

Essays in Information Economics and Market Structure

Andreas Uthemann

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
of
University College London.

Department of Economics
University College London

May 28, 2015

I, Andreas Uthemann, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

This thesis analyses three distinct economic problems through the common lens of informational asymmetries. In each chapter we show how market behaviour is best understood as the outcome of differentially informed market participants interacting with each other within the rules specified by the respective market under consideration.

Chapter 2 provides an explanation for the variety of contracts offered by competitive firms for seemingly identical products or services, e.g. in mobile communication or personal banking. We show that firms' menu of tariffs can be understood as screening devices for consumers with mistaken beliefs about their future demand. We furthermore show that while competition between firms prevents firms from exploiting their customers' limited cognitive ability, competition is not able to correct the inefficiency caused by customers making suboptimal choices.

Chapter 3 studies the effect of a financial transaction tax on the trading of a security. We construct a market microstructure model and estimate it using intraday transaction data for a stock traded on the NYSE (Ashland Inc.). The estimates are then used to simulate how a financial transaction tax would impact volume, spreads and informational efficiency in the asset market under consideration.

Chapter 4 constructs a model of observational learning with payoff externalities that provides a justification for the use of short term debt in the financing of investment projects. While financing with debt that is subject to roll over risk is often seen as a source of instability, potentially triggering investor runs on financially sound institutions, we show that it can play an important role in facilitating the revelation of privately held information about future performance of the investment.

Acknowledgements

First, I would like to thank my supervisor Antonio Guarino. His continuous encouragement and support have been essential to the completion of this thesis.

I would also like to thank Mark Armstrong for his guidance during the early stages of this work. His feedback has been of great help in the writing of the second chapter of this thesis. I am equally thankful to Nicola Pavoni both for his academic advice and for his help during the more challenging phases of my thesis.

The Department of Economics at UCL has not only given me the academic education necessary to undertake this work but has also been my second home over the last years. I particularly thank all my fellow doctoral students for the welcoming atmosphere and the stimulating intellectual climate they have helped to create.

I would like to acknowledge financial support from the *Economic and Social Research Council* (ESRC) in the form of a studentship awarded by the *Centre for Economic Learning and Social Evolution* (ELSE) based at UCL.

Lastly I would like to thank my parents, Anton and Hedwig Uthemann. Without their support I would have been unable to pursue my studies.

Contents

1	Introduction	10
2	Competitive Screening of Customers with Non-Common Priors	13
2.1	Introduction	13
2.2	Model Setup	18
2.2.1	Competition in the Utility Space	19
2.2.2	Hotelling Competition over a Single Type	22
2.2.3	Hotelling Competition with Multiple Unobservable Types	23
2.3	The Effects of Competition	24
2.3.1	Common Priors but Different Consumption Profiles	27
2.4	Conclusion	30
3	Financial Transaction Taxes and the Informational Efficiency of Financial Markets: A Structural Estimation	32
3.1	Introduction	32
3.2	The Model	35
3.3	Equilibrium Analysis	41
3.4	The Financial Transaction Tax	44
3.5	The Likelihood Function	49
3.6	Data	52
3.7	Results	54
3.7.1	Estimates	54
3.8	Simulating the Impact of a FTT	56

	<i>Contents</i>	6
3.8.1	Impact on Volume and Spreads	57
3.8.2	Impact of Informational Efficiency	60
3.9	Conclusion	62
4	Preemption and Delay: Debt Financing with Roll Over Risk	63
4.1	Introduction	63
4.2	Model Setup	66
4.3	The Liquidation Game	68
4.4	Efficient Liquidation	74
4.5	Optimal Maturity Structure	76
4.6	Discussion	78
4.7	Conclusion	80
	Appendices	82
A	Appendix to Chapter 2	82
B	Appendix to Chapter 3	87
B.0.1	Calibration of the Transaction Tax	87
B.0.2	Informational Cascades	89
B.0.3	Estimation Results for Simulated Data	93
C	Appendix to Chapter 4	95
C.0.4	Proofs	95
D	Note on Cojoint Work	98
	Bibliography	99

List of Figures

3.1	<i>Density of Signal.</i> The figure shows the signal's value-contingent density functions for different values τ	38
3.2	<i>Informed trader's decision.</i> The figure illustrates the signal realizations for which an informed trader decides to buy or sell when $V^d = v_h^d$	42
3.3	<i>Price elastic noise trader's decision.</i> The figure illustrates the shock realization for which a price elastic noise trader decides to buy or sell.	43
3.4	<i>Spreads by tax regime.</i> The figure displays average spreads on 'no event day' over course of trading day for three tax regimes.	60
4.1	<i>Share of short term debt and liquidation value</i>	78
B.1	<i>Determination of Roots.</i> Red curves correspond to signals $S = 1$, blue curves to $S = 0$. Dotted curves have $\tau = 0.3$, solid curves have $\tau = 0.6$. Intersections of the red curves with horizontal line occur at p_1 and p_2 , intersections of the blue curves with the horizontal line occur at p_3 and p_4	92

List of Tables

3.1	Estimation Results. Estimates for the six parameters of the model and their standard errors (s.e.).	55
3.2	Likelihood-Ratio Test. Log-likelihood value (LL) of the restricted and unrestricted model as well as the LR test statistic and corresponding p-value. . .	55
3.3	Estimation Results (Restricted Model). Estimates for the five parameters of the restricted model and their standard errors (s.e.).	56
3.4	Calibration of FTT	57
3.5	Trade volume by tax rates (all days). Fraction of trading times at which trade (buy or sell) takes place by trading period and tax rate.	58
3.6	Trade volume by tax rates (event days). Fraction of trading times at which trade (buy or sell) takes place by trading period and tax rate.	58
3.7	Bid-ask spread by tax rates (all days). Mean bid-ask spread by trading period and tax rate.	59
3.8	Bid-ask spread by tax rates (event days). Mean bid-ask spread by trading period and tax rate.	59
3.9	Bid-ask spread in relation to FTT (all days). Mean bid-ask spread divided by 2 as a fraction of the transaction tax ρ by trading period and tax rate.	59
3.10	Bid-ask spread in relation to FTT (event days). Mean bid-ask spread divided by 2 as a fraction of the transaction tax ρ by trading period and tax rate.	60
3.11	Price deviation from fundamental value. Absolute price deviation from fundamental value as a fraction of expected value by tax rates.	61

3.12	<i>Price deviation from price without FFT. Absolute price deviation from price without transaction tax as a fraction of expected value by tax rates.</i>	62
4.1	<i>Payoff matrix in case of liquidation</i>	68
B.1	<i>Calibration of FTT</i>	89
B.2	<i>Estimates for Simulation 1, Seed 1.</i>	93
B.3	<i>Estimates for Simulation 1, Seed 2.</i>	94
B.4	<i>Estimates for Simulation 2, Seed 1.</i>	94
B.5	<i>Estimates for Simulation 2, Seed 2.</i>	94
B.6	<i>Estimates for Simulation 3, Seed 1.</i>	94
B.7	<i>Estimates for Simulation 3, Seed 2.</i>	94

Chapter 1

Introduction

How informational asymmetries between market participants translate into market outcomes is strongly contingent on the structure the market imposes on the interactions between actors. Among the main determinants of market structure are the composition of the trading population, the actions these traders are allowed to undertake and the amount of information feedback they get about their and others' actions.

In this thesis we will consider three economic problems for which, in our opinion, market behaviour is best explained when perceived as the outcome of the interactions of differentially informed market participants. For each problem in turn we will show how asymmetric information together with important structural features of the market under consideration, such as the contract forms allowed for trade or the degree of competition between suppliers of goods, allow us to bring economic reasoning to bear on observed market outcomes and trace these back to the core economic forces at work. While chapters 2 and 4 rely on the tools of economic theory, in particular game theory and mechanism design, to perform this exercise, chapter 3 combines empirical and theoretical work by estimating a structural model of an asset market in order to quantify important underlying structural parameters of the market.

Chapter 2 provides an explanation for the variety of contracts offered by competitive firms for seemingly identical products or services, e.g. in mobile communication or personal banking. We show that two competing firms offering menus of nonlin-

ear price schedules to customers with mistaken priors about their future demand will be able to screen these customers on the basis of their priors. Firms' use of tariff menus can thus be understood as screening devices for boundedly rational consumers. However while such menus allow firms to screen their customers according to their ability to correctly forecast future demand, they do not allow them to extract additional rents provided the market is sufficiently competitive and firms' have identical priors concerning their customers types. Competition thus guarantees that each tariff only covers the fixed costs of the firm, but it cannot remedy the mistaken demand forecasts of customers.

Chapter 3 studies the effect the introduction of a financial transaction tax would have on the trading of a security. To this purpose we construct a market microstructure model and estimate it using intraday transaction data for a stock traded on the NYSE (Ashland Inc.). The estimates are then used to simulate how a financial transaction tax would impact volume, spreads and informational efficiency in the asset market under consideration. Our structural model includes price elastic liquidity traders as well as informed traders with signals of heterogeneous quality. Both features are essential for our analyses as the impact of the tax will crucially depend on how it affects the composition of the trading population in the market. Our results suggest that liquidity trade is relatively price inelastic and thus the tax predominantly affects informed trade. The reduction in informed trade is found to be substantial. A 0.2% tax on transactions, as has for example been implemented in France, would lead to a 10% drop in volume in our asset market. While spreads can be reduced by the tax as less informed trade in the market means that spreads have to provide less protection against traders with superior information, we find that convergence of the asset price to the fundamental value of the asset can be slowed down considerably. We find, for example, that a tax of 0.5% per transaction would lead to a complete crowding out of informed trade and thus the asset market would lose its ability to aggregate private information through asset prices.

The final chapter of this thesis constructs a model of observational learning that pro-

vides a justification for the use short term debt subject to roll over risk in the financing of investment projects. While such arrangements are often seen as a source of instability, potentially triggering investor runs on financially sound institutions, we show that they can play an important role in facilitating the revelation of privately held information about future performance. When investors can learn about the profitability of a project by observing other investors' funding decisions, waiting rather than withdrawing funding upon receiving negative information has a positive option value. Under equitable sharing rules for liquidation revenue this option value can outweigh the negative information received by individual investors and lead to the survival of unprofitable investment projects. Providing financial incentives for early withdrawers can counteract this option value effect and implement efficient liquidation decisions. Here the fragility of financing facilitates information aggregation, enables efficient liquidation decisions, and thereby raises the ex-ante value of investment projects.

Chapter 2

Competitive Screening of Customers with Non-Common Priors

2.1 Introduction

The multitude of tariffs for seemingly homogenous goods and services offered by firms in a wide range of competitive industries, such as in mobile communications or personal banking, constitutes a theoretical problem. While at first sight one is tempted to infer that firms use this variety of tariffs to price discriminate between customers with different consumption profiles, standard models of competitive price discrimination (Armstrong and Vickers [Armstrong and Vickers, 2001], Rochet and Stole [Rochet and Stole, 2002]) tell us that in sufficiently competitive markets with firms facing identical technological constraints and demand conditions, firms lose the ability to engage in price discrimination. Competition forces all firms to offer a single two-part tariff which consists of a fixed fee plus costs.

In this paper I will provide an explanation for these large menus of tariffs offered by competitive firms which is based on heterogeneity in cognitive abilities of the customer base. In doing so other possible causes which could potentially account for the phenomenon will be ignored, most prominently among them dynamic considerations like switching costs or asymmetries between firms both in technologies or in consumer preferences. At the center of the analysis will be the customer's ability to anticipate

his consumption profile for a product which is provided by firms in a competitive market. This ability is important as customers will have to choose between contracts specifying price-quantity schedules before knowing their exact demand for the good. Once signed up to a contract, they will find out about their exact valuation for the good and will choose their optimal consumption level from this specific contract. Crucially the firm can disagree with its customers at the time contracts are signed concerning the customer's consumption profile. One example would be that the firm believes its customer is too optimistic concerning his future intensity of usage of the service. In the technical term of the cognitive psychology literature, the firm thinks that its customer's beliefs are miscalibrated. I will show that if firms believe that some of their customers have miscalibrated beliefs but these beliefs are private information, firms will design their tariffs so as to screen customers with respect to these heterogeneous beliefs. Competition between firms will not prevent firms from using their menu design to screen customers in this manner.

The results of this model allow some insights into the question of what beneficial effects we can expect from market competition in the presence of boundedly rational market participants. It will be shown that the standard inefficiency that price discrimination creates by distorting optimal quantities for low types in order to extract more surplus from high types will be eliminated by competition even if some of the market participants commit mistakes in evaluating deals. However firms will not deliver the efficient quantity schedule to all customers, as some customers' perceived surplus is not their actual surplus. This will induce firms to create what in the behavioral economics literature is sometimes called *fictitious surplus*, surplus which only exists in the customers imagination, will never materialize and therefore can be costlessly provided by the firm. Competition will nevertheless ensure that firms do not make any additional profits by creating this kind of surplus. Standard customers, i.e. customers that shares common prior beliefs with the firms, will choose a cost plus fixed fee tariff. This tariff turns out to be the same tariff firms would offer to this customer type if no miscalibrated customers were present in the market. It follows that in this model the

presence of boundedly rational customers does not exert an externality on the fully rational customers.

Empirical Evidence on Miscalibration Calibration measures the agreement between the objective and subjective assessment of validity of a statement. A person is therefore perfectly calibrated if the subjective probability she assigns to any event matches the long run frequency of occurrences of this event. In our setting a perfectly calibrated customer can correctly quantify the probability that his marginal utility from consuming a certain quantity of the good is below a given level.¹

The notion of calibration has been extensively discussed in the experimental psychology literature (Lichtenstein et al. [Lichtenstein et al., 1982], McClelland and Bolger [McClelland and Bolger, 1994]). A robust result in this literature is that for a large variety of situations, experimental subjects show a significant degree of miscalibration. A well know example is the tendency of experimental subjects to overestimate their ability to correctly answer general knowledge question.

More recently economists have begun to elicit subjective probability distributions which matter for specific economic problems and check the goodness of calibration of people's beliefs. Dominitz and Manski [Dominitz and Manski, 2011] analyze probabilistic beliefs concerning equity returns using data from the Michigan Survey of Consumers and the Survey of Economic Expectations. They find significant interpersonal variation but intrapersonal stability in beliefs. When compared to the long run average of equity returns, subjects tend to be miscalibrated, on average overestimating both the mean and the volatility of returns. Similar studies have been conducted for job losses, eligibility for social security or income uncertainty with similar results. Manski [Manski, 2004] provides a survey of this literature.

Non-common Prior Approach At the time of contracting between customer and firm, the exact utility the customer will derive from the good is uncertain. This uncertainty will be indexed by a one-dimensional random variable which parameterizes the customer's utility function. Call this random variable θ and let its probability distribution

¹This obviously presumes the existence of an objectively valid probability distribution for these marginal utilities.

be given by $F(\theta)$. In the following model a person will be called miscalibrated if his prior beliefs concerning the random variable θ do not agree with the objective probability distribution $F(\theta)$.

It will be assumed that firms are perfectly calibrated, i.e. that their prior beliefs concerning θ are given by $F(\theta)$. Firms know that their customers are all identical and have a consumption profile which is determined by $F(\theta)$. Customers' prior beliefs can divert from $F(\theta)$. To solve the competitive model I will not have to restrict the way in which customers' beliefs divert from this probability distribution. All that is necessary is that firms know in which ways their customers' beliefs can differ from $F(\theta)$. In order to solve the monopolistic case, I will have to impose more structure on the set of possible deviations. I will use two criteria to order customers' beliefs. Customers' prior beliefs will be allowed to divert from $F(\theta)$ either in a first order stochastic dominance sense, or in a mean preserving spread sense. In psychological terms, customers will be allowed to be either overpessimistic (overoptimistic), or underconfident (overconfident).

The above description of the situation with calibrated firms and miscalibrated customers is only one possible interpretation one could give to this setup. What is crucial is that all firms share a common prior belief concerning θ and that firms and customers possibly disagree concerning the distribution of this variable. Another possibility would be to interpret firms as boundedly rational and customers to be perfectly calibrated. However it is not possible to interpret this in the sense that customers are better informed concerning their tastes. Such an interpretation would alter firms' strategies when compared to our analysis. Firms would now gain additional information through customers' choices which would have to be taken into account. Disagreement about beliefs is central to the analysis.

Related Literature There is a fast growing literature on optimal contract design in the presence of consumer biases. Two papers are particularly close in spirit to our analysis. Della Vigna and Malmendier [DellaVigna and Malmendier, 2006] analyse the profit-maximizing contract when customers have time-inconsistent preferences. They show that goods with immediate costs but delayed benefits (investment goods) are priced

below marginal cost, while goods with immediate benefits but delayed costs (leisure goods) are priced above marginal costs. Consumers that face self control problems and are (partially) naive about it will overestimate their demand in the former case and underestimate demand in the latter case. Firms will exploit these biases by readjusting the price profile, i.e. raising (lowering) the fixed fee and adjusting the per usage price accordingly in order to create what they call fictitious surplus. DellaVigna and Malmendier show that these pricing patterns are observable in a wide variety of industries ranging from health clubs to credit card companies.

Eliaz and Spiegelger [Eliaz and Spiegelger, 2006] extend the above analysis to a contracting environment in which consumers have time inconsistent preferences but vary in their degree of awareness concerning this inconsistency. While firms know that consumption of their good will induce a change of customers' preferences, customers assign a probability between zero and one to this shift, where this probability describes the customer's type unknown to the firm. Elias and Spiegelger show that three part tariffs are necessary to implement the optimal screening contract. While sufficiently naive customers will sign exploitative contracts, in the sense that under the correct prior the contract offers less utility than their outside option, the contract offers a commitment device for sufficiently sophisticated customers helping them to overcome their self-control problems. Our modeling approach differs from the above papers in that it focuses on the effects of varying degrees of demand uncertainty on the side of customers on contract design. In Della Vigna and Malmendier [DellaVigna and Malmendier, 2006] there is no demand uncertainty. Partially naive customers simply misconceive of their future demand, judging it too high or too low depending on the cost structure of the problem. Elias and Spiegelger [Eliaz and Spiegelger, 2006] introduce demand uncertainty into their framework as partially naive consumers are unsure about future changes in their preferences. However, as there are only two possible alternative preference structures, their model does not allow to disentangle the effects of overestimating or underestimating demand from increases in demand uncertainty without changes in the mean valuation

for the service provided.²

The structure of the model is very close to the setup of the Sequential Screening literature (Armstrong [Armstrong, 1996], Courty and Li [Courty and Li, 2000]). Here a monopolist tries to screen customers with differing demand patterns through contracts that are offered before customers know their actual demand. Types are indexed by their distribution functions over ex-post demand realization. Firms and customers ex-ante hold identical priors for these distribution functions. Thus customer types can be interpreted as high demand / low demand types if priors are ordered by first order stochastic dominance, or risky / safe customers if ordered by mean preserving spreads.

2.2 Model Setup

The Firms There are two firms A and B, which are situated at the opposite end of a segment of unit length. Each firm produces a single good. The costs of providing a quantity q of this good to a customer are $C(q)$, where we assume that $C'(q), C''(q) \geq 0$. Firms offer a menu of (non-linear) price schedules to their customers. Firm i 's menu of contracts J_i is given by

$$\{P^i(q, j)\}_{j \in J_i}. \quad (2.1)$$

A customer who has signed contract j at firm i pays $P^i(q, j)$ for a quantity q of the good.

Consumers Consumers are located on the line between firm A and B. A customer who has signed contract $j \in J_A$ at firm A and is situated at x on the line derives a utility of

$$u(q, \theta) - P^A(q, j) - \tau x \quad (2.2)$$

from consuming a quantity q of the good. The same consumer on contract $k \in J_B$ at firm B would derive a utility of

$$u(q, \theta) - P^B(q, k) - (1 - x)\tau \quad (2.3)$$

²The advantage of this approach is that the way in which preferences can change is completely unrestricted. In our model ex-post preferences are indexed by a one-dimensional parameter and in addition to that have to satisfy a single-crossing property in this parameter.

from consuming q .

$\theta \in [\theta_L, \theta_H]$ is a one-dimensional preference shifter. It is assumed that $u_q(q, \theta) > 0$, $u_\theta(q, \theta) > 0$, and that the utility function satisfies a standard single-crossing property in θ , i.e.

$$u_{q\theta}(q, \theta) > 0. \quad (2.4)$$

τ is the consumer's per-unit travel cost.

Information Structure Both consumer location x and the consumer's demand type θ are private information. While the consumer knows x at the time of signing a contract, he does not know θ . He only finds out about his exact demand once he has signed a contract. A consumer has prior beliefs concerning the possible realizations of θ which are given by the probability distribution $F(\theta, \alpha)$. α will be called the consumer's *ex-ante type*. For the moment no restrictions are imposed on the parameter α .³ Let \mathcal{A} designate the set of *ex-ante types* that are active in the market. There is a measure one of each type $\alpha \in \mathcal{A}$, and for each type α this mass is distributed uniformly on the segment between firm A and B .

Firms do not know the exact type of their customers. They only know the locational distribution of types, the set \mathcal{A} of active *ex-ante types*⁴, and all firms share a common prior belief about the distribution of θ which we will designate by $F(\theta)$.

Under my intended interpretation of the model $F(\theta)$ will be the "true" distribution of θ , and the *ex-ante type* $\hat{\alpha}$ such that $F(\theta, \hat{\alpha}) = F(\theta)$ will be called the *fully rational type*.

2.2.1 Competition in the Utility Space

Each price schedule in the firm's menu can be considered as a *deal* of a certain value that is offered by the firm to its customer. Firms compete over customers by trying to offer them better deals, i.e. higher utility levels, through their choice of price schedules. This idea of formalizing competition in multidimensional objects as competition in the utility space has been first put forward by Bliss [Bliss, 1988] and further developed by

³When solving the monopoly case, α will be one-dimensional and α will order the set of distributions $\{F(\theta, \alpha)\}$ by first order stochastic dominance or mean preserving spreads.

⁴For the monopolistic case, the firm will need to know the distribution of α not only its support.

Armstrong and Vickers [Armstrong and Vickers, 2001].

