

# Three Essays in Microeconomic Theory

by

Dimitrios Vayanos

Diplome d'Ingenieur de l'Ecole Polytechnique (1988)

Submitted to the Department of Economics  
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1993

©Dimitrios Vayanos.

The author hereby grants to MIT permission to reproduce and to  
distribute copies of this thesis document in whole or in part.

Author .....  
Department of Economics  
May 10, 1993

Certified by .....  
Jean Tirole  
Professor of Economics  
Thesis Supervisor

Accepted by .....  
Richard Eckaus  
Chairman, Departmental Committee on Graduate Students

MASSACHUSETTS INSTITUTE  
OF TECHNOLOGY

JUN 09 1993

LIBRARIES

# Three Essays in Microeconomic Theory

by

Dimitrios Vayanos

Submitted to the Department of Economics  
on May 10, 1993, in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy

## Abstract

The first two chapters of this dissertation study models of markets for financial assets which are not perfectly liquid. Trade in these markets is costly; agents would strictly prefer not to trade than buying one unit and then selling one unit. Agents' behavior and the prices of the assets are thus not determined according to the "frictionless" Walrasian model.

Chapter 1 considers a market with some large traders whose trades affect prices. These traders, willing to reduce the price impact of their trades, will behave strategically, spreading trades over time. Chapter 1 studies their behavior and shows that when most of the trading volume comes from smaller non-strategic traders, large traders trade quickly and do so at a decreasing rate over time. Otherwise they trade slowly at an increasing and then decreasing rate. If large traders account for most of the trading volume, risk-sharing is much more inefficient than what static double auction models predict (the gains from trade are realized very slowly) and the outcome departs significantly from the Walrasian outcome. Chapter 1 concludes by studying how large traders' behavior and welfare change when the market clears more frequently, and whether large traders are better off if the information regarding their trading needs is publicly known.

Chapter 2 (co-authored with Jean-Luc Vila) studies the impact of (exogenous) transactions costs on the prices of financial assets. It is assumed that there are two assets, the "liquid" asset and the "illiquid" asset. The liquid asset is traded without costs, while trading the illiquid asset entails proportional costs. Transactions costs naturally create a "liquidity premium": the illiquid asset becomes cheaper than the liquid asset. The main result of chapter 2 is that with transactions costs the liquid asset becomes more expensive (in the absolute) while the illiquid asset may become cheaper or more expensive. Chapter 2 also shows that the effects of transactions costs on the liquidity premium and on the price of the liquid asset are stronger if the fraction of illiquid assets in the economy is higher.

Chapter 3 (co-authored with Diego Rodriguez) is an essay on the theory of the firm. It considers a multidivisional firm whose divisions make substitute (and potentially competing) products and studies whether top management would find it optimal to decentralize decision-making authority to the divisions. It is shown that

when a division is given the right to market its product to customers of the other division, its manager has an incentive to increase product quality. At the same time divisions locate their products too close together. Top management may or may not want to restrict competition, by controlling the access of divisions to customers, but even if competition is desirable, it may not be credible due to top management's temptation to intervene "ex-post".

Thesis Supervisor: Jean Tirole

Title: Professor of Economics

**To my parents and my sisters**



# Contents

<b>Acknowledgments</b>	<b>9</b>
<b>Introduction</b>	<b>11</b>
References . . . . .	15
<b>1 A Dynamic Model of an Imperfectly Competitive Bid-Ask Market</b>	<b>16</b>
1.1 Introduction . . . . .	16
1.2 The Model . . . . .	23
1.3 The Two-Period Case . . . . .	26
1.3.1 Some Preliminary Analysis . . . . .	26
1.3.2 Equilibrium . . . . .	29
1.3.3 The Impact of Noise . . . . .	32
1.4 The Multiperiod Case . . . . .	34
1.4.1 Some Preliminary Analysis . . . . .	34
1.4.2 Equilibrium . . . . .	37
1.4.3 Trading Patterns . . . . .	38
1.5 Allocative Efficiency . . . . .	40
1.6 The Effects of More Frequent Market Clearing . . . . .	44
1.7 The Case where Large Traders' Hedging Demands are Public Information	46
1.8 Conclusion . . . . .	48
1.A Appendix: Proofs of Results of Sections 1.3 and 1.4 . . . . .	50
1.A.1 Notation and Preliminary Derivations . . . . .	50

1.A.2	Derivation of Traders' Value Functions and Characterization of Equilibrium . . . . .	51
1.A.3	A Numerical Algorithm for Solving the Model . . . . .	62
1.A.4	An Alternative Characterization of Equilibrium . . . . .	63
1.A.5	The Two-Period Case . . . . .	70
1.B	Appendix: Determination of Total Welfare . . . . .	71
1.B.1	Large traders' welfare . . . . .	72
1.B.2	Small Traders' Welfare and Total Welfare . . . . .	78
1.C	Appendix: The Public Information Case . . . . .	79
1.C.1	Proof of Proposition 1.7.1 . . . . .	79
1.C.2	Welfare . . . . .	82
	References . . . . .	85
<b>2</b>	<b>Equilibrium Interest Rate and Liquidity Premium Under Proportional Transactions Costs</b>	<b>91</b>
2.1	Introduction . . . . .	91
2.2	The Model . . . . .	97
2.3	The No Transactions Costs Case . . . . .	100
2.3.1	The consumer's problem . . . . .	100
2.3.2	Equilibrium . . . . .	101
2.4	Transactions Costs and Assets' Returns . . . . .	102
2.4.1	The consumer's problem . . . . .	104
2.4.2	Equilibrium . . . . .	108
2.4.3	Comparative Statics . . . . .	116
2.5	Numerical Examples . . . . .	117
2.6	Conclusion . . . . .	118
2.A	Appendix: The No Transactions Costs Case . . . . .	120
2.B	Appendix: Proof of Proposition 2.4.1 . . . . .	121
2.C	Appendix: Proofs of Proposition 2.4.2 and Lemmas 2.4.3 and 2.4.4 . .	135
2.C.1	Proof of Lemma 2.4.3 . . . . .	135

2.C.2	Proof of Proposition 2.4.2 . . . . .	136
2.C.3	Proof of Lemma 2.4.4 . . . . .	138
2.D	Appendix: Proof of Lemma 2.4.5 . . . . .	140
2.E	Appendix: The Case $k=1$ . . . . .	141
	References . . . . .	144
<b>3</b>	<b>Decentralization and the Management of Competition</b>	<b>153</b>
3.1	Introduction . . . . .	153
3.2	The Model . . . . .	157
3.2.1	Supply and Demand . . . . .	157
3.2.2	Information Structure and Contracts . . . . .	158
3.3	The First Best . . . . .	163
3.4	The Decentralized Firm . . . . .	164
3.5	The Centralized Firm . . . . .	167
3.6	Comparison of Organizational Forms . . . . .	169
3.A	Appendix: Proof of Proposition 3.3.1 . . . . .	171
3.B	Appendix: Proof of Proposition 3.4.1 . . . . .	173
3.C	Appendix: Proof of Proposition 3.5.1 . . . . .	176
3.D	Appendix: Comparison of Organizational Forms . . . . .	177
	References . . . . .	179

# List of Figures

1-1	Trading patterns for noise = 30 . . . . .	88
1-2	Trading patterns for noise = 4 . . . . .	89
1-3	Trading patterns for noise = 0.1 . . . . .	90
2-1	Holdings of the liquid and illiquid assets . . . . .	148
2-2	Rates of return as functions of $k$ . . . . .	149
2-3	Rates of return as functions of $k$ . . . . .	150
2-4	Rates of return as functions of $k$ . . . . .	151
2-5	The graph of $g(t)$ . . . . .	152

# Acknowledgments

First of all, I want to thank my two main advisors, Jean Tirole and Jean-Luc Vila.

I am very grateful to Jean for many things. Because of his amazingly broad and deep knowledge of economics, he gave me excellent advice over a wide range of topics. The many hours that he spent discussing ideas with me and his very detailed comments to my drafts (which were actually longer than my drafts) definitely influenced my views about economics and greatly improved my research abilities. His encouragement and involvement in my progress, from the first day I came to MIT, was also very valuable to me.

I was also very lucky having met Jean-Luc. I not only benefited from his excellent advice on my work but I learnt a lot from him during our joint work (chapter 2 of this dissertation) as well. The innumerable discussions that we had were critical to the way I think about financial economics. I am also very grateful to Jean-Luc for being so encouraging and supportive at many stages of this dissertation.

I want to thank Drew Fudenberg, Oliver Hart and Peter Diamond for the courses they taught me and for the advice they very generously gave me. They stimulated my interest in many exciting areas of economics and I consider myself very lucky having interacted with such great economists.

Chi-fu Huang, Jean-Jacques Laffont, Lones Smith, Xavier Vives and Jiang Wang also gave me valuable advice during my dissertation years.

During these years I interacted with many of my classmates too. I spent many hours discussing ideas with Rabi Abraham, Fernando Branco, Glenn Ellison, Héddi Kallal, David Martimort, Walter Novaes, Kazu Ōhashi, Matthew Rabin, Diego Rodriguez, Lars Stole, Luigi Zingales and Jeff Zwiebel and I want to thank them for

this. My interaction with Diego was particularly fruitful and resulted in chapter 3 of this dissertation.

