 2 characterizes all equilibria for information structures in $\mathcal{I}$ and provides conditions for existence and uniqueness. It also presents comparative statics with respect to the correlation among traders' signals. To my knowledge the analysis here is novel for two reasons. First, almost all papers in the literature assume additive signals with errors that are independent across traders whereas I allow for correlation in errors—indeed the worst case information structure features correlated errors. Second, even if errors are assumed to be independent, existing papers assume that the aggregate endowment of traders is common knowledge (Chen & Duffie, 2021; Du & Zhu, 2017) or that signals are informative of the *sum* of the private and common components of values (Rostek & Weretka, 2012). Thus, the characterization of symmetric linear equilibria of this setting is novel to my knowledge.

Theorem 2. *If a symmetric linear equilibrium exists, then it is unique and has demand coefficients a, b, and c characterized by equations* (18)*,* (19)*, and* (20) *in Appendix B. Letting $\rho := \mathrm{corr}(\zeta_i, \zeta_j)$, the following statements hold:*

1. *If $\rho = 0$, a symmetric linear equilibrium exists if and only if*

$$\frac{\sigma_\zeta^2}{\sigma_\epsilon^2 \left(\frac{1}{\kappa}\right)^2 + \sigma_\zeta^2} < \frac{N-2}{N-1}\frac{1}{2}.$$

   The parameter range of equilibrium existence expands when ρ increases.

2. *Conditional on equilibrium existence, allocative efficiency is increasing in ρ.*

3. *The unique worst-case information structure is such that $\zeta_1, \ldots, \zeta_N$ are independent. That is,*

$$\mathbb{E}\left[\pi | \vec{\zeta}\right] = \sum_{i=1}^{N} \zeta_i.$$

4. *Under the condition for existence in Part 1 of this theorem, the worst case allocative*

efficiency, which occurs when $\rho = 0$, is:

$$-N\left[(1 + a^2 - 2a)\sigma_\epsilon^2 + \kappa^2 a^2 \sigma_\zeta^2\right]$$

where

$$a = 1 - \frac{2(N-1)}{N-2} \frac{\sigma_\zeta^2}{\sigma_\epsilon^2 \left(\frac{1}{\kappa}\right)^2 + \sigma_\zeta^2}.$$

5. *Under the condition for existence in Part 1 of this theorem, the maximal price impact, which occurs when $\rho = 0$, is*

$$\Lambda = \frac{1}{N-2}\left(\frac{1}{\kappa} + N \frac{\kappa a \sigma_\zeta^2}{a^2 \sigma_\epsilon^2 + \kappa^2 a^2 \sigma_\zeta^2}\right)$$

where a is as in Part 5.

Remark. *I have assumed that the marginal of π is unknown (though its mean can be inferred from the mean of ζ_i). If $\sigma_\pi^2 \geq N\sigma_\zeta^2$, then this is without loss. If $\sigma_\pi^2 \leq N\sigma_\zeta^2$, then the worst-case information structure when σ_π^2 is known is determined by the minimal possible ρ such that $\pi \geq_{mps} \frac{1}{1+\rho(N-1)} \sum_{i=1}^N \zeta_i$. This ρ is defined by*

$$\frac{N\sigma_\zeta^2}{1 + \rho(N-1)} = \sigma_\pi^2.$$

When σ_π^2 is known and $\sigma_\pi^2 \leq N\sigma_\zeta^2$, under the worst-case information structure, all traders' signals together fully reveal π.

If the marginal of π is known but the marginals of $\{\zeta_i\}$ are unknown then the worst-case information structure is also fully revealing and features positive correlation, but, at least in numerical examples I have computed, this correlation is "small" and tends to zero as σ_π^2 increases.

Part 1 of Theorem 2 implies that if an equilibrium exists when $\rho = 0$, then it must exist for all $\rho > 0$. Parts 2 and 3, imply that the worst-case information structure is the case when $\rho = 0$. Parts 4 and 5 characterize allocative efficiency and price impact under this worst-case information structure, and in doing so provides bounds on these objects that hold across all information structures in $\mathcal{I}$.

Intuitively, $\rho = 0$ is worst-case becase it is harder for a trader to predict how others will trade when signals are independent. Though the result seems intuitive, the story is actually more subtle. Indeed, the proof is somewhat involved because the equilibrium is characterized in terms of a cubic equation which is unwieldy to work with. Moreover, as correlation ρ increases, traders put more weight on their private signals in their demand schedules as well as on their endowments. The net welfare impact depends on relative rates of increase. The ratio of these weights is also nonmonotone in ρ. The proof, in the Appendix, must contend with these complications.

In light of Theorem 2 several questions come to mind. Which mechanisms perform well for these information structures and can best shield against adverse selection? Is there a sense in which the information structure in Theorem 2 is worst case that does not depend on the assumed mechanism? Are there mechanisms that offer better guarantees of allocative efficiency than double auctions? What is the max guarantee? Is there a nontrivial guarantee for any given parameters? What about when the information structure is not necessarily Gaussian and if signals can live in arbitrary, potentially high-dimensional spaces? By studying optimal mechanisms for the worst-case information structures of the type in Theorem 2 it turns out that I will be able to offer answers to these questions.

5.3. Optimal Mechanisms

Throughout, I now assume that the marginal on trader i's posterior expectation ζ_i is some arbitrary distribution G with a finite variance. Note that I am not assuming Gaussianity. I continue to assume that each trader sees a single-dimensional signal that we may take to be ζ_i. I further assume that $\vec{\zeta} = \{\zeta_i\}$ are jointly independent and that

$$\mathbb{E}\left[\pi|\vec{\zeta}\right] = \sum_{i=1}^{N} \zeta_i.$$

It turns out that, for this information structure, I am able to derive optimal mechanisms despite the multiple dimensions of private information.

Lemma 1. *In any mechanism, trader i's optimal message in a Bayes Nash equilibrium is a function of $t_i := \frac{1}{\kappa}e_i - \zeta_i$.*

Proof. Let m^*_{-i} denote the messages chosen by the other traders in an equilibrium of the mechanism. Trader i selects a message m_i to maximize

$$\mathbb{E}\left[-\ \frac{1}{2\kappa}\left(e_i + X_i(m_i, m_{-i}^*)\right)^2\ +\ \sum_{j=1}^{N}\zeta_j\left(e_i\ +\ X_i(m_i, m_{-i}^*)\right)\ -\ T_i(m_i, m_{-i}^*)|\zeta_i, e_i\right].$$

Rearranging, the objective is equivalent to

$$\mathbb{E}\left[-\frac{1}{2\kappa}e_i^2 - \underbrace{\left[\frac{1}{\kappa}e_i - \zeta_i\right]}_{t_i} X_i(m_i, m_{-i}^*) - \frac{1}{2\kappa}X_i(m_i, m_{-i}^*)^2\right.$$

$$\left. + \sum_{j\neq i}\zeta_j X_i(m_i, m_{-i}^*) - T_i(m_i, m_{-i}^*)|\underbrace{\frac{1}{\kappa}e_i - \zeta_i}_{t_i}\right].$$

Thus, the optimal choice of message depends on only t_i. □

By Lemma 1, the mechanism design problem is effectively single dimensional. I solve the analog of (2) for this setting using the same methods as in Section 4. Let H and h denote the cdf and pdf of t_i respectively for any given trader i. Let

$$\hat{e}(t) := \mathbb{E}\left[e_i|t_i = t\right].$$

The following Theorem 3 characterizes the optimal mechanism in the presence of adverse selection.

Condition 2. *It holds that $\hat{e}(t) + \alpha\frac{H(t)}{h(t)}$ and $\hat{e}(t) - \alpha\frac{1-H(t)}{h(t)}$ are weakly increasing in t.*

Theorem 3. *Suppose that Condition 2 holds. Then the unique solution of optimization problem* (11) *for $\{X_i\}$ sets*

$$X_i(\vec{t}) = -v(t_i) + \frac{1}{N}\sum_{j=1}^{N} v(t_j)$$

for each trader i where

$$v(t) = \hat{e}(t) + \alpha\frac{H(t)}{h(t)}, \quad t < t_a$$
$$v(t) = \hat{e}(t) - \alpha\frac{1 - H(t)}{h(t)}, \quad t > t_b$$
$$v(t) = \mathbb{E}\left[v(t_i)\right], \quad t \in (t_a, t_b)$$

for t_a and t_b such that v is continuous: $\lim_{e \to e_b} v(e) = v(e_a) = v(e_b)$.

As in the basic model, optimal mechanisms are characterized explicitly given three unknowns $(t_a, t_b, \mathbb{E}[v(t_i)])$ and there are special cases where the unknowns can also be characterized almost in closed form.

Corollary 3.1. *Suppose that Condition 2 is satisfied and h is symmetric about its mean. Then under the optimal mechanism,*

$$\mathbb{E}[v(t_i)] = \mu_e,$$

$$\hat{e}(t_a) + \alpha \frac{H(t_a)}{h(t_a)} = \mu_e,$$

and

$$\hat{e}(t_b) - \alpha \frac{1 - H(t_a)}{h(t_a)} = \mu_e.$$

Using Corollary 3.1, I am able to prove the following comparative statics.

Proposition 4. *Suppose that e_i and ζ_i are Gaussian. Then the following comparative statics hold:*

1. *As σ_ζ^2 increases, the region of binding participation constraints expands in that $t_b - t_a$ increases without bound. Thus, as $\sigma_\zeta^2 \to \infty$, for any given $\alpha > 0$, the probability of trade vanishes in that*

$$\mathbb{P}\left\{ \sum_{i=1}^{N} |X_i(\vec{t})| > 0 \right\} \to 0.$$

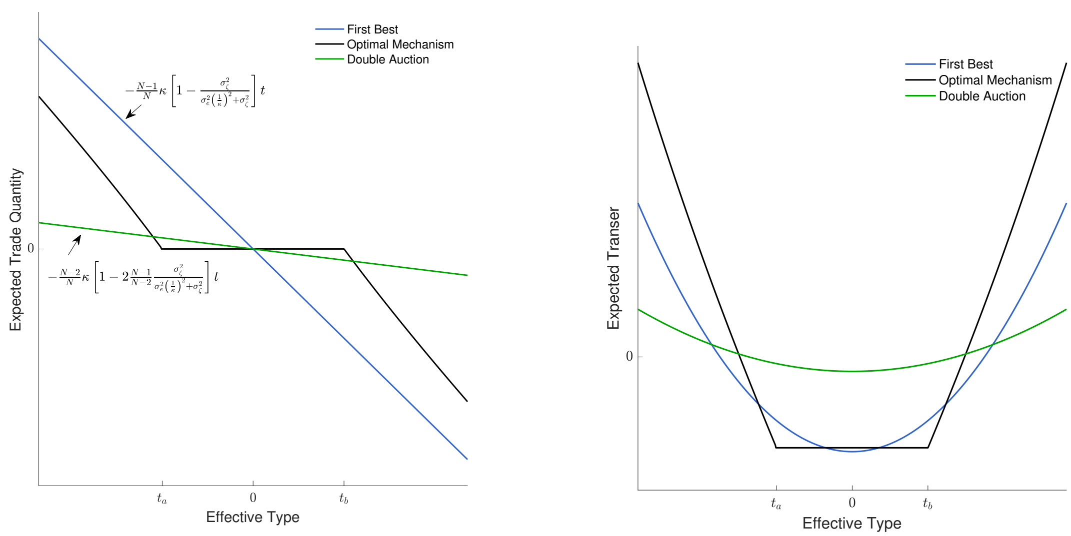
2. *As σ_ζ^2 increases the revenue-efficiency frontier shifts down and left.*

3. *For any given $\alpha > 0$, as σ_ζ^2 increases, the expected utility gain $\hat{U}_i(t_i) - U(t_i, 0)$ of any given trader i and the expected trade volume $\mathbb{E}|\sum_{i=1}^{N} X_i(\vec{t})|$ decrease.*

Proposition 4 implies that as adverse selection becomes worse in that σ_ζ^2 increases, the region of binding participation constraints increases, trade volume decreases and gradually vanishes, and the revenue-efficiency frontier shifts down and to the left. Notably, for any finite level of σ_ζ^2 there is always some probability of trade. This is contrast with the case of the double auction where (symmetric linear) equilibrium ceases to exist at a finite level of σ_ζ^2.

5.4. Illustrative Example

To further illustrate the properties of optimal mechanisms in the presence of adverse selection, I now present an example where $e_i \sim N(0, \sigma_e^2)$ and $\zeta_i \sim N(0, \sigma_\zeta^2)$. Panel (a) of Figure 4 plots the expected trade quantity as a function of effective type. As one can see, many qualitative properties are the same as the case without adverse selection. As σ_ζ^2 increases, the gap between the double auction and first best widens. The region of binding constraints also expands and eventually diverges.



(a) (b)

Figure 4: Interim allocation rule and interim transfer rule

Notes: Comparison of interim allocation rules and interim transfer rules when when $e_i \sim N(0, \sigma_e^2)$ and $\zeta_i \sim N(0, \sigma_\zeta^2)$.

As seen from panel (b), perhaps surprisingly, when σ_ζ^2 is sufficiently high, there is more curvature under first best than under the double auction which is in stark contrast with the case without adverse selection. This must be so because, as seen from panel (a), as σ_ζ^2 increases the ratio of the slopes of the blue and green curves diverges. That ratio is always bounded from above for the case without adverse selection.

Below in Figure 5, I plot the revenue-efficiency frontier as I increase σ_ζ^2 in increments of .5. As one can see the fronteir shifts down before gradually converging to a vertical line at autarky. The double auction reaches the autarky outcome much faster than under optimal mechanisms.

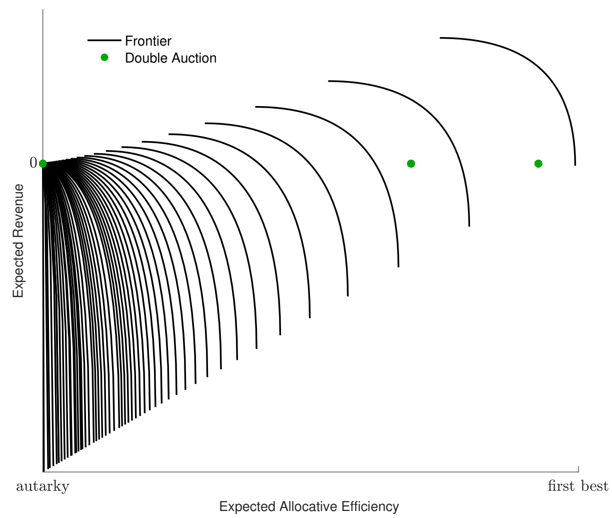


Figure 5: Revenue-Efficiency Frontier

Notes: The revenue-efficiency frontier for different values of σ_ζ^2 when $e_i \sim N(0, \sigma_e^2)$ and $\zeta_i \sim N(0, \sigma_\zeta^2)$.

5.5. Implementation: Double Auction with Transaction Fees

It turns out, that even with adverse selection, it is possible to bring the double auction to the revenue-efficiency frontier using transaction fees.

Proposition 5. *For any given $\alpha \in [0,1]$, the allocation rule $\{X_i\}_{i=1}^N$ in Theorem 1 is implementable by a double auction with transaction fee $\mathcal{T}$ characterized explicitly by equations of the Appendix.*

The proof is similar to the case without adverse selection except for the inclusion of another term in the transaction costs to offset the price inference effect about π. A large literature studies information aggregation in financial markets for a fixed mechanism. Corollary 3.2 implies that once one endogenizes the mechanism, the revenue-maximizing incentives of the designer worsen price informativeness.