Suppose firm A intended to offer a customer of type α a utility level of u . Obviously firm A would provide this utility level in a way that maximizes its own profits. Thus the following profit maximization problem implicitly determines the quantity schedule firm A would provide to a customer of type α if it wanted to guarantee him a utility level u net of travel costs:

$$\begin{aligned} \pi^A(u, \alpha) = \max_{P^A, \{q(\theta, \alpha)\}} & \int_{\theta_L}^{\theta_H} \left(P^A(q(\theta, \alpha), \alpha) - C(q(\theta, \alpha)) \right) f(\theta, \hat{\alpha}) d\theta \\ \text{subject to} & \int_{\theta_L}^{\theta_H} s^A(\theta, \alpha) f(\theta, \alpha) d\theta \geq u \end{aligned}$$

where $s^A(\theta, \alpha) = u(q(\theta, \alpha), \theta) - P^A(q(\theta, \alpha), \alpha)$.

Any chosen quantity schedule $\{q(\theta, \alpha)\}$ has to be ex-post incentive compatible which is equivalent to imposing the usual Envelope condition $s^A_\theta(\theta, \alpha) = u_\theta(q(\theta, \alpha), \theta)$ and requiring $q(\theta, \alpha)$ to be weakly increasing in θ .

This allows us to rewrite the constraint on consumers' expected utility under a quantity schedule $\{q(\theta, \alpha)\}$ as

$$s(\theta_L, \alpha) + \int_{\theta_L}^{\theta_H} u_\theta(q(\theta, \alpha), \theta) [1 - F(\theta, \alpha)] d\theta \geq u$$

Using the definition of consumer surplus to replace P^A in firm A 's objective function and integrating by parts we get

$$\begin{aligned} & \int_{\theta_L}^{\theta_H} \left(P^A(q(\theta, \alpha), \alpha) - C(q(\theta, \alpha)) \right) f(\theta, \hat{\alpha}) d\theta = \\ & \int_{\theta_L}^{\theta_H} \left\{ u(q(\theta, \alpha), \theta) - s(\theta_L, \alpha) - u_\theta(q(\theta, \alpha), \theta) \left(\frac{1 - F(\theta, \hat{\alpha})}{f(\theta, \hat{\alpha})} \right) - C(q(\theta, \alpha)) \right\} f(\theta, \hat{\alpha}) d\theta \end{aligned}$$

We can now substitute the constraint on consumers' expected utility into this expression upon noting that in any profit maximizing solution firm A will provide consumers with

an expected utility of exactly u .

$$\pi(u, \alpha) = \max_{\{q(\theta, \alpha)\}} \int_{\theta_L}^{\theta_H} \Lambda^C(q(\theta, \alpha), \theta, \alpha) f(\theta, \hat{\alpha}) d\theta - u \quad (2.5)$$

where

$$\Lambda^C(q, \theta, \alpha) = u(q, \theta) + u_\theta(q, \theta) \left[\frac{F(\theta, \hat{\alpha}) - F(\theta, \alpha)}{f(\theta, \hat{\alpha})} \right] - C(q) \quad (2.6)$$

We get a candidate solution for the profit-maximizing schedule by maximizing the objective function pointwise with respect to q for each θ :

$$q(\theta, \alpha) = \arg \max_{q \geq 0} \Lambda^C(q, \theta, \alpha) \quad (2.7)$$

If Λ^C is strictly quasi-concave in q , this candidate is given by the first order condition

$$\Lambda_q^C(q(\theta, \alpha), \theta, \alpha) = 0 \quad (2.8)$$

If furthermore Λ^C is supermodular in (q, θ) , $q(\theta, \alpha)$ will be (weakly) increasing in θ and the above candidate is indeed the profit maximizing quantity schedule.

Notice that this schedule is independent of the level of utility u provided to a customer of type α . Differences in utility are not provided by altering the quantity schedule but are undertaken through changes in $s(\theta_L, \alpha)$, i.e. through changes in the “fixed fee” associated with the deal.

To construct the price schedule $P(\cdot, \alpha)$ that implements the profit maximizing quantity schedule $\{q(\theta, \alpha)\}$ we can follow the usual steps. First note that

$$s(\theta_L, \alpha) = u - \int_{\theta_L}^{\theta_H} u_\theta(q(\theta, \alpha), \theta) d\theta \quad (2.9)$$

This allows us to recover the surplus function

$$s(\theta, \alpha) = s(\theta_L, \alpha) + \int_{\theta_L}^{\theta_H} u_\theta(q(x, \alpha), x) dx \quad (2.10)$$

Now from the definition of $s(\theta, \alpha)$ we can construct the optimal price schedule

$$P^*(q(\theta, \alpha), \alpha) = u(q(\theta, \alpha), \theta) - s(\theta, \alpha) \quad (2.11)$$

We will separate P^* into a usage charge $P(\cdot, \alpha)$ such that $P(q(\theta_L, \alpha), \alpha) = 0$ and a fixed fee $t(u, \alpha)$. This fee will adjust the value of the deal to the required utility level u .

For future reference let $c(\alpha)$ designate the expected cost of implementing the profit maximizing quantity schedule net of the fixed fee, i.e.

$$c(\alpha) = \int_{\theta_L}^{\theta_H} \{C(q(\theta, \alpha)) - P(q(\theta, \alpha), \alpha)\} f(\theta, \hat{\alpha}) d\theta \quad (2.12)$$

Note that for the fully rational type $\alpha = \hat{\alpha}$ this cost is obviously zero, as the profit maximizing price schedule equals actual costs.

Furthermore we designate by $S(\alpha)$ the total surplus excluding the fixed fee a customer of type α will receive from the profit maximizing quantity schedule $q(\theta, \alpha)$

$$S(\alpha) = \int_{\theta_L}^{\theta_H} \{u(q(\theta, \alpha), \theta) - P(q(\theta, \alpha), \alpha)\} f(\theta, \alpha) d\theta \quad (2.13)$$

2.2.2 Hotelling Competition over a Single Type

Suppose for the moment that only a single ex-ante type α is on the market. As described above a measure one of them is distributed uniformly between the two firms A and B . The following proposition characterizes the equilibrium of the standard Hotelling game:

Proposition 1. *Provided that $\tau \leq (2/3)[S(\alpha) - c(\alpha)]$, offering the fixed fee $t(\alpha) = \tau + c(\alpha)$ and the price schedule $P(q, \alpha)$ is an equilibrium of the above Hotelling game.*

Proof. Suppose firm B offers the suggested schedule. Then firm A 's market share, given that it offers a surplus of S (net of travel costs) to its customers, is

$$x = \min \left\{ \frac{1}{2} \left(1 + \frac{S - S(\alpha) + \tau + c(\alpha)}{\tau} \right), \frac{S}{t} \right\}$$

Obviously firm A will provide the surplus level S through the profit maximizing quantity

schedule. Customers' surplus from buying at firm A net of travel costs will then be $S(\alpha) - k$, where k is the fixed fee firm A will charge in addition to the price schedule $P(q, \alpha)$.

Under the assumption that $\tau \leq (2/3)[S(\alpha) - c(\alpha)]$ the market is fully covered and profits for firm A when charging a fixed fee k are

$$[k - c(\alpha)] \frac{1}{2} \left(1 + \frac{\tau + c(\alpha) - k}{\tau} \right)$$

The fixed fee that maximizes these profits is given by $k = \tau + c(\alpha)$ which proves the claim. \square

2.2.3 Hotelling Competition with Multiple Unobservable Types

Now let us go back to the initially described setting in which the two firms face a variety of ex-ante types α , where the exact type of each customer is unknown to the firm. A measure one of each α type is distributed uniformly between A and B .

The following proposition states that the competitive situation with multiple unobservable types is separable. Offering the equilibrium derived for the single type setting to every type α present in the market is an equilibrium of this more complex game. The incentive constraints that ensure self-selection will turn out to be non-binding.

Proposition 2. *Suppose $\tau \leq (2/3)[S(\alpha) - c(\alpha)]$ for all α present in the market. Then each firm offering a menu of contracts in which each α receives his contract from Proposition 1 is an equilibrium.*

Proof. The following proof adopts an argument from the proof of Proposition 5 in Armstrong and Vickers [Armstrong and Vickers, 2001]. Suppose firm B offered such a menu of contracts and that furthermore this menu was ex-ante incentive compatible in the sense that each α would indeed choose his contract from Proposition 1. Now suppose that firm A could actually observe the types α . This gives us an upper bound on the profits firm A could make. Then it is a best reply for firm A to offer each type α the contract derived in Proposition 1.

If we can show that this menu of contracts is also incentive compatible, that is that

no type α would like to deviate to any other contract originally intended for $\alpha' \neq \alpha$, then it is also a best reply by firm A if it cannot observe the types. Ex-ante incentive compatibility of contracts requires that

$$\begin{aligned} \int_{\theta_L}^{\theta_H} [u(q(\theta, \alpha), \theta) - P^*(q(\theta, \alpha), \alpha)] f(\theta, \alpha) d\theta \\ \geq \int_{\theta_L}^{\theta_H} [u(q(\theta, \alpha'), \theta) - P^*(q(\theta, \alpha'), \alpha')] f(\theta, \alpha) d\theta \end{aligned}$$

for all $\alpha' \neq \alpha$.

Upon substituting in $t(\alpha) = \tau + c(\alpha)$ and rearranging this is equivalent to

$$c(\alpha') - c(\alpha) \geq \int_{\theta_L}^{\theta_H} [u(q(\theta, \alpha'), \theta) - P(q(\theta, \alpha'), \alpha')] f(\theta, \alpha) d\theta - S(\alpha)$$

Now suppose there existed an $\alpha' \neq \alpha$ for which this inequality would be violated. Then the costs of moving a customer of type α from the quantity schedule $q(\theta, \alpha)$ to $q(\theta, \alpha')$ would be lower than the consumer surplus (real and fictitious) created through this reallocation. Thus a firm could raise its profits from a type α by such a move, raising or lowering the fixed fee in order to keep the type's expected utility from the contract unaltered. But this contradicts the condition that $q(\theta, \alpha)$ is the profit maximizing quantity schedule for type α . \square

The proof shows that incentive compatibility under competition is a direct consequence of profit maximization by firms. Furthermore the proof does not require us to impose any structure on the set of priors. Thus Proposition 2 holds for any kind of differing priors held by consumers in the market.

2.3 The Effects of Competition

To get an understanding of what competition can and cannot achieve in the presence of customers with mistaken priors, I will contrast the above result to two settings. Firstly, I will analyze the optimal tariff design of a monopolist where consumers are as described above.

Secondly, I will analyze the case where consumers do not have mistaken priors, but

differ in their true consumption profiles. That is for two ex-ante types $\alpha \neq \alpha'$, we will have $F(\theta, \alpha) \neq F(\theta, \alpha')$, but firms and consumers have common priors. In other words, I will analyze a competitive version of the standard Sequential Screening problem (Armstrong [Armstrong, 1996], Courty and Li [Courty and Li, 2000]).

Suppose that instead of having a duopoly, there is only a single firm providing the good to customers. Customers are as described in the above setting except that I will ignore travel costs here.

Again, the firm offers its customers a menu of price schedules $\{P(q, j)\}_{j \in J}$ before these customers know their type θ . Customers will have ex-ante types $\alpha \in \mathcal{A}$, but now in order to solve the problem, I will have to assume that the monopolist knows the population distribution of ex-ante types $G(\alpha)$. Here it will furthermore be necessary to impose more structure on the set of priors $\{F(\theta, \alpha)\}_{\alpha \in \mathcal{A}}$. While in the competitive case, the ex-ante incentive constraints were not binding, in any profit-maximizing solution of the monopolist, some ex-ante constraint will have to be binding. In order to know which these are, I will solve the problem making two alternative assumptions on the set of priors, Assumption 1 or alternatively Assumptions 2A and 2B.

The condition imposed under Assumption 1 is that high α types always associate a higher probability with having a high demand in the second period than low α types. Formally for every $\alpha' > \alpha$, the distribution $F(\theta, \alpha')$ first order stochastically dominates $F(\theta, \alpha)$.

Assumption 1. $F_\alpha(\theta, \alpha) \leq 0$ for all θ .

Under the conditions imposed by Assumption 2, the high α types will always be less confident to accurately predict their demand in the second period than the low α type. We will formalize this idea by imposing that for any $\alpha' > \alpha$, the distribution $F(\theta, \alpha')$ is a mean preserving spread of $F(\theta, \alpha)$.

Assumption 3 will only be used to derive sufficient conditions for an optimum. It implies that for all $\alpha' > \alpha$, $F(\theta, \alpha) < F(\theta, \alpha')$ if $\theta < z$, and $F(\theta, \alpha) > F(\theta, \alpha')$ if $\theta > z$.

Assumption 2. $\int_{\theta_L}^{\theta} F_{\alpha}(x, \alpha) dx \geq 0$ for all θ and $\int_{\theta_L}^{\theta_H} F_{\alpha}(x, \alpha) dx = 0$.

Assumption 3. The distribution functions $F(\theta, \alpha)$ cross in one and only one point $\theta = z$ for all α .

Define $\Lambda^m(q, \theta, \alpha)$ as follows

$$\Lambda^m(q, \theta, \alpha) = u(q, \theta) + u_{\theta}(q, \theta) \left[\frac{F(\theta, \hat{\alpha}) - F(\theta, \alpha)}{f(\theta, \hat{\alpha})} \right] - C(q) + u_{\theta}(q, \theta) \frac{F_{\alpha}(\theta, \alpha)}{f(\theta, \hat{\alpha})} \left(\frac{1 - G(\alpha)}{g(\alpha)} \right) \quad (2.14)$$

The following propositions characterizes the monopolist's profit-maximizing menu design.

Proposition 3. Suppose the set of priors $\{F(\theta, \alpha)\}$ satisfies Assumption 1. Then the optimal quantity schedule $\{\{q(\theta, \alpha)\}_{\theta \in \Theta}\}_{\alpha \in \mathcal{A}}$ is given by

$$q(\theta, \alpha) = \arg \max_q \Lambda^m(q, \theta, \alpha)$$

if $q(\theta, \alpha)$ is (weakly) increasing in both θ and α .

Proof. see Appendix □

Proposition 4. Suppose the set of priors $\{F(\theta, \alpha)\}$ satisfies Assumptions 2 and 3. Then the optimal quantity schedule $\{\{q(\theta, \alpha)\}_{\theta \in \Theta}\}_{\alpha \in \mathcal{A}}$ is given by

$$q(\theta, \alpha) = \arg \max_q \Lambda^m(q, \theta, \alpha)$$

if $q(\theta, \alpha)$ is (weakly) increasing in θ , (weakly) decreasing in α on $\theta \leq z$, and (weakly) increasing in α on $\theta > z$.

Proof. see Appendix □

These results allow us to compare the monopoly outcome to the menu of tariffs supplied by firms in a duopoly with differentiated brands. We have seen that the optimal quantity schedules $\{q(\theta, \alpha)\}$ can be derived by pointwise maximization of Λ^m as defined by

(2.14) in the monopoly case, and of Λ^c defined by (2.6) in the duopoly case. These two functions differ in one term

$$u_{\theta}(q, \theta) \frac{F_{\alpha}(\theta, \alpha)}{f(\theta, \hat{\alpha})} \left(\frac{1 - G(\alpha)}{g(\alpha)} \right)$$

which is absent in the competitive setting. This term reflects the usual inefficiency a monopolist creates by distorting the quantity schedules of the low α types in order to extract more surplus from the high α types. For the highest α type this deviation from the efficient quantity schedule disappears. As in the standard models of competitive price discrimination, this inefficiency is eliminated by competition. Firms lose the ability to distort the quantity schedules of some ex-ante type in order to extract more surplus from another ex-ante type. As we have seen, the competitive solution is separable, ex-ante incentive constraints are not binding.

2.3.1 Common Priors but Different Consumption Profiles

Now, I will analyze a setting with competitive firms which face consumers with real differences in their consumption profiles, e.g. frequent v. infrequent users. It will be shown that in such a setting firms will not be able to use the design of their menu in order to screen customers with different consumption profiles.

Let us return to the previous duopoly setting, but suppose that $F(\theta, \alpha)$ describes the “true” consumption profile of type α . Again, there will be a heterogeneous population of α types belonging to some set \mathcal{A} , but firms and consumers agree that consumption profiles are actually different. That is if a firm knew the α type of its customer, the firm and this customer would agree that the customer’s consumption profile is given by $F(\theta, \alpha)$. However, as before, α will be private information of the consumers and if firms want to discriminate between consumers, they will have to do so in an incentive compatible way.

Consider a firm that provides a utility level u (net of travel costs) to a customer

of type α . The solution to this problem is given by

$$\begin{aligned} \pi^A(u, \alpha) = \max_{P^A, \{q(\theta, \alpha)\}} \int_{\theta_L}^{\theta_H} \left(P^A(q(\theta, \alpha), \alpha) - C(q(\theta, \alpha)) \right) f(\theta, \alpha) d\theta \\ \text{subject to } \int_{\theta_L}^{\theta_H} s^A(\theta, \alpha) f(\theta, \alpha) d\theta \geq u \end{aligned}$$

Notice that the only difference to the situation with non-common priors is that the firm's expectation is taken over the same distribution as the consumer's evaluation of the deal. Carrying out the same steps as before we find that the firm's profits as a function of the type α and the provided utility level u are

$$\pi(u, \alpha) = \max_{\{q(\theta, \alpha)\}} \int_{\theta_L}^{\theta_H} [u(q(\theta, \alpha), \theta) - C(q(\theta, \alpha))] f(\theta, \alpha) d\theta - u \quad (2.15)$$

Now obviously the optimal quantity schedule is the efficient one, that is where marginal utility equals marginal cost, $u_q(q, \theta) = C_q(q)$. Clearly, the optimal quantity schedule does not depend on α . The firm provides the same quantity schedule $\{q(\theta)\}$ to all its customers irrespective of their type α . Furthermore, the single crossing property $u_{q\theta} > 0$ ensures that $q(\theta)$ is increasing in θ and thus implementable.

The above solution is obvious. The profit-maximizing deal maximizes total surplus and extracts any surplus that the firm wants to extract through a fixed fee ex-ante.

Let us now analyze competition by firms over (unobservable) types $\alpha \in \mathcal{A}$. The following Proposition follows the line of argument of Proposition 5 in Armstrong and Vickers [Armstrong and Vickers, 2001], and establishes that firms will only offer one contract. This contract will be a fixed fee plus cost contract. Thus firms will not be able to screen customers by the α type.

Proposition 5. *Suppose $\tau \leq (2/3)S(\alpha)$ for all α present in the market. Then it is an equilibrium for both firms to offer a single contract with a fixed fee equal to τ and usage charge $P(q) = C(q)$.*

Proof. Suppose firm B offered this contract. Assume firm A could actually observe α , which gives us an upper bound on its profits. Then firm A 's market share, given that it

offers a surplus of S (net of travel costs) to its customers, is

$$x = \min \left\{ \frac{1}{2} \left(1 + \frac{S - S(\alpha) + \tau}{\tau} \right), \frac{S}{t} \right\}$$

Obviously firm A will provide the surplus level S through the profit maximizing quantity schedule. Customers' surplus from buying at firm A net of travel costs will then be $S(\alpha) - k$, where k is the fixed fee firm A will charge in addition to the price schedule $P(q)$.

Under the assumption that $\tau \leq (2/3)S(\alpha)$ the market is fully covered and profits for firm A when charging a fixed fee k are

$$k \frac{1}{2} \left(1 + \frac{\tau - k}{\tau} \right)$$

The fixed fee that maximizes these profits is given by $k = \tau$ which proves the claim. \square

Thus, firms in this setting will provide all customers with the efficient quantity schedule irrespective of their consumption profile. If we compare this outcome to the outcome when firms compete over consumers with mistaken priors, we see that in this latter settings firms do not maximize total "real" surplus when designing the quantity schedule for type α , but total real and *fictitious surplus*. *Fictitious surplus* appears in the firm's design problem through an additional term in Λ^c

$$u_{\theta}(q, \theta) \left[\frac{F(\theta, \hat{\alpha}) - F(\theta, \alpha)}{f(\theta, \hat{\alpha})} \right]$$

This term reflects the disagreement between firms and customers concerning the surplus provided by any given quantity schedule. When designing the optimal contract, firms do so by providing customers with "what customers want" in a cost minimizing way. If customers misperceive the surplus implied by a deal, even competitive firms will not be able to correct this misperception. Firms will compete over customers by delivering what customers want, not by delivering what firms think is best for them.

However firms will not be able to raise their profits by dealing with mistaken customers. Firms make a profit which is equal to τ on each customer irrespective of their type α .

All additional profits that could be made by creating *fictitious surplus* are competed away. Firms do not have a comparative advantage in exploiting mistaken priors. They only have a comparative advantage in terms of brand preferences of customers. This is why their profit levels are linked to travel costs, i.e. the intensity of brand preferences. Finally, note that because the competitive problem with mistaken customers is separable in the ex-ante types, the fully rational customer $\hat{\alpha}$ gets the efficient quantity schedule. He does not suffer or enjoy any externalities from the presence of mistaken customers in the market.

2.4 Conclusion

This paper has shown that menus of contingent contracts by competitive firms can be understood as screening devices for customers whose beliefs differ from firms' beliefs. In a setting where consumers' tastes are private information but firms and customers agree on the value implied by any offered contract for a given value of the private information, firms facing competition will not be able to screen customers with respect to their tastes. Here however, firms and customers can disagree about the implied value of any given deal and this disagreement allows firms to screen customers through menu design even if the exact reason for disagreement is private information of the customer. It turns out that competition prevents firms from exploiting customers with mistaken beliefs through contract design. The profits a firm makes on a customer are independent of the beliefs this customer holds. Also flawed evaluation of deals by some customers does not influence the kind of deals offered to other customers. The presence of boundedly rational customers in the market does not exert an externality on other customers as long as the market is sufficiently competitive.

Empirically, these results are meaningful if there are a fixed number of psychological types present in the market. In this case this model would predict that all firms offer the same number of contracts, which should all be identical, and each tailored towards a specific psychological type. If this is the case, one could try to recover these types from the shape of contracts of firms and consumption data.⁵ Because of the separability

⁵Miravete [Miravete, 2004] shows how to recover the distribution of types in a standard non-linear pricing problem from the shape of tariffs. In our context the shape of tariffs alone would not be enough.

result, each contract schedule should only contain information on one specific type and hence such an exercise is possible.

One would also need information on individual consumption to recover the true distribution of ex-post types.

Chapter 3

Financial Transaction Taxes and the Informational Efficiency of Financial Markets: A Structural Estimation

3.1 Introduction

In August 2012 France introduced a financial transaction tax. Anyone buying shares of companies with market value exceeding €1 billion incurs a 0.2 percent tax on the value of the transaction. A similar tax was introduced in Italy in March 2013. A proposal by the European Council for the introduction of a 0.1 percent *ad valorem* transaction tax for the entire European Union by 2014 (later postponed to 2016) has been favoured by countries like Austria, France, Germany, Italy and Spain, and opposed by others, like Sweden and the UK. Proposals to introduce the tax have also been advanced in many other countries, especially after the 2007-8 financial crisis.