I thank Dominique Henriët, Claude Henry and Jean-Michel Lasry for getting me into economics.

I acknowledge financial support during the last year of my dissertation from the Olin foundation.

Finally I am especially grateful to my parents, Evangelos and Eleni, for the education they gave me and for motivating me to study abroad. I also want to thank them, as well as my sisters Eliana, Marina and Phivi (kouvali), for their continuing love and support.

# Introduction

A large body of research on financial markets assumes that these markets are perfectly liquid. Trade is costless in the sense that agents are indifferent between not trading, and buying one unit and then selling one unit. This assumption is significantly violated, however, in many financial markets.

Consider, for instance, the sale of a large block of stock. This sale will generally depress the stock price for two reasons. First, many agents simply do not know that a block is being sold and do not take the other side of the transaction. Second, even some of the agents who are aware of the sale do not know whether this sale is motivated by bad news about the stock's value or not. The seller of the block, willing to reduce its price impact, may incur additional costs such as searching for buyers and not completing the sale immediately. Finally, he has to pay transactions taxes and exchange fees.

With costly trade, agents' behavior and asset prices may be substantially different than what predicted by the Walrasian model of perfectly liquid markets. The first two chapters of this dissertation study how agents' behavior and asset prices are determined in illiquid markets.

Chapter 1 studies a market with some large traders whose trades affect prices. Its focus on the strategic behavior of large agents is shared with the literature on insider trading.<sup>1</sup> Papers in this literature have studied in detail the trading behavior of informed agents, but have considerably simplified the behavior of agents who trade because of non-informational reasons.<sup>2</sup> Chapter 1 focuses instead on the trading

---

<sup>1</sup>See Kyle (1985) and (1989) for seminal papers in this literature.

<sup>2</sup>There are some exceptions. See for instance Admati and Pfleiderer (1988).

behavior, in a dynamic setting, of agents who do not have an informational motive to trade but rather trade to hedge their positions.<sup>3</sup>

Large traders, willing to reduce the price impact of their trades will behave strategically and spread them over time. The rest of the market, knowing that large traders spread their trades over time, will use current prices to forecast their future trades as well as future prices. Large traders will, in turn, take this into account when choosing their trading strategy. Chapter 1 shows that when most of the trading volume comes from smaller non-strategic traders, large traders trade quickly at a decreasing rate over time. Otherwise they trade slowly at an increasing and then decreasing rate. These trading strategies are substantially different than the ones obtained in insider trading models. Chapter 1 also studies the efficiency of the market allocation (which is related to the speed at which large traders hedge their positions). It shows that when large traders account for most of the trading volume, risk sharing is much more inefficient than what static double auction models predict.<sup>4</sup>

Chapter 2 studies the impact of transactions costs on asset prices in a general equilibrium framework. It considers an infinite horizon economy with finitely-lived agents. Agents buy assets for life-cycle purposes. There are two otherwise identical assets: “liquid” assets which are traded without costs, and “illiquid” assets whose trade entails costs. The costs of trading the illiquid asset are exogenous to the model and, for simplicity, are assumed to be proportional to the quantity traded (they may be due to transactions taxes or exchange fees).

As it is well-known transactions costs create a “liquidity premium”: the illiquid asset becomes cheaper than the liquid asset. The main result of chapter 2 is that with transactions costs the liquid asset becomes more expensive (in the absolute). The price of the illiquid asset varies then in an ambiguous direction since it decreases

---

<sup>3</sup>Chapter 1 is also related to the literature on optimal portfolio rebalancing with transactions costs. (See, for instance, Dumas and Luciano (1991) and Fleming et al. (1992).) Indeed, papers in this literature also study optimal investor behavior in illiquid markets. These papers assume that illiquidity takes the form of proportional transactions costs (say a transactions tax) and consider an exogenous price process. By contrast, chapter 1 assumes that trading costs are quadratic (i.e. are market impact costs) and endogenizes the price process.

<sup>4</sup>See McAfee and McMillan (1987) and Wilson (1990) for surveys of the literature on double auctions.



relative to the price of the liquid asset, but the liquid asset becomes more expensive. There are cases where transactions costs actually *raise* the price of the illiquid asset. Chapter 2 also shows that the effects of transactions costs on the liquidity premium and on the price of the liquid asset are stronger if the fraction of the illiquid asset in the economy is higher. These results suggest that when studying how a change in transactions costs on a fraction of the assets in the economy<sup>5</sup> will affect the prices of these assets, one should not take the prices of the other assets as given, especially if the fraction of the first group of assets is high.

Chapter 3 is an essay on the theory of the firm. It studies how decision-making authority should be allocated in a multidivisional firm whose divisions make substitute (and potentially competing) products. In particular, should top management control the design and pricing of these products, thus centralizing decision making, or is it better that divisions decide which products to make and how to sell them? These questions are clearly of great concern to many large firms.

Chapter 3 considers a firm which is composed by top management and two divisions which design horizontally differentiated products, characterized by their "location" and their "quality", and then sell them to customers. It focuses, for simplicity, on the allocation of decision rights at the price setting stage. More precisely, top management can restrict the set of customers with whom a division can deal, thus prohibiting competition for customers, or it can allow both divisions to approach all customers.

It is shown that when top management decentralizes decision-making authority to the divisions, the latter, aiming also at "stealing" each other's customers, offer higher quality products. However, for the same reason they locate these products too close together. Top management may or may not want to restrict competition, by controlling the access of divisions to customers. Even if competition is desirable, it might not be credible due to top management's temptation to intervene "ex-post" and mandate the allocation of customers. Chapter 3 suggests that in this case top management would prefer to commit not to intervene by not collecting precise infor-

---

<sup>5</sup>Such as imposing transactions taxes in the stock market.

mation on customers' characteristics and thus being unable to allocate customers to divisions.

## References

- Admati, Anat R. and Paul Pfleiderer (1988), "*A Theory of Intraday Patterns: Volume and Price Variability*", *The Review of Financial Studies*, 1:3-40.
- Dumas, Bernard and Elisa Luciano (1991), "*An Exact Solution to the Portfolio Choice Problem Under Transactions Costs*", *Journal of Finance*, 46,2,June:577-595.
- Fleming, Wendell, Sanford Grossman, Jean-Luc Vila and Thaleia Zariphopoulou (1992), "*Optimal Portfolio Rebalancing with Transaction Costs*", mimeo, Massachusetts Institute of Technology.
- Kyle, Albert S. (1985), "*Continuous Auctions and Insider Trading*", *Econometrica*, 53:1315-1336.
- Kyle, Albert S. (1989), "*Informed Speculation and Imperfect Competition*", *Review of Economic Studies*, 56:517-556.
- McAfee, Preston R. and John McMillan (1987), "*Auctions and Bidding*", *Journal of Economic Literature*, 25:699-738.
- Wilson, Robert (1990), "*Two Surveys: Auctions and Entry Deterrence*", Technical Report No. 3, Stanford University.

# Chapter 1

## A Dynamic Model of an Imperfectly Competitive Bid-Ask Market

### 1.1 Introduction

In this chapter we study a dynamic model of a financial market. This market is imperfectly competitive: some traders are large and take into account their effect on prices.

Large traders play an increasingly important role in modern financial markets; institutional investors, large dealers and mutual fund managers are obvious examples. These agents execute large trades and, as it has widely been documented, their trades have a significant price impact.<sup>1</sup> These trades may be informationally motivated: large traders may trade in order to benefit from inside information about assets' fundamental values. Their trades thus affect prices since they signal some of this information to the rest of the market. The strategic behavior of large insiders has been very well studied since the pioneering work of Kyle (1985).

Large traders also have non-informational motives to trade: they may want to

---

<sup>1</sup>See Kraus and Stoll (1972), Holthausen, Leftwich and Mayers (1987), and Chan and Lakonishok (1992).

rebalance their portfolios for instance. Their trades may still affect prices if the market cannot distinguish them from informationally motivated trades. In addition, even if the market correctly does not attribute them to inside information, large non-informational trades can have a significant price impact. This will be the case if there is only a small number of (risk-averse) agents present in the market.<sup>2</sup> Indeed, by taking the other side of these trades, risk-averse market participants take large risky positions; they will do so only if the price changes in a way which is favorable to them.

It seems likely that large traders whose trades, although not informationally motivated affect prices (for one of the two reasons outlined above) will also behave strategically. In particular they will not find it optimal to complete their trades immediately, even though this is efficient for risk-sharing purposes: doing this would affect prices too much. They will find it more advantageous to spread their trades over time. The rest of the market, knowing that large traders follow such a strategy, will use current prices to forecast their future trades as well as future prices. Large traders will in turn, take this into account when choosing their trading strategy.

Chan and Lakonishok (1993) study the trading behavior of large investment management firms. Their finding that these firms trade in a correlated manner over time<sup>3</sup> implies that large orders are indeed split into smaller pieces, and is very consistent with the preceding discussion. They also find that the stocks in which these firms traded did not experience any abnormal returns after the firms completed their trades. This suggests (but does not imply) that these firms did not trade because of superior information.