Corollary 3.2. *Suppose that e_i and ζ_i are Gaussian. Whenever a symmetric linear equilibrium of the double auction without transaction fees exists, the price in that equilibrium is more Blackwell informative than in the equilibrium of the double auction with transaction fees set optimally for any $\alpha > 0$.*

5.6. Robust Mechanisms for Allocative Efficiency

Is there a sense in which the information structures that I consider are worst-case that does not depend on assuming the double auction mechanism? What is the maximal guarantee of allocative efficiency that can be achieved by a mechanism within a large class of information structures? I next address these questions.

Suppose now that each trader i sees a signal $s_i \in \mathbb{S}_i$ of the fundamental π where the signal space $\mathbb{S}_i$ is an arbitrary measurable set. Let

$$\mathcal{I} = (\{\mathbb{S}_i\}, F)$$

denote the information structure where $F \in \Delta(\pi \times \Pi_{i=1}^{N} \mathbb{S}_i)$ is the joint distribution of signals and the fundamental.

I assume that the designer knows some aspects of $\mathcal{I}$. Let $\zeta_i = \mathbb{E}[\pi|s_i]$ denote trader i's expectation of the common component of the payoff conditional on her signal s_i. I assume that $\zeta_i \sim G$ is known a and that $\{\zeta_i\}$ are pairwise stohastically monotone—that is, a higher realization of ζ_j leads to an increase in the conditional distribution of ζ_i given ζ_j in the sense of first-order stochastic dominance for any pair i and j. Further, for simplicity, I continue to assume that it is known that traders' endowments are independent of each other and of $\{\zeta_i\}$ and π.

That is, the designer may have some sense of traders' expectations about the fundamental of the asset, but does not know exactly how those expectations arose (the details of the signal structure) and in particular, how they are correlated across traders, and thus does not know how these signals combine to give the best estimate of the common value

$$\pi(s_1, ..., s_N) = \mathbb{E}\left[\pi|s_1, ..., s_N\right]$$

or even what spaces these signals live in. An alternative interpretation is not based on information necessarily, but that values are interdependent for other reasons and the designer simply has no idea how values are interdependent.

Here are two examples of information structures with independent signals in the ambiguity set (though there are of course many other examples with correlation among signals).

Example 2. *Suppose that signals are just traders' posterior expectations $\zeta_1, ..., \zeta_N$ and that these are independent random variables (and have mean zero as I had normalized* *earlier). Then both*

$$\pi(\zeta_1, .., \zeta_N) = \sum_{i=1}^{N} \zeta_i$$

and

$$\pi(\zeta_1, .., \zeta_N) = -1 + \Pi_{i=1}^{N}(\zeta_i + 1)$$

can be are consistent with information structures in the ambiguity set.

In the face of this ambiguity, the designer looks for a mechanism that is robust in the sense that it delivers the highest guarantee across all possible information structures consistent with the marginal G of signals.

To state the objective a bit more formally I first define some notation. Given a market design $\mathcal{M}$ and information structure $\mathcal{I}$, let the set of Bayes Nash Equilibria be denoted $\mathcal{E}(\mathcal{M}, \mathcal{I})$.

Let

$$\mathrm{AE}(\mathcal{M}, \mathcal{I}, \mathcal{E}) = -\sum_{i=1}^{N} \mathbb{E}_{\mathcal{I},\mathcal{E}}\left[\frac{1}{2\kappa}\left(e_i + X_i(\vec{m}^*)\right)^2\right]$$

where $\{X_i\}$ is the allocation rule in $\mathcal{M}$ and where the expectation is taken with respect to $\vec{m}^*$ according to the probability distribution induced by $\mathcal{I}$ and $\mathcal{E}$.

I adopt the solution concept of a strong maxmin solution defined in Brooks & Du (2021).

Definition 5.1. A strong maxmin solution $(\mathcal{M}, \mathcal{I}, \mathcal{E})$ consists of a mechanism $\mathcal{M}$, information structure $\mathcal{I}$ and Bayes-Nash equilibrium $\mathcal{E}$ such that

1. For any mechanism $\mathcal{M}'$ and any equilibrium $\mathcal{E}$ of $(\mathcal{M}', \mathcal{I})$, it holds that $\mathrm{AE}(\mathcal{M}, \mathcal{I}, \mathcal{E}) \leq \mathrm{AE}(\mathcal{M}', \mathcal{I}, \mathcal{E}')$

2. For any information structure $\mathcal{I}'$ and any equilibrium $\mathcal{E}'$ of $(\mathcal{M}, \mathcal{I}')$, it holds that $\mathrm{AE}(\mathcal{M}, \mathcal{I}', \mathcal{E}') \geq \mathrm{AE}(\mathcal{M}, \mathcal{I}, \mathcal{E})$

3. $\mathcal{E}$ is an equilibrium of $\mathcal{M}, \mathcal{I}$.

The following theorem assumes compact support but this is largely technical. For example, suppose it is known that $(\vec{\zeta}, \pi)$ is Gaussian, then the result applies.

Theorem 4. *Suppose that each ζ_i has compact support and G is such that Condition 2 holds when $\alpha = 0$. Then there is a strong maxmin solution where the designer sets the* *allocation rule in Theorem 3 with $\alpha = 0$ and where nature selects the information structure assumed in Subsection 5.3.*

Proof. Recall that trader i selects a message m_i to solve

$$\max_{m_i} \mathbb{E}\left[- \frac{1}{2\kappa}\left(e_i + X_i(m_i, m^*_{-i})\right)^2 + \pi\left(e_i + X_i\left(m_i, m^*_{-i}\right)\right) - T_i(m_i, m^*_{-i}) | \zeta_i, e_i \right].$$

Now suppose we set the transfer rule

$$T_i(\vec{m}) = -m_i \hat{X}_i(m_i) + \int_{\underline{t}}^{m_i} \hat{X}_i(s)\, \mathrm{d}s - \frac{1}{2\kappa_i} X_i(\vec{m})^2 + c \tag{7}$$

For the transfer rule in (7), all cross terms involving m_i and m^*_{-i} drop out from the objective. Thus, each trader submits the same message t_i under any information structure.

Now, I directly verify that allocative efficiency is at its lowest level under the candidate worst-case information structure where signals are independent. We have

$$\begin{aligned}
\mathbb{E}\left[\sum_{i=1}^{N}\left(e_i + X_i(\vec{t})\right)^2\right] &= N\sigma_e^2 - \frac{N-1}{N}\mathbb{E}\left[\sum_{i=1}^{N} e_i v(t_i)\right] + \mathbb{E}\left[\sum_{i=1}^{N} X_i^2\right] \\
&= N\sigma_e^2 - \frac{N-1}{N}\mathbb{E}\left[\sum_{i=1}^{N} e_i v(t_i)\right] \\
&+ \mathbb{E}\left[\sum\left(-\frac{N-1}{N} v(t_i) + \frac{1}{N}\sum_{j\neq i} v(t_i)\right)^2\right] \\
&= N\sigma_e^2 - (N-1)\mathbb{E}\left[e_i v(t_i) + \left(\frac{N-1}{N} + \frac{1}{N^2}\right) v(t_i)^2\right] \\
&- \frac{1}{N}\sum_{i=1}^{N}\sum_{j\neq i} \operatorname{cov}\left(v(t_i), v(t_j)\right).
\end{aligned}$$

Because the coefficient on covariances are negative and because $\{\zeta_i\}$ are pairwise stochastically monotone, it follows that independence is the worst-case outcome. □

Theorem 4 illustrates the sense in which the information structures we have studied so far are worst-case without fixing, at the outset, the double auction mechanism. The main weakness of the result is that it allows for a budget deficit. However, this is in alignment with the classical analysis of Vickrey (1961), Clarke (1971), and Groves (1973) on efficient implementation and also Bergemann & Välimäki (2002). In some cases, the government can enforce participation. Following the classic papers by Vickrey, Clark, and Groves for the private value environment, there has been a substantial literature investigating efficient implementation with interdependent values in the early to mid 2000s. Dasgupta & Maskin (2000) and Jehiel & Moldovanu (2001) show that achieving the ex-post efficient implementation is generally impossible with interdependent values. For my model setting, I take a different perspective from this literature by focusing instead on deriving maximal guarantees of efficiency that hold across a range of informational environments/interdependencies.

Let me also note that the proof of Theorem 4 shows that the mechanism is robust, not just to information, but to traders' higher order beliefs. In other words, it does not rely on common knowledge of the information structure among traders.

Corollary 4.1. *The strong maxmin mechanism in Theorem 4 is robust to traders' higher order beliefs about other traders' beliefs about the information structure.*

In general, deriving robust guarantees when the designer also cares about revenue is intractable. As seen from the proof of Theorem 4, the main difficulty is that traders may not find it individually rational to participate as we vary the information structure. Indeed, conditional on full participation of each trader i, one can always construct the transfer rule such that trader i always optimally reports t_i regardless of the information structure by the same argument as in the proof of Theorem 4. However, given that efficiency and revenue are typically quite "aligned," our analysis of optimal mechanisms when $\alpha > 0$ for worst case information structures for efficiency is still instructive of how to design trade to maximize revenue in the presence of severe adverse selection concerns.

6. Multiple Assets

Can the efficient allocation be implemented by double auctions with transaction fees when there are multiple assets? Is it more or less costly to implement the efficient allocation when there are more assets? Does this depend on the correlation among asset payoffs or among asset endowments? What do revenue-maximizing allocations look like? Can they be implemented by double auctions with transaction fees? In this section, I investigate these questions for a multi-asset version of the model.

6.1. Environment

There are now multiple assets $a \in A$. Trader i's utility is

$$U(\vec{e}_i + \vec{X}_i, T_i) = \vec{\mu}_\pi \cdot \vec{X}_i - \frac{1}{2\kappa}\left(\vec{e}_i + \vec{X}_i\right)^\top \Sigma \left(\vec{e} + \vec{X}_i\right) - T_i \tag{8}$$

where

- $\vec{\pi} = \{\pi_a\}_{a \in A}$ is the vector of asset payoffs
- $\vec{\mu}_\pi = \{\mu_{\pi_a}\}_{a \in A}$ is the mean of $\vec{\pi}$
- Σ is the covariance matrix of $\vec{\pi}$
- $\vec{e}_i = \{e_{ia}\}_{a \in A}$ is the vector of trader i's asset endowments
- $\vec{X}_i = \{X_{ia}\}_{a \in A}$ is the vector of trade quantities of the assets
- T_i is the net transfer
- $\kappa > 0$ is a constant.

I assume that endowments $\vec{e} = \{\vec{e}_i\}$ are independent across traders and are also independent of assets' payoffs $\vec{\pi}$. However, for a given trader i, her endowments of the different assets may be correlated. I retain all other aspects of the baseline model of Section 3.

6.2. Optimal Mechanisms

Proposition 6 states that it is possible to implement the efficient allocation by running $|A|$ separate double auctions with transaction fees. Moreover, it is possible to do so with ex-ante budget balance, regardless of correlation among endowments or among payoffs of the different assets. This result is perhaps surprising because traders can not make their trades in one auction contingent on the prices in the other auctions. One might have thought this would make implementation impossible.

Proposition 6. *The efficient allocation is implementable by $|A|$ separate double auctions with transaction fees. It is implementable with either ex-ante budget balance or in an ex-post equilibrium.*

The proof proceeds by constructing derivative securities with payoffs that are uncorrelated but span the payoffs of the underlying assets. Because the derivatives are uncorrelated, a trader's incentives to trade one derivative does not depend on the trader's endowment of any other derivative. We can therefore organize separate double auctions for each derivative and calibrate separate transaction fees to target the efficient allocation of each derivative (and thus asset) as in Proposition 3.

The analysis here is complementary to that of Rostek & Yoon (2023) who consider the design of derivative securities in an analogous environment but fix double auction mechanisms and Gaussian uncertainty. Rostek & Yoon (2023) show that optimal derivative securities are typically correlated and depend critically on the covariance of the underlying payoffs and are in general difficult to characterize sharply. Proposition 6 implies that if one can jointly design securities and transaction fees then it is optimal to host exchanges for uncorrelated securities. An added benefit of doing so is that traders have dominant strategies for their choices of demand schedules. In general, if securities are not uncorrelated and are traded in separate double auctions, there will be complicated cross-exchange cross-inference effects due to the fact that clearing price on one exchange is informative of the clearing price on other exchanges (Rostek & Yoon, 2021; Chen & Duffie, 2021; Rostek & Yoon, 2023). As a result of these effects, any equilibrium can not be in dominant strategies.

In general, solving for optimal mechanisms when the designer has a revenue-maximizing motive is intractable due to the multiple dimensions of private information. I am only able to derive optimal mechanisms under the following condition.

Condition 3. *The following hold:*

- $\vec{\pi} = \{\pi_a\}$ *are independent Gaussian random variables with variance* σ_π^2.
- $\vec{e} = \{e_{ia}\}$ *are independent Gaussian random variables with variance* σ_e^2.

Note that, besides Gaussianity, either one of the properties in Condition 3 is without loss on its own and is just a normalization, however the two together is restrictive. This is because one can always construct securities so that either payoffs or endowments are independent. On can also redefine securities by scaling them up or down so that the variances of all payoffs or all endowments are the same. However, one can not in general do this to *simultaneously* satisfy both properties of Condition 3. Note also that I make no restrictions on the mean $\vec{\mu}_e = \{\mu_{ea}\}_{a\in A}$ of $\vec{e}$ or the mean $\vec{\mu}_\pi = \{\mu_{\pi_a}\}_{a\in A}$ of $\vec{\pi}$.

Proposition 7. *Suppose that Condition 3 holds. Let* $r_i := \sqrt{\sum_{a\in A}(e_{ia}-\mu_{ea})^2}$ *be the length of* $\vec{e}_i - \vec{\mu}_e$ *for each trader* i*. For each* i *and* a*, let* $\tilde{e}_{ia} := (e_{ia}-\mu_{ea})/r_i$*. Then the unique solution to* (32) *for the optimal allocation rule* $\vec{X}$ *sets*

$$X_{ia}(\vec{e}) = -v(r_i)\tilde{e}_{ia} + \frac{1}{N}\sum_{j=1}^{N} v(r_j)\tilde{e}_{ja}$$

for each i *and* a *where*

$$v(r) = 0, \quad r \in [0, \sqrt{\alpha}\sigma_e),$$
$$v(r) = r - \frac{\alpha\sigma_e^2}{r}, \quad r \in [\sqrt{\alpha}\sigma_e, \infty)$$

and where $\alpha \in [0,1]$ *is the weight on revenue in the objective* (32) *stated formally in Appendix C. The optimal allocation rule is implemented by transfer rules of the form in Appendix equation* (34).

Proposition 7 shows that many properties of optimal mechanisms generalize when there are many assets, at least under Condition 3. To prove the proposition, I exploit the "radial symmetry" of the optimization problem by building on insights from Armstrong (1996) and Wilson (1993) who study a multiproduct monopolist.

Consider Figure 6 below which depicts the case of two assets. Under the optimal mechanism, the region of binding participation constraints is a closed disk of radius $\sqrt{\alpha}$ and is thus larger when the objective places more weight on revenue. Traders with types in this region expect to trade zero of both assets. Outside the region, the only binding incentive constraints lie along radial lines, are local, and are in the direction of the disk. Each traders' type is distorted only radially. That is, the optimal mechanism distorts only the magnitude of a trader's endowment $\vec{e}_i$. The proof of Proposition 7 is in Appendix C.