Of course, transaction taxes are not a recent discovery. They have been implemented in several countries, like Sweden, US and UK during the past decades. The resurgence of a vibrant debate on the optimality of transaction taxes, triggered by the recent financial crisis, has seen many proposers of the tax blaming financial markets, and in particular, financial speculation, for systemic risk. In this view, the tax would reduce excessive trading and volatility, and perhaps lower the probability of further crises. Opponents of

the tax, instead, argue that it would only lower market efficiency.

In this paper we propose a novel methodology to study the impact of a transaction tax on informational efficiency, liquidity and volatility. In a market microstructure framework, we analyze the process of price formation. Informed traders interact with noise traders, who provide liquidity to the market. Their trading activity reveals information that may be aggregated by the price. The question we ask is how the introduction of a tax affects market participation and decisions by these two classes of traders. It is clear that if the tax discouraged all informed traders from participating in the market, the price would stop aggregating information. On the other hand, if it only discouraged the participation of traders with very imprecise private information, the market could actually become more efficient. The way in which the tax affects the decisions of noise traders can also have ambivalent effects. Less noise traders in the market may make the trading flow more informative; on the other hand, lower liquidity may also make trading more costly for informed traders, and lower informational efficiency.

Similar considerations hold for price volatility. While one would like to avoid excess price volatility, we know that prices may be volatile because of informed traders' activity. Whether the tax reduces "good" or "bad" volatility again depends on the effect that it has on informed and noise traders' decisions and on their interaction.

The difficulty of identifying the effect of the taxes on different categories of traders has been one of the major problems in evaluating the impact of transaction taxes (see, e.g., Habermeier and Kirilenko [Habermeier and Kirilenko, 2001]). To identify these effects, we build a model with informed traders (who receive a private signal on the fundamental asset value) and noise traders (both price inelastic and price elastic). Through maximum likelihood, we estimate the proportion of these three groups of traders, as well as the precision of the private signals received by informed traders. Once we have obtained our estimates, we use them to simulate the model in the presence of a transaction tax. We compare trading decisions, bid and ask spreads and price levels in

the estimated and the simulated model, thus being able to disentangle and to quantify the effects of the tax.

Since the main aim of this paper is methodological, we estimate our model using high frequency data for only one stock traded in the NYSE. We find that in our asset market a transaction tax would mostly affect traders who trade on private information. Traders who trade for liquidity reasons are price inelastic and therefore, do not change their demand in response to a transaction tax. A transaction tax would only crowd out some informed trading and thus reduce the informational content of the order flow. In our asset market even moderate taxes in the suggested range for a financial transaction tax, namely 0.5 percent on the value of the transaction, would completely crowd out informed trade and thus destroy the market's ability to aggregate information. A 0.2 percent tax on transactions as implemented in France would lead to a 10 percent drop in volume in our asset market. We find an ambiguous effect of the tax on spreads as on the one hand the reduction of informed trading leads to narrower spreads (market makers need less protection against traders with superior information), but, on the other hand, private information is revealed less quickly and asymmetric information in the market persists for longer than in the absence of a tax.

There is a large academic debate on the transaction tax. A paper that studies the impact of the tax on informational efficiency and volatility in a market microstructure framework is the recent paper by Rosenthal and Thomas [Rosenthal and Thomas, 2012]. They consider a model in which traders strategically submit limit and market orders. Their analysis shows that the transaction tax widens the spread, increases volatility and lowers gains from trade. Dupont and Lee [Dupont and Lee, 2007] consider a quote driven market, building on Glosten and Milgrom [Glosten and Milgrom, 1985]. For tractability, however, they limit the analysis to a static, one period model. They show that the effect of the tax on the bid-ask spread depends on the level of asymmetric information in the market. Other papers related to ours are Mannaro et al. [Mannaro et al., 2008] and Pellizzari and Westerhoff [Pellizzari and Westerhoff, 2009].

Their analysis is very different from ours, in that they use heterogeneous agents models. In common with us, they are interested in studying the effect of the transaction tax on different groups of traders (who use technical and fundamental trading rules). They both find that transaction taxes may hamper market liquidity, although less so in the presence of a specialist. The effect of the transaction tax on price volatility has also been studied in a general equilibrium competitive economy by Song and Zhang [Song and Zhang, 2005]. In common with our work, they are also interested in the response of different classes of traders to the tax. They show that a transaction tax may lower volatility, by discouraging noise traders to trade, but can also have the opposite effect, since in a less liquid market the price impact of a trade can be higher. The result that volatility may increase with a tax was also shown in earlier theoretical work by Kupiec [Kupiec, 1996].

The rest of the paper is organized as follows. Section 2 and 3 describe the theoretical model without tax. Section 4 introduces a financial transaction tax into the theoretical model. Section 5 then presents the likelihood function and section 6 describes the data used to estimate the structural model through maximum likelihood. Section 7 presents estimates for the parameters of our structural model. Section 8 reports simulations quantifying the impact of a financial transaction tax on our asset market. Finally section 9 concludes. An appendix contains further estimation results and other supplementary material.

3.2 The Model

Building on Easley et al. [Easley et al., 1997] we consider an asset that is traded by a sequence of traders who interact with a market maker. Trading days are indexed by $d = 1, 2, 3, \dots$. Time within each day is discrete and indexed by $t = 1, 2, 3, \dots$

The asset

We denote the fundamental value of the asset on day d by V^d . The asset value does not change during the day, but can change from one day to the next. At the beginning of the day, with probability $1 - \alpha$ the asset value remains the same as on the previous

day ($V^d = v^{d-1}$), and with probability α it changes. On each day d , the value of the asset on the previous day $d - 1$, v^{d-1} , is known to all market participants.¹ As we will explain, when the value of the asset changes from one day to the other, there are informed traders in the market; for this reason, we say that an information event has occurred. If an information event occurs, with probability $1 - \delta$ the asset value decreases to $v^{d-1} - \lambda_L$ (bad informational event), and with probability δ it increases to $v^{d-1} + \lambda_H$ (good informational event), where $\lambda_L > 0$ and $\lambda_H > 0$. Informational events are independently distributed over the days of trading. To simplify the notation, we define $v_H^d := v^{d-1} + \lambda_H$ and $v_L^d := v^{d-1} - \lambda_L$. Finally, we assume that $(1 - \delta)\lambda_L = \delta\lambda_H$, which, as will become clear later, guarantees that the closing price is a martingale.

The market

The asset is exchanged in a specialist market. Its price is set by a market maker who interacts with a sequence of traders. At any time $t = 1, 2, 3, \dots$ during the day a trader is randomly chosen to act and can buy, sell, or decide not to trade. Each trade consists of the exchange of one unit of the asset for cash. The trader's action space is, therefore, $\mathcal{A} = \{buy, sell, no\ trade\}$. We denote the action of the trader at time t on day d by X_t^d and the history of trades and prices until time $t - 1$ of day d by h_t^d .

The market maker

At any time t of day d , the market maker sets the prices at which a trader can buy or sell the asset. When posting these prices, the market maker must take into account the possibility of trading with traders who (as we shall see) have some private information on the asset value. He will set different prices for buying and for selling, that is, there will be a bid-ask spread. We denote the ask price (the price at which a trader can buy) at time t by a_t^d and the bid price (the price at which a trader can sell) by b_t^d .

Due to (unmodeled) potential competition, the market maker makes zero expected profits. This implies that he sets the ask and bid prices equal to the expected value of

¹For more comments on this point, see footnote 18. Note that v^{d-1} is the realization of the random variable V^{d-1} . Throughout the text, we will denote random variables with capital letters and their realizations with lowercase letters.

the asset conditional on the information available at time t and on the chosen action. Moreover, due to potential competition, the ask is the smallest value satisfying this condition for a buy and the bid the highest value satisfying this condition for a sell, that is,

$$a_t^d = \min_a \left\{ a : a - \mathbb{E}(V | X_t^d = \text{buy}, a, h_t^d) = 0 \right\}$$

and

$$b_t^d = \max_b \left\{ b : \mathbb{E}(V | X_t^d = \text{sell}, b, h_t^d) - b = 0 \right\}.$$

We will sometimes refer to the market maker's expectation conditional on the history of trades only as the "price" of the asset, and we will denote it by $p_t^d = \mathbb{E}(V^d | h_t^d)$.²

The traders

There are a countable number of traders. Traders act in an exogenous sequential order. Each trader is chosen to take an action only once, at time t of day d . Traders are of two types, informed and noise. The trader's own type is private information.

On no event days, all traders in the market are noise traders. On information event days, at any time t an informed trader is chosen to trade with probability μ and a noise trader with probability $1 - \mu$, with $\mu \in (0, 1)$.³

Informed traders have private information about the asset value. They receive a private signal on the new asset value and observe the previous history of trades and prices, and the current prices.⁴

²Standard arguments show that $b_t^d \leq p_t^d \leq a_t^d$ (see Glosten and Milgrom [Glosten and Milgrom, 1985]).

³In other words, μ should be interpreted as the proportion of informed-based trading decisions in a day (and not as the proportion of informed traders in the population). Of course, in a no-event day, the proportion of informed-based trading decisions is zero.

⁴As we will explain later, in the model there is a one-to-one mapping from trades to prices. For this reason, in bringing the model to the data, we only need to assume that traders observe the history of past prices.

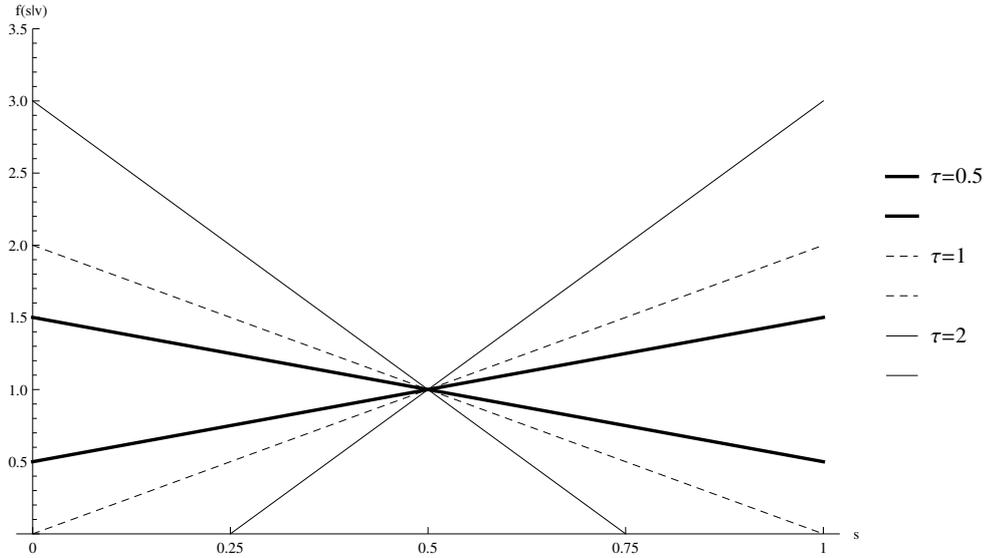


Figure 3.1: Density of Signal. The figure shows the signal's value-contingent density functions for different values τ

The private signal S_t^d has the following value-contingent densities:

$$f^H(s_t^d | V^d = v_H^d) = 1 + \tau(2s_t^d - 1),$$

$$f^L(s_t^d | V^d = v_L^d) = 1 - \tau(2s_t^d - 1),$$

with $\tau \in (0, \infty)$. (See Figure 3.1).

For $\tau \in (0, 1]$, the support of the densities is $[0, 1]$. In contrast, for $\tau > 1$, the support shrinks to $[\frac{\tau-1}{2\tau}, \frac{\tau-1+2\sqrt{\tau}}{2\tau}]$ for $f^H(\cdot | v_H^d)$ and to $[\frac{\tau+1-2\sqrt{\tau}}{2\tau}, \frac{\tau+1}{2\tau}]$ for $f^L(\cdot | v_L^d)$ (in order for the density functions to integrate to one). Note that, given the value of the asset, the signals S_t^d are i.i.d.⁵ The signals satisfy the monotone likelihood ratio property. At each time t , the likelihood ratio after receiving the signal, $\frac{\mathbb{P}(V^d = v_H^d | h_t^d, s_t^d)}{\mathbb{P}(V^d = v_L^d | h_t^d, s_t^d)} = \frac{1 + \tau(2s_t^d - 1)}{1 - \tau(2s_t^d - 1)} \frac{\mathbb{P}(V^d = v_H^d | h_t^d)}{\mathbb{P}(V^d = v_L^d | h_t^d)}$, is higher than that before receiving the signal if $s_t^d > 0.5$ and lower if $s_t^d < 0.5$. For this reason we refer to a signal larger than 0.5 as a “good signal” and to a signal smaller than 0.5 as a “bad signal.”

The parameter τ measures the informativeness of the signals. When $\tau \rightarrow 0$, the den-

⁵Conditional i.i.d. signals deliver a likelihood function that can be easily brought to the data. Nevertheless, the result that herding arises in equilibrium would also hold in an economy in which signals are not conditionally independent.

sities are uniform and the signals are completely uninformative. As τ increases, the signals become more and more informative. For any given $\tau \in (0, 1)$, the support of the distribution of the likelihood ratio is bounded away from 0 and infinity, while for $\tau \geq 1$ it is not. Following Smith and Sørensen [Smith and Sorensen, 2000], in the first case we say that beliefs are bounded and in the second case that they are unbounded. With bounded beliefs, no signal realizations (even the most extreme ones) reveal the asset value with probability one. With unbounded beliefs, in contrast, some high (low) signal realizations are possible only when the asset value is high (low), and therefore the signal can be perfectly informative.⁶ As τ tends to infinity, the measure of perfectly informative signals tends to one.

An informed trader knows that an information event has occurred and that, as a result, the asset value has changed with respect to the previous day. Moreover, his signal is informative on whether the event is good or bad. Nevertheless, according to the signal realization that he receives and the precision τ , he may not be completely sure of the effect of the event on the asset value. For instance, he may know that there has been a change in the investment strategy of a company, but not be sure whether this change will affect the asset value in a positive or negative way. The parameter τ can be interpreted as measuring the precision of the information that the trader receives, or the ability of the trader to process such private information. Finally, note that, given our signal structure, informed traders are heterogenous, since they receive signal realizations with different degrees of informativeness about the asset's fundamental value.

In addition to capturing heterogeneity of information in the market, a linear density function for the signal makes it possible to compute the traders' strategies and the market maker's posted prices analytically. As a result, we obtain a simple and tractable likelihood function. Moreover, in contrast to other specifications such as a discrete signal (e.g., a noisy binary signal), our choice avoids creating a discontinuity in the

⁶In particular, any signal greater than or equal to $\frac{\tau+1}{2\tau}$ reveals that the asset value is v_d^H , whereas a signal lower than or equal to $\frac{\tau-1}{2\tau}$ reveals that the asset value is v_d^L .

likelihood function, which would make estimation problematic.

An informed trader's payoff function, $U : \{v_L^d, v_H^d\} \times \mathcal{A} \times [v_L^d, v_H^d]^2 \rightarrow \mathbf{R}^+$, is defined as

$$U(v^d, X_t^d, a_t^d, b_t^d) = \begin{cases} v^d - a_t^d & \text{if } X_t^d = \text{buy} \\ 0 & \text{if } X_t^d = \text{no trade} \\ b_t^d - v^d & \text{if } X_t^d = \text{sell} \end{cases}$$

An informed trader chooses X_t^d to maximize $\mathbb{E}[U(V^d, X_t^d, a_t^d, b_t^d) | h_t^d, s_t^d]$ (i.e., he is risk neutral). Therefore, he finds it optimal to buy whenever $\mathbb{E}(V^d | h_t^d, s_t^d) > a_t^d$ and to sell whenever $\mathbb{E}(V^d | h_t^d, s_t^d) < b_t^d$. He chooses not to trade when $b_t^d < \mathbb{E}(V^d | h_t^d, s_t^d) < a_t^d$. Otherwise, he is indifferent between buying and not trading, or selling and not trading.

Noise traders trade for unmodeled reasons, such as liquidity or hedging needs. A fraction $1 - \psi$ of them are price inelastic: they buy with probability $\frac{\varepsilon}{2}$, sell with probability $\frac{\varepsilon}{2}$ and do not trade with probability $1 - \varepsilon$ (with $0 < \varepsilon < 1$). In other words, if a price inelastic noise trader is drawn to trade at time t , and he is willing to buy (which will occur with probability $\frac{\varepsilon}{2}$), he will do so independently of the posted quotes. The remaining fraction ψ is, instead, price elastic: their decision to buy, sell or not to trade does depend on the ask and bid prices. Specifically, each of these noise traders receives a shock N_t^d , with shocks distributed independently across time according to the uniform distribution on the interval $[0, 1]$. Upon receiving a shock n_t^d , a price elastic noise trader's valuation of the asset at time t of day d is $\mathbb{E}(V^d | h_t^d, n_t^d)$. This expectation is computed *as if* the shock (which is actually distributed independently of the asset value) were distributed on $[0, 1]$ according to the following value-contingent densities:

$$\begin{aligned} g^H(n_t^d | V^d = v_H^d) &= 2n_t^d, \\ g(s_t^d | V^d = v^{d-1}) &= 1, \\ g^L(s_t^d | V^d = v_L^d) &= 2(1 - n_t^d). \end{aligned}$$

Essentially, noise traders can be thought of as traders who, in addition to the common value V^d , have a private value from holding the asset. Whereas price inelastic noise traders have a private value so large that they are forced to buy or sell independently of the price (e.g., because they receive a liquidity shock), price elastic noise traders have smaller private values that make them inclined to buy, to sell or not to trade. The specific modeling of the price elastic noise traders that we have chosen has some useful features. First, it guarantees that the noise trader's valuation of the asset is always (strictly) between v_L^d and v_H^d , so that at any time t these noise traders' demand is, indeed, always price elastic.⁷ Second, our choice has the convenient property that by modeling both informed and noise traders in a similar fashion we are able to solve our model analytically. Finally, note that while the shock affects the trader's asset valuation in a time-varying fashion, the expectation at time t of the trader's valuation at time $t + 1$ is $\mathbb{E}(V_a^d | h_t^d)$, that is, the noise trader's valuation respects the martingale property. This way of modelling noise traders was first proposed by Glosten [Glosten, 2009]. One interpretation he offers is that noise traders are boundedly rational agents wrongly attributing information to uninformative signals.

3.3 Equilibrium Analysis

At each time t , the trading decision of an informed trader can simply be characterized by two thresholds, σ_t^d and β_t^d , satisfying the equalities

$$\mathbb{E} \left[V^d | h_t^d, \sigma_t^d \right] = b_t^d \quad (3.1)$$

and

$$\mathbb{E} \left[V^d | h_t^d, \beta_t^d \right] = a_t^d. \quad (3.2)$$

An informed trader will sell for any signal realization smaller than σ_t^d and buy for any signal realization greater than β_t^d . Obviously, the thresholds at each time t depend on

⁷This would not be the case if, for instance, we had modeled the noise traders as valuing the asset $\mathbb{E}(V^d | h_t^d) + \rho_t$ or $\rho_t \mathbb{E}(V^d | h_t^d)$, where ρ_t would represent an i.i.d. additive or multiplicative private component. In such cases, even if we chose the support of ρ_t such that at time 1 of each day the noise traders' demand would be price elastic, still, it would become price inelastic as $\mathbb{E}(V^d | h_t^d)$ converges to the extremes of the support.

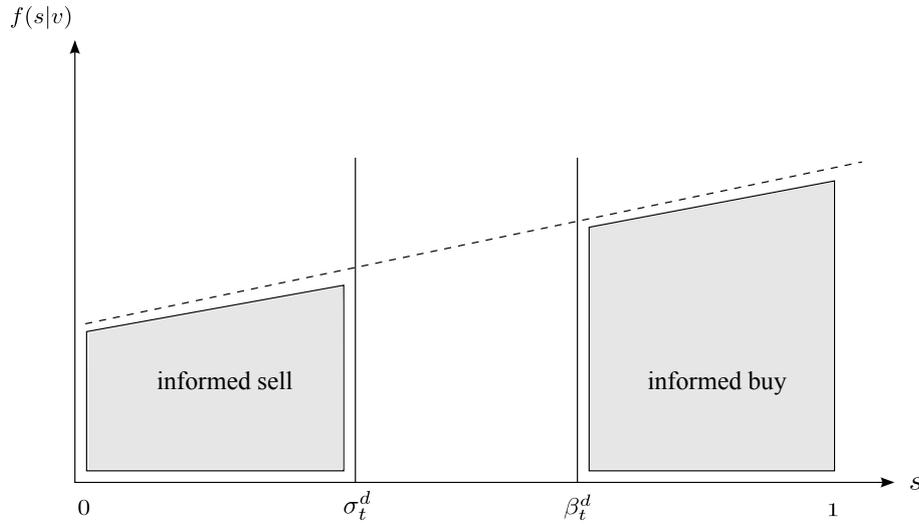


Figure 3.2: Informed trader's decision. The figure illustrates the signal realizations for which an informed trader decides to buy or sell when $V^d = v_h^d$.

the history of trades until that time and on the parameter values.⁸

Figure 3.2 (drawn for the case of a good informational event) illustrates the decision of informed traders. An informed trader buys the asset with a signal higher than the threshold value β_t^d , sells it with a signal lower than σ_t^d , and does not trade otherwise. The measure of informed traders buying or selling is equal to the areas (labeled as “informed buy” and “informed sell”) below the line representing the signal density function.

A similar analysis holds for price elastic noise traders. At each trading time t on day d their behavior can be characterized by a buy threshold κ_t^d and a sell threshold γ_t^d . The buy threshold κ_t^d for a noise trader is implicitly given by

$$\mathbb{E} \left(V^d | h_t^d, \kappa_t^d \right) = a_t^d \quad (3.3)$$

⁸Since noise traders buy and sell with probabilities bounded away from zero, standard arguments prove that the bid and ask prices exist and are unique (see, e.g., Cipriani and Guarino [Cipriani and Guarino, 2008]). Similar arguments can be used to prove existence and uniqueness of the thresholds (see Cipriani and Guarino [Cipriani and Guarino, 2014]).