In this chapter we study the dynamic behavior of large traders who have non-informational motives to trade. We assume for simplicity that no insiders operate in the market so that their trades are not attributed to inside information. Rather, large

---

<sup>2</sup>Although there may be a very large number of risk-averse agents in the economy, fixed costs of information gathering and market participation will imply that only a small number of them will be present in the market to take the other side of large trades, at least in the short-run. See Grossman and Miller (1988).

<sup>3</sup>These firms may be trading in the same direction for more than seven days.

trades affect prices because there is a small number of risk-averse market participants to take the other side of these trades. We address questions such as: What determines large traders' strategies and what are the equilibrium patterns of trade? Should we expect to observe larger quantities to be exchanged at the beginning of the trading session or does trade start slowly and most of it occurs well after the market opens? How *efficient* is the market allocation? In other words do large traders trade quickly so that most of the gains from risk-sharing are realized, or are we very far from the competitive outcome (where the gains from trade are realized immediately)?

In view of the increasing importance of large traders in modern financial markets, an understanding of their dynamic behavior may be useful for studying issues concerning the design and operation of these markets. We examine two such issues.<sup>4</sup>

The first issue is the frequency of market clearing. Many exchanges face the trade-off of whether to operate continuously over the trading day or whether to hold auctions at discrete points in time, at a given frequency. Increasing the frequency of market clearing has some costs such as operating costs but would probably attract more trade. An analysis of large traders' dynamic behavior is helpful in answering the complementary question of how the efficiency of the market allocation<sup>5</sup> varies with the market clearing frequency.

The second issue is the extent to which information about large trades is made public. Large traders sometimes voluntarily "preannounce" their trades, i.e. publicly reveal information about the trades that they wish to do in the future. It is of interest to know whether they are always hurt if information regarding their future trades is publicly known (since prices immediately adjust to this information) or whether they would ever benefit if the market has this information.<sup>6</sup>

In our model, a risky asset is exchanged for the numéraire. Trade takes place at a

---

<sup>4</sup>See Cohen *et al* (1986) for a description of the the operation of different financial exchanges. See also Pagano and Roëll (1990) for a discussion of the costs and benefits of recent reforms in many European stock markets.

<sup>5</sup>Efficiency is measured by the ratio of realized gains from trade over total potentially realizable gains from trade.

<sup>6</sup>If insiders operate in the market, large non-information traders may benefit if the rest of the market is aware of their trades because these trades will not be attributed to insiders. See Admati and Pfleiderer (1991).

discrete number of dates.<sup>7</sup> As it was said before, we assume that no market participant has inside information about asset payoffs; rather this information is *publicly* revealed over time. There are two types of traders. *Large* traders are risk-averse and initiate trades for hedging (or portfolio rebalancing) purposes. Their hedging demands are private information. They also act as speculators absorbing trades initiated by other large traders or by small traders. Since there is a finite number of large traders, the risk-absorption capacity of the market is finite: any trade affects prices even if no market participant has private information about asset payoffs. Large traders willing to minimize the price impact of their trades will not complete them immediately. They will thus bear the cost of holding unhedged positions since information about asset payoffs is gradually revealed. In addition the allocation of risk in the market will not be efficient. *Small* traders are also present in the market. We assume for simplicity that at each date they submit an exogenous, random, independent over time quantity of market orders. Because of these orders, market inferences about trades coming from large traders are imperfect. Small traders are assumed to behave exactly as the noise traders in the models of trade between asymmetrically informed agents. It is important to note that in contrast to the above models, the quantity of trade in our model does not go to zero as the quantity of market orders coming from small traders goes to zero. The reason why we introduce small traders is rather because it is very realistic to assume that the market does not make perfect inferences about the quantity of trades coming from large traders. We also get a much richer model. We finally assume that all traders can buy or sell many units of the risky asset and that the trading mechanism at each date is a Walrasian auction, where large traders submit limit orders (demand functions) and small traders submit market orders.

The brief description of the model given above, makes it clear that our work is closely related to the literature on double auctions with private values (i.e. auctions where there are many buyers and sellers, each of whom having a privately known willingness to trade).<sup>8</sup> Most of the papers in that literature (Satterthwaite and Williams

---

<sup>7</sup>We also consider the continuous-time limit.

<sup>8</sup>For surveys of this literature see McAfee and McMillan (1987) and Wilson (1990).

(1989) and Wilson (1985) among others) have studied the one-shot double auction where risk-neutral buyers (sellers) have 0-1 demands (supplies). These papers have studied the importance of the welfare loss due to imperfectly competitive behavior and the speed at which this loss goes to zero as the number of participants becomes large. The main difference between our work and these papers is that we consider a dynamic model. As it is well-known, conclusions of static models about the efficiency of trade may change substantially in a more realistic (at least in our case) dynamic framework.

There are very few papers that study dynamic models of double auctions. Wilson (1986) considers a continuous-time model of a bid-ask market where risk-neutral buyers and sellers have 0-1 demands. (See also the discussion in Wilson (1987).) He conjectures an equilibrium where buyers and sellers trade successively in the order of their respective valuations and costs, and derives the equations characterizing it. Our work is different than his in many respects. First, our traders are risk-averse, trade for hedging purposes and have multi-unit demands. Second, we introduce the small traders who add some “noise” to the inference process. These assumptions, which are more realistic in a financial market framework, enable us to examine some interesting questions that cannot be addressed in Wilson’s framework. In particular, we can study how large traders spread their trades and how the noise in the inference process (which is related to the ratio of trades coming from small traders to trades coming from large traders) affects the patterns of trade and the efficiency of the market allocation. A third difference is that we also consider the welfare implications of alternative trading systems. Finally we get a more tractable model that has a simple structure and that can be solved by a straightforward numerical algorithm.

Our work is also related to the literature on dynamic insider trading. Following Kyle’s work (1985), papers in this literature (for example Foster and Viswanathan (1992) and Holden and Subrahmanyam (1992)) have studied how quickly private information about an asset’s payoff gets incorporated into prices. In those papers a small number of risk-neutral insiders<sup>9</sup> benefit from their information, trading with

---

<sup>9</sup>See also Vives (1992) and Wang (1989), (1991) for models where there is a very large number



some risk-neutral uninformed market-makers. Market-makers accept to trade only because there is some chance that the orders they receive come instead from uninformed “noise” traders who trade for liquidity purposes. These papers study the strategic behavior of insiders while we study the behavior of agents who do not have an informational motive to trade. In addition, the assumptions of these papers, by allowing them to focus on the speed of information dissemination, make the question of the efficient realization of the gains from trade (which is the main focus of our work) trivial. Indeed, the risk-neutrality of the insiders and the market-makers implies that the only gains from trade in the market come from the liquidity traders. Since the latter are always assumed to have inelastic demands, the market outcome is efficient in the sense that all the gains from trade are realized.

There are however some similarities between the dynamic behavior of our large traders and, say, Kyle’s (1985) monopolistic insider. Indeed, both have private information about future prices and choose their current trades knowing that the market makes (noisy) inferences from current prices about that information. At the same time there are two important differences: First, our large traders have private information about future prices only because they decide to carry out some of their trades in the future. The magnitude of the adverse selection problem is thus endogenous; it depends on large traders’ strategies. By contrast, Kyle’s insider has an informational advantage because he is the first to learn some news about the asset’s payoffs. Therefore, the magnitude of the adverse selection problem is exogenously determined by the impact that these news will have on the asset’s price. The second difference is that our risk-averse large traders, bearing the cost of holding unhedged positions,<sup>10</sup> are impatient. Their impatience endogenously determines the speed at which trade takes place and information is revealed. On the contrary, Kyle’s risk-neutral insider can costlessly delay his trades until the date at which his information is made public. Therefore the speed of information revelation crucially depends on that exogenously given date. As it will be clear later, trading patterns in the two models are indeed

---

of informed agents who do not take into account their effect on prices.

<sup>10</sup>Since information about the risky asset’s payoffs is gradually revealed over time.

very different.

Our main results are the following. First, we find that the noise in inference process (which is related to the importance of small traders in the market) is a major determinant of large traders' behavior, and by consequence of trading patterns and welfare. As one might expect by extrapolating Kyle's analysis, if there are many small traders in the market, large traders trade faster. The result that noise is an important determinant of large traders' strategies is not true in a static model. Indeed, in the static version of our model large traders' strategies are independent of noise.<sup>11</sup> This illustrates well the interest of studying a dynamic model instead.

We find that when small traders account for a substantial fraction of the trading volume, large traders trade at a decreasing rate over time, while in the opposite case they start trading slowly, with most trade taking place well after the beginning of the trading session and decreasing afterwards. By contrast, Kyle's insider keeping market liquidity constant, trades larger and larger quantities as time passes. We also find that when there are few small traders, the outcome can be substantially more inefficient than what static double auctions models would predict.