Unfortunately, it is not possible to implement the optimal allocation in dominant demand-submission strategies using separate double auctions with transaction fees. However, *it is possible* to do so with cross-exchange transaction fees, that is, transaction fees that can be made contingent on trades and prices on all exchanges.

Definition 6.1. A double auction with cross-exchange transaction fee $\mathcal{T} : \mathbb{R}^{2|A|} \Rightarrow \mathbb{R}$ is an exchange mechanism that operates in three stages:

1. Each trader i submits a demand schedule $q_{ia} : \mathbb{R} \Rightarrow \mathbb{R}$ for each asset a.

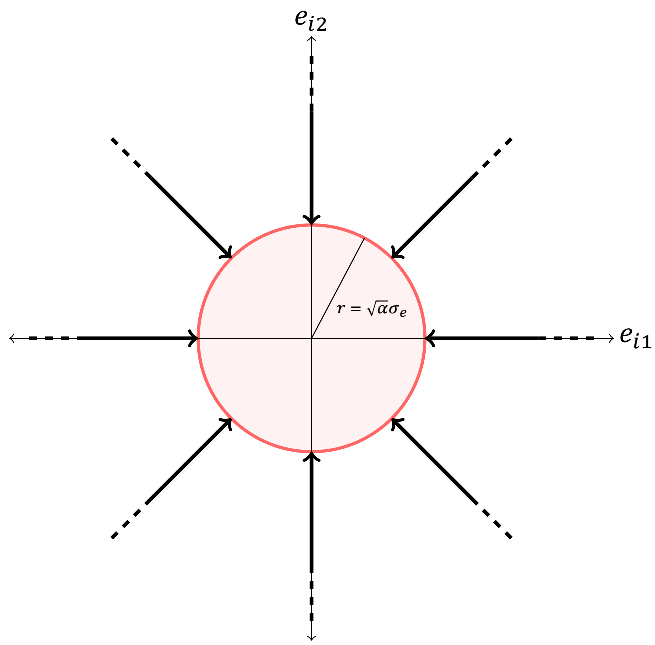


Figure 6: Binding participation and incentive constraints

Notes: The figure illustrates the binding participation and incentive constraints for the case of two assets. In the general case with $|A| \geq 2$ assets, the region of binding participation constraints is a ball in $\mathbb{R}^{|A|}$ with radius $\sqrt{\alpha}$. The binding incentive constraints continue to point in the direction of the ball and bind only radially.

2. The clearing price p_a for each asset a is computed: $\sum_{i=1}^{N} q_i(p) = 0$. If there does not exist a unique clearing price for each asset a, then no trades or transfers are executed.

3. If a unique clearing price p_a exists for each asset a then trader i pays in total $\sum_{a \in A} p_a q_i(p_a) + \mathcal{T}(\{q_{i_a}(p_a), p_a\}_{a \in A})$ in return for $q_{ia}(p_a)$ units of each asset a.

Note that though cross-exchange transaction fees depend on trade on all exchanges, each exchange nonetheless clears separately in that the demand for any given asset by any given trader is contingent on only the price of that asset. In this sense, the implementation is still relatively simple.[12]

Proposition 8. *Suppose that Condition 1 holds. Then the optimal allocation in Proposition 7 for any given* $\alpha \in [0, 1]$ *can be implemented by a double auction with a cross-*

[12]Most exchanges clear separately, though there are exceptions (Rostek & Yoon, 2021; Chen & Duffie, 2021; Budish et al., 2023).

exchange transaction fee $\mathcal{T}$ defined by

$$\mathcal{T}\left(\{q_{i_a}(p_a), p_a\}_{a\in A}\right) = b\left(\sum_{a\in A}\left(q_{ia}(p) + \frac{1}{N}p_a\right)^2\right) + \frac{1}{2}c\sum_{a\in A} q_{ia}(p)^2 + d$$

where $b : \mathbb{R} \to \mathbb{R}$ satisfies

$$b(x) = \frac{N-1}{N}\frac{\sigma_\pi^2}{\kappa}\left[\int_0^{\sqrt{x}} v^{-1}(s)\,\mathrm{d}s - \frac{1}{2}x\right], \quad x \geq 0,$$

$$c = -\frac{\sigma_\pi^2}{\kappa(N-1)},$$

and d is a constant set sufficiently high so that participation is individually rational for all types of all traders.

Though Condition 3 is restrictive, Proposition 8 further demonstrates the remarkable effectiveness of transaction fees even in a setting where there are multiple assets and when the designer may care about revenue. For this setting at least, there is no need for cross-exchange market clearing. That is, there is no need to allow a trader's demand in one exchange to be made contingent on the prices in other exchanges. This offers a complementary perspective to Budish et al. (2023) who argue for a market design that features cross-asset market clearing. Though cross-asset clearing can be very beneficial for welfare, Propositions 6 and 8 show it may not be needed in all instances, especially when there is freedom to set transaction fees or to design derivatives.

7. Heterogeneity

Our analysis so far has assumed that traders are ex-ante symmetric both with respect to the probability distributions of their endowments and their holding capacities. In reality, these likely differ across traders. Though symmetry is often assumed in the market microstructure literature, some papers allow for heterogeneity which can give rise to new and interesting phenomenona. By allowing for heterogeneity, we will be able to investigate questions such as: How should trade be designed in the presence of heterogeneity? Should retail traders with low holding capacity receive better or worse terms of trade than sophisticated institutional traders who typically have higher holding capacity? Are double auctions with transaction fees still optimal?

7.1. Environment

I retain all aspects of the baseline model of Section 3 except I now allow the cdfs $\{F_i\}$ of endowments and also holding capacities $\{\kappa_i\}$ to differ across traders. The analog of the designer's objective (2) with these changes is stated formally in the Appendix (see (38)).

7.2. Optimal Mechanisms

The derivation of optimal mechanisms is analogous to that of Section 3 except now I require a stronger technical condition.

Condition 4. *It holds that $e - \alpha\frac{1-F_i(e)}{f_i(e)}$ and $e + \alpha\frac{F_i(e)}{f_i(e)}$ are weakly increasing in e and f_i has full support on $\mathbb{R}$ for each trader i.*

The full support assumption is needed to ensure that there exist types of each agent who expect to trade zero.[13] If it is violated, we can simply perturb each f_i to have full support, calculate the optimal mechanism using Theorem 5, and then take limits as the mass in the tails vanish.

Theorem 5. *Suppose that Condition 4 holds. Then the unique solution of the optimization problem* (38) *for $\{X_i\}$ sets*

$$X_i(\vec{e}) = -v_i(e_i) + \frac{\kappa_i}{\sum_{j=1}^N \kappa_j}\sum_{j=1}^N v_i(e_j)$$

for each trader i where

$$\begin{aligned} v_i(e) &= e + \alpha\frac{F_i(e)}{f_i(e)}, \quad e < e_{ia} \\ v_i(e) &= e - \alpha\frac{1-F_i(e)}{f_i(e)}, \quad e > e_{ib} \\ v_i(e) &= \frac{\kappa_i}{\sum_{j\neq i}\kappa_j}\sum_{j\neq i}\mathbb{E}\left[v_j(e_j)\right], \quad e \in [e_{ia}, e_{ib}] \end{aligned}$$

for e_{ia} and e_{ib} such that v is continuous: $\lim_{e\to e_{ib}} v(e) = v(e_{ia}) = v(e_{ib})$.

Now, each trader unloads her virtual endowment in the auction, but absorbs a fraction of the aggregate virtual endowment that is proportional to her risk capacity. Computing equilibria is now significantly more difficult because there are $3N$ unknowns $\{e_{ia}, e_{ib}, \mathbb{E}\left[v_i(e_i)\right]\}_{i=1}^N$.

[13]These types need not exist if traders are sufficiently heterogeneous in their risk capacities.

However, importantly, there continue to be settings where the optimal mechanism can be computed almost in closed-form for which I am able to derive comparative statics.

Corollary 5.1. *Suppose that Condition 4 holds. If $\vec{\mu}_e = [\mu_{1e}, ..., \mu_{Ne}]$ and $\vec{\kappa} = [\kappa_1, ..., \kappa_N]$ are colinear and each F_i is symmetric about its mean, then $\mathbb{E}\left[v_i(e_i)\right] = \mu_{ie}$ and thus,*

$$e_{ia} + \alpha \frac{F_i(e_{ia})}{f_i(e_{ia})} = \mu_{ie},$$

and

$$e_{ib} - \alpha \frac{1 - F_i(e_{ib})}{f_i(e_{ib})} = \mu_{ie}$$

for each i.

At first thought the colinear requirement seems quite restrictive. It is certainly with loss but it holds whenever each F_i has a zero mean or, if each F_i has a common mean (but may differ in all higher-order moments) and all traders have the same holding capacity $\kappa_i = \kappa_j$ for each i, j. The symmetry assumption is restrictive but holds whenever $\{F_i\}$ are Gaussian, a common assumption in the market microstructure literature.

Let me note that in the case when each F_i is uniform, we can often compute the optimal mechanism in closed form, just as in the basic model.

Remark. *If each F_i is uniformly distributed, as long as there exists interior types who trade zero in expectation, then the optimal mechanism is computed in closed form by solving a system of linear equations.*

7.3. Comparative Statics

Using Corollary 5.1, I am able to derive the following comparative statics. Proposition 9 focuses on the special case when each F_i is Gaussian with a zero mean. A nonzero mean ex-ante would be an ex-ante source of heterogeneity between prospective buyers and sellers which I seek to shut down.

Proposition 9. *Suppose that the cdf F_i of trader i's endowment is Gaussian with mean zero and variance σ_{ie}^2. Then the following comparative statics hold:*

1. *The regions of binding participation constraints $\{[e_{ia}, e_{ib}]\}$ do not depend on $\{\kappa_i\}$.*

2. *Suppose that σ_{ie}^2 increases for trader i. Then, if $\alpha > 0$:*

(a) *The range* $[e_{ia}, e_{ib}]$ *of binding participation constraints expands for trader* i *but has no effect on the range* $[e_{ja}, e_{jb}]$ *for any trader* $j \neq i$.

(b) *The expected utility gain of trader* i *decreases.*

(c) *The expected utility gain of trader* $j \neq i$ *is unaffected.*

3. *If* κ_i *increases for trader* i *then:*

(a) *The expected utility gain decreases for trader* i.

(b) *The expected utility gain increases for trader* $j \neq i$.

Proposition 9 shows how a trader's utility gains from trade depends on the other traders in the market as well as her own characteristics under the optimal mechanism. Perhaps surprisingly, as seen from Part 2c, for any given trader, an increase in the variability of the other traders' endowments has no effect on her expected gains from trade. Part 3 of the proposition shows that a trader benefits when the other traders in the economy have lower holding capacities. This is because she will be better compensated by the exchange.

7.4. Implementation: Double Auction with Transaction Fees

One might wonder whether a double auction with transaction fees can still implement the revenue-efficiency frontier when there is heterogeneity. It turns out that this is possible as long as the transaction fees can be tailored to each trader.

Proposition 10. *A double auction with trader-specific transaction fees can implement any outcome on the revenue-efficiency frontier.*

The intuition is the same as for the basic model. One can compute an explicit characterization of the implementing transaction fees which are of a similar in form to that of Proposition 3 except with coefficients that depend on the individual traders' holding capacities $\{\kappa_i\}$. Thus, even in the presence of heterogeneity, transaction fees remain a powerful tool. With heterogeneity, exchanges have incentives to tailor trading fees to individual traders whether it be to promote efficiency or revenue. In practice,"effective" trading fees often differ across traders—large instutitutional traders can often negotiate their trading fees.[14]

[14]For example, see https://optiver.com/insights/a-little-understood-cost-of-trading-options-in-the-us/.

8. Dynamics

The majority of this paper's analysis is conducted in a static setting. However, there is a sense in which the results extend to dynamic settings.

Suppose that trader i's utility is of the form

$$U(\{X_i(t), T_i(t); t \geq 0\}) = \int_0^{\tau} e^{-\delta t}\left[-\frac{1}{2\kappa}\,(e_i + X_i(t))^2 - T_i(t)\right] \mathrm{d}t$$

where $X_i(t)$ denotes the quantity that a trader holds in her inventory at time t in excess of her initial endowment and where $\tau \in \Delta\,(\mathbb{R} \cup \{\infty\})$ is the terminal time.

The designer's problem is to set a path of allocations and transfers for each trader i to maximize a convex combination of the net-present-values of revenue and allocative efficiency subject to the constraints that 1) it is individually rational for each trader to participate in the mechanism *at each instant in time*, 2) it is incentive compatible for each trader to report her endowment truthfully, 3) the mechanism does not absorb or supply any net quantity of the asset at any time. A formal statement of the objective is in Appendix E.

Proposition 11 implies that the dynamic problem simplifies to a static problem and so the analysis of the of the basic model of Section 3 applies straightforwardly.

Proposition 11. *In any solution to* (39)*, all trades are executed at* $t = 0$. *That is,* $X_i(t) = X_i(0)$ *and* $T_i(t) = 0$ *for each* $t > 0$ *for each i.*

Of course, whenever the weight α on revenue is positive, this result relies on 1) the strong commitment power of the designer who must resist the temptation to execute trades in the future and 2) the absence of renewed endowment shocks over time. Let me note that several papers study the setting with a single initial shock which seems realistic for shorter time horizons (see e.g., Vayanos (1999); Almgren & Chriss (2001)).[15] Though the reliance on commitment is an important limitation, my analysis nevertheless yields an upper bound on what is achievable with weaker forms of commitment. Full commitment is a conventional assumption in the literature on dynamic mechanism design (Bergemann & Välimäki, 2010; Pavan et al., 2014; Athey & Segal, 2013).

When there are renewed endowment shocks over time, I am more limited in what I can say. However, I am able to derive conditions when the ex-post efficient allocation is implementable.

[15] Vayanos (1999) allows for renewed endowments shocks in the model setup but much of the analysis focuses attention on the limiting case as the variance of future shocks vanishes.

Proposition 12. *Suppose that each trader i has an endowment process of the form*

$$e_{it} = e_{i0} + Z_{it}, \quad t \in [0, \tau]$$

where $\{Z_{it}; t \in [0, \tau]\}_{i=1}^{N}$ *are independent Brownian Motions and* $\{e_{i0}\}_{i=1}^{N}$ *are independent Gaussian random variables with common mean* μ_e*. Then the efficient allocation can be implemented with an ex-ante budget surplus by a continuously run double auction with transaction fees.*[16]

Thus, even in a dynamic setting with renewed endowment shocks, a double auction with transaction fees can be a powerful policy tool. In fact, it is possible to use them to achieve efficiency with a budget surplus which was not possible in the static model. This is because gains from future trades can be used to relax participation constraints at $t = 0$. Indeed, as a trader grows patient $\delta \to 0$, her initial private information $\{e_{i0}\}$ becomes irrelevent for her gains from participation because of the stationarity and ergodicity of the endowment process. Thus, it is as though interim participation constraints are relaxed to ex-ante participation constraints.

This result is in contrast with Vayanos (1999) who shows that, fixing the double auction, as the frequency of trade increases, welfare decreases because traders choose to trade less aggressively.[17] Proposition 12 implies that opposite can be true if the designer can introduce transaction fees, which are often employed in practice, into the mechanism. Thus, whether the result in Vayanos (1999) holds when the exchange can set transaction fees is an open question and may depend on the strength of the exchange's motives for revenue maximization.