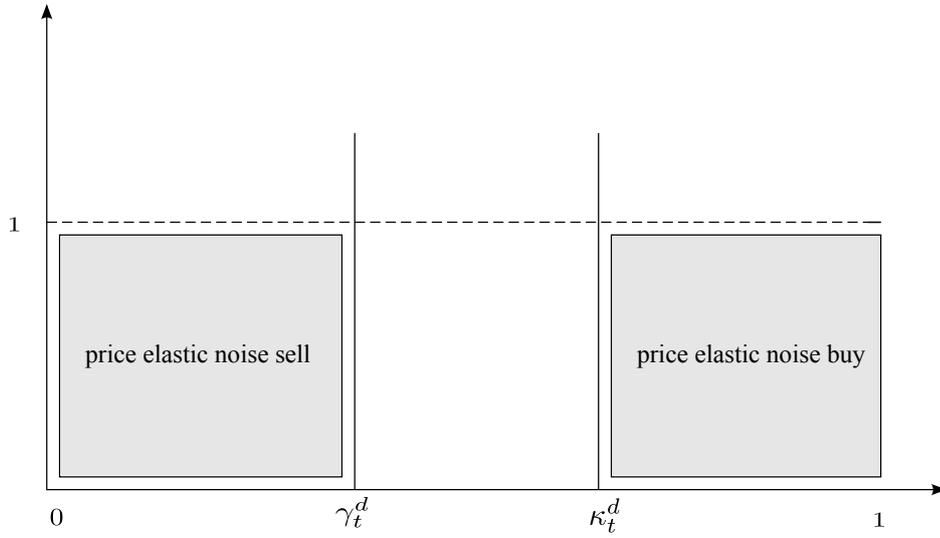


Figure 3.3: Price elastic noise trader's decision. The figure illustrates the shock realization for which a price elastic noise trader decides to buy or sell.

and the sell threshold γ_t^d by

$$\mathbb{E}\left(V^d | h_t^d, \gamma_t^d\right) = b_t^d. \quad (3.4)$$

Figure 3.3 illustrates the resulting trading probabilities for price elastic noise traders. Notice that since the private value shocks N_t^d are drawn from a uniform distribution on $[0, 1]$ irrespective of the type of event day, these probabilities only depend on the thresholds γ_t^d and κ_t^d but not on the realization of V^d unlike in the case of the informed traders.

Let us now derive the equilibrium buy thresholds. Obviously, (3.2) and (3.3) together imply that

$$\mathbb{E}\left(V^d | h_t^d, \kappa_t^d\right) = E\left(V^d | h_t^d, \beta_t^d\right), \quad (3.5)$$

from which one can derive a linear relationship between the two thresholds,

$$\kappa_t^d = c_0(h_t^d) + c_1(h_t^d) \beta_t^d, \quad (3.6)$$

where $c_0(h_t^d)$ and $c_1(h_t^d)$ are functions of the history h_t^d only.

As we know, the ask price has to yield zero expected profits to the market maker, given his beliefs upon observing a buy order. It is thus given by the following expression:

$$a_t^d = v^{d-1} + \Pr\left(v_H^d | \text{buy}_t^d, h_t^d, a_t^d, b_t^d\right) \lambda_H - \Pr\left(v_L^d | \text{buy}_t^d, h_t^d, a_t^d, b_t^d\right) \lambda_L, \quad (3.7)$$

where the probabilities of a good and bad event will be functions of the thresholds β_t^d and κ_t^d . By (3.6) we can express κ_t^d as a function of β_t^d .

Similarly, the expectation of an informed trader with signal β_t^d can be expressed as

$$\mathbb{E}\left(V^d | h_t^d, \beta_t^d\right) = v^{d-1} + \Pr\left(v_H^d | h_t^d, \beta_t^d\right) \lambda_H - \Pr\left(v_L^d | h_t^d, \beta_t^d\right) \lambda_L. \quad (3.8)$$

After substituting (3.6) into (3.7) and then both (3.7) and (3.8) into (3.2), one obtains a quadratic equation in β_t^d . This solution corresponds to the smallest equilibrium ask price that is consistent with zero expected profits for the market maker. The threshold β_t^d is then the smallest solution of the quadratic equation for the buy threshold such that $\beta_t^d \in [0, 1]$. An analogous approach yields a quadratic equation for the equilibrium sell thresholds σ_t^d ; σ_t^d is the largest solution in the interval $[0, 1]$, that is, it corresponds to the largest equilibrium bid price that is consistent with zero expected profits.⁹ Once we have solved for β_t^d and σ_t^d we can obtain the buy and sell thresholds for the price elastic noise traders, κ_t^d and γ_t^d , which are functions of the informed traders' thresholds.

3.4 The Financial Transaction Tax

We now consider the introduction of a transaction tax into the market. Whenever a trader buys or sells the asset, he has to pay a tax ρ . Therefore, if he buys, he pays $a_t^d + \rho$, and if he sells, he receives $b_t^d - \rho$. As a result, an informed trader with signal s_t^d finds it optimal to buy when $\mathbb{E}(V^d | h_t^d, s_t^d) > a_t^d + \rho$, and to sell when $\mathbb{E}(V^d | h_t^d, s_t^d) < b_t^d - \rho$; he chooses not to trade when $b_t^d - \rho < \mathbb{E}(V^d | h_t^d, s_t^d) < a_t^d + \rho$,

⁹The nature of the solutions guarantees that the equilibrium sell threshold is smaller than the equilibrium buy threshold, that is, $\sigma_t^d \leq \beta_t^d$.

and is indifferent between trading and not when an equality holds. Similarly, a price elastic noise trader with private value shock n_t^d buys if $\mathbb{E}(V^d|h_t^d, n_t^d) > a_t^d + \rho$, sells if $\mathbb{E}(V^d|h_t^d, n_t^d) < b_t^d - \rho$, chooses not to trade if $b_t^d - \rho < \mathbb{E}(V^d|h_t^d, n_t^d) < a_t^d + \rho$ and is indifferent between two actions if an equality holds.

Even in the presence of a tax, the equilibrium can still be characterized in terms of buy and sell thresholds for informed traders and price elastic noise traders. Of course, such thresholds will differ from those of the case without tax, since the introduction of a transaction tax can make trading unprofitable for some informed and noise traders who would have traded without a tax. Specifically, the buy thresholds for informed and price elastic noise traders are now implicitly given by the following system of equations

$$\begin{aligned}\mathbb{E}(V^d|h_t^d, \beta_t^d) &= a_t^d + \rho, \\ \mathbb{E}(V^d|h_t^d, \kappa_t^d) &= a_t^d + \rho.\end{aligned}$$

Note that from this system one can immediately derive (3.5) and (3.6), as in the previous analysis. After equating $\mathbb{E}(V^d|h_t^d, \beta_t^d)$ to the ask price and after some manipulations, one obtains

$$\begin{aligned}\Pr(v_H^d|h_t^d, \beta_t^d) - \Pr(v_H^d|h_t^d, \text{buy}_t^d) - \frac{\rho}{\lambda_H} \\ = \left(\frac{\delta}{1-\delta}\right) \left[\Pr(v_L^d|h_t^d, \beta_t^d) - \Pr(v_L^d|h_t^d, \text{buy}_t^d)\right],\end{aligned}\quad (3.9)$$

which turns out to be a cubic equation in β_t^d ; its smallest root in $[0, 1]$ gives the equilibrium buy threshold for informed traders. From this and (3.6) it is immediate to obtain the equilibrium buy threshold for the price elastic noise trader. An identical analysis finds the equilibrium sell thresholds.

From this analysis we derive two conclusions: informed traders who receive less informative signals and price elastic noise traders who receive weaker shocks may

find it optimal to abstain from trade to avoid paying the tax; at the same time, the tax changes the equilibrium bid and ask prices, since the measure of informed and noise traders on both sides of the market may be different.

This model, as any Glosten and Milgrom [Glosten and Milgrom, 1985] type of model, is not ideal to derive analytical results about comparative statics. This is not an issue for us, since we are interested in quantifying the effects of the tax, rather than proposing a new theoretical analysis. Nevertheless, it is important to highlight the main forces at work. The transaction tax can affect the informational efficiency of the market in different ways. To see this, consider a simple example. Suppose in the absence of the tax at time 1 in equilibrium, informed traders would buy for a signal greater than 0.55, sell for a signal lower than 0.45, and would not trade otherwise. Furthermore, suppose that, instead, in the presence of the tax, in equilibrium, they would buy for a signal greater than 0.6, sell for a signal lower than 0.4, and would not trade otherwise. In this example, a signal of 0.57 would be (partially) revealed in the absence of the tax, but not in its presence. Indeed, without the tax, the informed trader receiving such a signal would buy and the market maker would update the prices up. With the tax, instead, the trader would not trade and the market maker would be unable to infer whether the informed trader potentially trading at that time received a good or bad signal. All this suggests that the tax may lower the informational efficiency. On the other hand, note that in the presence of the tax, the buy order may become more informative, in that informed traders buy only after receiving stronger signals (a signal higher than 0.6, in our example). One can find parameter values for which the first effect is dominant, and others for which the second effect is. Note also that in our example, because of the symmetry, a no trade is completely uninformative about the type of (good or bad) signal that an informed trader received. There may be cases in which, even in the presence of the tax, the no trade does reveal information about the signal, since it is chosen for a different measure of good and bad signals. In that case, the tax, although creating a disincentive for traders to participate in the market for some signals, does not prevent these signals to be partially learned in the market; they are revealed, although

obviously in a noisy way, by the no trade decisions. When we consider times greater than 1, there may be an additional effect. As we know (e.g. Cipriani and Guarino [Cipriani and Guarino, 2014]) informed traders may herd in the market. In periods of herding, informational efficiency is lowered. The transaction tax lowers the incentive of traders to herd, since they may find it optimal not to trade with a specific signal, rather than herding. This is an extra channel through which the transaction tax can be beneficial for informational efficiency.

All this analysis concerns informed traders. As we explained above, the transaction tax also affects noise traders, who can find it optimal not to trade if they have to pay a tax. This can increase the informativeness of trades, since they may now be more likely to come from informed traders. At the same time, the lower liquidity in the market can enlarge the bid-ask spread, which in turn can affect information aggregation negatively.

It is important to observe that this analysis holds for time t when the tax is introduced at that time. For future times, the effects of a tax on the informativeness of trades (and on the spread) are even more ambiguous. More informative trades at time t may mean a larger spread at that time. On the other hand, they also imply faster learning, which may mean a lower expected spread (and more informative trades) at future times.

Overall, this analysis shows that a transaction tax affects informational efficiency in various ways, and the net effect is, a priori, ambiguous. What is not ambiguous is the asymptotic effect of the tax. We know that in the absence of a tax, the price converges almost surely to the true asset value (see, e.g., Avery and Zemsky [Avery and Zemsky, 1998]). If informed traders have bounded beliefs, this convergence result does not hold when a tax is introduced. In the presence of a tax, asymptotically, as t goes to infinity, all informed traders will stop trading. When they do so, trading becomes completely uninformative on which of the two (good or bad) events has occurred. Eventually, the market maker will stop updating the price. In the social learning literature such a situation is referred to as an informational cascade, a concept

that we now define formally:

Definition 1. *An informational cascade arises at time t of day d when*

$$\Pr(V^d = v_H^d | h_t^d, X_t^d = x) = \Pr(V^d = v_H^d | h_t^d),$$

for all $x \in \{\text{buy}, \text{sell}, \text{no trade}\}$.

The next proposition states our result formally:

Proposition 6. *In equilibrium, as $t \rightarrow \infty$, on an event day an informational cascade arises almost surely if and only if beliefs are bounded (i.e., $\tau < 1$). In the cascade, all informed traders decide not to trade; that is, for almost all histories there exists a time T , such that, for any $t > T$,*

$$\Pr(X_t^d = \text{no trade} | s_t^d, h_t^d) = 1,$$

for any signal realization s_t^d .

Proof. See Cipriani and Guarino [Cipriani and Guarino, 2008], Proposition 2. □

Over time, as the price aggregates private information, the informational content of a given signal becomes relatively less important than that of the history of trades. After a long enough sequence of trades, the valuations of the traders (for any possible signal) become so close to the bid and ask prices that the expected gain from acting upon private information becomes smaller than the tax. At this point, an informed trader prefers not to trade independently of the signal, and an informational cascade arises.

A cascade only happens when the informed traders' belief that $V_d = v_d^H$ has reached . The following proposition makes this statement more precise by deriving limits for beliefs $q_t^d = \Pr(V^d = v_h^d | V^d \neq v^{d-1}, h_t^d)$ such that as long as those limits are not reached, informed traders will continue to trade and private information will continue to be aggregated by prices:

Proposition 7. *If $\rho < \left(\frac{2\tau\delta}{1-\tau|2\delta-1|}\right)\lambda_h$, on an event day an informational cascade occurs only if $\Pr(V^d = v_h^d|h_t^d)$ has reached either an lower threshold r^* or an upper threshold $1 - r^*$, where r^* is the smallest root of the quadratic equation*

$$\frac{2(1-r)p\tau}{1-(2r-1)\tau} = (1-\delta)\frac{\rho}{\lambda_h}.$$

The root r is increasing in the transaction tax ρ and decreasing in τ , δ , and λ_h .

If $\rho > \left(\frac{2\tau\delta}{1-\tau|2\delta-1|}\right)\lambda_h$, on an event day, informed traders never trade for all t . The probability of a good event converges almost surely to $\Pr(V^d = v_h^d|h_t^d) = \delta$.

Proof. See Appendix. □

Of course, in our dataset the number of trades is finite, therefore there is no guarantee that a cascade occurs on a given day. We will come back to this point when studying the effect of the transaction tax on informational efficiency. To reach this aim, we need first to estimate the structural parameters of our model. To do so we now derive the likelihood function of the model.

3.5 The Likelihood Function

Since we use data for stocks that could be traded without paying a tax, we aim to estimate the parameters of the model without a financial transaction tax. We have characterized equilibrium behavior of market participants for all possible histories of trades. Therefore, we can write the likelihood function for the history of trades only, disregarding bid and ask prices.¹⁰

Let us denote the history of trades at the end of a trading day by $h^d := h_{T_d}^d$, where

¹⁰The likelihood function is a function of trades only and does not depend on bid and ask prices. Bid and ask prices in our model are uniquely determined by the trade sequence and thus do not contain any additional information once we condition on the order flow. This would not be true if in addition to private information, public information was revealed during a trading day. Our likelihood function however would still be valid, as the probability of trade by any type of trader and after any given history of trades would not be affected by the arrival of public information given our specification of informed and noise trader demand.

T_d is the number of trading times on day d . We denote the likelihood function by

$$\mathcal{L}(\Phi; \{h^d\}_{d=1}^D) = \Pr\left(\{h^d\}_{d=1}^D | \Phi\right)$$

where $\Phi := \{\alpha, \delta, \mu, \tau, \psi, \varepsilon\}$ is the vector of parameters.

Next recall that on day d all market participants know v^{d-1} and the occurrence of information events is independent across days. Thus the sequence of trades on day d only depends on the realization of V^d and not on any trading data from days other than d . We can therefore write the likelihood function as the product of the likelihoods of daily trading sequences

$$\mathcal{L}(\Phi; \{h^d\}_{d=1}^D) = \Pr\left(\{h^d\}_{d=1}^D | \Phi\right) = \prod_{d=1}^D \Pr(h^d | \Phi).$$

Now consider the likelihood of a sequence of trades for a given day. Unlike in the standard market microstructure model of Easley et al. [Easley et al., 1997] where only the total number of buys and sells matter for the probability of a given history of trades, in our model the sequence of trades is important. Informed and price elastic traders update their valuations depending on the trading sequence and thus their probability of trading depends on the observed history of trades up to the time in which they act. We have, therefore, to characterize the likelihood function for the history of trades on day d starting at time 1 and then working in a recursive way up to time T_d . At trading time t the probability of a given action x_t^d depends on the sequence of previous trades h_t^d and we have that

$$\Pr\left(h_{t+1}^d | \Phi\right) = \prod_{s=1}^t \Pr\left(x_s^d | h_s^d, \Phi\right) = \Pr(x_t^d | h_t^d, \Phi) \Pr\left(h_t^d | \Phi\right). \quad (3.10)$$

To compute $\Pr(x_t^d | h_t^d, \Phi)$, we express it in terms of the value-contingent trading probabilities

$$\begin{aligned} \Pr(x_t^d | h_t^d, \Phi) &= \Pr(x_t^d | h_t^d, v_H^d, \Phi) \Pr(v_H^d | h_t^d, \Phi) + \\ &\Pr(x_t^d | h_t^d, v_L^d, \Phi) \Pr(v_L^d | h_t^d, \Phi) + \Pr(x_t^d | h_t^d, v^{d-1}, \Phi) \Pr(v^{d-1} | h_t^d, \Phi). \end{aligned}$$

We now illustrate how to compute the value-contingent probabilities of a trade x_t^d . Consider period $t = 1$ and suppose, for instance, that there has been a buy order, $x_1^d = \text{buy}$. The probability of such an order for a given asset value depends on the buy thresholds for informed and price elastic noise traders, β_1^d and κ_1^d , which are functions of the market maker's prior beliefs α and δ and of the other parameters. Having obtained the value-contingent probabilities of a buy order in period 1, we can then update the market makers' beliefs concerning the asset value V^d using Bayes' rule. Hence, consider period t and suppose again that x_t^d is a buy order. The equilibrium buy thresholds, β_t^d and κ_t^d , will be functions of the market maker's beliefs given the trading history up to (but not including) period t as well as the parameters of the model.

Once we have solved for β_t^d and κ_t^d , we can compute the probability of a buy order on a good event day as follows:

$$\begin{aligned} \Pr(x_t^d = \text{buy} | h_t^d, v_H^d, \Phi) &= \\ \mu \left[1 - F^H(\beta_t^d | v_H^d) \right] &+ (1 - \mu) \left[\psi(1 - \kappa_t^d) + (1 - \psi) \left(\frac{\varepsilon}{2} \right) \right], \end{aligned}$$

where $F^H(\cdot | v_H^d)$ is the cumulative distribution function of $f^H(\cdot | v_H^d)$. Recall that a trader active at time t is an informed trader with probability μ and a noise trader with probability $1 - \mu$. An informed trader buys if his signal is above the buy threshold β_t^d which happens with probability $1 - F^H(\beta_t^d | v_H^d)$. With probability $1 - \psi$ a noise trader is price inelastic, in which case he buys with probability $\varepsilon/2$. A price elastic noise trader, in turn, buys if his shock is larger than κ_t^d , which happens with probability $1 - \kappa_t^d$ as these shocks are uniformly distributed.

Similarly, on a bad event day we have that

$$\Pr(x_t^d = \text{buy} | h_t^d, v_L^d, \Phi) = \mu \left[1 - F^L(\beta_t^d | v_L^d) \right] + (1 - \mu) \left[\psi (1 - \kappa_t^d) + (1 - \psi) \left(\frac{\varepsilon}{2} \right) \right].$$

Finally, in a no event day ($V^d = v^{d-1}$) a market order can only come from a noise trader with either elastic demand (a fraction ψ of traders) or inelastic demand ($1 - \psi$).

Therefore, for a buy order we have that

$$\Pr(x_t^d = \text{buy} | h_t^d, v^{d-1}, \Phi) = \psi (1 - \kappa_t^d) + (1 - \psi) \frac{\varepsilon}{2}.$$

By following an analogous procedure we obtain the value-contingent probabilities for a sell order at t by computing σ_t^d and γ_t^d . The probability of a sell order on a good event day, for instance, is

$$\Pr(x_t^d = \text{sell} | h_t^d, v_H^d, \Phi) = \mu F^H(\sigma_t^d | v_H^d) + (1 - \mu) \left[\psi \gamma_t^d + (1 - \psi) \left(\frac{\varepsilon}{2} \right) \right];$$

the probability of a no-trade is just the complement to the probabilities of a buy and of a sell.

Finally, to compute $\Pr(x_t^d | h_t^d, \Phi)$, we need the conditional probabilities of V^d given the history until time t , that is, $\Pr(V^d = v | h_t^d, \Phi)$ for $v \in \{v_L^d, v^{d-1}, v_H^d\}$. These can also be computed recursively by using Bayes's rule.

3.6 Data

Given that the purpose of this work is methodological, we perform our empirical analysis on one stock only, Ashland Inc., traded on the New York Stock Exchange and already used in the seminal paper by Easley et al. [Easley et al., 1997].

We obtained the data from the TAQ (Trades and Quotes) dataset.¹¹ The dataset contains the posted bid and ask prices (the “quotes”), the prices at which the transactions occurred (the “trades”), and the time when the quotes were posted and when the transactions occurred. We used transactions data on Ashland Inc. in 1995, for a total of 252 trading days. The data refer to trading on the New York Stock Exchange, the American Stock Exchanges, and the consolidated regional exchanges.

The TAQ dataset does not sign the trades, that is, it does not report whether a transaction was a sale or a purchase. To classify a trade as a sell or a buy order, we used the standard algorithm proposed by Lee and Ready [Lee and Ready, 1991]. We compared the transaction price with the quotes that were posted just before a trade occurred.¹² Every trade above the midpoint was classified as a buy order, and every trade below the midpoint was classified as a sell order; trades at the midpoint were classified as buy or sell orders according to whether the transaction price had increased (uptick) or decreased (downtick) with respect to the previous one. If there was no change in the transaction price, we looked at the previous price movement, and so on.¹³

TAQ data do not contain any direct information on no-trades. We used the established convention of inserting no-trades between two transactions if the elapsed time between them exceeded a particular time interval (see, e.g., Easley et al. [Easley et al., 1997]). We obtained this interval by computing the ratio between the total trading time in a day and the average number of buy and sell trades over the 252 days (see, e.g., Chung et al. [Chung et al., 2005]). In our 252 trading-day window, the average number of trades per day was 90.2. We divided the total daily trading time (390 minutes) by 90.2 and obtained a unit-time interval of 259 seconds (i.e., on average, a trade occurred every 259 seconds). If there was no trading activity for 259 seconds or more, we inserted

¹¹Hasbrouck [Hasbrouck, 2004] provides a detailed description of this dataset.

¹²Given that transaction prices are reported with a delay, we followed Lee and Ready’s [Lee and Ready, 1991] suggestion of moving each quote ahead in time by five seconds. Moreover, following Hasbrouck [Hasbrouck, 1991], we ignore quotes posted by the regional exchanges.