Our result that when there are few small traders the outcome is much more inefficient than what a static model would predict, has a natural counterpart when studying how the frequency of market clearing affects efficiency. As one might expect, when there are few small traders, large traders trade more slowly if the market clears more frequently. By contrast, if there are many small traders and the market clearing frequency increases, large traders complete their trades within a shorter time after the market opening.

We finally find that if large traders account for a substantial fraction of the trading volume, they are better off if information regarding their hedging demands is publicly known.<sup>12</sup>

The rest of the chapter is structured as follows: In section 1.2, we present the model. In section 1.3, we study the case where there are only two trading dates. This

---

<sup>11</sup>The static version of our model is a particular case of Madhavan's (1990) model where there is no private information about the asset's payoff. See also Kyle (1989).

<sup>12</sup>This holds for *any* value of their hedging demands.

case is relatively simple to analyse and helpful for exhibiting the basic structure of the model. The more interesting multiperiod case is studied in section 1.4 and the efficiency of the market allocation is examined in section 1.5. In section 1.6, we study the effects of varying the market clearing frequency and give a heuristic analysis of the continuous-time limit. Section 1.7 examines whether large traders benefit if the market has exact information about their trading needs and section 1.8 concludes. All proofs are in the appendix.

## 1.2 The Model

We consider an economy where a risky asset is exchanged for the numéraire. Time, denoted by  $t$ , is continuous and runs from 0 to 1.<sup>13</sup> Trade takes place at a finite number of equally spaced dates between these two times. There are  $L$  such dates denoted by  $l$  ( $l = 1 \dots L$ ). Trading date  $l$  corresponds to time  $t = \frac{l}{L+1}$ . Times will always be in parentheses, while dates will be in superscripts.

The risky asset (that in our finite-horizon economy can be best thought of as a futures contract) pays a liquidating dividend  $d(1)$  at time 1. At time 0,  $d(1)$  is normally distributed with mean  $\bar{d}$  and variance  $\Sigma^2$ . Without loss of generality we will assume that  $\bar{d} = 0$ . Information about  $d(1)$  is *publicly* revealed over time at a constant rate. In other words,  $d(t) \equiv E(d(1)|\Omega(t))$  is a Brownian motion, where  $\Omega(t)$  represents information known as of time  $t$ . (The drift of this Brownian motion is 0 and its variance  $\Sigma^2$ .) For convenience, we define  $d^l \equiv d(\frac{l}{L+1})$  and  $\sigma^2 \equiv \frac{\Sigma^2}{L+1}$ . Holdings of the numéraire are reinvested at an interest rate equal to 0 after each trading date and are consumed at time 1 together with the risky asset's payoffs.

There are two types of traders: large and small.

- Large traders are present in the market at all trading dates. Their number is  $N$ . They are risk-averse and their preferences over time 1 consumption,  $\tilde{C}(1)$ ,

---

<sup>13</sup>Time 0 can be naturally interpreted as the time when the market opens at a given day.

are described by a negative exponential utility function

$$- \exp(-\alpha \tilde{C}(1)).$$

They trade because of a hedging motive revealed to them at time 0. It is assumed that at time 0 large trader  $i$  is endowed with  $e_i$  units of the asset.<sup>14</sup> The fact that  $e_i$  is different for each  $i$  implies that the initial allocation is not optimal and that there are gains from trade. Traders with high endowments will want to sell and traders with low endowments will want to buy, for hedging purposes.  $e_i$  is private information to trader  $i$ . The  $e_i$ 's are assumed to be normally distributed, independent of each other, each with mean 0 and variance  $\Sigma_e^2$ .<sup>15</sup> Indexes concerning traders will always be in subscripts.

- *Small* traders submit at each period an exogenous, random, independent over time quantity of market orders. The existence of small traders implies that information about the trading needs of the large traders is not revealed very quickly. As stated in the introduction, trade can occur in our model even in the absence of small traders. We introduce them because it is very realistic to assume that the market does not make perfect inferences about the quantity of trades coming from large traders and because the model is much richer. It is assumed that the cumulative quantity of market orders (executed or not) due to small traders follows a Brownian motion with drift 0 and variance  $\Sigma_u^2$ . Therefore, the quantity of market orders to be executed at date  $l$ , which are due to small traders, is normally distributed with mean 0 and variance  $\sigma_u^2 \equiv \frac{\Sigma_u^2}{L+1}$ . This quantity will be denoted by  $u^l$ , with the convention that  $u^l > 0$  corresponds to sell orders.

Finally, all traders can buy or sell many units of the risky asset and the trading mechanism at each date is a Walrasian auction, where large traders submit limit

---

<sup>14</sup> $e_i$  may also be interpreted as trader  $i$ 's holdings of a different asset which is correlated with the asset which is traded in the market we are considering.

<sup>15</sup>Negative values for the  $e_i$ 's can be interpreted as short positions.

orders (demand functions) and small traders submit market orders.

Before leaving this section we discuss some of the assumptions. An important assumption is that the parameter  $e_i$  describing the portfolio preferences of large trader  $i$  does not change over time. In other words, large traders do not have constantly changing demands; rather they experience an once and for all change to their portfolio preferences. Therefore in the beginning of the trading session there is a large quantity of trades to be done, and this quantity decreases over time. This assumption stands in contrast to the corresponding one for small traders: the latter have constantly changing demands. One could envision constructing a model where large traders have constantly changing demands, too. This model might also have the advantage of being *stationary*. We believe that many of the results of this model would be true in that model too. Moreover, market openings are natural examples of the situation that we describe (i.e. where a large quantity of trades to be done has accumulated). Finally, a model where large traders have constantly changing demands is much more difficult to solve than this one.

We also briefly discuss the somewhat rudimentary way to capture the behavior of the small traders. It was assumed for tractability, that small traders submit an exogenous quantity of market orders which is independent over time. Therefore small traders are not allowed to provide liquidity to the market, helping to absorb trades coming from large traders or from other small traders, and cannot postpone their own trades. If their objectives were more fully specified, for example if they were assumed to be risk-averse traders who traded for hedging purposes, the risk-absorption capacity of the market would simply be bigger. Moreover, given that each of them has a cost of postponing his own trades and a very small benefit, since these trades affect prices very little, he would not postpone them. It is thus clear that the results would not be very different.

## 1.3 The Two-Period Case

In this section we assume that there are only two trading dates. The analysis is then relatively simple and illustrates well why large traders do not carry all their trades at the first date and how the rest of the market responds to their strategy. The impact of the noise in the inference process, due to the small traders, can also be studied in this simplified framework.

We will restrict attention to linear symmetric Markov (LSM) equilibria. In subsection 1.3.1 we describe LSM equilibria in order to introduce some notation. In subsection 1.3.2 we show that there exists a unique such equilibrium and we discuss the equations characterizing it. Finally, in subsection 1.3.3 we study the impact of noise on large traders' strategies.

### 1.3.1 Some Preliminary Analysis

In a LSM equilibrium, large trader  $i$ 's date 1 demand has the form

$$x_i^1(p^1) = A^1 - B^1 p^1 - a^1 e_i^1 \quad (1.3.1)$$

where we set  $e_i^1 = e_i$  for trader  $i$ 's endowment before trade at date 1 and  $p^1$  for the date 1 price. In words, trader  $i$ 's date 1 demand depends only on the date 1 price,  $p^1$ , and on his initial endowment,  $e_i^1$ .  $A^1$ ,  $B^1$ ,  $a^1$  are parameters to be determined.

Market clearing requires that

$$\sum_{j=1}^N x_j^1(p^1) - u^1 = 0. \quad (1.3.2)$$

Equations 1.3.1 and 1.3.2 imply that trader  $i$ 's purchase at date 1 is

$$x_i^1 = a^1 \left( \frac{\sum_{j=1}^N e_j^1}{N} - e_i^1 \right) + \frac{u^1}{N}. \quad (1.3.3)$$

Therefore in equilibrium, after trade at date 1, trader  $i$ 's new asset holdings are

$$e_i^1 + x_i^1 = (1 - a^1)e_i^1 + \frac{a^1 \sum_{j=1}^N e_j^1 + u^1}{N}. \quad (1.3.4)$$

Equation 1.3.4 shows that after date 1 trade, trader  $i$ 's asset holdings are a fraction  $(1 - a^1)$  of his initial asset holdings plus  $1/N$ 'th of the total quantity of the asset that has been supplied at that date:  $a^1 \sum_{j=1}^N e_j^1 + u^1$ .

We will set

$$e_i^2 = (1 - a^1)e_i^1 \quad (1.3.5)$$

for the part of trader  $i$ 's asset holdings which are "*idiosyncratic*" to him, and

$$\bar{e}^2 = \frac{a^1 \sum_{j=1}^N e_j^1 + u^1}{N} \quad (1.3.6)$$

for the part of trader  $i$ 's asset holdings which are *common* to all traders.

In a hypothetical competitive world where traders do not take into account their impact on price,  $a^1 = 1$ . All large traders would then have exactly the same asset holdings after trade at date 1 and the allocation of risk would be efficient. However, because of imperfectly competitive behavior it will be the case that  $0 < a^1 < 1$ . Although the dispersion of large traders' asset holdings decreases, it does not become 0. The parameter  $a^1$  measures how much this dispersion decreases and how efficient is trade at date 1. Since  $a^1 < 1$ , large traders' asset holdings are still different after date 1 trade and there are incentives to further share risk at date 2.