9. Conclusion

The objective of this paper has been to investigate the optimal design of exchange under a variety of market conditions for a workhorse model in the literature on finance market microstructure. There is only a relatively small literature at the intersection of finance market microstructure and mechanism design of trade. This paper has made some progress toward filling this gap.

[16]A version of this result first appears in the permanent working paper Chen & Zhang (2020) of which I am a coauthor.

[17]A low frequency of trade approximates a static environment.

A contribution of this paper is to identify the quadratic holding cost model, which has become a workhorse in market microstructure in just the past fifteen years or so, as a highly tractable environment for mechanism design analysis. The model is more tractable than CARA utility models and also can be studied without noise traders which is critical for mechanism design. The environment is flexible enough that I am able to develop a unified treatment of optimal exchange design that allows for adverse selection, multiple assets, heterogeneity, and to a more limited degree dynamics. Though in some of these environments, I require more restrictive assumptions than others, the analysis in a common framework is valuable because it enables direct comparisons of results.

Some implications of my analysis include: 1) optimal mechanisms may distort both ex-post and interim allocations for traders with extreme types; 2) double auctions where trades clear separately can often implement outcomes on the efficiency-revenue frontier if they can be augmented with transaction fees and this is so even when there are multiple assets, heterogeneous traders, or dynamics; 3) worst-case information structures (within a large class) for allocative efficiency entail traders receiving independent single-dimensional signals about the asset's payoff; 4) a robust mechanism for allocative efficiency features allocation rules that are additively separable across traders; 5) prices may be less informative about fundamentals if exchanges have profit-maximizing motives.

There are two promising directions for future work. One direction would be to apply the characterization of optimal mechanisms in this paper to investigate conditions when existing results in the literature hold or are altered when the mechanism itself is allowed to vary endogenously. As a proof of concept, this paper has done this for analysis of information aggregation in prices, design of derivative securities or innovation in market clearing, and to a more limited extent, the optimal frequency of trade. Results like these offer a complementary perspective to existing work by helping us to understand the particular role that the chosen trading mechanism plays in the analyses.

A second promising direction would be to investigate a model with competing mechanisms in the sense of Biais et al. (2000). My preliminary investigations suggest that a model of competing *exchange* mechanisms often leads to tipping onto a single exchange. Developing a cohesive analysis with competing exchanges that coexist in a stable equilibrium would have great value. A plausible conjecture is that outcomes in a model with competing exchanges may appear similar to outcomes in this paper when the designer places a lower weight on revenue. Investigating whether this is so and conditions for when competition among exchanges exacerbates or mitigates inefficiencies would be a natural next step to explore that is of clear relevance for policy.

Online Appendix

A. Omitted Material for Section 4

Lemma 2. *It is without loss of generality for the designer to restrict attention to deterministic mechanisms. That is, the designer can not achieve a higher value for the objective* (2) *by selecting a stochastic mechanism.*

Proof. Because utility is quasilinear in transfers it is obvious that we may without loss take transfers to be deterministic. I shall now show it is also without loss to consider deterministic allocation rules.

Let $\vec{X} = \{X_i\}$ be an implementable allocation rule that is potentially stochastic. Let $\vec{X}(\vec{e}, \omega) = \{X_i(\vec{e}, \omega)\}$ denote the random vector of trade quantities conditional on truthful reporting and given the state $\omega \in \Omega$. We can write

$$\vec{X}(\vec{e}, \omega) = \mathbb{E}\left[\vec{X}(\vec{e}, \omega)|\vec{e}\right] + \vec{Z}(\omega)$$

where $\mathbb{E}\left[\vec{Z}(\omega)|\vec{e}\right] = 0$.

Define the deterministic allocation rule $\vec{Y} = \{Y_i\}$ by

$$Y_i(\vec{e}) = \mathbb{E}\left[\vec{X}(\vec{e}, \omega)|\vec{e}\right], \quad \forall \vec{e}.$$

Clearly $\vec{Y}$ is implementable because $\hat{\vec{Y}}$ is weakly decreasing because $\hat{\vec{Y}} = \hat{\vec{X}}$ and $\hat{\vec{X}}$ is implementable by assumption. Moreover, under $\vec{Y}$ expected transfers and allocative efficiency are both higher than under $\vec{X}$ because they are concave in the allocation rule (see (1) and also (4) which applies even when mechanism are stochastic). Moreover, participation is individually rational under $\vec{Y}$ if it is so under $\vec{X}$ because the utility gain from participation depends on only the interim allocation rule and $\hat{\vec{Y}} = \hat{\vec{X}}$ (see (3) which applies even when the mechanism is stochastic).

It thus follows that any stochastic mechanism can be improved on by a deterministic mechanism. □

The rest of this appendix gives a formal proof of Theorem 1 and its corollaries. I first prove the following auxilary Lemma 3.

Lemma 3. *If* $\alpha > 0$, *then any solution to* (2) *is unique up to measure zero differences. If* $\alpha = 0$, *then all solutions have the same allocation rules up to measure zero differences.*

Proof. Using equation (4) for the transfer rule and integrating by parts, I rewrite the optimization problem (2) as

$$\max_{\{X_i,\hat{U}_i(\mu_e)\}_{i=1}^N} \sum_{i=1}^{N} \mathbb{E}\left[-\frac{1}{\kappa}\left(e_i - \alpha\frac{\mathbb{1}_{\{e_i\geq\mu_e\}} - F(e_i)}{f(e_i)}\right)X_i(\vec{e}) - \frac{1}{2\kappa}X_i(\vec{e})^2\right] - \alpha\hat{U}_i(\mu_e) \quad (9)$$

such that

$$-\frac{1}{\kappa}\int_{\mu_e}^{e_i}\hat{X}_i(s)\,\mathrm{d}s + \hat{U}_i(\mu_e) - U\big(\mu_e, 0\big) \geq 0, \quad \forall e_i, \forall i$$

$$\hat{X}_i \text{ is nondecreasing}, \quad \forall i$$

$$\sum_{i=1}^{N} X_i(\vec{e}) = 0, \quad \forall \vec{e}$$

where recall the notation convention $\int_{\mu}^{e_i} \equiv -\int_{e_i}^{\mu}$ whenever $e_i < \mu$. It is easy to see that any solution must necessarily be unique up to measure zero differences when $\alpha > 0$ because the objective function is strictly concave in the allocation rule $\vec{X}$ and because the set of $(\vec{X}, \vec{\hat{U}})$ satisfying the constraints is convex. When $\alpha = 0$ these same observations imply that the solution for the allocation rule is necessarily unique though there are multiple values for $\hat{U}_i(\mu_e)$ that are optimal. □

Consider the problem

$$\min f(x) \text{ such that } x \in \Omega, G(x) \leq \theta \quad (10)$$

where Ω is a convex subset of a vector space X, f is a real-valued convex functional on Ω, and G is a convex mapping from Ω into a normed space Z. Let P be a convex cone in Z. For $z, y \in Z$ write $z \geq y$ if $z - y \in P$. The cone, P, which defines this relation, is called the positive cone in Z. To complete the proof of Theorem 1, it will be useful to apply the following restatement of Theorem 2 which appears in Chapter 8 of Luenberger (1997).

Theorem 6 (Luenberger (1997))**.** *Let P denote the positive cone of Z and suppose P is closed. Suppose there exists $z_0^* \in Z^*$, $z_0^* \geq \theta$, and an $x_0 \in \Omega$ such that the Lagrangian*

$L(x, z^*) = f(x) + \langle G(x), z^* \rangle$ *possess a saddle point at* x_0, z_0^*. *That is,*

$$L(x_0, z^*) \leq L(x_0, z_0^*) \leq L(x, z_0^*)$$

for all $x \in \Omega$, $z_0^* \geq \theta$. *Then* x_0 *solves* (10).

Using Lemma 3 and Theorem 6 we can now complete the proof of Theorem 1.

Proof of Theorem 1. The relaxed version of problem (2) which omits the monotonicity constraint fits into the framework (10) with the following translation.

1. Let

$$f = \mathbb{E}\left[\frac{1}{\kappa}\left(e_i - \alpha\frac{\mathbb{1}_{\{e_i \geq \mu_e\}} - F(e_i)}{f(e_i)}\right)X_i(\vec{e}) + \frac{1}{2\kappa}X_i(\vec{e})^2\right] + \alpha\hat{U}_i(\mu_e).$$

2. Let Ω be the set of all $(\{X_i\}_{i=1}^N, \{U_i\}_{i=1}^N)$ such that $\sum_{i=1}^N X_i = 0$ and each X_i is a measurable function mapping from $[\underline{e}, \overline{e}]$ to $\mathbb{R}$ and each U_i lives in $\mathbb{R}$.

3. Let $G : \Omega \Rightarrow C([\underline{e}, \overline{e}])$ be the function that maps

$$(\{X_i\}_{i=1}^N, \{U_i\}_{i=1}^N) \mapsto \left\{\int_{\mu_e}^{(\cdot)} \frac{1}{\kappa_i}\mathbb{E}\left[X_i(s, Z_{-i})\right] \mathrm{d}s - U_i(\mu_e) - \frac{1}{2\kappa}\mu_e^2\right\}_{i=1}^N.$$

   Note that G is convex.

4. Let Z be the set $\left(C[\underline{e}, \overline{e}]\right)^N$.

5. Let $\theta \in Z$ be the be the function that is equal to zero everywhere.

6. The set Z^* is the set of bounded variation measures on $[\underline{e}, \overline{e}]$.

7. Let P be the set of nonnegative continuous functions.

It is clear that P is nonempty, closed in Z, and contains an interior point in Z. Thus we can apply Theorem 6 to verify the optimality of the allocation rule in Theorem 1 for the relaxed problem which was obtained by solving for a saddle point of the Lagrangian. Under the sufficient conditions the monotonicity constraint is not binding and so the allocation rule in Theorem 1 must also be a solution of the original problem (2). □

Proof of Corollary 1.2. The proof is immediate from Theorem 1. □

Proof of Corollary 1.1. We verify that the proposed solution in the statement of the corollary does indeed satisfy the conditions of Theorem 1.

We have

$$e_b - \mu_e = \alpha \frac{1 - F(e_b)}{f(e_b)},$$

and so, using the symmetry of F about μ_e, we have

$$e_b - \mu_e = \alpha \frac{F(\mu_e - (e_b - \mu_e))}{f(\mu_e - (e_b - \mu_e))}$$

$$\Leftrightarrow$$

$$\mu_e - (e_b - \mu_e) + \alpha \frac{1 - F(\mu_e - (e_b - \mu_e))}{f(\mu_e - (e_b - \mu_e))} = \mu_e$$

$$\Leftrightarrow$$

$$e_a - \mu_e = \alpha \frac{F(e_a)}{f(e_a))}$$

as desired.

I now verify that $\mathbb{E}\left[v(e_j)\right] = \mu_e$. We have that

$$v(e_j) - \mu_e = e - \mu_e - \alpha \frac{1 - F(e)}{f(e)}$$

when $e \geq e_b$ and

$$v(e_j) - \mu_e = e - \mu_e + \alpha \frac{F(e)}{f(e)}$$

whenever $e \leq e_a$. Everywhere else we have $v(e_j) - \mu_e = 0$.

Because $v(e - \mu_e) = -v(\mu_e - e)$ and we indeed have

$$\mathbb{E}\left[v(e_j)\right] = \mu_e$$

where I have used symmetry of f and symmetry of e_b and e_a about μ_e. □

Proof of Proposition 2. To prove Part 1 recognize that

$$\mathbb{E}\left[X_i(\vec{e})^2\right] = \left(\frac{(N-1)^2}{N^2} + \frac{N-1}{N^2}\right) \mathrm{var}[v(e_i)] = \frac{N-1}{N} \mathrm{var}[v(e_i)].$$

Moreover,

$$\hat{X}_i(e_i) = -\frac{N-1}{N} v(e_i) + \frac{N-1}{N} \mathbb{E}[v(e_i)]$$

where by, Corollary 1.2, $\mathbb{E}[v(e_i)]$ does not depend on N. Thus by (4) and (3), the expected transfer and expected utility gain are both proportional to $\frac{N-1}{N}$ as claimed because $\hat{U}_i(\mu_e) - U(\mu_e, 0) = 0$ under the optimal mechanism.

Similarly the expected holding costs relative to autarky are

$$\mathbb{E}\left[\frac{1}{2\kappa}X_i(\vec{e})^2 + \frac{1}{\kappa}e_i(X_i)\right]$$

which also is proportional to $\frac{N-1}{N}$.

I now prove part 2. Clearly, when α increases revenue increases and allocative efficiency decreases. Because the expected utility gain of any trader is equal to the expected change allocative efficiency and expected transfer from trading, it follows that the expected utility gain must decrease.

To show that the total expected trading volume decreases it suffices to show that $X_i(\vec{e})$ decreases in the mean-preserving spread order when α increases whenever F is symmetric about its mean. This is straightforward because $v_\alpha(e_i)$ is a mean-preserving contraction of $v_{\alpha'}(e_i)$ simply by using the mean-preserving spread definition in terms of cdfs and the symmetry of F. □

Proof of Proposition 3. See the proof of Proposition 5 which proves the more general case where there may be adverse selection.

□

Example 3. *Suppose that* $e_i \sim \text{Exp}(1)$ *and* $\alpha = 1$. *Then*

$$\mathbb{E}\left[v(e_i)\right] = .86814$$

and so under the optimal mechanism, a trader of type zero purchases

$$-\frac{N-1}{N}.86814$$

units in expectation rather than

$$-\frac{N-1}{N}$$

as she would under the efficient mechanism. Under the double auction she trades in expectation

$$-\frac{N-2}{N}.$$

Thus, the distortion can be higher under the optimal mechanism.

Expected trade volume is trickier.

B. Omitted Material for Section 5

Formal Statement of the Designer's Objective. Given $\alpha \in [0,1]$, the objective is to solve

$$\max_{\{X_i\},\{T_i\}} \mathbb{E}\left[\alpha \underbrace{\sum_{i=1}^{N} T_i(\vec{t})}_{\text{Revenue}} + (1-\alpha) \underbrace{\sum_{i=1}^{N} -\frac{1}{2\kappa}\left(e_i + X_i(\vec{t})\right)^2}_{\text{Allocative Efficiency}}\right] \tag{11}$$

such that,

$$\begin{aligned}
\text{(P)}: &\quad \mathbb{E}\left[U\left(e_i + X_i(\vec{t}), T_i(\vec{t})\right)\middle| e_i\right] \geq U(e_i, 0), \quad \forall e_i, \forall i \\
\text{(IC)}: &\quad t_i \in \arg\max_{m_i} \mathbb{E}\left[U\left(e_i + X_i(m_i, \vec{t}_{-i}), T_i(m_i, \vec{t}_{-i})\right)\middle| e_i\right], \quad \forall e_i, \forall i \\
\text{(E)}: &\quad \sum_{i=1}^{N} X_i(\vec{t}) = 0, \quad \forall \vec{e}
\end{aligned}$$

where $\vec{t} = \{t_i\}_{i=1}^N$ is the vector of traders' effective types.

The rest of this Appendix contains all omitted proofs for Section 5. I first prove the auxilary Lemmas 4 and 5 which will be used in the proof of Theorem 2.