¹³We classified all trades with the exception of the opening trades, since these trades result from a trading mechanism (an auction) substantially different from the mechanism of trading during the day (which is the focus of our analysis).

one or more no-trades to the sequence of buy and sell orders. The number of no-trades that we inserted between two consecutive transactions was equal to the number of 259-second time intervals between them.

Our sample of 252 trading days contained on average 149 decisions (buy, sell, or no-trade) per day. The sample was balanced, with 30 percent of buys, 31 percent of sells, and 40 percent of no-trades.

Finally, remember that in our theoretical model we assume that the closing price is a martingale. For the case of Ashland Inc. during 1995, the data support the hypothesis that the closing price is a martingale with respect to the history of past prices (i.e., the information available in our dataset): the autocorrelogram of price changes is not significantly different from zero, at all lags and at all significance levels (see Cipriani and Guarino [Cipriani and Guarino, 2014]).

3.7 Results

We start by presenting the parameter estimates for the model without a financial transaction tax. We will then illustrate the impact of the transaction tax by simulating an economy with our parameter estimates and different tax rates.

3.7.1 Estimates

We estimate the parameters through maximum likelihood, using the Nelder-Mead simplex algorithm.¹⁴ ¹⁵ The optimisation routine converges to the same parameter set starting from a large set of initial conditions. Table 3.1 presents these estimates and their corresponding standard errors.¹⁶

Given our underlying structural model, the market maker believes that an information

¹⁴We use an implementation of the algorithm by the Numerical Algorithm Group (NAG) for the NAG Matlab Toolbox.

¹⁵We report estimations of simulated data for our model in Appendix B.0.2. We create artificial transaction data using our model for a range of parameters and manage to recover the parameters using our estimation routine with satisfactory precision.

¹⁶Standard errors are computed numerically with the BHHH estimator (see Berndt et al. [Berndt et al., 1974])

<i>parameters</i>	<i>estimates</i>	<i>s.e.</i>
α	0.2785	0.026
δ	0.6252	0.067
μ	0.4134	0.013
τ	0.4284	0.027
ψ	0.3165×10^{-8}	0.022
ε	0.5698	0.006

Table 3.1: Estimation Results. Estimates for the six parameters of the model and their standard errors (*s.e.*).

event occurs on approximately one out of four trading days ($\alpha = 0.279$) and if there is a change in the fundamental value of the asset, it is more likely to be an upward movement than a downward movement ($\delta = 0.625$). On event days about 40 percent of traders are informed ($\mu = 0.413$), the rest being noise traders. Price inelastic noise traders are slightly more likely to trade than to refrain from trading ($\varepsilon = 0.5698$). Recall that if they do trade, they buy and sell with equal probability by assumption. We estimate the share of price elastic noise traders to be negligible ($\psi < 10^{-8}$). Thus, essentially all trading activity that is not due to private information about the fundamental asset value is price inelastic. More precisely, we tested the hypothesis that the share of price elastic noise traders is zero, by performing a likelihood ratio test where we set $\psi = 0$ for the restricted model. The outcome of this test is reported in Table 3.2.

<i>LL (unrestricted)</i>	<i>LL (restricted)</i>	<i>LR test</i>	<i>p-value</i>
-40573.87	-40574.32	0.8859	0.35

Table 3.2: Likelihood-Ratio Test. Log-likelihood value (*LL*) of the restricted and unrestricted model as well as the *LR* test statistic and corresponding *p-value*.

We cannot reject the hypothesis that there are no price elastic noise traders in our asset market at any conventional level of significance. In our simulations for the impact of a financial transaction tax, we therefore use estimates obtained for a restricted version of our original model with ψ set to zero. These are the the estimates found by Cipriani and Guarino [Cipriani and Guarino, 2014], The five parameter estimates are reported in Table 3.3. The estimates for the restricted model are almost identical to those obtained for the unrestricted model.

<i>parameters</i>	<i>estimates</i>	<i>s.e.</i>
α	0.2803	0.026
δ	0.6238	0.067
μ	0.4167	0.012
τ	0.4526	0.026
ε	0.5691	0.002

Table 3.3: Estimation Results (Restricted Model). Estimates for the five parameters of the restricted model and their standard errors (s.e.).

3.8 Simulating the Impact of a FTT

We will now simulate how trading in our stock (Ashland Inc.) would be influenced by the introduction of a financial transaction tax. To perform simulations of our model under different tax scenarios we will use the estimates for the parameter vector $\{\alpha, \delta, \mu, \tau, \varepsilon\}$ reported in Table 3.3. To calculate the equilibrium trading thresholds for our informed and price elastic noise traders in the presence of the tax ρ , we will need an estimate for λ_H which parameterizes the size of the possible movements in the fundamental value of the asset.¹⁷ It is the size of the tax in relation to the possible gains and losses from holding the asset that determine the impact of the tax (recall expression (3.9)). We use the difference between daily closing and opening prices of Ashland Inc. for our sample period to obtain an estimate for λ_H . Unlike in our theoretical model, observed price changes for our stock across days are due to two kinds of events. First, those that were private information to traders and were revealed to the market during the course of trading. These changes are what λ_H and λ_L in our model refer to. Second, the events that became public information during or after the trading day. To obtain a measure of λ_H we must disentangle the two components. To do so, we performed a variance decomposition suggested by Hasbrouck [Hasbrouck, 1991] that allows us to decompose stock price changes into a trade-correlated component, interpreted as the component driven by private information, and a trade-uncorrelated component. This procedure is discussed in more detail in the Appendix. Table B.1 reports the results of this decomposition exercise. It gives a measure for the size of the tax in relation to

¹⁷Recall that by assumption the size of a potential fall in the fundamental value is given by $\lambda_L = [\delta/(1-\delta)]\lambda_H$.

potential changes in the fundamental value, where the tax rates have been chosen to lie in the range of tax rates for a financial transaction tax proposed or already recently implemented.

<i>tax</i>	0.1 %	0.2 %	0.5 %
ρ/λ_h	0.1526	0.3053	0.7632

Table 3.4: Calibration of FTT

These estimates suggest that if we imposed, for example, a 0.1 percent tax, the amount of tax paid would be approximately equal to 15 percent of the upside risk in the fundamental value due to private information. This cost would go up to 75 percent for a 0.5 percent tax.

We simulate the model for 10,000 trading days each having 149 trading times.¹⁸ The simulations yield a trade sequence and the corresponding ask and bid prices for each trading day.

3.8.1 Impact on Volume and Spreads

We start by investigating the impact of the tax on the trading volume. Tables 3.5 and 3.6 report the percentage of trading times that result in a trade (either buy or sell) with and without tax. The results are reported by type of event day and for different trading periods. As one can see, if we consider all days and all trading activity during the course of a day, the introduction of a 0.5 percent transaction tax leads, on average, to a drop in the trading volume of 9 percentage points (from trading (buy or sell) at 59 percent of the possible trading times to 50 percent). If we consider event days only, this impact increases to a 33 percentage point drop in volume. A smaller tax of 0.1 percent would still lead to a drop of 15 percentage points. As our estimate for the share of price elastic noise traders was zero, this drop is purely due to informed traders, who now prefer not to trade on their private information in order to avoid paying the tax.

¹⁸The transaction data for Ashland Inc. have an average of 149 trading times per trading day given our chosen no-trade interval.

	1 – 149	1 – 50	51 – 100	101 – 149
0%	0.59	0.60	0.59	0.59
0.1%	0.55	0.57	0.55	0.53
0.2%	0.53	0.55	0.53	0.52
0.5%	0.50	0.50	0.50	0.50

Table 3.5: Trade volume by tax rates (all days). Fraction of trading times at which trade (buy or sell) takes place by trading period and tax rate.

	1 – 149	1 – 50	51 – 100	101 – 149
0%	0.66	0.68	0.65	0.64
0.1%	0.51	0.58	0.50	0.56
0.2%	0.45	0.49	0.44	0.40
0.5%	0.33	0.33	0.33	0.33

Table 3.6: Trade volume by tax rates (event days). Fraction of trading times at which trade (buy or sell) takes place by trading period and tax rate.

Note that for a 0.5 percent tax the percentage of trades in event days is 33. This value corresponds to no informed trader trading and all trades coming from price inelastic noise traders (since in this case a trade happens with probability $(1 - \mu)\epsilon = 0.332$). Essentially this means that in our asset market a transaction tax of 0.5 percent would crowd out all informed trade.

Tables 3.7 and 3.8 show the impact of the tax on the bid-ask spread. Clearly, since a tax of 0.5 percent crowds out all informed traders, the spread for this tax level drops to zero. This provides a note of caution should one want to use bid-ask spreads to evaluate the impact of a transaction tax on market liquidity. Narrower spreads can be a consequence of informed traders being driven out of the market. We note that while spreads are monotonically decreasing in the tax on event days, the relationship is non-monotonic if we pool all days, first increasing and then decreasing with the tax. To understand this, note that in our economy (with no price elastic noise traders) the tax has two opposing effects on the spread. On the one hand, a higher tax crowds out informed traders and thus the market maker needs to protect himself less against private information. This

leads to a lower spread. On the other hand, fewer informed traders also means that the market maker learns the value less quickly and, therefore, spreads converge to zero at a slower pace. Figure 3.4 illustrates the behaviour of spreads over the course of a trading day across different tax regimes.

	1 – 149	1 – 50	51 – 100	101 – 149
0%	0.025	0.037	0.024	0.016
0.1%	0.027	0.038	0.026	0.016
0.2%	0.023	0.033	0.023	0.014
0.5%	0	0	0	0

Table 3.7: Bid-ask spread by tax rates (all days). Mean bid-ask spread by trading period and tax rate.

	1 – 149	1 – 50	51 – 100	101 – 149
0%	0.054	0.061	0.057	0.043
0.1%	0.045	0.049	0.048	0.036
0.2%	0.042	0.046	0.045	0.033
0.5%	0	0	0	0

Table 3.8: Bid-ask spread by tax rates (event days). Mean bid-ask spread by trading period and tax rate.

Finally, it is interesting to study how the cost of the tax compares to the cost due to the bid-ask spread (which in our framework is caused by the adverse selection due to private information). We use half the spread as a measure of this cost. Tables 3.9 and 3.10 report the ratio of half the spread to the amount of tax paid per transaction ρ . For all tax levels the transaction cost caused by the tax is larger than the cost induced by the spread.

	1 – 149	1 – 50	51 – 100	101 – 149
0.1%	0.40	0.56	0.38	0.24
0.2%	0.17	0.24	0.17	0.10
0.5%	0	0	0	0

Table 3.9: Bid-ask spread in relation to FTT (all days). Mean bid-ask spread divided by 2 as a fraction of the transaction tax ρ by trading period and tax rate.

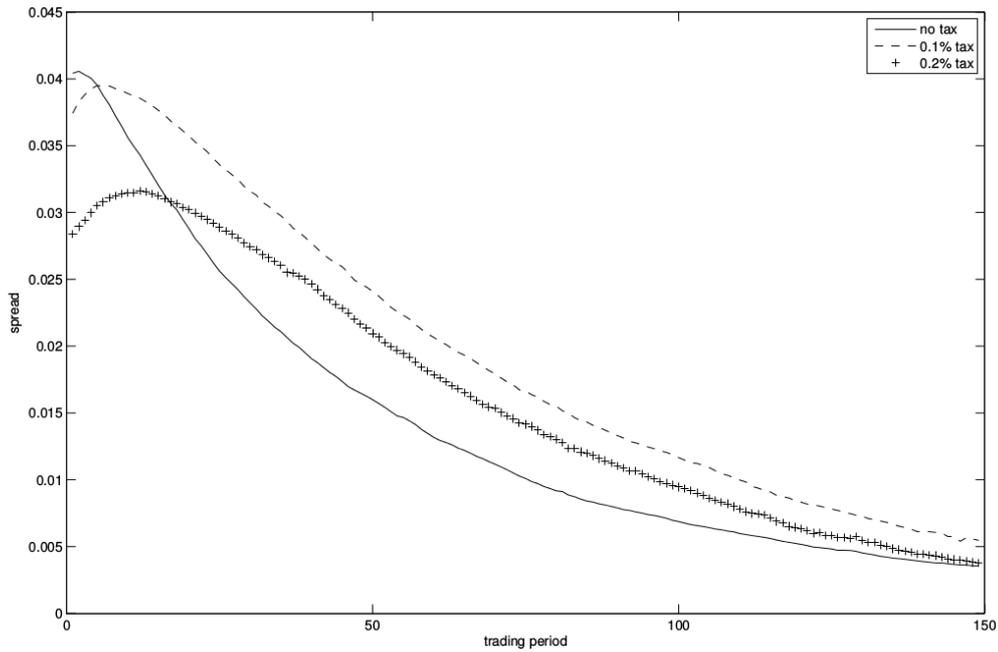


Figure 3.4: Spreads by tax regime. The figure displays average spreads on 'no event day' over course of trading day for three tax regimes.

	1 – 149	1 – 50	51 – 100	101 – 149
0.1%	0.66	0.72	0.71	0.53
0.2%	0.31	0.34	0.33	0.24
0.5%	0	0	0	0

Table 3.10: Bid-ask spread in relation to FTT (event days). Mean bid-ask spread divided by 2 as a fraction of the transaction tax ρ by trading period and tax rate.

3.8.2 Impact of Informational Efficiency

We now investigate how the transaction tax influences the ability of the market to aggregate private information into the price. We focus on the evolution of the market maker's expected value of the fundamental value, $p_t^d = \mathbb{E}(V^d | h_t^d)$ (which we have defined, for convenience, as the price of the asset) over the course of the trading day. We analyse to what extent this price deviates from the true fundamental value v^d over the course of the trading day across different tax rates. In particular, we focus on the average absolute deviation of this price from the fundamental value normalized by the

expected value of the asset at the start of trading, which is δ , that is,

$$\frac{1}{T} \sum_{t=1}^T \left| \frac{p_t^d - v^d}{\delta} \right|.$$

Table 3.11 reports this average deviation from the fundamental value for the different tax rates. We focus on event days only, as our focus is on how well the price is able to reflect private information about asset values. On days without events, no such information is present in the market.

First we note that, as a 0.5 percent tax crowds out all informed trade, the price under such a tax regime remains constant at the expected value of the fundamental asset at $t = 1$ which is δ . On event days the deviations of the price from the fundamental values are given by $(1 - \delta)/\delta$ and $(\delta - 0)/\delta$ for good and bad days respectively.

Increasing the tax unambiguously reduces the informational efficiency of the asset market. Going from a market without transaction tax to a market where each transaction is taxed at 0.2 percent increases the average deviation of the price from the fundamental value by 8 percentage points on good event days and by 14 percentage points on bad event days.

	<i>good event</i>	<i>bad event</i>
0%	0.37	0.61
0.1%	0.40	0.67
0.2%	0.45	0.75
0.5%	0.61	1

Table 3.11: Price deviation from fundamental value. Absolute price deviation from fundamental value as a fraction of expected value by tax rates.

Next we consider by how much prices in a market with transaction taxation would deviate from the price that would obtain in a market without a transaction tax. Table 3.12 reports the average absolute deviation over a trading day normalized by the expected value of the asset δ . We see that this deviation can be as large as 46 percent on *bad event days* for a tax of 0.5%.

	<i>good event</i>	<i>bad event</i>
0.1%	0.12	0.16
0.2%	0.17	0.23
0.5%	0.33	0.46

Table 3.12: *Price deviation from price without FFT. Absolute price deviation from price without transaction tax as a fraction of expected value by tax rates.*

3.9 Conclusion

We have developed a novel methodology to quantify the impact of a financial transaction tax on informational efficiency and market liquidity. We have extended a standard model of sequential trading by including both price elastic liquidity traders and traders with private information of heterogeneous quality. To illustrate our methodology, we have estimated the model for a stock traded on the NYSE and then used these estimates to simulate the effect a transaction tax would have on trading in this specific stock.

For the specific stock used in the analysis (Ashland Inc.) we found that noise trading (e.g. trades due to liquidity needs) is relatively unresponsive to prices and, thus, the major effect of a transaction tax is on informed trading. Therefore, the financial transaction tax has a negative effect on informational efficiency.

Whether this result is peculiar to our stock or holds across many types of stocks and for various time horizons is an important topic for future research.

Chapter 4

Preemption and Delay: Debt Financing with Roll Over Risk

4.1 Introduction

When investing under uncertainty investment decisions undertaken by other investors carry double weight. Firstly, they may directly impact the returns of the common investment project. Failure to roll over short term debt by one party, for example, can inflict considerable costs on the remaining investors if it triggers costly liquidation. Investment decision with several parties involved will generally involve such payoff externalities. Secondly, when investment decisions are based on private information, observing other investors' actions reveals additional information about the investment's payoff. Such informed actions thus create an informational externality by reducing the payoff uncertainty of those who are able to observe them.

The strategic implications of these two externalities for dynamic interactions between investors can be quite distinct. Informational externalities are well-known to cause strategic delays (e.g. Chamley and Gale [Chamley and Gale, 1994]). Waiting has an option value as it allows to gather additional information by observing other investors' actions. This option value reduces the incentives to act immediately upon receiving private information. Investors try to outwait each other. Payoff externalities, on the other hand, can create preemption motives when late movers' payoffs are neg-

actively affected by previously taken actions. Banks runs (e.g. Diamond and Dybvig [Diamond and Dybvig, 1983]) are a prominent consequence of such negative payoff externalities.

We study the interplay of such informational and payoff externalities in a model where several parties invest in a common project with risky returns. We show that the pre-emption motive created by negative payoff externalities can counteract the incentives to delay actions that arise from the ability to observe other investors. This counteracting force is shown to facilitate information revelation and thereby improve the efficiency of investment decisions.

In our model two investors hold claims, consisting of a mix of short term and long term debt, to the future risky payoffs of an investment project and can receive private information about these payoffs. Short term debt has to be rolled over at intermediate stages of the project and failure to do so by either investor triggers liquidation with a fixed liquidation value. Thus the decision to liquidate by one investor impacts the payoff of the other investor. This is the source of the payoff externality in our model. Short term debt which has not been rolled over is senior to both long term debt and rolled-over short term debt in the case of liquidation. The ratio of short term to long term debt thus controls the size of the payoff externality. The ability to observe the roll over decisions of the other investor allows inference about his information and thereby creates informational externalities. We solve for the unique perfect Bayesian equilibrium in symmetric strategies of this game and show that an appropriately chosen mix of short term and long term debt can guarantee efficient liquidation decisions. We show that the optimal level of the payoff externalities created through short term debt crucially depends on the liquidation value of the project and the quality of the private information received by investors.

Related Literature The topic of observational learning has triggered a large literature. Important early contributions are Chamley and Gale [Chamley and Gale, 1994], Gul and Lundholm [Gul and Lundholm, 1995] and Bikhchandani et al. [Bikhchandani et al., 1998]. A key insight of these models is that the ability to observe others' behaviour generates

an option value from waiting and can thereby impede the revelation of privately held information. Gale [Gale, 1996] provides an overview.

Our model is closely related to Weeds [Weeds, 2002] who analyses a model of investment under uncertainty where preemption motives interact with incentives to delay caused by option values. However, in her work the option value of waiting does not arise from the ability to observe other players' actions but from a commonly observed stochastic process that reveals information about the profitability of investment over time. Frisell [Frisell, 2003] also models a setting where these two forces are at work, analysing a product placement game. However his focus is on understanding the timing of moves rather than on the issue of information revelation. Gu [Gu, 2011] analyses a model of fundamental bank runs in the tradition of Allen and Gale [Allen and Gale, 2002]. In his model depositors receive both information about their consumption preferences and the payoff of an illiquid asset their bank is invested in. Depositors then have to decide whether to withdraw or keep their savings in the bank. Gu's setup includes both payoff externalities arising through the possibility of bank runs as well as informational externalities coming from the ability to observe other depositors' withdrawal decision. However his agents are atomistic and only a measure zero set of them has private information. Additionally informed agents do not move simultaneously which weakens preemption motives. Our setup with a finite number of actors and simultaneous moves allows for richer strategic interactions and consequently yields equilibrium behaviour that differs from Gu's analysis. We will see that in our model mixed strategies play an important role, whereas Gu's equilibrium is in pure strategies. Furthermore our focus is on the optimal financial structure of the investment in the presence of payoff and informational externalities, a topic Gu does not deal with.

Technically, the equilibrium construction in this paper is closely related to the approach used in Murto and Välimäki [Murto and Välimäki, 2011] which looks at information aggregation in an exit game when privately informed players can observe each other exit decisions. However in their model exit decisions do not cause any payoff externalities, which facilitates the equilibrium analysis considerably.

The rest of the paper is organised as follows. Section 2 introduces the setup of our investment game, sections 3 solves for the unique perfect Bayesian equilibrium of the game. Section 4 derives the efficient investment policy and section 5 shows how it can be implemented through an appropriate mix of short term and long term debt claims. Section 6 discusses assumptions of the model. Section 7 concludes.

4.2 Model Setup

Project Characteristics We consider an investment project with uncertain payoff. The project can either be successful and pay out Y units at maturity, or it is unsuccessful in which case its payoff at maturity is zero. Whether the project is successful or not depends on the underlying state of the world $\theta \in \{0, 1\}$ where state 1 ($\theta = 1$) implies success and state 0 ($\theta = 0$) failure. The prior probability of success is $0 < \mu_0 < 1$.

Time evolves in discrete periods $t = 0, 1, 2, \dots$. At $t = 0$ nature chooses the state of the world θ with $\mathbb{P}(\{\theta = 1\}) = \mu_0$. At each point in time $t > 0$, if still active, the investment project matures with probability $0 < \gamma < 1$ and pays out in the successful state only. If the project does not mature in period t it can be liquidated. Liquidation yields $0 < L < Y$ units irrespective of the state of the world θ . If the project has neither matured nor been liquidated in period t it carries on into the next period $t + 1$.

Investors Investors are risk neutral and have a discount rate of zero. Each investor's claims consist of a mix of short term and long term debt with face value d^s and d^l respectively. Long term debt entitles its owner to a payment of d^l when the project matures. Short term debt can be claimed in every period t with a promised payment of d^s . Alternatively an investor can decide to roll over his short term debt in which case he is entitled to a payment of d^s in the next period $t + 1$. Roll over decisions for short term debt are made simultaneously by all active investors.