After date 1 trade all traders know the realization of  $\bar{e}^2$  i.e. the total quantity of the asset that has been supplied. Each of them uses this information as well as his knowledge of his own date 1 trade to form expectations about the future trades coming from the other large traders.  $\bar{e}^2$  is obviously correlated with these future trades but not perfectly, because of the date 1 trades due to small traders. We now describe how these expectations are formed.

To fix ideas consider trader  $i$ . As the rest of the market, he observes  $\bar{e}^2$ . His expectations about  $\sum_{j \neq i} e_j^2 / N$ , which is related to the future trades coming from the other large traders are

$$E\left(\frac{\sum_{j \neq i} e_j^2}{N} \mid \bar{e}^2, e_i^1\right) = (1 - a^1)\beta\left(\frac{\bar{e}^2}{a^1} - \frac{e_i^1}{N}\right) = (1 - a^1)\beta\left(\frac{\sum_{j \neq i} e_j^1}{N} + \frac{u^1}{a^1 N}\right) \quad (1.3.7)$$

where  $\beta \in (0, 1)$  is defined by

$$\beta = \frac{(a^1)^2 \Sigma_e^2 (N - 1)}{(a^1)^2 \Sigma_e^2 (N - 1) + \sigma_u^2} \quad (1.3.8)$$

and is related to the signal-to-noise ratio of  $\bar{e}^2/a^1 - e_i^1/N$  i.e. of  $\sum_{j \neq i} e_j^1/N + u^1/a^1 N$ .

Denoting  $(1 - a^1)\beta\bar{e}^2/a^1$  by  $s^2$ , ( $s$  for “signal”) and using equation 1.3.5, we can rewrite equation 1.3.7 as

$$E\left(\frac{\sum_{j \neq i} e_j^2}{N} \mid \bar{e}^2, e_i^1\right) = s^2 - \beta \frac{e_i^2}{N}. \quad (1.3.9)$$

Equation 1.3.9 shows that each trader’s expectations depend on his own “idiosyncratic” asset holdings and the “signal”  $s^2$  which is common to all traders. Therefore in a LSM equilibrium, trader  $i$ ’s date 2 demand will have the following form

$$x_i^2(p^2) = A^2 - B^2 p^2 - a^2 e_i^2 - A_s^2 s^2 - A_e^2 \bar{e}^2. \quad (1.3.10)$$

His demand is a function of the date 2 price,  $p^2$ , his “idiosyncratic” asset holdings,  $e_i^2$ , the “signal”  $s^2$  which is related to his beliefs about other large traders’ hedging needs, and the “common” asset holdings  $\bar{e}^2$ .  $A^2$ ,  $B^2$ ,  $a^2$ , and  $A_s^2$  and  $A_e^2$  are parameters to be determined.<sup>16</sup>

Equations 1.3.1 and 1.3.10 describe traders’ behavior in equilibrium. It is important to note that if trader  $i$  deviates from equilibrium at date 1 by, say, buying one more unit, the other traders will still believe that he traded according to equilibrium but that the total hedging demand was different. Therefore at date 2 each of them will still submit a demand function determined by equation 1.3.10 (but using different

---

<sup>16</sup>Since  $s^2$  and  $\bar{e}^2$  are perfectly correlated, it might seem that the parameters  $A_s^2$  and  $A_e^2$  in equation 1.3.10 cannot be identified. We can, however, identify  $A_e^2$  by computing the change in traders’ demands in the hypothetical case where the asset holdings of each of them, after date 1 trade, increase by the same amount for an exogenous reason.



values of  $s^2$  and  $\bar{e}^2$ ).

### 1.3.2 Equilibrium

We are now ready to characterize equilibrium, i.e. determine the parameters of equations 1.3.1 and 1.3.10. In Lemma 1.3.1 (proven in appendix 1.A) we characterize equilibrium at date 2.

**Lemma 1.3.1** *There exists a unique date 2 equilibrium in which traders' demands have the form given by equation 1.3.10. In this equilibrium, the parameters  $A^2$ ,  $B^2$ ,  $a^2$ ,  $A_s^2$  and  $A_e^2$  are given by*

$$\frac{A^2}{B^2} = d^2 \quad (1.3.11)$$

$$\frac{A_e^2}{B^2} = \alpha\sigma^2. \quad (1.3.12)$$

$$\frac{N-2}{N-1} \frac{1}{B^2} = \alpha\sigma^2 \quad (1.3.13)$$

$$A_s^2 = 0 \quad (1.3.14)$$

and

$$a^2 = \frac{N-2}{N-1}. \quad (1.3.15)$$

Moreover, the price is given by

$$p^2 = d^2 - \alpha\sigma^2 \frac{\sum_{j=1}^N e_j^2}{N} - \alpha\sigma^2 \bar{e}^2 - \alpha\sigma^2 \frac{N-1}{N-2} \frac{u^2}{N}. \quad (1.3.16)$$

This equilibrium is a special case of the equilibrium derived by Madhavan in (1990) when there is no private information. (See also Kyle (1989).) As in (1990), a linear symmetric equilibrium exists for  $N > 2$ . If  $N = 2$  there is not enough competition, and the only candidate equilibrium has large traders submitting price-inelastic demands (i.e.  $B^2 = 0$ ), so no market clearing price exists. From now on it will be assumed that  $N > 2$ .

Equations 1.3.11 to 1.3.13 are easy to interpret. The constant term in traders'

demands,  $A^2$ , is related to the expected asset payoff and is such that the constant term in the price,  $A^2/B^2$ , is exactly this expected payoff. The coefficient of  $\bar{e}^2$  in traders' demands,  $A_{\bar{e}}^2$ , is such that the coefficient of  $\bar{e}^2$  in the price,  $A_{\bar{e}}^2/B^2$ , is exactly the same as in the hypothetical case where traders did not take into account their effect on the price (the competitive case). This is clear:  $\bar{e}^2$  is an endowment *common* to all traders, therefore if it changes trades will remain the same and the price change will reflect the change in the marginal valuation of any trader. Finally, the sensitivity of demands to price,  $B^2$ , is inversely related to traders' risk-aversion and asset risk (term  $\alpha\sigma^2$ ), and is smaller than its value in the competitive case, by the term  $(N - 2)/(N - 1)$ .

We now come to equation 1.3.14. The coefficient of  $s^2$ ,  $A_s^2$ , is equal to 0. In other words, traders when forming their demands do not use their information concerning other traders' hedging needs. The reason for this is that we are in the last trading date and there is no more trade to take place. Large traders are then treated exactly in the same way as small traders. This will obviously not be the case in the multiperiod model studied in section 1.4.

Finally, the coefficient of  $e_i^2$ ,  $a^2$ , which is related to the efficiency of trade at date 2, is smaller than 1 by the term  $(N - 2)/(N - 1)$  as equation 1.3.15 shows. The reason is that large traders, like textbook monopsonists (or monopolists), take into account their effect on the price and restrict their trades.

We characterize equilibrium at date 1 in Proposition 1.3.1 (proven in appendix 1.A).

**Proposition 1.3.1** *There exists a unique date 1 equilibrium in which traders' demands have the form given by equation 1.3.1. In this equilibrium, the parameters  $A^1$ ,  $B^1$  and  $a^1$  are given by*

$$\frac{A^1}{B^1} = d^1 \tag{1.3.17}$$

$$\frac{N - 2}{N - 1} \frac{1}{B^1} = 2\alpha\sigma^2 + (1 - a^1) \frac{\beta}{a^1} Q_3 \tag{1.3.18}$$

and

$$\left(a^1 \frac{N-1}{N-2} - 1\right) 2\alpha\sigma^2 = -(1-a^1)Q_3 \left(\frac{N-1}{N} + \frac{2(N-1)}{N(N-2)}\beta\right) \quad (1.3.19)$$

where  $Q_3$  is a positive constant determined in appendix 1.A. In particular  $0 < a^1 < \frac{N-2}{N-1}$ .

Equation 1.3.17 corresponds to equation 1.3.11 of Lemma 1.3.1. However, equation 1.3.18 is not the exact equivalent of equation 1.3.13 because of the term  $(1-a^1)(\beta/a^1)Q_3$ . If large traders completed all their trades at date 1 (i.e. if  $a^1 = 1$ ) this term would be zero. The price-sensitivity of demands,  $B^1$ , would then be inversely related to traders' risk-aversion and asset risk in a straightforward way, as in equation 1.3.13. However large traders do not complete their trades at date 1, this term is positive, and demands are less price-sensitive (i.e. the market is less liquid). Indeed, suppose that a higher than expected quantity is supplied. The market will require a larger price fall to absorb this quantity because it may have come not only from the small traders but from a large seller as well, who is likely to sell more in the future. Equation 1.3.18 also shows that if market inferences are not very noisy (i.e.  $\beta$  is large), demands are less price-sensitive and the market is less liquid, as intuition suggests.