Lemma 4. *If a symmetric linear equilibrium exists then it is unique and has demand coefficients* a, b, *and* c *that are characterized by the following equations*

$$\frac{b}{a} = \kappa\left(1 - \frac{(b/a)^2(N-1)\rho\alpha\hat{\sigma}_\zeta^2}{\sigma_\epsilon^2 + (b/a)^2\hat{\sigma}_\zeta^2}\right) \tag{12}$$

$$b = \kappa\left(\frac{N-2}{N-1} - 2\frac{1+\frac{N-2}{2}\rho}{1+(N-1)\rho}\frac{\left(\frac{b}{a}\right)^2\hat{\sigma}_\zeta^2}{\sigma_\epsilon^2 + \left(\frac{b}{a}\right)^2\hat{\sigma}_\zeta^2}\right) \tag{13}$$

$$c = \frac{N-1}{N-2}\left(\frac{1}{\kappa} + N\frac{b\alpha\hat{\sigma}_\zeta^2}{a^2\sigma_\epsilon^2 + b^2\hat{\sigma}_\zeta^2}\right)^{-1}. \tag{14}$$

where $\hat{\sigma}_\zeta^2 := \sigma_\zeta^2\left[1 + (N-2)\rho - \rho^2(N-1)\right]$.

Proof. I conjecture an equilibrium of the form

$$q_i(p) = -ae_i + b\zeta_i - cp.$$

It follows from the rules of conditional Gaussian random variables that

$$\mathbb{E}\left[\pi|\zeta_1, \dots, \zeta_N\right] = \alpha \sum_{i=1}^{N} \zeta_i$$

for some α which we shall now determine.

Because $\zeta_i = \mathbb{E}\left[\pi|\zeta_i\right]$, it must be that

$$\alpha\zeta_i + \alpha(N-1)\rho\zeta_i = \zeta_i$$

which implies that

$$\alpha = \frac{1}{1+(N-1)\rho}.$$

By market clearing, if trader i submits the demand schedule $q_i(\cdot)$ when all other traders follow the conjectured strategy then the price must satisfy

$$p = \frac{-a\sum_{j\neq i} e_j + b\sum_{j\neq i}\zeta_j + q_i(p)}{(N-1)c}$$

$$\Leftrightarrow$$

$$p = \frac{-a\sum_{j\neq i} e_j + b\frac{1}{\alpha}\hat{\pi} - b\frac{1}{\alpha}\zeta_i + q_i(p)}{(N-1)c}. \tag{15}$$

In order for $q_i(\cdot)$ to be optimal, the first-order condition, the solution to

$$\sup_{\Delta} \mathbb{E}\left[\pi|p, q_i(p), \zeta_i\right](q_i(p)+\Delta) - \frac{1}{2\kappa}(e_i + q_i(p) + \Delta)^2 - (p + \Lambda\Delta)(q_i(p)+\Delta) \tag{16}$$

must be to set $\Delta = 0$, for any given $p \in \mathbb{R}$ where $\Lambda := 1/[(N-1)c]$ denotes price impact. This can be shown formally by a straightforward application of the calculus of variations.

The first-order condition is

$$-\frac{1}{\kappa}(e_i + q_i(p)) + \mathbb{E}\left[\pi|p, q_i(p), \zeta_i, e_i\right] = \Lambda q_i(p) + p. \tag{17}$$

In order to compute the conditional expectation, I start by computing the variance of

$\sum_{j\neq i}\zeta_j$ conditional on ζ_i. We have

$$\mathrm{var}\left[\sum_{j\neq i}\zeta_j\right] = (N-1)\sigma_\zeta^2\left[1+(N-2)\rho\right]$$

Then, by the rules of conditional Gaussian random variables,

$$\begin{aligned}\mathrm{var}\left[\sum_{j\neq i}\zeta_j|\zeta_i\right] &= (N-1)\sigma_\zeta^2\left[1+(N-2)\rho\right] - \frac{\left(\rho(N-1)\sigma_\zeta^2\right)^2}{\sigma_\zeta^2} \\ &= \sigma_\zeta^2(N-1)\left[1+(N-2)\rho-\rho^2(N-1)\right].\end{aligned}$$

Therefore,

$$\pi|\zeta_i \sim N\left(\zeta_i, \alpha^2\hat{\sigma}_\zeta^2(N-1)\right)$$

where

$$\hat{\sigma}_\zeta^2 \equiv \sigma_\zeta^2\left[1+(N-2)\rho-\rho^2(N-1)\right].$$

Again, using the rules for conditional Gaussians, we have

$$\begin{aligned}\mathbb{E}\left[\pi|\zeta_i, -a\sum_{j\neq i}e_i + b\frac{1}{\alpha}\pi\right] &= \zeta_i + \frac{b\alpha\hat{\sigma}_\zeta^2}{a^2\sigma_\epsilon^2+b^2\hat{\sigma}_\zeta^2}\left[-a\sum_{j\neq i}e_i + b\frac{1}{\alpha}\pi - b\frac{1}{\alpha}\zeta_i\right] \\ &= \zeta_i + \frac{b\alpha\hat{\sigma}_\zeta^2}{a^2\sigma_\epsilon^2+b^2\hat{\sigma}_\zeta^2}\left[\frac{p-\Lambda q_i(p)}{\Lambda} - b(N-1)\rho\zeta_i\right].\end{aligned}$$

Substituting into the first-order condition (17) and matching coefficients yields:

$$\left(\Lambda+\frac{1}{\kappa}+\frac{b\alpha\hat{\sigma}_\zeta^2}{a^2\sigma_\epsilon^2+b^2\hat{\sigma}_\zeta^2}\right)a = \frac{1}{\kappa} \tag{18}$$

$$\left(\Lambda+\frac{1}{\kappa}+\frac{b\alpha\hat{\sigma}_\zeta^2}{a^2\sigma_\epsilon^2+b^2\hat{\sigma}_\zeta^2}\right)b = 1-\frac{b^2(N-1)\rho\alpha\hat{\sigma}_\zeta^2}{a^2\sigma_\epsilon^2+b^2\hat{\sigma}_\zeta^2} \tag{19}$$

$$-\left(\Lambda+\frac{1}{\kappa}+\frac{b\alpha\hat{\sigma}_\zeta^2}{a^2\sigma_\epsilon^2+b^2\hat{\sigma}_\zeta^2}\right)c = -1+\frac{b\alpha\hat{\sigma}_\zeta^2}{a^2\sigma_\epsilon^2+b^2\hat{\sigma}_\zeta^2}c(N-1) \tag{20}$$

where recall that $\Lambda := 1/[(N-1)c]$. Equations (18), (19), and (20) are necessary conditions for a symmetric linear equilibrium.

Rearranging (20) yields

$$\frac{N-2}{N-1}=\left(\frac{1}{\kappa}+N\frac{b\alpha\hat{\sigma}_\zeta^2}{a^2\sigma_\epsilon^2+b^2\hat{\sigma}_\zeta^2}\right)c$$

or

$$\Lambda=\frac{1}{N-2}\left(\frac{1}{\kappa}+N\frac{b\alpha\hat{\sigma}_\zeta^2}{a^2\sigma_\epsilon^2+b^2\hat{\sigma}_\zeta^2}\right). \tag{21}$$

Substituting into (19) and (20) yields equations (13) and (14)

Dividing (19) by (18) yields (12). By inspection there is a unique solution to (12). Given the solution to (12) for b/a, coefficients b and c are characterized uniquely by (13) and (14). Thus there can be at most one symmetric linear equilibrium.

□

Lemma 5. *An increase in correlation ρ leads to an increase in the demand coefficients a, b, and c.*

Proof. I first show that b is increasing in ρ. Let $y := b/a$. Using (19) we have

$$\frac{N-1}{N-2}\frac{1}{\kappa}b+\frac{2(N-1)}{N-2}\frac{1+\frac{N-2}{2}\rho}{1+(N-1)\rho}\frac{y^2\hat{\sigma}_\zeta^2}{\sigma_\epsilon^2+y^2\hat{\sigma}_\zeta^2}=1 \tag{22}$$

Consider the second term on the LHS:

$$\frac{2(N-1)}{N-2}\left(1+\frac{N-2}{2}\rho\right)\frac{y^2(1-\rho)\sigma_\zeta^2}{\sigma_\epsilon^2+y^2\sigma_\zeta^2\left[1+(N-2)\rho-(N-1)\rho^2\right]}=1 \tag{23}$$

Suppose for contradiction that the above term is ever increasing in ρ. Suppose further that $y^2\hat{\sigma}_\zeta^2$ is increasing. Then (12) implies that y is decreasing. But because

$$\left(1+\frac{N-2}{2}\rho\right)(1-\rho)$$

is decreasing in ρ, it follows that (23) is decreasing which is a contradiction. Now suppose that $y^2\hat{\sigma}_\zeta^2$ is decreasing. Because

$$\frac{1+\frac{N-2}{2}\rho}{1+(N-1)\rho}$$

is also decreasing it follows that (23) must again be decreasing. We thus have a contradiction.

I shall now prove that a is likewise increasing in ρ. For the reader's ease, I have reproduced equation (18) below

$$\left(\Lambda + \frac{1}{\kappa} + \frac{b\alpha\hat{\sigma}_\zeta^2}{a^2\sigma_\epsilon^2 + b^2\hat{\sigma}_\zeta^2}\right)a = \frac{1}{\kappa}.$$

Suppose for contradiction that a is ever decreasing in ρ. Then it must be

$$\frac{y\alpha\hat{\sigma}_\zeta^2}{\sigma_\epsilon^2 + y^2\hat{\sigma}_\zeta^2}$$

is increasing in ρ. If y was increasing we would have a contradiction because (12) would show y must be decreasing. But we know $y = b/a$ and b is increasing so if a was ever decreasing then y must be increasing. Thus it must be that a is increasing.

Finally I now show that c is increasing in ρ or equivalently that $\Lambda := 1/[(N-1)c]$ is decreasing in ρ. Recall that

$$\left(\Lambda + \frac{1}{\kappa} + \frac{b\alpha\hat{\sigma}_\zeta^2}{a^2\sigma_\epsilon^2 + b^2\hat{\sigma}_\zeta^2}\right)a = \frac{1}{\kappa}.$$

Suppose for contradiction that Λ is increasing in ρ. Then by inspecting (21), it must be that that the term in parenthesis must be increasing in ρ. But then a must be decreasing in ρ, a contradiction. □

Proof of Theorem 2. To prove part 1 of the theorem I recognize that in the special case when $\rho = 0$ equations (12), (13), and (14) can all be solved in closed form. In particular, a and b, in terms of primitives are given by

$$a = \frac{1}{\kappa}b = \frac{N-2}{N-1}\left[1 - \frac{2(N-1)}{N-2}\frac{\sigma_\zeta^2}{\sigma_\epsilon^2\left(\frac{1}{\kappa}\right)^2 + \sigma_\zeta^2}\right].$$

Second-order conditions for optimality for the problem (16) are satisfied if and only if

$$\frac{1}{2\kappa} + \Lambda \geq 0.$$

Let us consider first the case when $\rho = 0$. If $b < 0$, then in order for (19) to be satisfied it must be that

$$\Lambda + \frac{1}{\kappa} \leq 0.$$

This implies that

$$\frac{1}{2\kappa} + \Lambda < 0$$

which implies second-order conditions are not satisfied. On the other hand, as long as $b > 0$ then $\Lambda > 0$ and the second-order conditions are always satisfied.

Thus $b \geq 0$ is necessary and sufficient for equilibrium existence in that case.

What about when $\rho > 0$? Suppose $b < 0$. Consider again (19) which implies

$$\left(\Lambda + \frac{1}{\kappa}\right) b = 1 - \frac{b^2 \hat{\sigma}_\zeta^2}{a^2 \sigma_\epsilon^2 + b^2 \hat{\sigma}_\zeta^2} > 0.$$

Thus, once again we see that if $b > 0$ is not satisfied then we do not have second-order conditions satisfied. Since b is monotone increasing in ρ it follows that the range of equilibrium existence expands.

To prove part 2, I proceed as follows. To start, I derive an expression for expected allocative efficiency in terms of only the parameters a and b.

By straightforward computations, one can show that the expected sum of holding costs is simply

$$(N-1)\left[(1 + a^2 - 2a)\sigma_e^2 + b^2(1-\rho)\sigma_\zeta^2\right].$$

I shall now prove that the above is decreasing in ρ.

To start, let me rewrite the above as

$$(N-1)\left[(1 - 2a + a^2)\sigma_e^2 + y^2 a^2 (1-\rho)\sigma_\zeta^2\right]. \tag{24}$$

I shall minimize the above with respect to a taking y as given. The first-order condition is

$$(2a-1)\,\sigma_e^2 + 2y^2(1-\rho)\sigma_\zeta^2 a = 0$$

$$\Rightarrow$$

$$a^* = \frac{\sigma_e^2}{y^2(1-\rho)\sigma_\zeta^2 + \sigma_e^2}$$

$$\Rightarrow$$

$$\left[1+\frac{\sigma_\zeta^2 y^2(1-\rho)}{\sigma_\epsilon^2}\right]a^* = 1$$

I shall now prove that in any equilibrium $a < a^*$ meaning a is to the left of the bliss point. In equilibrium we have

$$\left(\frac{N-1}{N-2}+\frac{y\hat{\sigma}_\zeta^2}{y^2\hat{\sigma}_\zeta^2+\sigma_e^2}\frac{2N-2}{N-2}\alpha\kappa\right)a = 1.$$

Thus to prove $a < a^*$ it suffices to show

$$\frac{\sigma_\zeta^2 y^2(1-\rho)}{\sigma_\epsilon^2} \leq \frac{y\hat{\sigma}_\zeta^2}{y^2\hat{\sigma}_\zeta^2+\sigma_e^2}\frac{2N-2}{N-2}\alpha\kappa$$

$$\Leftrightarrow$$

$$\frac{\sigma_\zeta^2 y^2(1-\rho)}{\sigma_\epsilon^2} \leq \frac{y^2\sigma_\zeta^2(1-\rho)}{y^2\hat{\sigma}_\zeta^2+\sigma_e^2}\frac{2N-2}{N-2}\frac{\kappa}{y}$$

$$\Leftrightarrow$$

$$\frac{y}{\kappa} \leq \frac{\sigma_\epsilon^2}{y^2\hat{\sigma}_\zeta^2+\sigma_e^2}\frac{2N-2}{N-2}$$

$$\Leftrightarrow$$

$$1-\frac{y^2(N-1)\rho\alpha\hat{\sigma}_\zeta^2}{\sigma_\epsilon^2+y^2\hat{\sigma}_\zeta^2} \leq \frac{\sigma_\epsilon^2}{y^2\hat{\sigma}_\zeta^2+\sigma_e^2}\frac{2N-2}{N-2}$$

$$\Leftrightarrow$$

$$\sigma_\epsilon^2 + y^2\hat{\sigma}_\zeta^2\left(1-(N-1)\rho\alpha\right) \leq \sigma_\epsilon^2\frac{2N-2}{N-2}$$

$$\Leftrightarrow$$

$$y^2\hat{\sigma}_\zeta^2\left(1-(N-1)\rho\alpha\right) \leq \sigma_\epsilon^2\frac{N}{N-2}. \tag{25}$$

To do this, observe that

$$y^2\hat{\sigma}_\zeta^2\left(1-(N-1)\rho\alpha\right) \leq \kappa^2\sigma_\zeta^2\left(1+\rho(N-2)-\rho^2(N-1)\right)\left(1-(N-1)\rho\alpha\right)$$

$$\Leftrightarrow$$

$$y^2\hat{\sigma}_\zeta^2\left(1-(N-1)\rho\alpha\right) \leq \kappa^2\sigma_\zeta^2\left(1+\rho(N-2)-\rho^2(N-1)-(N-1)\rho(1-\rho)\right)$$

$$\Leftrightarrow$$

$$y^2\hat{\sigma}_\zeta^2\,(1-(N-1)\rho\alpha) \leq \kappa^2\sigma_\zeta^2\left(1+\rho(N-2)-\rho^2(N-1)-(N-1)\rho+(N-1)\rho^2)\right)$$

$$\Leftrightarrow$$

$$y^2\hat{\sigma}_\zeta^2\,(1-(N-1)\rho\alpha) \leq \kappa^2\sigma_\zeta^2(1-\rho).$$

To prove (25) it suffices to show

$$\kappa^2\sigma_\zeta^2(1-\rho) \leq \sigma_\epsilon^2\frac{N}{N-2}.$$

This holds because we know that

$$1-\frac{2(N-1)}{N-2}\frac{\sigma_\zeta^2}{\sigma_\epsilon^2\left(\frac{1}{\kappa}\right)^2+\sigma_\zeta^2} \geq 0$$

$$\Leftrightarrow$$

$$1 \geq \frac{2(N-1)}{N-2}\frac{\sigma_\zeta^2}{\sigma_\epsilon^2\left(\frac{1}{\kappa}\right)^2+\sigma_\zeta^2} \geq$$

$$\Leftrightarrow$$

$$\frac{\sigma_\epsilon^2}{\kappa^2\sigma_\zeta^2} \geq \frac{2(N-1)}{N-2}-1$$

$$\Leftrightarrow$$

$$\frac{\sigma_\epsilon^2}{\kappa^2\sigma_\zeta^2} \geq \frac{N}{N-2}.$$

Thus we have shown that $a < a^*$. Suppose that $y^2(1-\rho)$ is decreasing in ρ. Since a is always to the left of the bliss point and is increasing in ρ and since the minimum of (24) is always lower when ρ is bigger, it follows that allocative efficiency is increasing in ρ.