The number of active investors depends on the state of the world θ . In state 0 there are two active investors, each of whom has to make a roll over decision for his short term debt in every period. In state 1 there is at most one investor with a roll over decision to make. Given that the project is unsuccessful in state 0, the presence of two active

investors then indicates failure of the project. Hence having an active roll over decision to make is a negative signal for an individual investor. His beliefs concerning state 1 are given by $\mu < \mu_0$. A given investor does not know whether there is a second investor present who has a roll over decision to make.¹

Liquidation of project If at least one investor refuses to roll over his short term debt, the project has to be liquidated with liquidation revenue L . Claimed short term debt is senior to both long term debt and rolled over short term debt. The latter two are of equal seniority. Hence in case of a unilateral withdrawal by one investor, the investor who refuses to roll over receives

$$\min\{d^s, L\} + \max\left\{0, \frac{L - d^s}{2}\right\}$$

whereas the investor who has rolled over his short term debt receives

$$\max\left\{0, \frac{L - d^s}{2}\right\}.$$

If both investors refuse to roll over their short term debt in a given period, both receive half of the liquidation revenue, that is $L/2$.

In what follows we assume that a unilaterally withdrawing investor will receive the full face value of his short term debt, that is $d^s \leq L$. Furthermore to simplify notation we define

$$\phi = d^s + \frac{L - d^s}{2} \tag{4.1}$$

¹To generate such beliefs one could, for example, assume that there are two potentially active investors and that in state 0 both of them are active. In state 1 however at most one of them is chosen to be active. An investor is chosen to be active with probability $0 < q \leq 1$ and each investor has an equal probability to be active in state 1. An individual investor can only observe that he has an active decision to make, but does not know whether the same is true for the second investor. By Bayes' Rule we would have

$$\mu = \frac{(1/2)q\mu_0}{(1/2)q\mu_0 + (1 - \mu_0)} < \mu_0$$

The lower q , the more informative is the fact that one has to make a roll over decision. We can think of q as the probability of any investor receiving a negative signal about the outlook of the project, when in fact the project is successful.

	exit	roll over
exit	l, l	$\phi, 2l - \phi$
roll over	$2l - \phi, \phi$	

Table 4.1: *Payoff matrix in case of liquidation*

which stands for the amount received by a unilaterally withdrawing investor. Also define as $l = L/2$ the per-capita liquidation revenue. Payments in case at least one investor refuses to roll over his short term debt are then given by the payoff matrix in Table 4.1.

If the project matures in period t , an investor is paid $d^s + d^l$ in state 1 and zero in state 0. In the following we assume that the final payoff of the project in case of success is $Y = 2$ and each active investor in state 1 receives half of this, that is $d^s + d^l = 1$.

4.3 The Liquidation Game

An individual investor does not know whether he faces an active co-investor or not. He can however draw inference about the state of the world from the absence of liquidation up to the current period. If active investors do not roll over their short term debt with probability 1, the absence of liquidation in previous periods constitutes good news. It increases the likelihood of state 1 and thus the probability that the investment project has a positive payoff at maturity. However to the extent that observing the survival of the project provides valuable information on its eventual success and discount rates are zero, an individual investor has an incentive to wait in order to gather additional information. Waiting might dominate the withdrawal of short term financing and in equilibrium nothing can be learnt about the state of the world by observing the survival of the project. Investors would ignore their private information when making investment decisions. We show that if the cost of being preempted is sufficiently high and active investors' information is sufficiently precise, that is they are sufficiently pessimistic about the success of the project, such a situation cannot occur in equilibrium. Investors will always be able to learn about the state of the project from the absence of

its liquidation in previous periods.

We now analyse the dynamic game played by active investors, starting in period $t = 1$ and lasting until either the investment project matures or at least one investor withdraws his short term funding. The equilibrium concept will be *perfect Bayesian equilibrium* (PBE) and we will restrict our attention to symmetric strategies. An investor's information set at time t only consists of whether the project is still active or not. As the withdrawal of short term funding by either investor terminates the game, an individual investor's strategy σ then simply specifies an exit probability $\sigma_t \in [0, 1]$ for any time period t in which the game is still active. Given an equilibrium strategy σ beliefs μ are then calculated using Bayes' Rule

$$\mu_{t+1} = \frac{\mu_t}{1 - \sigma_t(1 - \mu_t)} \quad (4.2)$$

with $\mu_1 = \mu$. The pair (σ, μ) is a PBE of the game if for any time t in which the game is still active, σ_t is a best response to $\{\sigma_j\}_{j \geq t}$ given beliefs μ_t and beliefs are updated using the equilibrium strategy according to (4.2).

We will now show that this game has a unique perfect Bayesian equilibrium in symmetric strategies. Whether liquidation occurs in equilibrium depends on investors' initial beliefs μ . For sufficiently optimistic beliefs investors will always roll over their short term debt and the project will be carried out until maturity. In particular if the payoff from unilaterally withdrawing short term financing, ϕ , is lower than the expected eventual payoff of the project, μ , rolling over until maturity is preferred to withdrawing if the other investor does so, too. For sufficiently pessimistic beliefs the only equilibrium outcome is immediate refusal to roll over. Here even if the opponent's behaviour is maximally informative, that is he liquidates with probability 1 and thus the state of the world could be learnt by waiting one more period, an active investor prefers to withdraw short term funding. Waiting would yield $\mu + (1 - \mu)(2l - \phi)$ whereas withdrawing would result in a payoff of $\mu\phi + (1 - \mu)l$. Equating these two

payoff determines the critical value $\hat{\mu}$ for beliefs below which immediate withdrawal is the only equilibrium. For intermediate values of beliefs μ every active investor rolls over short term funding with strictly positive probability less than 1. Proposition 1 characterises the unique perfect Bayesian equilibrium in symmetric strategies, which is then proven to be an equilibrium of this game in a sequence of lemmata.

Proposition 8. *The liquidation game has a unique perfect Bayesian equilibrium in symmetric strategies. There exists a $\hat{\mu} \in (0, \phi)$ such that an informed player refuses to roll over short term debt if $\mu < \hat{\mu}$, refuses to do so with probability $0 < \sigma_1 < 1$ in the first period and never thereafter if $\mu \in (\hat{\mu}, \phi)$ and continues to finance the project until maturity with probability 1 if $\mu > \phi$ where*

$$\hat{\mu} = \frac{\phi - l}{1 - l} \text{ and } \sigma_1 = \frac{\phi - \mu}{(1 - \mu)l}$$

$V(\mu_t; \sigma)$ will designate the equilibrium value of the subgame starting at t for an active player given that the equilibrium strategy is σ and current beliefs of the player are μ_t . We have

$$V(\mu_t; \sigma) = \max \{ [1 - (1 - \mu_t)\sigma_t]\phi + (1 - \mu_t)\sigma_t l, \\ (1 - \mu_t)\sigma_t(2l - \phi) + [1 - (1 - \mu_t)\sigma_t][(1 - \gamma)\mu_{t+1} + \gamma V(\mu_{t+1}; \sigma)] \}$$

where the first term is the expected payoff of liquidation in t and the second term is the expected payoff of continuation in t given the equilibrium strategy σ .

We start by showing that if the expected payoff at maturity of the investment is sufficiently high, it will never be liquidated in equilibrium.

Lemma 1. *If $\mu_t > \phi$ no liquidation occurs in the subgame starting in period t .*

Proof. Suppose there exists a period $\tau \geq t$ such that $\sigma_\tau > 0$. Then the player has to

weakly prefer liquidation to continuation in period τ , that is

$$\begin{aligned} & [1 - (1 - \mu_\tau)\sigma_\tau]\phi + (1 - \mu_\tau)\sigma_\tau l \\ & \geq (1 - \mu_\tau)\sigma_\tau(2l - \phi) + [1 - (1 - \mu_\tau)\sigma_\tau][(1 - \gamma)\mu_{\tau+1} + \gamma V(\mu_{\tau+1}; \boldsymbol{\sigma})] \end{aligned}$$

where $V(\mu_{\tau+1}; \boldsymbol{\sigma})$ is the equilibrium payoff of the game in period $\tau + 1$. But as $V(\mu_{\tau+1}; \boldsymbol{\sigma}) \geq \mu_{\tau+1}$ ² this requires that

$$\phi \geq (1 - \mu_\tau)\sigma_\tau l + \mu_\tau$$

which cannot hold as $\mu_\tau > \phi$ was assumed. \square

Next we show that for sufficiently pessimistic beliefs liquidation has to occur with strictly positive probability.

Lemma 2. *If $\mu_t < \phi$ then in equilibrium we have $\sigma_t > 0$.*

Proof. We consider two possibilities in turn: firstly continuation forever from t onwards and, secondly, continuation with probability 1 up to some period $k > t$ and liquidation with strictly positive probability thereafter.

To see that the first scenario cannot be an equilibrium, note that the equilibrium payoff of continuing forever would be μ_t whereas a deviation to exit with probability 1 in period t would yield $\phi > \mu_t$.

Next suppose in equilibrium continuation is played with probability 1 from period t until $k > t$ and $k + 1$ is the first period with strictly positive exit probability. For $\sigma_{k+1} > 0$ we need $V(\mu_{k+1}; \boldsymbol{\sigma}) = [1 - (1 - \mu_{k+1})\sigma_{k+1}]\phi + (1 - \mu_{k+1})\sigma_{k+1}l$, which is the expected payoff from exiting in period $k + 1$. As $\sigma_{t+j} = 0$ for all $0 \leq j \leq k - t$ we have $\mu_{k+1} = \mu_k = \mu_t$ by (4.2). Nothing is learned from observing continuation of the project. Thus in k continuation is a best response to $\sigma_k = 0$ if

$$\phi \leq (1 - \gamma)\mu_t + \gamma\{[1 - (1 - \mu_t)\sigma_{k+1}]\phi + (1 - \mu_t)\sigma_{k+1}l\}$$

²Playing “always continue” from $\tau + 1$ onwards has an expected payoff of at least $\mu_{\tau+1} + (1 - \mu_{\tau+1})(2l - \phi)$.

But this inequality cannot hold if $\mu_t < \phi$. It follows from the two above observations that in any equilibrium we need $\sigma_t > 0$ as long as $\mu_t < \phi$. \square

The next lemma establishes that there cannot be two consecutive periods such that a player mixes between exiting and rolling over in one period, and then exits with strictly positive probability in the following period.

Lemma 3. *If $\mu_t < \phi$ it cannot be the case that $0 < \sigma_t < 1$ and $\sigma_{t+1} > 0$.*

Proof. In this case the player has to be indifferent between exiting and rolling over in period t . We thus need

$$V(\mu_t; \boldsymbol{\sigma}) = [1 - (1 - \mu_t)\sigma_t]\phi + (1 - \mu_t)\sigma_t l$$

Furthermore we have

$$V(\mu_t; \boldsymbol{\sigma}) = (1 - \mu_t)\sigma_t(2l - \phi) + [1 - (1 - \mu_t)\sigma_t][(1 - \gamma)\mu_{t+1} + \gamma V(\mu_{t+1}; \boldsymbol{\sigma})]$$

Combining these two conditions yields

$$(1 - \mu_t)\sigma_t(\phi - l) = [1 - (1 - \mu_t)\sigma_t][(1 - \gamma)(\mu_{t+1} - \phi) + \gamma(V(\mu_{t+1}; \boldsymbol{\sigma}) - \phi)]$$

the left hand side of which is strictly positive. However for $\sigma_{t+1} > 0$ we need $\mu_{t+1} \leq \phi$ or else continuation with probability 1 would be the only possible equilibrium of the subgame starting in period $t + 1$. For the equality to hold it must thus be the case that $V(\mu_{t+1}; \boldsymbol{\sigma}) > \phi$. But this contradicts the fact that liquidation with positive probability requires $V(\mu_{t+1}; \boldsymbol{\sigma})$ to equal the payoff from liquidation in period $t + 1$, which is less than ϕ . \square

For intermediate values of beliefs the project cannot be liquidated with probability 1. Here the option value of waiting and learning the true state in the next period if the opposing player exits exceeds the benefits of early exit.

Lemma 4. *The project is liquidated with probability 1 in period t if and only if $\mu_t + (1 - \mu_t)l < \phi$.*

Proof. We start with the *only if* part of the statement. Suppose $\sigma_t = 1$. If no exit occurs in period t the player knows that $\theta = 1$ and will continue in all subsequent periods given that $\phi < 1$. Rolling over in period t thus yields $(1 - \mu_t)(2l - \phi) + \mu_t$. Immediate exit yields $\mu_t\phi + (1 - \mu_t)l$. It follows that rolling over in period t is strictly preferred to exit as long as $\phi < \mu_t + (1 - \mu_t)l$.

We now prove the *if* part of the statement. By Lemma 2 we know that in equilibrium we need $\sigma_t > 0$ and by Lemma 3 we also know that this implies $\sigma_{t+j} = 0$ for all $j > 0$. Thus for any $\sigma_t > 0$ exiting is strictly preferred to rolling over if

$$[1 - (1 - \mu_t)]\phi + (1 - \mu_t)\sigma_t l > (1 - \mu_t)\sigma_t(2l - \phi) + \mu_t$$

where the left hand side is the expected payoff from exiting in period t and the right hand side is the expected payoff from continuation in t . It follows that exiting is strictly preferred to rolling over if $\phi > (1 - \mu_t)\sigma_t l + \mu_t$. But this is true for any $\sigma_t > 0$ and consequently in equilibrium we necessarily have $\sigma_t = 1$. \square

Define $\hat{\mu}$ as the belief such that $\hat{\mu} + (1 - \hat{\mu})l = \phi$. The above lemmata establish that if the initial belief μ is strictly lower than $\hat{\mu}$, then the only equilibrium in symmetric strategies is immediate exit with probability 1 in period 1. If $\mu > \phi$, then in equilibrium the project is never liquidated. For intermediate levels of beliefs ($\hat{\mu} < \mu < \phi$) the project has to be liquidated with strictly positive probability $0 < \sigma_1 < 1$ in period 1 in any equilibrium in symmetric strategies. Conditional on not having been liquidated then, it is never liquidated thereafter.

Lastly we derive the equilibrium exit probability σ_1 for $\hat{\mu} < \mu < \phi$. Continuation with probability 1 from period 2 onwards implies that $V(\mu_2; \sigma) = \mu_2$. For the player to be indifferent between exiting and rolling over in period 1 given this continuation value we need

$$\phi = (1 - \mu)\sigma_1 l + \mu \Rightarrow \sigma_1 = \frac{\phi - \mu}{(1 - \mu)l}$$

As $0 < \phi - \mu < (1 - \mu)l$ we have $0 < \sigma_1 < 1$ as required. It remains to check that

$\mu_2 \geq \phi$ or else continuation from period 2 onwards would not be an equilibrium. But by (4.2) and the assumption that $\phi \geq l$ we have

$$\mu_2 = \frac{\mu l}{\mu - (\phi - l)} \geq \phi$$

This concludes the proof of Proposition 8. \square

4.4 Efficient Liquidation

As a benchmark for efficient liquidation of the project we consider the roll over policy a planner would choose for an active investor, not knowing whether another active investor is present or not. For each period t the roll over decision can only be based on the number of periods the project has been active without being liquidated. We thus restrict the planner to using the same information available to an individual active investor when making roll-over decisions. We also restrict the planner to symmetric roll over policies, that is he cannot specify policies that vary with the identity of the investor. These two restrictions on the planner's choice of policy appear natural if we want to find a benchmark against which to judge the economic efficiency of the previously derived symmetric PBE.

Given the above restrictions a roll over policy is then a vector $\{\lambda_t\}_{t=1}^{\infty}$ where $\lambda_t \in [0, 1]$ specifies the probability with which an active investor refuses to roll over short term debt in period t given that the investment project has not been liquidated previously. The planner's problem can be formulated as a dynamic programming problem where the investor's belief μ_t constitutes the state variable in period t which evolves according to

$$\mu_{t+1} = \frac{\mu_t}{1 - \lambda_t(1 - \mu_t)} \quad (4.3)$$

using Bayes' Rule and the fact that an active co-investor would follow the suggested roll over policy. A given λ_t then implies an expected liquidation probability of $\mu_t \lambda_t + (1 - \mu_t)[1 - (1 - \lambda_t)^2]$ in period t . Either there is only one active investor (μ_t) in which case liquidation occurs with probability λ_t or the other investor has a roll over choice

as well $(1 - \mu_t)$ and the project is liquidated if at least one of the investors exits. This happens with probability $1 - (1 - \lambda_t)^2$. In case of continuation the project pays out with probability $1 - \gamma$. With probability γ the project continues with updated beliefs μ_{t+1} . The Bellman Equation for this problem is then given by

$$W(\mu_t) = \max_{\lambda_t \in [0,1]} \{ [\lambda_t \mu_t + [1 - (1 - \lambda_t)^2](1 - \mu_t)]l + \\ [(1 - \lambda_t)\mu_t + (1 - \lambda_t)^2(1 - \mu_t)] [(1 - \gamma)\mu_{t+1} + \gamma W(\mu_{t+1})] \} \quad (4.4)$$

where μ_{t+1} is obtained from (4.3).

Proposition 9. (*Efficient Liquidation*) *There exists a $\mu^* \in (0, 1)$ such that an investor with initial belief $\mu > \mu^*$ never withdraws short term funding. An investor with initial belief $\mu < \mu^*$ exits with probability $\lambda(\mu) \in (0, 1)$ in the first period. If no liquidation occurs in the first period the project is continued until maturity. We have*

$$\mu^* = \frac{2l}{1+l} > l \text{ and } \lambda(\mu) = \frac{1}{2} \left[1 - \frac{\mu - l}{(1 - \mu)l} \right]$$

Proof. see Appendix □

To gain intuition for this result, consider an active investor with belief μ . Suppose the symmetric roll over policy prescribes continuation with probability $1 - \lambda$ in the current period. The benefits from continuation are given by $\mu(1 - \lambda)(1 - l)$ and derive from receiving 1 rather than the liquidation value l in the high payoff state. The opportunity cost of continuation arises from the potential loss of liquidation value l in the low payoff state which occurs if both investors continue. It is thus $(1 - \mu)[1 - (1 - \lambda)^2]l$. The optimal continuation probability $\lambda(\mu)$ equalizes the marginal benefit and cost of continuation. For $\mu > \mu^*$ the marginal benefit of continuation exceeds the marginal cost for any $\lambda \in [0, 1]$. Thus liquidation is never optimal. Indeed, if $\mu > l$ the only reason for liquidating with positive probability is the creation of an option value of waiting for the co-investor. This option value is maximal if the co-investor learns the state with probability 1 by waiting one more period. Thus the maximal option value

is $(1 - \mu)l$ as the state is bad with probability $1 - \mu$ in which case liquidation yields l rather than 0. The opportunity cost of liquidation in a given period is $\mu - l$. Thus liquidation can never be efficient if

$$\mu - l > (1 - \mu)l \Rightarrow \mu > \mu^*$$

4.5 Optimal Maturity Structure

We now show how the efficient roll over policy derived in the previous section can be implemented as a symmetric PBE of the liquidation game by endowing investors with an appropriate mix of short term and long term debt claims to the investment project's payoff. Increasing the share of short term debt in an investor's portfolio increases his incentive to withdraw financing as the reward in case of unilateral exit goes up. The share of short term debt is thus a key design parameter in order to control an investor's roll over probability. It should be chosen in such a way that the incentives of an individual investor are aligned with the objectives of the planner. Here two conflicting aspects of the liquidation decision have to be traded off against each other. On the one hand a higher probability of withdrawal by one investor increases the option value of waiting for the other investor. The observation that the investment project is carried on for another period becomes more informative about the state of the world. On the other hand increasing the probability of liquidation lowers the payoff in the successful state 1. This cost is higher, the higher is the probability of success μ and lower, the higher is the liquidation value l . Proposition 3 characterises the optimal share of short term debt d^s .

Proposition 10. *For any initial belief $\mu \in (0, 1)$ there exists a face value d^s such that the equilibrium strategy σ of the liquidation game coincides with the efficient liquidation policy. We have*

$$d^s = \begin{cases} (1 - l)\mu & \text{if } \mu < \mu^* \\ 0 & \text{if } \mu \geq \mu^* \end{cases}$$

where $\mu^* = 2l/(1 + l)$.

Proof. Recall that from (4.1) we have $d^s = 2(\phi - l)$. ϕ needs to be such that for initial beliefs $\mu > \mu^*$ players never exit the project and for beliefs $\mu < \mu^*$ players exit with probability $\sigma_1 = \lambda(\mu) \in (0, 1)$ in the first period and continue with probability 1 if no liquidation has occurred in the first period.

First consider $\mu > \mu^*$. Here continuation until maturity is the unique symmetric equilibrium outcome if $\mu > \phi$ or equivalently if $d^s < 2(\mu - l)$. If we set $d^s = 0$ this inequality is satisfied for all $\mu > \mu^*$.

Next consider $\mu < \mu^*$. In this case the efficient liquidation policy prescribes exit with probability $\lambda(\mu)$ in the first period and continuation until maturity in the absence of immediate liquidation. For this to be consistent with equilibrium behaviour in a symmetric PBE we need $0 < \phi - \mu < (1 - \mu)l$ and $\sigma_1 = \lambda(\mu)$. The latter equality implies $\phi = l + (1/2)(1 - l)\mu$ or equivalently $d^s = (1 - l)\mu$ and it is easily verified that for this choice of ϕ the former inequalities hold for all $\mu \in (0, \mu^*)$. \square

We note that for sufficiently optimistic beliefs (i.e. imprecise private information) the benefits of creating option value are outweighed by the expected costs of liquidating a successful project and thus the optimal debt structure does not involve any short term debt. The sole purpose of short term debt in this model is to create a preemption motive, which for $\mu > \mu^*$ is not desirable. If however beliefs fall below the critical value μ^* and creating option values through positive withdrawal probabilities by active investors become optimal, the more optimistic investors are, the higher has to be the optimal share of short term debt in their portfolio. More optimistic investors need stronger incentives to be induced to withdraw short term funding.