As in Lemma 1.3.1 we find that trade at date 1 is inefficient, i.e.  $a^1 < 1$ . Even if no future trading opportunities existed, large traders would restrict their trades in order to execute them at better prices, exactly as they do at date 2. The existence of one more trading opportunity (date 2) induces them to restrict their trades *even more* ( $a^1 < (N-2)/(N-1)$ ). Indeed, large traders' benefit of trading less at date 1, which is that their remaining date 1 trades will be executed at better prices, remains unchanged. By contrast the cost of restricting their date 1 trades is now smaller since they can still cover their unhedged positions at date 2.

We now briefly discuss the determination of large traders' strategies. A more detailed discussion as well as the actual derivation of these strategies can be found in appendix 1.A.

For concreteness we consider trader  $i$  and we assume that he is a buyer. As it

was said before, he completes only a fraction of his purchases at date 1, and carries out some more at date 2. Suppose now that he reduces his date 1 purchases by a small quantity  $dx$  and increases his date 2 purchases by the same quantity. Then the following will happen:

- He will benefit since the price at which the rest of his date 1 purchases are executed will decrease. This is why he does not complete all his purchases at date 1 in the first place. The price at date 2 will rise, due to the imperfectly competitive behavior of the other traders, but this effect is much smaller.<sup>17</sup>
- He will bear the cost of holding a larger unhedged position between dates 1 and 2.
- He will bear an additional cost which is the following: Since he already carries out some purchases at date 2, he knows that the expected price at that date will be higher than the expected price at date 1. If he thus postpones some additional purchases for date 2, these purchases will be executed at a higher price.

This basic tradeoff determines the strategy of large buyer  $i$ .

### 1.3.3 The Impact of Noise

In this subsection we will study how the noise in the inference process, which is related to the importance of small traders in the market, affects large traders' strategies.

In appendix 1.A we show that if there are fewer small traders large traders trade less at date 1. There are three reasons for this. First, with fewer small traders, an additional trade at date 1 will affect prices more since it will be mostly attributed to large traders and will signal many more future trades in the same direction. Therefore if a large trader postpones some additional trades for date 2, the price at which the rest of his date 1 trades are to be executed will significantly improve since the market

---

<sup>17</sup>If the other traders were behaving as perfect competitors and trader  $i$  bought one unit less at date 1 and one unit more at date 2, the expected price at date 2 would remain the same.

will believe that there are fewer similar trades to follow. This trader will thus have a larger benefit to postpone trades for date 2. Second if there are fewer small traders the cost of postponing trades is smaller. Indeed, since there is less noise the market has better information about large traders' hedging demands and date 2 trades and the date 1 price better reflects these trades. Therefore, a large trader will not find the date 1 price much more advantageous than the date 2 price and postponing trades for date 2 will not be so costly. The third reason is that with fewer small traders, date 2 prices are less volatile and the risk of trading at date 2 instead of date 1 is smaller.

This reasoning is very similar to the one that explains why an insider trades less aggressively when there are fewer noise traders in the market.<sup>18</sup> As it was stated in the introduction, here the importance of large traders' information (which is related to their date 2 trades) is endogenous. In particular if there are fewer small traders, large traders trade less at date 1. The price may then become more sensitive to an additional trade since this trade signals many more future trades in the same direction. This decrease in market liquidity may induce large traders to trade even less. Liquidity may then further decrease and so on.

We should also note that in a static model large traders' strategies would not depend on the importance of small traders in the market. Indeed, in Lemma 1.3.1 we found that  $a^2 = (N - 2)/(N - 1)$  independently of how many trades come from small traders. By contrast if trade takes place over time, large traders will find it profitable to spread their trades. It will thus make a difference for them whether the rest of the market can easily "detect" them.

The analysis of the two-period case is helpful for understanding why large traders spread their trades over time and how noise affects their strategies. It is however clear that any application of the model involves more than two dates, in fact a very large number of them. Therefore by studying the multiperiod case (i.e. increasing the market-clearing frequency) we can get a more realistic idea of how quickly trade takes place, and how efficient is the market allocation. In addition, in a multiperiod model we can carry a more detailed analysis of large traders' strategies and of the

---

<sup>18</sup>The two-period version of our model is analogous to a *static* model of insider trading.

process by which prices adjust to reflect demand and supply, i.e. the price discovery process. In the next section we study the multiperiod case.

## 1.4 The Multiperiod Case

In this section we study the general case where there is any number of trading dates. We first generalize the notation of the previous section in subsection 1.4.1. In subsection 1.4.2 we characterize the equilibrium outcome and in subsection 1.4.3 we describe traders' strategies.

### 1.4.1 Some Preliminary Analysis

As trade proceeds over time, the dispersion of large traders' asset holdings decreases. As the discussion at the beginning of the previous section suggests, in a LSM equilibrium before date  $l$  trade, trader  $i$ 's asset holdings can be decomposed into two parts. The first part,  $e_i^l$ , is "idiosyncratic" to him and the second part,  $\bar{e}^l$ , is common to all traders. In addition as time passes the market obtains better information about large traders' hedging needs and remaining trades. This information is summarized in a signal  $s^l$ . In a LSM equilibrium, trader  $i$ 's date  $l$  demand has the form

$$x_i^l(p^l) = A^l - B^l p^l - a^l e_i^l - A_s^l s^l - A_{\bar{e}}^l \bar{e}^l. \quad (1.4.1)$$

His demand is thus a function of the date  $l$  price,  $p^l$ , his "idiosyncratic" asset holdings,  $e_i^l$ , the "common" asset holdings,  $\bar{e}^l$ , and the "signal"  $s^l$ .  $A^l$ ,  $B^l$ ,  $a^l$ , and  $A_s^l$  and  $A_{\bar{e}}^l$  are parameters to be determined.<sup>19</sup>

As in the previous section we find that after trade at date  $l$ , trader  $i$ 's asset holdings are

$$e_i^l + x_i^l = (1 - a^l)e_i^l + \frac{a^l \sum_{j=1}^N e_j^l + u^l}{N}. \quad (1.4.2)$$

It is thus clear that

---

<sup>19</sup>Equation 1.4.1 is also valid for  $l = 1$ , since  $s^1 = 0$  and  $\bar{e}^1 = 0$ .

$$e_i^{l+1} = (1 - a^l)e_i^l \quad (1.4.3)$$

and

$$\bar{e}^{l+1} = \bar{e}^l + \frac{a^l \sum_{j=1}^N e_j^l + u^l}{N}. \quad (1.4.4)$$

We now describe how traders form expectations about others' hedging needs. After trade at date  $l' < l$ , all traders observe the total quantity of the asset that has been supplied at that date,  $a^{l'} \sum_{j=1}^N e_j^{l'} + u^{l'}$ . Therefore, before trade at date  $l$  they obviously know  $\bar{e}^{l'}$  which is the sum of these quantities over all dates  $l' < l$  divided by  $N$ .<sup>20</sup> Moreover, each of them forms expectations about others' hedging needs at date  $l$ , i.e. their "idiosyncratic" asset holdings, using all these quantities as well as his knowledge of his own trades.

In appendix 1.A we show that trader  $i$ 's expectations about the "idiosyncratic" asset holdings of the other large traders are given by

$$E\left(\frac{\sum_{j \neq i} e_j^l}{N} \mid \left[a^{l'} \frac{\sum_{j=1}^N e_j^{l'}}{N} + \frac{u^{l'}}{N}\right]_{l' < l}, e_i^l\right) = s^l - (\beta^{l-1,1} + \beta^{l-1,2} + \dots + \beta^{l-1,l-1}) \frac{e_i^l}{N}. \quad (1.4.5)$$

Equation 1.4.5 generalizes equation 1.3.9. As before, we find that trader  $i$ 's expectations depend on his "idiosyncratic" asset holdings at date  $l$ , and on a signal  $s^l$ .  $s^l$  is given by

$$s^l = (1 - a^1)(1 - a^2) \dots (1 - a^{l-1}) [\beta^{l-1,1} \eta^1 + \beta^{l-1,2} \eta^2 + \dots + \beta^{l-1,l-1} \eta^{l-1}] \quad (1.4.6)$$

and is common to all traders. The expressions for the  $\eta^{l'}$ 's and the  $\beta^{l-1,l'}$ 's ( $l' < l$ ) are given below.