I now show that $y^2(1-\rho)$ is indeed decreasing in ρ. Consider again the equation for y:

$$y = \kappa\left(1-\frac{y^2(N-1)\rho\alpha\hat{\sigma}_\zeta^2}{\sigma_\epsilon^2+y^2\hat{\sigma}_\zeta^2}\right)$$

$$\Leftrightarrow$$

$$y = \kappa\left(1-\frac{y^2(N-1)\rho(1-\rho)\sigma_\zeta^2}{\sigma_\epsilon^2+y^2(1-\rho)(1+\rho(N-1))\sigma_\zeta^2}\right)$$

$$\Leftrightarrow$$

$$y = \kappa\left(1 - \frac{y^2(N-1)(1-\rho)\sigma_\zeta^2}{\frac{\sigma_\epsilon^2}{\rho} + y^2(1-\rho)(\frac{1}{\rho} + N - 1)\sigma_\zeta^2}\right)$$

If $y^2(1-\rho)$ is increasing in ρ, it follows from the above equation that y is decreasing in ρ, a contradiction.

Parts 3, 4, and 5 follow straightforwardly. □

Proof. Fixing a direct mechanism $\{X_i, T_i\}$ the expected utility gain from participating and reporting truthfull is

$$\hat{U}_i(t_i) - U_i(t_i, 0) = \mathbb{E}\left[\left(\pi_i - \frac{1}{\kappa_i}e_i\right)X_i(t) - \frac{1}{2\kappa_i}X_i(t)^2 - T_i(t)\middle| t_i\right].$$

if all other traders do so as well.

By the envelope integral formual, it is easy to show that

$$\hat{U}_i(t_i) - U_i(t_i, 0) = -\int_{\mu_t}^{t_i} \mathbb{E}_i[X_i(s, t_{-i})]\,\mathrm{d}s + \hat{U}_i(\mu_t) - U_i(\mu_t, 0)$$

under any incentive compatible mechanism.

This implies that

$$\hat{T}(t_i, t_{-i}) = \mathbb{E}_i\left[\left(-t_i + \sum_{j\neq i}\zeta_j\right)X_i(t_i, t_{-i}) + \int_{\mu_t}^{t_i} X_i(s, t_{-i})\,\mathrm{d}s - \frac{1}{2\kappa}X_i(t_i, t_{-i})^2\middle| t_i\right] - \left[\hat{U}_i(\mu_t) - U_i(\mu_t, 0)\right].$$

The designer's objective is the following

$$\sup_{\{X_i\}\{\hat{U}_i(\mu_t)\}} \alpha\sum_{i=1}^N \mathbb{E}\Bigg[\left(-t_i + \sum_{j\neq i}\zeta_j\right)X_i(t) + \int_{\mu_t}^{t_i} X_i(s, t_{-i})\,\mathrm{d}s - \frac{1}{2\kappa}X_i(t)^2 - \left(\hat{U}_i(\mu_t) - U_i(\mu_t, 0)\right)\Bigg] + (1-\alpha)\sum_{i=1}^N \mathbb{E}\left[\frac{1}{2\kappa}(X_i(t) + e_i)^2\right]. \quad (26)$$

such that

$$\sum_{i=1}^{N} X_i(t) = 0$$

and for each $i = 1, 2, \dots, N$ it holds that

$$\hat{U}_i(t_i) - U_i(t_i, 0) = -\int_{\mu_t}^{t_i} \mathbb{E}_i[X_i(s, t_{-i})]\,\mathrm{d}s + U_i(\mu_t) - U_i(\mu_t, 0) \geq 0$$

for all t_i and

$$\mathbb{E}_i[X_i(s, t_{-i})]$$

is nondecreasing in s.

Simplifying and omitting constant terms, the objective becomes:

$$\sup_{\{X_i\}\{\hat{U}_i(\mu_t)\}} \alpha\mathbb{E}\left[\sum_{i=1}^{N} -\frac{1}{\kappa}\hat{e}(t_i)X_i(t) - \int_{\mu_t}^{t_i} X_i(s, t_{-i})\,\mathrm{d}s - \frac{1}{2\kappa_i}X_i(t)^2 + U_i(\mu_t)\right]$$
$$+(1-\alpha)\mathbb{E}\left[\sum_{i=1}^{N} -\frac{1}{2\kappa_i}X_i(t)^2 - \frac{1}{\kappa}\hat{e}(t_i)X_i(t)\right]$$

subject to the same constraints.

Let Ω_i be the Lagrange multiplier on the participation constraint. The Lagrangian is

$$\sup_{\{X_i\}\{\hat{U}_i(\mu_t)\}} \mathbb{E}\left[\sum_{i=1}^{N} \hat{e}(t_i)X_i(t) - \alpha\int_{\mu_t}^{t_i} X_i(s, t_{-i})\,\mathrm{d}s - \frac{1}{2\kappa_i}X_i(t)^2 + \alpha U_i(\mu_t)\right]$$
$$+\sum_{i=1}^{N}\int_{\mu_t}^{\bar{t}}\left[\int_{\mu_t}^{t_i} \mathbb{E}\left[X_i(s, t_{-i})\right]\,\mathrm{d}s - U_i(\mu_t)\right]\mathrm{d}\Omega_i(t_i)$$

where I will ignore the monotonicity constraint on the interim-allocation rule for now.

Integrating by parts, we can rewrite the Lagrangian as

$$\mathbb{E}\left[\sum_{i=1}^{N} \hat{e}(t_i)X_i(t) + X_i(t)\frac{\alpha H(t_i) - \Omega_i(t_i)}{h(t_i)} - \frac{1}{2\kappa_i}X_i(t)^2\right] - \left(1 - \lim_{t\to\infty}\Omega_i(t)\right)\hat{U}_i(\mu_t)$$

I seek to maximize the above subject to

$$\sum_{i=1}^{N} X_i(t) = 0.$$

We define the virtual value as

$$v(t_i) = -\hat{e}(t_i) + \alpha \frac{H(t_i) - \Omega_i(t_i)}{h(t_i)}. \tag{27}$$

Ignoring the monotonicity constraint, by Jensen's inequality, the optimal allocation is

$$X_i(t) = -v_(t_i) + \frac{1}{N} \sum_{j=1}^{N} v(t_j) \tag{28}$$

as desired. The last step which is to verify that the method of Lagrange multipliers is valid follows the same steps as in the proof of Theorem 1. I omit this for brevity. □

Proof. Given $\epsilon > 0$, let $\hat{v}(e_i)$ denote

$$\hat{v}(e_i) = v(e_i) + \epsilon \max\{\min\{e_{ib}, e_i\} - e_{ia}, 0\}.$$

This is the virtual value of trader i with type e_i, but perturbed over the interval $[e_{ia}, e_{ib}]$ of binding participation constraints so that the function is strictly increasing. Let $\hat{X}_i(e_i, e_{-i})$ denote

$$\hat{X}_i(e_i, e_{-i}) = -\hat{v}(e_i) + \frac{1}{N} \sum_{j=1}^{N} \hat{v}(e_j). \tag{29}$$

Then as $\epsilon \to 0$, $\hat{X}_i \to X_i$ pointwise. Also, note that by construction $\sum_{i=1}^{N} \hat{X}_i = 0$ so that the auctioneer does not retain any net position in the asset.

Suppose that each trader i submits the demand schedule given by

$$\hat{q}_i(p) = -\hat{v}_i(Z_i) - \kappa p \tag{30}$$

to the double auction. Then the resulting market clearing price is

$$p = -\frac{1}{N\kappa} \sum_{j=1}^{N} \hat{v}(e_j).$$

One can verify that the resulting trade quantities coincide with the allocation (29). Moreover, the inverse residual demand curve facing trader i can be computed as follows. By market clearing,

$$\hat{q}_i + \sum_{j \neq i} -\hat{v}(e_j) - (N-1)\,\kappa p = 0.$$

This implies that

$$p = \left(\hat{q}_i + \sum_{j \neq i} -\hat{v}(e_j) \right) \frac{1}{(N-1)\,\kappa}.$$

The price impact is therefore $\Lambda = 1/[(N-1)\kappa]$.

Next, I constructre the transaction fee such that it is optimal for trader i to submit the demand schedule (30). I consider a side payment rule of the form

$$\mathcal{T}(p, q_i(p)) = \frac{\hat{c}}{2} q_i(p)^2 - \hat{T}\left(-q_i(p) - \kappa p\right) + \mathbb{E}\left[\sum_{j \neq i} \zeta_j | p - \Lambda q_i \right] q_i(p) + \hat{d}$$

where the constants $\hat{c}$ and $\hat{d}$ and function $\hat{T}$ are to be derived.

To derive them, we consider trader i's demand submission problem. For each price p, it must be optimal to purchase $\hat{q}_i(p)$ units: a deviation to purchasing $\hat{q}_i(p) + \Delta$ units must be suboptimal for all $\Delta \neq 0$. Thus $\Delta = 0$ must solve

$$\max_{\Delta} -\frac{1}{2\kappa_i}(e_i + \hat{q}_i + \Delta)^2 - \left(p + \Delta \frac{1}{(N-1)\,\kappa} \right)(\hat{q}_i + \Delta) \\ + \hat{T}\left(-\hat{q}_i - \Delta - \kappa \left(p + \Delta \frac{1}{(N-1)\,\kappa} \right) \right) - \frac{\hat{c}}{2}(\hat{q}_i + \Delta)^2$$

where to ease notation I have omitted the argument in $\hat{q}_i$. Taking a first derivative,

$$-\frac{1}{\kappa_i}(e_i + \hat{q}_i + \Delta) - \hat{c}(\hat{q}_i + \Delta) - p - \Delta \frac{1}{(N-1)\,\kappa} - \frac{N}{N-1}(\hat{q}_i + \Delta) \\ + \hat{T}'\left(-\hat{q}_i - \Delta - \kappa \left(p + \Delta \frac{1}{(N-1)\,\kappa} \right) \right)\left(1 + \frac{1}{N-1} \right) = 0.$$

This must hold at $\Delta = 0$ when $\hat{q}_i$ is as in (30). That is,

$$-\frac{1}{\kappa}\left(e_i - \hat{v}(e_i) - \kappa p\right) - \hat{c}\left(-\hat{v}(e_i) - \kappa p\right) =$$

$$p + \frac{N}{N-1}(-\hat{v}(e_i) - \kappa p) - \hat{T}'(\hat{v}(e_i))\left(1 + \frac{1}{N-1}\right).$$

Gathering the terms involving only p gives

$$\hat{c}\kappa p = -\frac{N\kappa}{N-1}p.$$

Thus we may ensure that these terms are consistent by setting

$$\hat{c} = -\frac{N}{N-1}.$$

We next gather only terms involving e_i:

$$\hat{v}(e_i)\frac{1}{\kappa} - \frac{1}{\kappa}e_i = -\hat{T}'(\hat{v}(e_i))\left(1 + \frac{1}{N-1}\right).$$

Writing l in place of $\hat{v}(e_i)$ and integrating both sides we have

$$\hat{T}(l) = \left(\frac{l^2}{2} - \int_{\mu_e}^{l} \hat{v}^{-1}(s)\,\mathrm{d}s\right)\frac{N-1}{N\kappa} + d$$

for some constant d. Above, $\hat{v}^{-1}$ is well-defined as we have perturbed the allocation rule to ensure that $\hat{v}$ is strictly monotone.

To ensure global optimality of $\Delta = 0$ under $\hat{T}$ and $\hat{c}$ it suffices to show that the objective is globally concave. Notice that the integral term is convex since $\hat{v}$ is increasing. Thus, ignoring this term, by a simple computation taking a second derivative of the objective ignoring this term, we can show that the objective is globally concave in Δ.

Thus we have achieved an indirect implementation of the perturbed allocation rule in (29). To derive an exact implementation, we take limits as $\epsilon \to 0$. One can show that in the limit $\hat{T}$ converges to T pointwise and so by relabeling $\hat{c}$ by c we obtain the side payment rule in the statement of the theorem. Incentive compatibility must hold at the limit. If this were not true, for ϵ sufficiently small, one can show that incentive compatibility is also violated for the perturbed allocation rule under the perturbed side payment rule which is a contradiction. □

Proof of Proposition 4. To prove part 1, using the rules of conditional Gaussian random variables we have

$$\frac{\frac{1}{\kappa^2}\sigma_e^2}{\frac{1}{\kappa^2}\sigma_e^2+\sigma_\zeta^2}t_a+\alpha\frac{H(t_a)}{h(t_a)}=\mu_e$$

which is equivalent to

$$\frac{\frac{1}{\kappa^2}\sigma_e^2}{\frac{1}{\kappa^2}\sigma_e^2+\sigma_\zeta^2}t_a+\alpha\frac{\Phi\left((t_a-\mu_t)/\sqrt{\frac{1}{\kappa^2}\sigma_e^2+\sigma_\zeta^2}\right)}{\phi\left((t_a-\mu_t)/\sqrt{\frac{1}{\kappa^2}\sigma_e^2+\sigma_\zeta^2}\right)}=\mu_e$$

where Φ and ϕ are the CDF and PDF of a standard Gaussian random variable. Let me define $z_a=(t_a-\mu_t)/\sqrt{\frac{1}{\kappa^2}\sigma_e^2+\sigma_\zeta^2}$. Then we have

$$\frac{\frac{1}{\kappa^2}\sigma_e^2}{\frac{1}{\kappa^2}\sigma_e^2+\sigma_\zeta^2}\left(z_a+\frac{\mu_t}{\sqrt{\frac{1}{\kappa^2}\sigma_e^2+\sigma_\zeta^2}}\right)\sqrt{\frac{1}{\kappa^2}\sigma_e^2+\sigma_\zeta^2}+\alpha\frac{\Phi\,(z_a)}{\phi(z_a)}=\mu_e$$

which is equivalent to

$$\frac{\frac{1}{\kappa^2}\sigma_e^2}{\sqrt{\frac{1}{\kappa^2}\sigma_e^2+\sigma_\zeta^2}}z_a+\alpha\frac{\Phi\,(z_a)}{\phi(z_a)}=0.$$

Clearly, becuase we know $z_a<0$, when σ_ζ^2 increases z_a must decrease without bound. Because we know that $t_a-\mu_t<0$ it follows that $t_a-\mu_t$ must also decrease without bound. Because t_b and t_a are symmetric about μ_t it follows that t_b-t_a increases as σ_ζ^2 increases and eventually diverges.