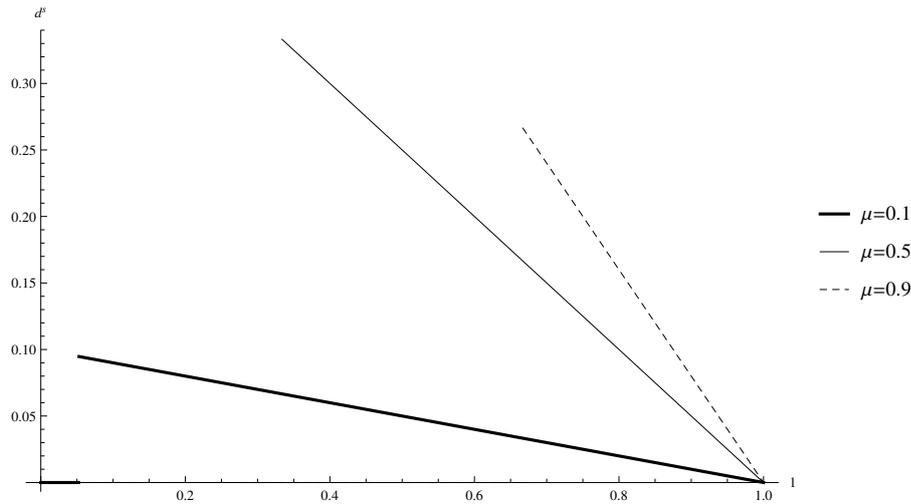


Figure 4.1: *Share of short term debt and liquidation value*

Figure 4.1 plots optimal short term debt shares d^s against liquidation values l for different levels of beliefs. Firstly, it can be seen that the more optimistic investors are about the success of the project, the higher has to be the liquidation value in order to justify the use of short term debt as a means to incentivise information revelation. Secondly, the optimal share of short term debt becomes more sensitive to variations in the liquidation value the more optimistic investors are. Lastly, we note that the optimal maturity structure is not necessarily monotonous in investors' beliefs. For a given liquidation value l the optimal short term debt share can first increase in the level of investors' beliefs about success, but then drop off to zero as liquidation becomes too costly.³

4.6 Discussion

Communication between Investors We have ruled out direct communication between investors. If we allowed investors to pool their information, they would be able to learn the state of the world with certainty. Firstly, ruling out this possibility is motivated by desire to analyse a market where investors are anonymous and the absence of liquidation of investment projects is the only way through which privately held information can be communicated. Secondly, suppose we allowed investors to directly exchange in-

³In figure 4.1 such a situation could for example occur for $l = 0.5$ where $d^s = 0$ for $\mu = 0.9$ but the optimal short term debt share is higher for $\mu = 0.5$ than for $\mu = 0.1$.

formation. Would investors have an incentive to reveal their private information truthfully? Not necessarily. Assume that short term debt face values are such that $\phi > l$ and consider two investors, A and B. Suppose B reveals his information truthfully and agreed behaviour is such that in state 0 both investors exit while in state 1 both roll over. Now A does have an incentive to convince B that he has not received any bad news about the project. In this case B will roll over and A will exit receiving ϕ rather than l , which is what he would get if he revealed his information truthfully. Obviously, if preemption motives are absent, that is if $\phi = l$, truthful revelation of information becomes unproblematic.

Asymmetric Strategies In our analysis we have focused on equilibria in symmetric strategies. The equilibrium analysed above was found to be unique within this class of equilibria. But what about perfect Bayesian equilibria in asymmetric strategies? There are indeed other equilibria of the game when we allow for asymmetries in the players' behaviour. Consider for example a setting where the active players' priors satisfy the condition

$$\mu < \phi < \mu + (1 - \mu)l.$$

Label players as investor A and investor B. Now consider the following equilibrium strategies: investor B, when active, always rolls over. Investor A, when active, always exits. As investor B always continues, investor A cannot learn anything from waiting. Thus as long as his payoff from exiting, which is ϕ given that B rolls over in period 1, is higher than his payoff from continuing until maturity, which is μ , he prefers to exit in period 1. Now consider investor B: if he waits in period 1, he knows the state of the project in period 2. Investor A's behaviour perfectly reveals this state. B's payoff from rolling over in period 1 is therefore $\mu + (1 - \mu)(2l - \phi)$, his payoff from exiting is $\mu\phi + (1 - \mu)l$. The former exceeds the latter if $\phi < \mu + (1 - \mu)l$. This proves that the above strategies for A and B, which depend on the identity of the investor, indeed constitute a perfect Bayesian equilibrium of the game.

So why the focus on symmetric strategies? We want to think of our model as applying to a market with anonymous investors without the ability to communicate other than

through their investment decisions. Within such a setting it is difficult to perceive how investors would be able to coordinate on the use of asymmetric strategies.

4.7 Conclusion

Financing investment projects with claims whose maturity does not match the maturity of the project itself bears the risk of premature liquidation through withdrawal of funding. This risk is particularly severe if the promised payoff of the claims is not state contingent, as such payoff structures create strong preemption motives for individual investors in the event of bad news about the project's outlook.

Here we have shown that such roll over risk, which in our model arises from a maturity mismatch due to short term debt financing of a long term project with risky payoffs, can improve investment decisions by making investors' actions more informative. We have seen that the ability to learn about the risky payoffs of the project by observing other investors' funding decisions creates an option value for waiting in order to gather additional information. This informational externality induces investors to roll over their debt rather than to act on private information about payoffs. The informational content of funding decisions is reduced. Rewarding the early withdrawal of funding through preferential allocation of liquidation revenue counteracts this force. This is what short term debt with its fixed face value achieves in our model. However if these preemption motives are too strong, investors exit immediately upon receiving bad news and nothing can be learned from other investors' funding decisions. We have derived the optimal share of short term debt to counterbalance such informational and payoff externalities. We have seen that if the private information received by investors is not sufficiently precise, the optimal share of short term debt financing is zero. The opportunity cost of liquidating a profitable project outweighs the option value created by the ability to observe funding decisions based on private information. For sufficiently precise private information about the project's payoff, the optimal share of short term debt is strictly positive. The preemption motive it creates for investors facilitates information revelation and enables the efficient liquidation of unprofitable investments. Optimal short term debt shares are found to decrease in the precision of the private information re-

ceived by investors and the liquidation value of the project. Better informed investors need weaker incentives to act on their private information. Equally, incentives to withdraw funding prematurely can be weakened if the opportunity costs of liquidation are low.

Appendix A

Appendix to Chapter 2

A1 Monopolistic Price Discrimination with Non-Common Priors

The following section closely follows Armstrong [Armstrong, 1996] who analyzes a very similar setting in which customers differ in their probability distributions for future demand, but the firm and customers have identical priors. Armstrong solves for the monopolist's optimal menu of tariffs under Assumption 1.

The monopolist offers his customers a menu of price schedules $\{P(q, j)\}_{j \in J}$ before these customers know their type θ . Customers will therefore have to forecast their type and thus their demand for the good in order to pick the individually optimal price schedule $P(q, j)$ from the offered menu J .

Consumers differ in their prior over the ex-post type θ . In particular the prior of a consumer of ex-ante type α is the distribution function $F(\theta, \alpha)$. The firm has a prior $F(\theta, \hat{\alpha})$ over the ex-post types θ which is identical for each of its customers. It cannot observe the ex-ante types α but knows their population distribution $G(\alpha)$, where $\alpha \in [\alpha_L, \alpha_H]$.

The firm chooses a menu of price schedules $\{P(q, j)\}$ that maximizes expected profits given that each customer picks the price schedule which maximizes her expected utility (where expectations are taken over the respective priors). Appealing to the revelation principle the problem is to maximize profits over a menu of price schedules

$\{P(q(\theta, \alpha), \alpha)\}$, such that these schedules are both ex-ante and ex-post incentive compatible, and that they are ex-ante individually rational.

A consumer's surplus with ex-post type θ , who has chosen the price schedule $P(q, \alpha)$ will be defined as

$$s(\theta, \alpha) = \max_{q \geq 0} u(q, \theta) - P(q, \alpha)$$

The ex-post implementability of each price schedule $P(q, \alpha)$ is ensured by the usual Envelope condition $s_\theta(\theta, \alpha) = u_\theta(q(\theta, \alpha), \theta)$ and the requirement that $q(\theta, \alpha)$ be non-decreasing in θ .

These conditions allow us to write the expected utility of a type α consumer from choosing the tariff $P(q, \tilde{\alpha})$ as

$$v(\alpha, \tilde{\alpha}) = s(\theta_L, \tilde{\alpha}) + \int_{\theta_L}^{\theta_H} u_\theta(q(\theta, \tilde{\alpha}), \theta) [1 - F(\theta, \alpha)] d\theta \quad (\text{A.1})$$

Define the maximum of this function as

$$V(\alpha) = \max_{\alpha_L \leq \tilde{\alpha} \leq \alpha_H} v(\alpha, \tilde{\alpha}) \quad (\text{A.2})$$

Then to implement the menu of tariffs we need to ensure that $V(\alpha)$ is maximized at $\alpha = \tilde{\alpha}$. This yields a second Envelope condition

$$V'(\alpha) = - \int_{\theta_L}^{\theta_H} u_\theta(q(\theta, \alpha), \theta) F_\alpha(\theta, \alpha) d\theta \geq 0 \quad (\text{A.3})$$

The fact that $V'(\alpha)$ is non-negative follows from the the condition that $u_\theta > 0$ and $F_\alpha \leq 0$ in case of Assumption 1.

Under Assumption 2 it follows from the fact that $q(\theta, \alpha)$ is non-decreasing in θ , the single-crossing property $u_{q\theta} > 0$, and the conditions imposed on the integral of F_α .

Now as $V(\alpha)$ is weakly increasing in α , if the participation constraint is satisfied

for the lowest type α_L , it is necessarily satisfied for all other $\alpha > \alpha_L$. Thus it is optimal for the firm to set $V(\alpha_L)$ equal to the outside option which we will normalize to zero, i.e. $V(\alpha_L) = 0$.

Therefore the rent of a type α consumer under any incentive compatible scheme will be

$$V(\alpha) = - \int_{\alpha_L}^{\alpha_H} \int_{\theta_L}^{\theta_H} u_{\theta}(q(\theta, \tilde{\alpha}), \theta) F_{\alpha}(\theta, \tilde{\alpha}) d\theta d\tilde{\alpha} \quad (\text{A.4})$$

and by (A.1)

$$s(\theta_L, \alpha) = V(\alpha) - \int_{\theta_L}^{\theta_H} u_{\theta}(q(\theta, \alpha), \theta) [1 - F(\theta, \alpha)] d\theta \quad (\text{A.5})$$

Lemma 5. (*Armstrong (1996)*) *Under Assumption 1, if the function $s(\theta_L, \alpha)$ in (A.1) is given by (A.5), then the type α consumer will choose $\tilde{\alpha} = \alpha$ in (A.1) provided that $q(\theta, \alpha)$ is (weakly) increasing in α .*

Proof. Using the expression in (A.1) and differentiating with respect to $\tilde{\alpha}$ yields

$$v_{\tilde{\alpha}}(\alpha, \tilde{\alpha}) = \int_{\theta_L}^{\theta_H} u_{q\theta}(q(\theta, \tilde{\alpha}), \theta) q_{\alpha}(\theta, \tilde{\alpha}) [F(\theta, \tilde{\alpha}) - F(\theta, \alpha)] d\theta$$

Under Assumption 1, $v(\alpha, \tilde{\alpha})$ is increasing for all $\tilde{\alpha} < \alpha$ and decreasing for all $\tilde{\alpha} > \alpha$ as long as $q_{\alpha} \geq 0$. Thus a sufficient condition for ex-ante implementability under Assumption 1 is that $q(\theta, \alpha)$ be weakly increasing in α . \square

Lemma 6. *Under Assumption 2 and 3, if the function $s(\theta_L, \alpha)$ in (A.1) is given by (A.5), then the type α consumer will choose $\tilde{\alpha} = \alpha$ in (A.1) provided that $q(\theta, \alpha)$ is weakly decreasing in α for all $\theta \leq z$, and weakly increasing in α for all $\theta \geq z$.*

Proof. To see this first look at the case $\tilde{\alpha} < \alpha$. In this case $F(\theta, \tilde{\alpha}) - F(\theta, \alpha)$ is less or equal to zero for all $\theta \leq z$ and greater or equal to zero for all $\theta \geq z$. Thus under the assumptions imposed on q_{α} , $q_{\alpha}(\theta, \tilde{\alpha})[F(\theta, \tilde{\alpha}) - F(\theta, \alpha)]$ will be greater or equal to zero on $\theta \in [\theta_L, \theta_H]$, and therefore $v(\alpha, \tilde{\alpha})$ will be weakly increasing in $\tilde{\alpha}$ for all $\tilde{\alpha} < \alpha$. The reverse holds true for $\tilde{\alpha} > \alpha$, i.e. $v(\alpha, \tilde{\alpha})$ will be weakly decreasing on

this interval. It follows that $v(\alpha, \tilde{\alpha})$ reaches its maximum at $\tilde{\alpha} = \alpha$. \square

Firm's profits from a customer of type α are

$$\int_{\theta_L}^{\theta_H} [u(q(\theta, \alpha), \theta) - C(q(\theta, \alpha))] f(\theta, \hat{\alpha}) d\theta - \int_{\theta_L}^{\theta_H} u_\theta(q(\theta, \alpha), \theta) [1 - F(\theta, \hat{\alpha})] d\theta - s(\theta_L, \alpha)$$

Add and subtract $V(\alpha)$ from this expression to get

$$\int_{\theta_L}^{\theta_H} \left\{ u(q(\theta, \alpha), \theta) + u_\theta(q(\theta, \alpha), \theta) \left[\frac{F(\theta, \hat{\alpha}) - F(\theta, \alpha)}{f(\theta, \hat{\alpha})} \right] - C(q(\theta, \alpha)) \right\} f(\theta, \hat{\alpha}) d\theta - V(\alpha)$$

Total profits are

$$\pi = \int_{\alpha_L}^{\alpha_H} \left\{ \int_{\theta_L}^{\theta_H} \left[u(q(\theta, \alpha), \theta) + u_\theta(q(\theta, \alpha), \theta) \left[\frac{F(\theta, \hat{\alpha}) - F(\theta, \alpha)}{f(\theta, \hat{\alpha})} \right] - C(q(\theta, \alpha)) \right] f(\theta, \hat{\alpha}) d\theta - V(\alpha) \right\} dG(\alpha)$$

Integrating by parts using (A.3) we have

$$\int_{\alpha_L}^{\alpha_H} V(\alpha) dG(\alpha) = - \int_{\alpha_L}^{\alpha_H} \int_{\theta_L}^{\theta_H} u_\theta(q(\theta, \alpha), \theta) F_\alpha(\theta, \alpha) [1 - G(\alpha)] d\theta d\alpha$$

Substituting this expression back into the firm's profits yields

$$\int_{\alpha_L}^{\alpha_H} \int_{\theta_L}^{\theta_H} \Lambda^m(q(\theta, \alpha), \theta, \alpha) f(\theta, \hat{\alpha}) g(\alpha) d\theta d\alpha \quad (\text{A.6})$$

where

$$\Lambda^m(q, \theta, \alpha) = u(q, \theta) + u_\theta(q, \theta) \left[\frac{F(\theta, \hat{\alpha}) - F(\theta, \alpha)}{f(\theta, \hat{\alpha})} \right] - C(q) + u_\theta(q, \theta) \frac{F_\alpha(\theta, \alpha)}{f(\theta, \hat{\alpha})} \left(\frac{1 - G(\alpha)}{g(\alpha)} \right) \quad (\text{A.7})$$

To maximize profits, we can maximize Λ^m for each combination of (θ, α) pointwise

with respect to q , i.e.

$$q(\theta, \alpha) = \arg \max_{q \geq 0} \Lambda^m(q, \theta, \alpha) \quad (\text{A.8})$$

If Λ^m is strictly quasi-concave in q , the optimal quantity schedule is given by the first order condition

$$\Lambda_q^m(q(\theta, \alpha), \theta) = 0 \quad (\text{A.9})$$

If furthermore Λ^m is supermodular in (q, θ) , then $q(\theta, \alpha)$ will be (weakly) increasing in θ .

For optimality we also require $q_\alpha(\theta, \alpha) \geq 0$ under Assumption 1. Under Assumptions 2 and 3 we require that $q_\alpha(\theta, \alpha) \leq 0$ for all $\theta < z$ and $q_\alpha(\theta, \alpha) \geq 0$ for all $\theta > z$.

The quantity schedule defined by (A.9) deviates from the efficient quantity schedule which equates marginal utility to marginal cost, i.e. $u_q(q(\theta, \alpha), \theta) = C'(q(\theta, \alpha))$, in two terms. The term $u_{q\theta}[F(\theta, \hat{\alpha}) - F(\theta, \alpha)]/f(\theta, \hat{\alpha})$ arises from the firm's attempt to create *fictitious surplus*. The second source of inefficiency is the usual distortion imposed on low α types in order to extract more surplus from high α types.

While the first distortion disappears for the fully rational type $\hat{\alpha}$, the second only disappears for the highest type α_H . This obviously implies that the quantity schedules $q(\theta, \alpha)$ will be inefficient for all types α , unless $\alpha_H = \hat{\alpha}$ in which case there will be no distortions at the top.

Appendix B

Appendix to Chapter 3

B.0.1 Calibration of the Transaction Tax

To simulate prices and trades in our model with transaction taxes we first need to quantify the size of the potential upward movement in the fundamental value λ_h as, unlike in the model without taxes, this quantity enters the calculations for the equilibrium thresholds.

Unlike in our theoretical model, observed price changes for our stock across days will be due to two kinds of events. Firstly, those that were private information of traders and were revealed to the market during the course of trading. These changes are what λ_h and $\lambda_l = \delta/(1 - \delta)\lambda_h$ in our model refer to. Secondly, there will be price changes that are due to events that became public information during or after the trading day without previously having been private information of any of our traders. To get a measure of λ_h we need to “clean” stock prices of such public information events. To do so we will perform a variance decomposition suggested by Hasbrouck [Hasbrouck, 1991] that allows us to decompose stock price changes into a trade-correlated component, interpreted as the component driven by private information, and a trade-uncorrelated component.

We start by assuming that the difference between opening and closing price of the security on any given trading day d , ΔP_d , consists of two components : a private com-

ponent X_d and a public component Y_d . Furthermore let us assume that X_d and Y_d are independent random variables. Then we have

$$\text{Var}(\Delta P_d) = \text{Var}(X_d + Y_d) = \text{Var}(X_d) + \text{Var}(Y_d)$$

We know the theoretical distribution of X_t . We have

$$X_t = \begin{cases} \lambda_h & \text{with probability } \alpha\delta \\ 0 & \text{with probability } 1 - \alpha \\ \left(\frac{\delta}{1-\delta}\right) \lambda_h & \text{with probability } \alpha(1 - \delta) \end{cases}$$

From the martingale property of prices we have $\mathbb{E}(X_t) = 0$. It follows that

$$\text{Var}(X_t) = \alpha\delta\lambda_h^2 + \alpha(1 - \delta) \left(\frac{\delta}{1 - \delta}\right)^2 \lambda_h^2 = \left(\frac{\alpha\delta}{1 - \delta}\right) \lambda_h^2$$

We now perform Hasbrouck's variance decomposition [Hasbrouck, 1991] to determine what fraction of the total variance of price changes $\text{Var}(\Delta P_d)$ is due to private information. To do so we use the intraday quotes and trades data for the Ashland stock described in the Data Section, where trades are signed using the Lee-Ready algorithm. We estimate Hasbrouck's vector autoregressive system (using a lag length of 5 trading periods) for the evolution of quote revisions and signed trades. From our estimates we recover Hasbrouck suggested measure of relative trade informative $R_w^2 = \sigma_{w,x}^2 / \sigma_w^2$ (see p.577 [Hasbrouck, 1991]). We have

$$\sigma_{w,x}^2 = 1.0344 \times 10^{-7}, \quad \sigma_w = 6.0169 \times 10^{-7} \quad \text{and} \quad R_w^2 = 0.1705$$

Thus this decomposition suggests that 17.05% of the variance of intraday price changes $\text{Var}(\Delta P_t)$ is due to private information, that is

$$\text{Var}(X_d) = R_w^2 \text{Var}(\Delta P_d) = 0.1705 \text{Var}(\Delta P_d)$$

Using the difference between daily closing and opening prices of Ashland Inc. for our sample period we estimate the standard deviation of the daily price changes to be $\sigma(\Delta P_t) = 0.3691$. Our estimates for the parameters are $\alpha = 0.2803$ and $\delta = 0.6238$.

The variance decomposition together with the previously obtained expression for the variance of X_t now allows us to get a value for λ_h as follows

$$\lambda_h = \sigma(\Delta P_d) \sqrt{R_w^2 \left(\frac{1 - \delta}{\alpha \delta} \right)} = 0.2236$$

In our simulations, we simulate the effect of a 0.1%, 0.2% and 0.5% financial transaction tax. As our tax ρ applies to one unit of the stock traded at a time, we calculate ρ as a percentage of the average opening price of Ashland Inc. over our sample period which is $\bar{P} = 34.13$. The value of ρ/λ_h for different tax rates, an expression which is needed for calculation of the equilibrium trading thresholds of our model, is then given in Table B.1.

<i>tax</i>	0.1 %	0.2 %	0.5 %
ρ/λ_h	0.1526	0.3053	0.7632

Table B.1: Calibration of FTT

We see that for tax rates in the suggested range for the FTT, our estimations suggest that for Ashland Inc. the transaction costs caused by such a tax would be between 15% (0.01% tax) and 75% (0.05% tax) of the maximum possible gain from private information (λ_h).

B.0.2 Informational Cascades

To lighten notation define as $\hat{\alpha}_t = \Pr(V^d \neq v^{d-1} | \mathcal{H}_t)$ the posterior probability concerning the occurrence of an information event where $\{\mathcal{H}_t\}$ is the filtration induced by the trading process. Also define as $\hat{q}_t = \Pr(V^d = v_h^d | V^d \neq v^{d-1}, \mathcal{H}_t)$ the posterior probability of a good event given that an informational event has occurred. α_t and q_t denote realizations of these random variables, that is posteriors for a given trading history $h_t \in \mathcal{H}_t$.

Lemma 1 $\hat{\alpha}_t$ converges almost surely to $I_{\{V^d \neq v^{d-1}\}}$.

Sketch of Proof The proof follows standard steps in establishing convergence for Bayesian learning: The posterior probability $\hat{\alpha}_t$ is a martingale with respect to the filtration $\{\mathcal{H}_t\}$ induced by the trading process. As posterior probabilities are bounded, we know by the Martingale Convergence Theorem that the posterior probability will converge almost surely to some random variable Y_∞ . As the probability of observing a trade differs across type of event day, it has to be the case that $Y_\infty = I_{\{V^d \neq v^{d-1}\}}$.