When forming his expectations, a trader uses his knowledge of the total quantities

---

<sup>20</sup>This is a direct consequence of equation 1.4.4.

supplied at each date,  $a'' \sum_{j=1}^N e_j'' + u''$ . He has to normalize these quantities (divided by  $N$ ), in terms of the first period asset holdings, for instance. To do this normalization, he has to divide the quantity supplied at date  $l'$ , by  $(1-a^1)(1-a^2)\dots(1-a^{l'-1})a^{l'}$ . Indeed it a straightforward implication of equation 1.4.3 that

$$e_j'' = (1-a^1)(1-a^2)\dots(1-a^{l'-1})e_j^1. \quad (1.4.7)$$

The  $\eta^{l''}$ 's are these normalized quantities, and are given by

$$\eta^{l''} = \frac{\sum_{j=1}^N e_j^1}{N} + \frac{u''}{(1-a^1)(1-a^2)\dots(1-a^{l'-1})a^{l'}N}. \quad (1.4.8)$$

These normalized quantities are weighted by the  $\beta^{l-1,l''}$ 's. These are given by

$$\beta^{l-1,l''} = \frac{(1-a^1)^2(1-a^2)^2\dots(1-a^{l'-1})^2(a^{l'})^2\Sigma_e^2(N-1)}{\sum_{l''<l}[(1-a^1)^2(1-a^2)^2\dots(1-a^{l''-1})^2(a^{l''})^2]\Sigma_e^2(N-1) + \sigma_u^2}. \quad (1.4.9)$$

$\beta^{l-1,l''}$  measures the importance of date  $l'$  trades for forming expectations at date  $l$ . It is large if the trades coming from large traders at date  $l'$ , which are a fraction  $a^{l'}$  of their current asset holdings and a fraction  $(1-a^1)(1-a^2)\dots(1-a^{l'-1})a^{l'}$  of their initial asset holdings, are important relative to all trades until date  $l$ .

Finally, these normalized quantities are multiplied by  $(1-a^1)(1-a^2)\dots(1-a^{l-1})$  to reflect asset holdings before date  $l$  trade. This gives us equation 1.4.6.

It is easy to see that  $s^{l+1}$  is obtained recursively as a function of  $s^l$  and the quantity traded at date  $l$  in the following simple manner

$$s^{l+1} = (1-a^l)[(1-\beta^{l,l})s^l + \beta^{l,l}(\frac{\sum_{j=1}^N e_j^l}{N} + \frac{u^l}{a^l N})]. \quad (1.4.10)$$

Equation 1.4.10 shows that  $s^{l+1}$  is a weighted average of  $s^l$  and  $\sum_{j=1}^N e_j^l/N + u^l/a^l N$ , multiplied by  $1-a^l$ , the fraction of the unrealized trades coming from large traders at date  $l$ . The weights reflect the importance of what is learned at date  $l$  relative to what was known before that date.



## 1.4.2 Equilibrium

We are now ready to characterize the equilibrium outcome. This is done in Proposition 1.4.1.

**Proposition 1.4.1** *In a LSM equilibrium, the parameters  $A^l$ ,  $B^l$ ,  $a^l$ ,  $A_s^l$  and  $A_e^l$  of equation 1.4.1 are given by*

$$\frac{A^l}{B^l} = d^l \quad (1.4.11)$$

$$\frac{A_s^l}{B^l} = (L + 1 - l)\alpha\sigma^2 \quad (1.4.12)$$

$$\frac{N - 2}{N - 1} \frac{1}{B^l} = (L + 1 - l)\alpha\sigma^2 + (1 - a^l) \frac{\beta^{l,l}}{a^l} (Q_2^{l+1} (1 - a^l) \frac{\beta^{l,l}}{a^l} \frac{1}{N - 1} + Q_3^{l+1}) \quad (1.4.13)$$

$$\frac{A_s^l}{B^l} = (1 - a^l)(1 - \beta^{l,l})(Q_2^{l+1} (1 - a^l) \frac{\beta^{l,l}}{a^l} \frac{1}{N - 1} + Q_3^{l+1}) \quad (1.4.14)$$

and

$$\begin{aligned} (a^l \frac{N - 1}{N - 2} - 1)(L + 1 - l)\alpha\sigma^2 &= -(1 - a^l)(Q_2^{l+1} (1 - a^l) \frac{\beta^{l,l}}{a^l} \frac{1}{N - 1} + Q_3^{l+1}) \\ &\quad (1 - \frac{1}{N}(1 - \beta^{l,1} - \dots - \beta^{l,l-1})) + \frac{2(N - 1)}{N(N - 2)}\beta^{l,l} \end{aligned} \quad (1.4.15)$$

where  $Q_2^{l+1}$  and  $Q_3^{l+1}$  are determined in appendix 1.A.

Proposition 1.4.1 (proven in appendix 1.A) generalizes Lemma 1.3.1 and Proposition 1.3.1.

The interpretation of the first two equations (1.4.11 and 1.4.12) is straightforward. It is important to note that the price impact of the “common” asset holdings ( $A_e^l/B^l$ ) decreases with  $l$ . Indeed, as we approach time 1, there is less risk to be borne.

As in Proposition 1.3.1 we find that demands are less price-sensitive than what they would be if large traders completed all their trades at date  $l$ . Indeed, as equation

1.4.13 shows,  $B^l$  is highest when  $a^l = 1$ . It is also clear that if trade at date  $l$  is very informative about large traders' hedging needs, i.e.  $\beta^{l,l}$  is bigger, demands are less price-sensitive and the market is less liquid.

Equation 1.4.14 shows that traders when forming their demands at date  $l$  would not use their information concerning others' hedging needs if trade at that date were efficient, (i.e.  $A_s$  would be 0 if  $a^l = 1$ ). Indeed, in this case large traders are treated exactly as small traders. However information about large traders' hedging needs is used at each date (except the last), since the latter do not complete all their trades at that date and postpone some for future dates. If for instance, it is believed that there are some large sellers in the market, each trader decreases his demand for a *given* price since he expects to be buying in the future.

Large traders' strategies are determined by considerations similar to those of the previous section. More precisely, if at date  $l$  large trader  $i$  who for concreteness is a buyer postpones some of his purchases for date  $l + 1$ , his payoff will change because of the three reasons outlined in the previous section. In the multiperiod case his payoff will change because of an additional reason, namely that from date  $l + 1$  on, the beliefs of the other traders about the total hedging demand will in general be different. Their demands will change too and so will the prices at which the remaining purchases of large trader  $i$  will be executed. Suppose, for example, that large traders are expected in equilibrium to trade more at date  $l$  than at date  $l + 1$ . If trader  $i$  postpones some of his purchases for date  $l + 1$ , then from that date on the other traders will believe that there is less buying pressure. Their demands will thus be lower and trader  $i$ 's remaining purchases will be executed at lower expected prices.

### 1.4.3 Trading Patterns

In this subsection we describe large traders' strategies. At this stage this description is not supported by formal proofs. It is rather based on the intuition that an analysis of the equations characterizing equilibrium provides and on many numerical simulations.

We first note that the impact of noise is the same as in the two-period case: with fewer small traders, large traders trade more slowly. Indeed, their benefit of postponing trades at a given date is larger since by doing so they give a strong signal to the market that fewer trades in the same direction will follow. As a result the price at which the rest of their trades at that date are executed significantly improves. In addition their cost of postponing trades is smaller since the market has better information about their hedging needs and the current date's price reflects more accurately the impact of their trades on next date's price. Therefore trading at the current date is not much more advantageous than trading at the next date.

We now describe large traders' strategies and the time variation of the trading volume. Suppose first that large traders account for a small fraction of the trading activity. Then the benefit of postponing trades is small. Indeed, the market does not regard current trades as good signals of future ones and prices do not respond much to these trades. In addition the cost of trading later is important since the market receives little information about large traders' hedging needs, and the current date's price does not reflect the impact of their trades on next date's price. This cost of postponing trades together with the cost due to worse hedging outweigh the benefit of trading later due to signalling and large traders trade at a *decreasing* rate over time. Suppose now that large traders account for most of the trading activity. Then prices adjust very quickly to reflect their hedging needs. However, even if the market can predict relatively accurately the total hedging demand, current trades remain valuable signals of future ones for a longer period.<sup>21</sup> During that period, the benefit of postponing trades due to signalling outweighs the cost of doing so due to worse hedging and large traders trade at a *increasing* rate. As time passes and more information is accumulated, the signalling value of an additional trade decreases and so does the benefit of postponing trades. The hedging cost of trading later dominates then and large traders trade at a *decreasing* rate.

---

<sup>21</sup>The fact that the market puts significant weight on trades even though it can already predict relatively accurately the total hedging demand may seem counterintuitive. In fact, it is not. If one successively receives good signals about an unknown random variable, he can quickly predict it accurately, but he puts substantial weight on future signals.

The above trade patterns indeed appear in numerical simulations.<sup>22</sup> These simulations are made under a simplifying assumption. It is assumed, quite realistically, that price movements coming from fundamentals (information about the liquidating dividend) are much larger than price movements due to the trades of the different market participants. Price movements due to fundamentals have a variance of the order of  $\Sigma^2$  while price movements due to agents' trades have a variance of the order of  $\alpha^2 \Sigma^4 \max(\Sigma_e^2/N, \Sigma_u^2/N^2)$ .<sup>23</sup> Therefore, price movements due to fundamentals dominate if large traders' risk-aversion is small ( $\alpha$  small), or fundamentals not too volatile ( $\Sigma^2$  small), or trade needs not too large ( $\Sigma_u^2$  and  $\Sigma_e^2$  small) or  $N$  big.