I now prove Part 3. The expected utility gain of a trader is

$$\hat{U}_i(t_i)-U_i(t_i,0)=-\int_{\mu_t}^{t_i}\hat{X}_i(s)\,\mathrm{d}s+\hat{U}_i(\mu_t)-U_i(\mu_t,0). \tag{31}$$

Note that $\hat{U}_i(\mu_t)-U_i(\mu_t,0)$ is always set to zero whenever $\alpha>0$. We have that

$$\hat{X}_i(t_i)\frac{N}{N-1}=-\left[\frac{\frac{1}{\kappa^2}\sigma_e^2}{\frac{1}{\kappa^2}\sigma_e^2+\sigma_\zeta^2}t_i+\alpha\frac{\Phi\left((t_i-\mu_t)/\sqrt{\frac{1}{\kappa^2}\sigma_e^2+\sigma_\zeta^2}\right)}{\phi\left((t_i-\mu_t)/\sqrt{\frac{1}{\kappa^2}\sigma_e^2+\sigma_\zeta^2}\right)}\right]+\mu_e$$

whenever $t_i < t_a$. Clearly when $t_i < t_a$ this gets less positive. When σ_ζ^2 increases the term inside the hazard rate increases and thus becomes more positive and so the term in brackets becomes more negative. On the other hand, when σ_ζ^2 increases the negative of the first term in brackets added to μ_e becomes less positive. We can deal analogously for $t_i > t_b$. Therefore the expected utility gain of a trader decreases.

Now I must argue that the expected trade volume decreases. To do this I argue that $X_i(\vec{t})$ is a mean-preserving contraction of what it once was.

Recall that we can have the transformation

$$\frac{\frac{1}{\kappa^2}\sigma_e^2}{\sqrt{\frac{1}{\kappa^2}\sigma_e^2 + \sigma_\zeta^2}} z_i + \alpha \frac{\Phi(z_i)}{\phi(z_i)} = 0.$$

Now recall the definition of a mean-preserving contraction. Using that definition we see that when σ_ζ^2 is higher this is a mean-preserving contraction.

I now prove Part 2. Consider reducing σ_ζ^2 but mantaining the same allocation and transfer rules. Then each trader submits the same message as before. Then I think I can argue based on mean-preserving contraction that allocative efficiency improves. I now argue that the expected utility gain of all traders summed together decreases. Because allocative efficiency improves, this can only happen if total revenue increases. To see this simply note that (31) is convex in t_i. Thus, reducing the variance of t_i must reduce the expected utility gain.

$\square$

C. Omitted Material for Section 6

Proof of Proposition 6. Construct derivative securities that are linear combinations of the underlying assets such that the payoffs of these derivatives are uncorrelated. Without loss of generality, suppose that the derivatives are the primitive securities. Then the utility of trader i is

$$U(\vec{e}_i + \vec{X}_i, T_i) = \sum_{a \in A} \mu_{\pi_a} X_{ia} - \frac{1}{2\kappa} \sum_{a \in A} \sigma_a^2 \left(e_{ia} + X_{ia}\right)^2 - T_i$$

where $\sigma_a^2 = \text{var}[\pi_a]$.

Consider running a separate double auction with transaction fees of the form in Proposition 3 for each asset. Because the utility function is additively separable across assets, each trader's optimal strategy will be, for each asset, just as in the basic model of Section 3. Thus, in equilibrium, the allocation of each asset is efficient. Moreover, as long as participation fees for each double auction are set sufficiently high, the budget is ex-ante balanced. □

Formal Statement of Designer's Objective. Given $\alpha \in [0, 1]$, the objective is to solve

$$\max_{\{X_i\},\{T_i\}} \mathbb{E}\left[\alpha \underbrace{\sum_{i=1}^{N} T_i(\vec{e})}_{\text{Revenue}} +(1-\alpha) \underbrace{\sum_{i=1}^{N}\sum_{a\in A} -\frac{1}{2\kappa}\sigma_\pi^2\Big(e_i + X_{ia}(\vec{e})\Big)^2}_{\text{Allocative Efficiency}}\right] \tag{32}$$

such that,

$$\begin{aligned}
(\text{P}):&\quad \mathbb{E}\left[U\Big(\vec{e}_i + \vec{X}_i(\vec{e}), T_i(\vec{e})\Big)\Big|\vec{e}_i\right] \geq U\left(\vec{e}_i, 0\right), \quad \forall e_i, \forall i\\
(\text{IC}):&\quad \vec{e}_i \in \arg\max_{m_i} \mathbb{E}\left[U\Big(\vec{e}_i + \vec{X}_i(m_i, \vec{e}_{-i}), T_i(m_i, \vec{e}_{-i})\Big)\Big|\vec{e}_i\right], \quad \forall \vec{e}_i, \forall i\\
(\text{E}):&\quad \sum_{i=1}^{N} \vec{X}_i(\vec{e}) = \vec{0}, \quad \forall \vec{e}
\end{aligned}$$

where $\vec{e} = \{\vec{e}_i\}_{i=1}^N$ is the vector of traders' endowment vectors.

Proof of Proposition 7. I prove the Proposition 7 for the special case when $\vec{\mu}_a = \vec{0}$ to ease the exposition. The general case is analogous.

Trader i selects messages $\vec{m}_i = \{m_{ia}\}_{a\in A}$ to solve

$$\sup_{\vec{m}_i} \sum_{a\in A}\left[-\frac{1}{2\kappa}X_{ia}(m_{ia}, \vec{e}_{-ia})^2 - \frac{1}{\kappa}e_{ia}X_{ia}(m_{ia}, \vec{e}_{-i1}) + \mu_{\pi_a}X_{i1}(m_{ia}, \vec{e_{-ia}})\right] - T_i(\vec{m}_i, \vec{m}_{-i}).$$

I conjecture that only local incentive constraints bind along the rays. Thus we suppose that the designer knows $\{\tilde{e}_{ia}\}_{a\in A}$ and only has to elicit r_i. I will verify later, truthful reporting is indeed incentive compatible under the resulting candidate optimal mechanism.

Let $\vec{\tilde{e}}_i = \{\tilde{e}_{ia}\}_{a\in A}$. By the envelope integral formula,

$$\hat{U}(r_i, \vec{\tilde{e}}_i) - U(r_i, \vec{\tilde{e}}_i, 0) = -\frac{1}{\kappa}\int_0^{r_i} \sum_{a\in A} \tilde{e}_{ia}\hat{X}_{ia}(s, \vec{\tilde{e}}_i)\, \mathrm{d}s + \hat{U}_i(0, \vec{\tilde{e}}_i) - U(0, \vec{\tilde{e}}_i, 0) \tag{33}$$

where with a transparent abuse of notation $\hat{U}(r_i, \vec{\tilde{e}}_i)$ denotes the expected utility from truthful reporting for a trader with type $r_i, \vec{\tilde{e}}_i$ and where $U(r_i, \vec{\tilde{e}}_i, 0)$ for a trader of that type under autarky.

This implies that a trader i's expected transfer is

$$\hat{T}_i(r_i, \vec{\tilde{e}}_i) = \sum_{a \in A} \left(\mu_{\pi_a} - \frac{1}{\kappa} e_{ia} \right) \hat{X}_{ia}(r_i, \vec{\tilde{e}}_i) + \frac{1}{2\kappa} X_{ia}(r_i, \vec{\tilde{e}}_i)^2 + \int_0^{r_i} \frac{1}{\kappa} \left[\tilde{e}_{ia} \hat{X}_{ia}(s, \vec{\tilde{e}}_i) \right] \mathrm{d}s$$
$$- \left[\hat{U}_i(0, \vec{\tilde{e}}_i) - U(0, \vec{\tilde{e}}_i, 0) \right]. \quad (34)$$

Using (34) we can write down the planner's objective as

$$\mathbb{E}\Bigg[- \frac{1}{\kappa} \sum_{a \in A} e_{ia} \hat{X}_{ia}(r_i, \vec{\tilde{e}}_i) + \frac{1}{2\kappa} \sum_{a \in A} X_{ia}(r_i, \vec{\tilde{e}}_i)^2 + \alpha \frac{1}{\kappa} \int_0^{r} \sum_{a \in A} \tilde{e}_{ia} \hat{X}_{ia}(s, \vec{\tilde{e}}_i) \, \mathrm{d}s$$
$$- \alpha \left[\hat{U}_i(0, \vec{\tilde{e}}_i) - U(0, \vec{\tilde{e}}_i, 0) \right] \Bigg]$$

such that (33) is nonnegative and such that $\sum_{i=1}^{N} \vec{X}_i = \vec{0}$.

Let $\Omega(r|\vec{\tilde{e}}_i)$ denote the Lagrange multiplier on trader i's participation constraint that $\hat{U}(r, \vec{\tilde{e}}_i) - U(r, \vec{\tilde{e}}_i, 0) \geq 0$. Forming the Lagrangian, integrating by parts, and rearranging yields the objective

$$\sup_{\vec{X}_i} \mathbb{E}\Bigg[\sum_{a \in A} \left(\frac{1}{2\kappa} X_{ia}(r_i, \vec{\tilde{e}})^2 - \left(r - \alpha \frac{\Omega(r_i|\vec{\tilde{e}}_i) - F(r|\vec{\tilde{e}}_i)}{f(r_i|\vec{\tilde{e}}_i)} \right) \frac{1}{\kappa} \tilde{e}_{ia} X_{ia}(r_i, \vec{\tilde{e}}) \right)$$
$$- \left(\alpha - \Omega(r_i|\vec{\tilde{e}}_i) \right) \left[\hat{U}_i(0, \vec{\tilde{e}}_i) - U(0, \vec{\tilde{e}}_i, 0) \right] \Bigg] \quad (35)$$

such that $\sum_{i=1}^{N} \vec{X}_i = \vec{0}$ where $f(\cdot|\vec{\tilde{e}}_i)$ denotes the PDF of r_i and $F(r|\vec{\tilde{e}}_i)$ denotes the CDF or r_i, both conditional on $\vec{\tilde{e}}_i$.

The solution to (35) yields

$$X_{ia}(r_i, \vec{\tilde{e}}_i) = - \left(r_i - \alpha \frac{\Omega(r_i|\vec{\tilde{e}}_i) - F(r_i|\vec{\tilde{e}}_i)}{f(r|\vec{\tilde{e}}_i)} \right) \tilde{e}_{ia}$$
$$+ \frac{1}{N} \sum_{j=1}^{N} \left(r_i - \alpha \frac{\Omega(r_j|\vec{\tilde{e}}_i) - F(r_j|\vec{\tilde{e}}_j)}{f(r_j|\vec{\tilde{e}}_j)} \right) \tilde{e}_{ja}.$$

Note that F is the Rayleigh distribution:

$$F(r|\vec{\tilde{e}}_i) = 1 - e^{-\frac{r^2}{2\sigma_a^2}}$$

$$f(r|\vec{\tilde{e}}_i) = \frac{r}{\sigma_a^2} e^{-\frac{r^2}{2\sigma_a^2}}.$$

Because F and f do not depend on $\vec{\tilde{e}}_i$ it is logical to conjecture that Ω also does not depend on $\vec{\tilde{e}}_i$ and the region of binding participation constraints is an interval $[0, r]$ for some $r > 0$. A trader with a type r_i in this interval will expect to trade zero of both types. These observations yield the allocation rule in the statement of the Proposition.

There are two remaining steps. The first step is to verify that the Lagrangian approach taken is valid. This is is straightforward and analogous to that used in the proof of Theorem 1. The last step is to verify that only the incentive constraints bind radially. This follows from the proof of Proposition 8. There, I show that a double auciton with cross-exchange transaction fee can indeed implement the allocation rule in Proposition 7.

□

Proof of Proposition 8. Note that, given preferences of the form (8), a change in σ_π^2 is equivalent to a change in κ. To reduce notation, I prove Propositon 8 for the special case when $\sigma_\pi^2 = 1$.

If each trader i submits the demand schedule

$$q_{ia}(p_a) = -v(r_i)\tilde{e}_{ia} - \kappa p_a \tag{36}$$

for asset a, then the unique market clearing price is

$$p_a = \frac{1}{N\kappa} \sum_{j=1}^{N} v(r_j)\tilde{e}_{ja}$$

and so the resulting allocation is of the desired form in Proposition 7.

I now engineer the transaction fee so that (36) is indeed optimal for each asset a. I conjecture that the implementing transaction fee is of the form

$$\mathcal{T}(\{q_{i_a}(p_a), p_a\}_{a\in A}) = b\left(\sum_{a=1}^{N}\left(q_{ia} + \frac{1}{N}p_a\right)^2\right) + \frac{1}{2}cq_{ia}^2 + d$$

for some function $b : \mathbb{R} \Rightarrow \mathbb{R}$ and constants c and d to be determined.

If it is optimal for trader i to purchase $\vec{q}_i = \{q_{ia}\}_{a\in A}$ units when the vector of asset prices is $\vec{p} = \{p_a\}_{a\in A}$, if $\vec{q}_i = \{q_{ia}\}_{a\in A}$ then the solution to

$$\sup_{\{\Delta_a\}_{a\in A}} \left[\sum_{a\in A} -\frac{1}{2\kappa}\left(r_i\tilde{e}_{ia} + q_{ia} + \Delta_a\right)^2 - (p_a + \Lambda_a\Delta_a)(q_{ia} + \Delta_a) - \frac{1}{2}c\left(q_{ia} + \Delta_a\right)^2\right] \\ - b\left(\sum_{a=1}^{N}\left(q_{ia} + \kappa p_a + \Delta_a\left(1 + \kappa\Lambda_a\right)\right)^2\right) \quad (37)$$

is to set $\Delta_a = 0$ for each a. Above, $\Lambda_a = \frac{1}{(N-1)\kappa}$ is trader i's price impact when the other traders submit a demand schedule of the form in (36). The first-order optimality conditions are

$$-\frac{1}{\kappa}\left(r_i\tilde{e}_{ia} + q_{ia}\right) = p_a + \Lambda_a q_{ia} + cq_{ia} \\ + 2b'\left(\sum_{a=1}^{N}\left(q_{ia} + \kappa p_a\right)^2\right)\left(q_{ia} + \kappa p_a\right)\left(1 + \kappa\Lambda_a\right)$$

for each $a \in A$. Using that $\Lambda_a = \frac{1}{\kappa(N-1)}$, we equivalently have

$$-\frac{1}{\kappa}\left(r_i\tilde{e}_{ia} + q_{ia}\right) = p_a + \frac{1}{\kappa(N-1)}q_{ia} + cq_{ia} \\ + 2b'\left(\sum_{a=1}^{N}\left(q_{ia} + \kappa p_a\right)^2\right)\left(q_{ia} + \kappa p_a\right)\left(1 + \frac{1}{N-1}\right)$$

for each $a \in A$.