Proposition 7 If $\rho < \left(\frac{2\tau\delta}{1-\tau|2\delta-1|}\right) \lambda_h$, on an event day an informational cascade occurs only if $\Pr(V^d = v_h^d | h_t^d)$ has reached either a lower threshold r^* or an upper threshold $1 - r^*$, where r^* is the smallest root of the quadratic equation

$$\frac{2(1-r)p\tau}{1-(2r-1)\tau} = (1-\delta) \frac{\rho}{\lambda_h}.$$

The root r is increasing in the transaction tax ρ and decreasing in τ , δ , and λ_h .

If $\rho > \left(\frac{2\tau\delta}{1-\tau|2\delta-1|}\right) \lambda_h$, in an event day, informed traders never trade for all t . The probability of a good event converges almost surely to $\Pr(V^d = v_h^d | h_t^d) = \delta$.

Proof Informed traders do not trade after trading history h_t if (and only if) the following two inequalities hold:

$$\mathbb{E}(V^d | S_t = 1, h_t) < a_t + \rho$$

$$\mathbb{E}(V^d | S_t = 0, h_t) > b_t - \rho,$$

that is an informed trader with the best possible signal ($S_t = 1$) does not want to buy, and an informed trader with the worst possible signal ($S_t = 0$) does not want to sell given ask and bid prices and the transaction tax. If informed traders do not trade, equilibrium ask and bid prices are equal to $\mathbb{E}(V^d | h_t)$. It follows that informed traders do not trade after history h_t if (and only if)

$$\mathbb{E}(V^d | S_t = 1, h_t) - \mathbb{E}(V^d | h_t) < \rho$$

$$\mathbb{E}(V^d | h_t) - \mathbb{E}(V^d | S_t = 0, h_t) < \rho.$$

Now consider the first trading opportunity on a given day. We have $\Pr(V^d = v_h^d | \emptyset) = \alpha \cdot \delta$ and $\mathbb{E}(V^d | \emptyset) = v^{d-1}$. Thus informed traders do not trade iff

$$\begin{aligned} & \left[\frac{(1 + \tau)\delta - (1 - \tau)(1 - \delta) \left(\frac{\delta}{1 - \delta} \right)}{(1 + \tau)\delta + (1 - \tau)(1 - \delta)} \right] \lambda_h < \rho \\ - & \left[\frac{(1 - \tau)\delta - (1 + \tau)(1 - \delta) \left(\frac{\delta}{1 - \delta} \right)}{(1 - \tau)\delta + (1 + \tau)(1 - \delta)} \right] \lambda_h < \rho. \end{aligned}$$

These inequalities are jointly satisfied iff $\rho > \left(\frac{2\tau\delta}{1 - \tau|2\delta - 1|} \right) \lambda_h \equiv \bar{\rho}$. Lastly, note that as informed traders do not trade, no matter what action is observed at the first trading opportunity the market maker's beliefs in the next period are $q_1 = \delta$ and therefore $\mathbb{E}(V^d | h_1) = v^{d-1}$ for any history h_1 . Again, informed traders will not trade at the second trading opportunity and so on.

Now consider taxes such that $\rho < \bar{\rho}$. By Lemma 1 we know that $\Pr(V^d = v_h^d | \mathcal{H}_t)$ converges almost surely to 0 on days without informational event. On days with informational event $\hat{\alpha}_t$ converges to 1 almost surely and consequently $\mathbb{E}(V^d | \mathcal{H}_t)$ and $\mathbb{E}(V^d | V^d \neq v^{d-1}, \mathcal{H}_t)$ converge almost surely to the same random variable.¹ Thus for almost all trading histories on event days there exists a T such that if $t > T$ informed traders do not trade if and only if

$$\begin{aligned} \mathbb{E}(V^d | S_t = 1, h_t) - \mathbb{E}(V^d | V^d \neq v^{d-1}, h_t) &< \rho \\ \mathbb{E}(V^d | V^d \neq v^{d-1}, h_t) - \mathbb{E}(V^d | S_t = 0, h_t) &< \rho. \end{aligned}$$

These two inequalities are equivalent to

$$\begin{aligned} \frac{2(1 - q_t)q_t\tau}{1 - (2q_t - 1)\tau} &< (1 - \delta) \frac{\rho}{\lambda_h} \\ \frac{2(1 - q_t)q_t\tau}{1 - (1 - 2q_t)\tau} &< (1 - \delta) \frac{\rho}{\lambda_h}. \end{aligned}$$

¹Convergence is guaranteed by the Martingale Convergence Theorem and the fact that $\Pr(V^d = v_h^d | \mathcal{H}_t)$ is a bounded martingale.

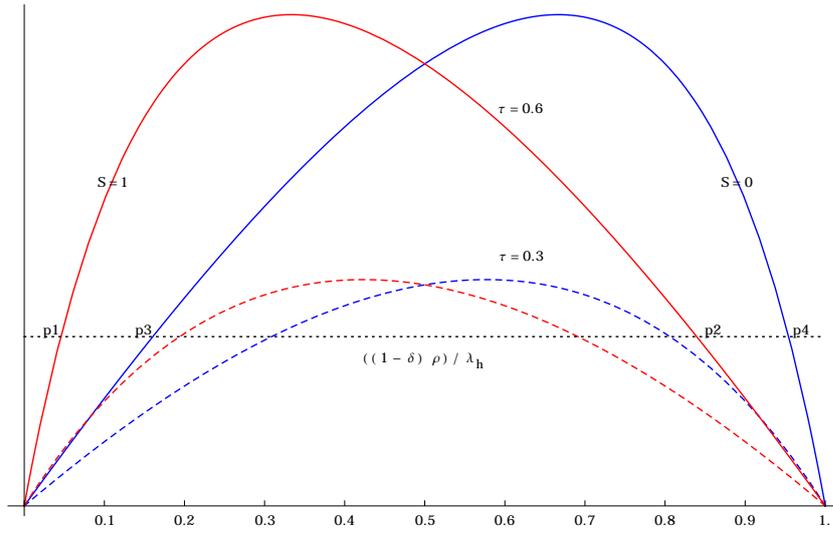


Figure B.1: Determination of Roots. Red curves correspond to signals $S = 1$, blue curves to $S = 0$. Dotted curves have $\tau = 0.3$, solid curves have $\tau = 0.6$. Intersections of the red curves with horizontal line occur at p_1 and p_2 , intersections of the blue curves with the horizontal line occur at p_3 and p_4 .

Figure B.1 illustrates these inequalities graphically. The first inequality is satisfied for values of q_t on the dotted line falling outside the interior of the red hyperbola. The second inequality is satisfied for q_t on the dotted line falling outside the interior of the blue hyperbola. p_1 and p_2 are the intersections of the dotted line with the red hyperbola. They are given by the two roots of the quadratic equation

$$\frac{2(1-p)p\tau}{1-(2p-1)\tau} = (1-\delta)\frac{\rho}{\lambda_h}$$

p_3 and p_4 are the intersections of the dotted line with the blue hyperbola. They are given by the two roots of the quadratic equation

$$\frac{2(1-p)p\tau}{1-(1-2p)\tau} = (1-\delta)\frac{\rho}{\lambda_h}$$

Note that if p solves the first equation, $1-p$ solves the second, so we have $p_1 = 1-p_4$ and $p_2 = 1-p_3$. Simple calculations show that the maximum of both hyperbolas is

$\frac{1-\sqrt{1-\tau^2}}{\tau}$ and that we have

$$(1-\delta)\frac{\bar{\rho}}{\lambda_h} \leq \frac{1-\sqrt{1-\tau^2}}{\tau}$$

Thus if $\rho < \bar{\rho}$ both equations will have two real roots, graphically the dotted line will intersect both hyperbolas.

It follows for almost all histories h_t and large enough t that informed traders do not trade if and only if $q_t \neq (p_1, p_2) \cup (1-p_2, 1-p_1)$. By the Martingale Convergence Theorem we know that \hat{q}_t converges almost surely. If informed traders trade, this posterior would change after observing a trade. Thus the posterior has to converge to a random variable X_∞ with $Pr(X_\infty \in (p_1, p_2) \cup (1-p_2, 1-p_1)) = 0$. ■

B.0.3 Estimation Results for Simulated Data

The following tables report estimates for simulated data generated by our basic 6 parameter model without transaction taxes. We simulated data for three sets of parameters and then estimated these parameters by maximum likelihood using the likelihood function derived in section 3.5. We maximised the likelihood function numerically using the Nelder-Mead algorithm as implemented by the Numerical Algorithm Group (NAG) for their Matlab toolbox. The following tables reports the best estimates we have obtained. For each set of parameters we have simulated data for two different seeds and for 220, 500, and 1000 trading days each having 149 trading opportunities per day.

	α	δ	μ	τ	ψ	ε
<i>simulation</i>	0.28	0.62	0.42	0.45	0.2	0.8
<i>220 days</i>	0.32	0.58	0.42	0.51	0.26	0.80
<i>500 days</i>	0.29	0.65	0.41	0.46	0.33	0.79
<i>1000 days</i>	0.29	0.59	0.42	0.43	0.17	0.80

Table B.2: *Estimates for Simulation 1, Seed 1.*

	α	δ	μ	τ	ψ	ε
<i>simulation</i>	0.28	0.62	0.42	0.45	0.2	0.8
<i>220 days</i>	0.28	0.52	0.44	0.42	0.16	0.80
<i>500 days</i>	0.28	0.59	0.43	0.40	0.23	0.79
<i>1000 days</i>	0.30	0.60	0.42	0.44	0.20	0.80

Table B.3: Estimates for Simulation 1, Seed 2.

	α	δ	μ	τ	ψ	ε
<i>simulation</i>	0.60	0.52	0.16	0.80	0.35	0.6
<i>220 days</i>	0.43	0.54	0.18	0.65	0.06	0.67
<i>500 days</i>	0.40	0.63	0.21	0.53	0.14	0.68
<i>1000 days</i>	0.56	0.50	0.16	0.78	0.37	0.59

Table B.4: Estimates for Simulation 2, Seed 1.

	α	δ	μ	τ	ψ	ε
<i>simulation</i>	0.60	0.52	0.16	0.80	0.35	0.6
<i>220 days</i>	0.44	0.54	0.17	0.99	0.09	0.68
<i>500 days</i>	0.51	0.60	0.16	0.78	0.21	0.66
<i>1000 days</i>	0.69	0.53	0.18	0.77	0.47	0.53

Table B.5: Estimates for Simulation 2, Seed 2.

	α	δ	μ	τ	ψ	ε
<i>simulation</i>	0.30	0.72	0.26	0.60	0.40	0.35
<i>220 days</i>	0.33	0.50	0.29	0.53	0.39	0.35
<i>500 days</i>	0.29	0.69	0.27	0.64	0.36	0.39
<i>1000 days</i>	0.30	0.69	0.26	0.62	0.39	0.36

Table B.6: Estimates for Simulation 3, Seed 1.

	α	δ	μ	τ	ψ	ε
<i>simulation</i>	0.30	0.72	0.26	0.60	0.40	0.35
<i>220 days</i>	0.33	0.73	0.24	0.59	0.37	0.36
<i>500 days</i>	0.32	0.80	0.26	0.53	0.44	0.30
<i>1000 days</i>	0.32	0.71	0.25	0.62	0.40	0.35

Table B.7: Estimates for Simulation 3, Seed 2.

Appendix C

Appendix to Chapter 4

C.0.4 Proofs

Proof of Proposition 1

Proof. We need to prove that liquidation only occurs in the first period and only does so with strictly positive probability for beliefs $\mu < \mu^* \equiv 2l/(1+l)$.

First consider a simplified setting in which liquidation can only occur in the first period. If the project is not liquidated by either party in the first period it will continue until maturity. In this case the planner's objective function is

$$\{\mu\lambda + (1-\mu)[1 - (1-\lambda)^2]\}l + [\mu(1-\lambda) + (1-\mu)(1-\lambda)^2]\mu'$$

where λ is the symmetric liquidation probability. By Bayes' Rule we have

$$\mu' = \frac{\mu}{\mu + (1-\mu)(1-\lambda)}$$

The objective function thus simplifies to

$$\lambda [(2-\lambda)(1-\mu) + \mu]l + (1-\lambda)\mu$$

which is strictly concave in λ . Here the optimal liquidation probability, which we

designate by $\lambda(\mu)$, is

$$\lambda(\mu) = \begin{cases} \frac{1}{2} \left[1 - \frac{\mu-l}{(1-\mu)l} \right] & \text{if } \mu < \mu^* \\ 0 & \text{otherwise.} \end{cases}$$

Now consider the Bellman Equation of the original planner problem (4.4). To simplify the analysis we subtract μ from both sides and define $\tilde{W}(\mu) = W(\mu) - \mu$. Together with $\mu = [1 - \lambda(1 - \mu)]\mu'$ which follows from Bayes' Rule (4.3) we obtain

$$\tilde{W}(\mu) = \max_{\lambda \in [0,1]} \left\{ \lambda [(2 - \lambda)(1 - \mu)l - (1 - l)\mu] + (1 - \lambda)(1 - \lambda[1 - \mu])\gamma\tilde{W}(\mu') \right\}$$

Suppose the policy of liquidating with probability $\lambda(\mu)$ in the initial period upon receiving a bad signal and never thereafter is also the optimal policy for the original problem. Then the value function corresponding to this policy, namely

$$G(\mu) = \begin{cases} \lambda(\mu) [(2 - \lambda(\mu))(1 - \mu)l - (1 - l)\mu] & \text{if } \mu < \mu^* \\ 0 & \text{otherwise} \end{cases}$$

should also satisfy the above Bellman Equation.

To see that this is the case we substitute G into the Bellman Equation and verify that it indeed solves the functional equation, that is

$$G(\mu) = \max_{\lambda \in [0,1]} \left\{ \lambda [(2 - \lambda)(1 - \mu)l - (1 - l)\mu] + (1 - \lambda)(1 - \lambda[1 - \mu])\gamma G\left(\frac{\mu}{1 - \lambda(1 - \mu)}\right) \right\}$$

The function inside the curly brackets is strictly concave in λ with first derivative given by

$$\begin{cases} (1 - \gamma)[(1 - \mu)l - (\mu - l) - 2\lambda(1 - \mu)l] & \text{for } \lambda < \lambda(\mu) \\ (1 - \mu)l - (\mu - l) - 2\lambda(1 - \mu)l & \text{for } \lambda \geq \lambda(\mu) \end{cases}$$

and second derivative

$$\begin{cases} -2(1-\gamma)(1-\mu)l < 0 & \text{if } \lambda < \lambda(\mu) \\ -2(1-\mu)l < 0 & \text{if } \lambda \geq \lambda(\mu) \end{cases}$$

For $\mu < \mu^*$ the maximum is achieved by an interior solution characterized by the first order condition

$$(1-\mu)l - (\mu-l) - 2\lambda(1-\mu)l = 0 \Leftrightarrow \lambda(\mu) = \frac{1}{2} \left[1 - \frac{\mu-l}{(1-\mu)l} \right]$$

For $\mu \geq \mu^*$ the maximum is achieved by the corner solution $\lambda(\mu) = 0$. Hence G indeed satisfies the Bellman Equation. \square

Appendix D

Note on Cojoint Work

The research presented in chapter 3 of this thesis, *Financial Transaction Taxes and the Informational Efficiency of Financial Market: A Structural Estimation*, is based on joint research with Marco Cipriani and Antonio Guarino. Each co-author has contributed equally to this project. I thank both of them for the fruitful cooperation.

Bibliography

- [Allen and Gale, 2002] Allen, F. and Gale, D. (2002). Optimal financial crises. *Journal of Finance*, 53(4):1245–1284.
- [Armstrong, 1996] Armstrong, M. (1996). Nonlinear Pricing with Imperfectly Informed Consumers.
- [Armstrong and Vickers, 2001] Armstrong, M. and Vickers, J. (2001). Competitive Price Discrimination. *RAND Journal of Economics*, 32(4):579–605.
- [Avery and Zemsky, 1998] Avery, C. and Zemsky, P. (1998). Multidimensional Uncertainty and Herd Behavior in Financial Markets. *American Economic Review*, 88(4):724–48.
- [Berndt et al., 1974] Berndt, E. K., Hall, B. H., Hall, R. E., and Hausman, J. A. (1974). Estimation and inference in nonlinear structural models. In *Annals of Economic and Social Measurement*, volume 4, pages 103–116.
- [Bikhchandani et al., 1998] Bikhchandani, S., Hirshleifer, D., and Welch, I. (1998). Learning from the behavior of others: Conformity, fads, and informational cascades. *Journal of Economic Perspectives*, 12(3):151–170.
- [Bliss, 1988] Bliss, C. (1988). A Theory of Retail Pricing. *Journal of Industrial Economics*, 36(4):375–91.
- [Chamley and Gale, 1994] Chamley, C. and Gale, D. (1994). Information revelation and strategic delay in a model of investment. *Econometrica*, 62(5):1065–1085.

- [Chung et al., 2005] Chung, K. H., Li, M., and McInish, T. H. (2005). Information-based trading, price impact of trades, and trade autocorrelation. *Journal of Banking & Finance*, 29(7):1645–1669.
- [Cipriani and Guarino, 2008] Cipriani, M. and Guarino, A. (2008). Herd Behavior and Contagion in Financial Markets. *The B.E. Journal of Theoretical Economics*, 8(1):1–56.
- [Cipriani and Guarino, 2014] Cipriani, M. and Guarino, A. (2014). Estimating a Structural Model of Herd Behavior in Financial Markets. *American Economic Review*, 104(1):224–51.
- [Courty and Li, 2000] Courty, P. and Li, H. (2000). Sequential Screening. *Review of Economic Studies*, 67(4):697–717.
- [DellaVigna and Malmendier, 2006] DellaVigna, S. and Malmendier, U. (2006). Paying Not to Go to the Gym. *American Economic Review*, 96(3):694–719.
- [Diamond and Dybvig, 1983] Diamond, D. and Dybvig, P. (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy*, 91(3):401–419.
- [Dominitz and Manski, 2011] Dominitz, J. and Manski, C. F. (2011). Measuring and interpreting expectations of equity returns. *Journal of Applied Econometrics*, 26(3):352–370.
- [Dupont and Lee, 2007] Dupont, D. and Lee, G. (2007). Effects of Securities Transaction Taxes on Depth and Bid-Ask Spread. *Economic Theory*, 31(2):393–400.
- [Easley et al., 1997] Easley, D., Kiefer, N. M., and O’Hara, M. (1997). One Day in the Life of a Very Common Stock. *Review of Financial Studies*, 10(3):805–35.
- [Eliaz and Spiegel, 2006] Eliaz, K. and Spiegel, R. (2006). Contracting with Diversely Naive Agents. *Review of Economic Studies*, 73(3):689–714.
- [Frisell, 2003] Frisell, L. (2003). On the interplay of informational spillovers and payoff externalities. *RAND Journal of Economics*, 34(3):582–592.

- [Gale, 1996] Gale, D. (1996). What have we learned from social learning? *European Economic Review*, 40(3):617–628.
- [Glosten, 2009] Glosten, L. R. (2009). The Welfare Cost of Informed Trade.
- [Glosten and Milgrom, 1985] Glosten, L. R. and Milgrom, P. R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14(1):71–100.
- [Gu, 2011] Gu, C. (2011). Herding and bank runs. *Journal of Economic Theory*, 146(1):163–188.
- [Gul and Lundholm, 1995] Gul, F. and Lundholm, R. (1995). Endogenous timing and the clustering of agents' decisions. *Journal of Political Economy*, 103(5):1039–1066.
- [Habermeier and Kirilenko, 2001] Habermeier, K. F. and Kirilenko, A. (2001). Securities Transaction Taxes and Financial Markets. IMF Working Papers 01/51, International Monetary Fund.
- [Hasbrouck, 1991] Hasbrouck, J. (1991). The Summary Informativeness of Stock Trades: An Econometric Analysis. *Review of Financial Studies*, 4(3):571–95.
- [Hasbrouck, 2004] Hasbrouck, J. (2004). Working with the WRDS (TAQ and CRSP) Data.
- [Kupiec, 1996] Kupiec, P. H. (1996). Noise traders, excess volatility, and a securities transactions tax. *Journal of Financial Services Research*, 10(2):115–129.
- [Lee and Ready, 1991] Lee, C. M. C. and Ready, M. J. (1991). Inferring Trade Direction from Intraday Data. *Journal of Finance*, 46(2):733–46.
- [Lichtenstein et al., 1982] Lichtenstein, S., B., F., and Phillips, L. D. (1982). Calibration of probabilities: The state of the art to 1980. In Kahneman, D., Slovic, P., and Tversky, A., editors, *Judgment under Uncertainty*, pages 306–334. Cambridge University Press.

- [Mannaro et al., 2008] Mannaro, K., Marchesi, M., and Setzu, A. (2008). Using an artificial financial market for assessing the impact of Tobin-like transaction taxes. *Journal of Economic Behavior & Organization*, 67(2):445–462.
- [Manski, 2004] Manski, C. F. (2004). Measuring Expectations. *Econometrica*, 72(5):1329–1376.
- [McClelland and Bolger, 1994] McClelland, A. and Bolger, F. (1994). The calibration of subjective probability: Theories and models 1980-94. In Wright, G. and Ayton, P., editors, *Subjective probability*. John Wiley & Sons.
- [Miravete, 2004] Miravete, E. (2004). Are all those Calling Plans Really Necessary? The Limited Gains From Complex Tariffs.
- [Murto and Välimäki, 2011] Murto, P. and Välimäki, J. (2011). Learning and information aggregation in an exit game. *The Review of Economic Studies*, 78(4):1426–1461.
- [Pellizzari and Westerhoff, 2009] Pellizzari, P. and Westerhoff, F. (2009). Some effects of transaction taxes under different microstructures. *Journal of Economic Behavior & Organization*, 72(3):850–863.
- [Rochet and Stole, 2002] Rochet, J.-C. and Stole, L. A. (2002). Nonlinear Pricing with Random Participation. *Review of Economic Studies*, 69(1):277–311.
- [Rosenthal and Thomas, 2012] Rosenthal, D. W. and Thomas, N. D. M. (2012). Transact taxes in a price maker/taker market. MPRA Paper 40556, University Library of Munich, Germany.
- [Smith and Sorensen, 2000] Smith, L. and Sorensen, P. (2000). Pathological Outcomes of Observational Learning. *Econometrica*, 68(2):371–398.
- [Song and Zhang, 2005] Song, F. M. and Zhang, J. (2005). Securities transaction tax and market volatility*. *The Economic Journal*, 115(506):1103–1120.

- [Weeds, 2002] Weeds, H. (2002). Strategic delay in a real options model of R&D competition. *The Review of Economic Studies*, 69(3):729–747.