Substituting in (36) yields

$$-\frac{1}{\kappa}r_i\tilde{e}_{ia} = p_a - \left(c + \frac{1}{\kappa(N-1)} + \frac{1}{\kappa}\right)\left(v(r_i)\tilde{e}_{ia} + \kappa p_a\right) \\ - 2b'\left(v(r_i)^2\right)v(r_i)\tilde{e}_{ia}\left(1 + \frac{1}{N-1}\right).$$

In order for this to hold for all r_i, p_a, $\{\tilde{e}_{ia}\}_{a\in A}$ we require that

$$2b'\left(v(r_i)^2\right)v(r_i)\left(1+\frac{1}{N-1}\right)=\frac{1}{\kappa}r_i-\left(c+\frac{1}{\kappa(N-1)}+\frac{1}{\kappa}\right)v(r_i)$$

and

$$\kappa\left(c+\frac{1}{\kappa(N-1)}+\frac{1}{\kappa}\right)=1.$$

Thus

$$c=-\frac{1}{\kappa(N-1)}$$

and we have a differential equation for b:

$$2b'\left(x^2\right)x\left(1+\frac{1}{N-1}\right)=\frac{1}{\kappa}v^{-1}(x)-\frac{1}{\kappa}x,\quad x>0.$$

A technical issue is that v is invertible over a range of values. However, this is can be easily dealt with as in the proof of Proposition 3 by perturbing v to be strcitly monotone (an thus invertible). Following the arguments above we can implement the allocation with the perturbed v and then take limits to show that v itself is also implementable. I omit these steps.

Let me define $\tilde{b}$ such that $\tilde{b}(x) := b(x^2)$, $x\in\mathbb{R}$. Then we have

$$\left(1+\frac{1}{N-1}\right)\tilde{\mathcal{T}}'(x)=\frac{1}{\kappa}v^{-1}(x)-\frac{1}{\kappa}x,\quad x\in\mathbb{R}.$$

Integrating both sides yields

$$\tilde{b}(x)=\frac{1}{1+\frac{1}{N-1}}\left[\frac{1}{\kappa}\int_0^x v^{-1}(s)\,\mathrm{d}s-\frac{1}{2\kappa}x^2\right]+d.$$

Note that that though $v^{-1}(0)$ is not well-defined, the value of the integral does not depend on $v^{-1}(0)$.

It is easy to verify that because v is increasing, $\mathcal{T}$ is convex and thus the objective (37) is also convex. Therefore, the first-order necessary conditions for optimality are also sufficient and $\Delta_a=0$ for each a indeed solves (37) when trader i follows strategy (36). To show rigorously that this is in turn sufficient for trader i to have no deviations from (36) to an alternative set of demand schedules is a straightforward but tedious application of the calculus of variations. I omit this step for brevity. □

D. Omitted Material for Section 7

Formal Statement of Designer's Objective. Given $\alpha \in [0,1]$, the objective is to solve

$$\max_{\{X_i\},\{T_i\}} \mathbb{E}\left[\alpha \underbrace{\sum_{i=1}^{N} T_i(\vec{e})}_{\text{Revenue}} + (1-\alpha) \underbrace{\sum_{i=1}^{N} -\frac{1}{2\kappa_i}\Big(e_i + X_i(\vec{e})\Big)^2}_{\text{Allocative Efficiency}}\right] \tag{38}$$

such that,

$$\begin{aligned}
&(\mathrm{P}): \quad \mathbb{E}\left[U_i\Big(e_i + X_i(\vec{e}), T_i(\vec{e})\Big)\Big|e_i\right] \geq U_i\,(e_i, 0)\,, \quad \forall e_i, \forall i \\
&(\mathrm{IC}): \quad e_i \in \arg\max_{m_i} \mathbb{E}\left[U_i\Big(e_i + X_i(m_i, \vec{e}_{-i}), T_i(m_i, \vec{e}_{-i})\Big)\Big|e_i\right], \quad \forall e_i, \forall i \\
&(\mathrm{E}): \quad \sum_{i=1}^{N} X_i(\vec{e}) = 0, \quad \forall \vec{e}
\end{aligned}$$

where $\vec{e} = \{e_i\}_{i=1}^{N}$ is the vector of traders' endowments.

Proof of Proposition 9. Part 1 follows immediately from Corollary 5.1.

Part 2a follows form Corollary 5.1 and uses the same argument as in the proof of Part 1 of Proposition 4.

To prove Part 2b and 2c I observe that $\hat{X}_i$ increases when it is negative and $\hat{X}_i$ decreases when it is positive when σ_i^2 increases. Thus from the formula for the expected utility gain,

$$\hat{U}_i(e_i) - U_i\Big(e_i, 0\Big) = -\frac{1}{\kappa_i}\int_{\mu_e}^{e_i} \hat{X}_i(s)\,\mathrm{d}s + \hat{U}_i(\mu_e) - U_i\Big(\mu_e, 0\Big),$$

trader i's expected utility gain must decrease. On the other hand an increase in σ_i^2 has no effect on $\hat{X}_j$ for any $j \neq i$ and thus the expected utility gain of all traders $j \neq i$ must be unaffected.

To prove part 3a, I recognize that the expected trade quantity of trader i, as a function of κ_i is proportional to

$$\frac{1}{\kappa_i + \sum_{j \neq i} \kappa_j}.$$

Thus, inspecting again the equation for the expected utility gain we see that it must decrease for trader i. Because $\hat{X}_j$ does not depend on κ_i there is no effect on any traders $j \neq i$ whch proves Part 3b.

□

Proof of Proposition 10. The proof is analogous to that of Proposition 5.

□

E. Omitted Material for Section 8

Formal Statement of Designer's Objective. The designer's problem is to select a path of allocations and transfers $\{X_i(\vec{e},t), T_i(\vec{e},t); t \geq 0, \vec{e} \in \mathbb{I}^N\}_{i=1}^N$ to maximize

$$\mathbb{E}\left[\alpha \sum_{i=1}^{N} \int_0^\infty e^{-\delta t}\left[-\frac{1}{2\kappa}\left(e_i + X_i(\vec{e},t)\right)^2\right] \mathrm{d}t + (1-\alpha)\sum_{i=1}^{N}\int_0^\infty e^{-\delta t} T_i(\vec{e},t)\,\mathrm{d}t\right] \tag{39}$$

subject to

$$\begin{aligned}
\text{(P)}: &\quad \mathbb{E}\left[U\Big(\{X_i(\vec{e},t), T_i(\vec{e},t); t \geq 0\Big)\Big|e_i\right] \geq U\left(e_i, 0\right), \quad \forall e_i, \forall i \\
\text{(IC)}: &\quad e_i \in \arg\max_{m_i} \mathbb{E}\left[U\Big(\{X_i(m_i, \vec{e}_{-i}), T_i(m_i, \vec{e}_{-i}); t \geq 0\}\Big)\Big|e_i\right], \quad \forall e_i, \forall i \\
\text{(E)}: &\quad \sum_{i=1}^{N} X_i(\vec{e},t) = \sum_{i=1}^{N} e_i, \quad \forall \vec{e}, \forall t
\end{aligned}$$

where, with abuse of notation, on the RHS of (P), $U\left(e_i, 0\right)$ indicates trader i's utility from not participating in trade in which case she retains her endowment at all points in time and makes no transfers.

Proof of Proposition 11. Let $X_i(\vec{e},t)$ be an implementable allocation rule. Local incentive compatibility of truth telling implies that the transfer rule must be of the form

$$\begin{aligned}
\hat{T}_i(\vec{e},t) = \mathbb{E}&\left[\sum_{i=1}^{N}\sum_{t=0}^{\infty} -e^{-\delta t}\frac{1}{2\kappa}\Big(e_i + X_i(\vec{e},t)\Big)^2\Big|e_i\right] \\
&- \frac{1}{\kappa}e_i\hat{X}_i(e_i) + \frac{1}{\kappa}\int_{\underline{e}}^{e_i} \mathbb{E}\left[\overline{X}_i(s,\vec{e}_{-i})\right]\mathrm{d}s - \left[\hat{U}(\underline{e}) - U\Big(\underline{e}, 0\Big)\right]
\end{aligned}$$

where

$$\overline{X}_i(\vec{e},t) \equiv \sum_{t=0}^{\infty} e^{-\delta t} X_i(\vec{e},t).$$

In order for global incentive compatibility to be satisfied $\mathbb{E}\left[\overline{X}_i(s,\vec{e}_{-i})\right]$ must be nonincreasing. Consider replacing X_i with its average value. Construct a new allocation rule

$$X_i(\vec{e},t) = \frac{1}{\delta}\sum_{t=0}^{\infty} e^{-\delta t} X_i(\vec{e},t).$$

Since $X_i(\vec{e},t)$ is implementable so too must $X_i(\vec{e},t)$ since it's interim expectation must be decreasing. By Jensen's inequality this can only slacken participation constraints. By Jensen's inequality both allocative efficiency and transfers must increase.

□

References

Almgren, R. & Chriss, N. (2001). Optimal execution of portfolio transactions. *Journal of Risk*, 3, 5–40.

Andreyanov, P., Park, J., & Sadzik, T. (2023). *Optimal Robust Double Auctions*. Technical report, Working paper.

Andreyanov, P. & Sadzik, T. (2017). *Robust Mechanism Design of Exchange*. Working Paper.

Armstrong, M. (1996). Multiproduct nonlinear pricing. *Econometrica: Journal of the Econometric Society*, (pp. 51–75).

Athey, S. & Segal, I. (2013). An efficient dynamic mechanism. *Econometrica*, 81(6), 2463–2485.

Babus, A. & Parlatore, C. (2022). Strategic fragmented markets. *Journal of Financial Economics*, 145(3), 876–908.

Bergemann, D., Brooks, B., & Morris, S. (2017). First-price auctions with general information structures: Implications for bidding and revenue. *Econometrica*, 85(1), 107–143.

Bergemann, D., Heumann, T., & Morris, S. (2015). *Information and market power*. Working Paper.

Bergemann, D. & Morris, S. (2013). Robust predictions in games with incomplete information. *Econometrica*, 81(4), 1251–1308.

Bergemann, D. & Välimäki, J. (2002). Information acquisition and efficient mechanism design. *Econometrica*, 70(3), 1007–1033.

Bergemann, D. & Välimäki, J. (2010). The dynamic pivot mechanism. *Econometrica*, 78(2), 771–789.

Biais, B., Glosten, L., & Spatt, C. (2005). Market microstructure: A survey of microfoundations, empirical results, and policy implications. *Journal of Financial Markets*, 8(2), 217–264.

Biais, B., Martimort, D., & Rochet, J.-C. (2000). Competing mechanisms in a common value environment. *Econometrica*, 68(4), 799–837.

Brooks, B. & Du, S. (2021). Optimal auction design with common values: An informationally robust approach. *Econometrica*, 89(3), 1313–1360.

Budish, E., Cramton, P., Kyle, A. S., Lee, J., & Malec, D. (2023). *Flow trading*. Technical report, National Bureau of Economic Research.

Chen, D. & Duffie, D. (2021). Market fragmentation. *American Economic Review*, 111(7), 2247–74.

Chen, D. & Zhang, A. L. (2020). Subsidy schemes in double auctions. *Available at SSRN 3668394*.

Clarke, E. H. (1971). Multipart pricing of public goods. *Public choice*, (pp. 17–33).

Colliard, J.-E. & Foucault, T. (2012). Trading fees and efficiency in limit order markets. *The Review of Financial Studies*, 25(11), 3389–3421.

Crémer, J. & McLean, R. P. (1988). Full extraction of the surplus in bayesian and dominant strategy auctions. *Econometrica: Journal of the Econometric Society*, (pp. 1247–1257).

Dasgupta, P. & Maskin, E. (2000). Efficient auctions. *The Quarterly Journal of Economics*, 115(2), 341–388.

Du, S. & Zhu, H. (2017). What is the Optimal Trading Frequency in Financial Markets? *The Review of Economic Studies*, 84(4), 1606–1651.

Foucault, T., Kadan, O., & Kandel, E. (2013). Liquidity cycles and make/take fees in electronic markets. *The Journal of Finance*, 68(1), 299–341.

Glosten, L. R. (1994). Is the electronic open limit order book inevitable? *The Journal of Finance*, 49(4), 1127–1161.

Groves, T. (1973). Incentives in teams. *Econometrica: Journal of the Econometric Society*, (pp. 617–631).

Jantschgi, S., Nax, H. H., Pradelski, B., & Pycia, M. (2022). Double auctions and transaction costs.

Jehiel, P. & Moldovanu, B. (2001). Efficient design with interdependent valuations. *Econometrica*, 69(5), 1237–1259.

Jullien, B. (2000). Participation constraints in adverse selection models. *Journal of Economic Theory*, 93(1), 1–47.

Kyle, A. S. (1985). Continuous Auctions and Insider Trading. *Econometrica*, 53(6), 1315–1335.

Kyle, A. S. (1989). Informed Speculation with Imperfect Competition. *The Review of Economic Studies*, 56(3), 317–355.

Lambert, N. S., Ostrovsky, M., & Panov, M. (2014). Strategic trading in informationally complex environments. In *Proceedings of the fifteenth ACM conference on Economics and computation* (pp. 3–4).

Lu, H. & Robert, J. (2001). Optimal trading mechanisms with ex ante unidentified traders. *Journal of Economic Theory*, 97(1), 50–80.

Luenberger, D. G. (1997). *Optimization by vector space methods*. John Wiley & Sons.

Malamud, S. & Rostek, M. (2017). Decentralized Exchange. *American Economic Review*, 107(11), 3320–3362.

Malinova, K. & Park, A. (2015). Subsidizing liquidity: The impact of make/take fees on market quality. *The Journal of Finance*, 70(2), 509–536.

McAfee, R. P. (1991). Efficient allocation with continuous quantities. *Journal of Economic Theory*, 53(1), 51–74.

Myerson, R. B. (1981). Optimal auction design. *Mathematics of operations research*, 6(1), 58–73.

Myerson, R. B. & Satterthwaite, M. A. (1983). Efficient mechanisms for bilateral trading. *Journal of economic theory*, 29(2), 265–281.

Pavan, A., Segal, I., & Toikka, J. (2014). Dynamic mechanism design: A myersonian approach. *Econometrica*, 82(2), 601–653.

Rostek, M. & Weretka, M. (2012). Price Inference in Small Markets. *Econometrica*, 80(2), 687–711.

Rostek, M. & Yoon, J. H. (2021). Exchange design and efficiency. *Econometrica*, 89(6), 2887–2928.

Rostek, M. J. & Yoon, J. H. (2023). Financial product design in decentralized markets. *Available at SSRN 3631479*.

Sannikov, Y. & Skrzypacz, A. (2016). Dynamic trading: Price inertia and front-running. *Preprint*.

Vayanos, D. (1999). Strategic Trading and Welfare in a Dynamic Market. *The Review of Economic Studies*, 66(2), 219–254.

Vickrey, W. (1961). Counterspeculation, auctions, and competitive sealed tenders. *The Journal of finance*, 16(1), 8–37.

Vives, X. (2011). Strategic Supply Function Competition With Private Information. *Econometrica*, 79(6), 1919–1966.

Wilson, R. B. (1993). *Nonlinear pricing*. Oxford University Press, USA.

Wittwer, M. (2021). Connecting disconnected financial markets? *American Economic Journal: Microeconomics*, 13(1), 252–82.