We make this assumption in order to abstract away from the complicated interaction between traders' risk-aversion and the price uncertainty due to their trades. The equations then simplify considerably and become exactly the same as in the fictitious case where traders are risk-neutral and have a constant per-period cost of holding the asset which is quadratic in the amount that they are holding. (This cost corresponds to their aversion to fundamental risk).<sup>24</sup>

Under this assumption the equilibrium outcome depends only on  $N$ , the number of large traders,  $L$ , the number of trading dates and  $\Sigma_u^2/N\Sigma_e^2$ , the ratio of trades coming from small traders to trades coming from large traders. (It does not depend on  $\alpha$  and  $\Sigma^2$  for instance.) Figures 1.1 to 1.3 illustrate our results for  $L = 120$  trading dates and for various values of  $N$  and  $\Sigma_u^2/N\Sigma_e^2$ , (denoted by "noise").

## 1.5 Allocative Efficiency

In this section we study the efficiency of the market allocation. We first define total welfare as well as a measure of efficiency. We then present some numerical results

---

<sup>22</sup>The algorithm that we use to solve the (polynomial) equations characterizing equilibrium is similar in spirit to the one followed in Kyle (1985). See Appendix 1.A for details.

<sup>23</sup>To understand why this is so, consider, for example, equation 1.3.16 giving the price at date 2. The variance of the first term, which corresponds to fundamentals is of the order of  $\Sigma^2$ , while the variance of the second, third and fourth term, which correspond to different agents' trades is of the order of  $\alpha^2 \Sigma^4 \max(\Sigma_e^2/N, \Sigma_u^2/N^2)$ .

<sup>24</sup>The simplification of the equations does not matter much for the simulations. It rather facilitates the analysis of these equations so that the simulation results are easier to understand.

and compare them to the results obtained in static double auction models.

To derive the welfare of a large trader we proceed in two steps. First we take his initial endowment of risky asset as given, and find his (“interim”) expected utility at time 0. We then find his “ex-ante” utility taking expectations with respect to his initial endowment and we calculate his welfare in certainty equivalent terms. This computation is particularly simple if we assume, as before, that price movements coming from fundamentals dominate price movements due to the trades of the different market participants. All the results of this subsection are derived under this assumption.

The derivation of small traders’ welfare is trickier since we have not specified their objectives. Following papers in the market microstructure literature that examine welfare issues, for example Admati and Pfleiderer (1991), we assume that small traders care only about the expectation of their proceeds from selling the asset and their expenses incurred in buying it. (We refer to these as “trading proceeds”.) We believe that a model in which small traders’ objectives were more fully specified<sup>25</sup> would not produce a very different expression for their welfare. However in such a model small traders would also care about the variability of their trading proceeds and about the speed at which they can trade. (This speed is related to the market-clearing frequency.) It is not clear how taking into account these factors would affect welfare *comparisons* when the market-clearing frequency is kept constant. However, small traders’ preferences for fast execution would certainly play a role in the comparison of trading systems that operate at different frequencies.

Under the above assumptions we get a very simple expression for total welfare. In appendix 1.B we show that welfare,  $W$ , is given by

$$W = -\frac{1}{2}\alpha\Sigma^2\Sigma_e^2\left[1 + \frac{N-1}{L+1}(1 + (1-a^1)^2 + \dots + (1-a^1)^2 \dots (1-a^L)^2) + \frac{\Sigma_u^2}{N\Sigma_e^2} \frac{L}{2(L+1)}\right] \quad (1.5.1)$$

---

<sup>25</sup>If for instance they were assumed to be risk-averse traders who traded for hedging purposes.

where the  $a^l$ 's ( $l \in [1, L]$ ) have their equilibrium values.

It is easy to see that welfare is always smaller than the value  $W_c$  that the above expression would take if  $a^1$  were equal to 1 i.e. large traders completed all their trades at date 1. Simple algebra<sup>26</sup> shows that in the competitive case where large traders indeed complete their trades at date 1, total welfare is equal to  $W_c$ . We thus naturally find that welfare is highest in the competitive case and use this case as a benchmark. Moreover, equation 1.5.1 shows that welfare decreases if trade takes place more slowly, (and  $(1 - a^1)^2 \dots (1 - a^l)^2$  is higher for each  $l$ ) as intuition suggests.

To determine how efficient is the market outcome we need to subtract total welfare given in equation 1.5.1 from its value in the competitive case and see how small the difference is. We also need some measure of the total realizable gains from trade in order to express this difference in relative terms. We define the total realizable gains from trade as the difference between  $W_c$  and the value of welfare given by equation 1.5.1 if all the  $a^l$ 's were equal to 0, i.e. large traders did not trade between them.<sup>27</sup>

In the following table we present some values of the unrealized gains from trade as a percentage of total realizable gains from trade, obtained in our model, as well as analogous values obtained in static double auctions models. In the first three rows we report results that we get in our model for  $L = 120$  trading dates and for various values of  $\Sigma_u^2/N\Sigma_e^2$ , the ratio of trades coming from small traders to trades coming from large traders (denoted by "noise"). In the fourth row we report the results of our static model (where large traders' strategies and market efficiency are independent of noise). Finally, in the fifth row we present results obtained in static double auctions models where traders are risk-neutral, have 0-1 demands or supplies and there is no noise.

---

<sup>26</sup>It is straightforward to solve the model in the competitive case.

<sup>27</sup>Since we do not know small traders' surpluses we can not define total realizable gains from trade as the difference between  $W_c$  and total welfare if there is no trade at all. The normalization chosen is not important for the following discussion since our point is that the market outcome is not very efficient when there are *few* small traders in the market.

	N=4	N=8	N=16	N=24
dynamic, noise 30	4.21	2.55	1.62	1.23
dynamic, noise 4	8.04	5.17	3.31	2.51
dynamic, noise 0.1	25.8	15.7	9.13	6.47
static (our model)	11.11	2.04	0.44	0.18
static (other models)	5.6	1.5	0.4	0.2

As our previous analysis suggests, the market outcome is more efficient if there are more large traders in the market, for a given amount of “noise”, and if there is more “noise”. Moreover, looking at the above table we can make two important observations.

First, when there is little noise, the outcome is substantially more inefficient than what static models predict. (This is true for both the static double auction models with risk-neutral traders and 0-1 demands as well as for the static version of our model.)<sup>28</sup> As the analysis of section 1.3 suggested, if the market is open for more dates than one, large traders will trade less at the first date. Indeed, by doing so they get a better price for their remaining trades at that date and are still able to cover their unhedged positions trading at later dates. When there is little noise the market strongly reacts to their strategy. Date 1 demands become less price-sensitive since an additional trade signals many more future ones. As a result large traders want to postpone their trades even more, this reinforces the above market reaction, and so on.

Second, the rate of convergence to the competitive outcome is much smaller in a dynamic model than in a static model. (This is, again, true for both static models.) If, for example,  $N$  increases from 4 to 16, unrealized gains from trade as a percentage of total realizable gains from trade are divided by at most 3 in the dynamic model and by more than 12 in both static models. As Satterthwaite and Williams (1989)

---

<sup>28</sup>Note, however, that at the end of the trading session much more trade has taken place in a dynamic model than in a static model. This result is analogous to Kyle's (1985) result that when trade ends, prices reflect all the insider's information in a continuous-time model and only half of it in a static model.

have indeed shown, in the one-shot buyer's bid double auction (where risk-neutral traders have 0-1 demands and supplies) the rate of convergence to the competitive outcome is of the order of  $\frac{1}{N^2}$ <sup>29</sup> We conjecture that the rate of convergence in our dynamic model (at least in its continuous-time version) is of the order of  $\frac{1}{N}$  and hope to establish this result in future work.

## 1.6 The Effects of More Frequent Market Clearing

In this brief section we study how the efficiency of the outcome, or equivalently the difference in relative terms between total welfare,  $W^c$ , in the competitive case and total welfare,  $W$ , in equilibrium, varies with the frequency of market clearing. Obviously, if the market clears more frequently ( $L$  is higher)  $W^c$  will be higher since large traders will start trading earlier (at time  $1/(L + 1)$ ).<sup>30</sup> Here we address the complementary question of how the difference between  $W^c$  and  $W$  will change.

Although at this stage we have not proven a general result, the numerical simulations that we have performed suggest that if the market clears more frequently large traders trade faster if there are many small traders and vice-versa. The reason why increasing the market clearing frequency can imply that large traders trade more slowly is similar to why a dynamic model can predict a less efficient outcome than a static model. If the market clears more frequently, large traders spread their trades more over the larger number of dates. By doing so, they obtain a better price at each of the initial dates but can still hedge their positions through more frequent trading. Since large traders trade less at a given date and more in the future, market liquidity may be lower at that date since an additional trade signals many more future ones. Moreover, all demands are adjusted taking into account the fact that more trades will

---

<sup>29</sup>Gresik and Satterthwaite (1989) perform a analysis for optimal trading mechanisms with similar results.

<sup>30</sup>There will be some additional welfare gains because small traders will be able to trade more quickly after they experience a need to do so. However, as it was stated in the previous section our measure for welfare does not capture this effect.



follow and prices at the initial dates are not very different.<sup>31</sup> Because of the above market reactions, large traders will postpone their trades even more and so on.

To understand why an increase of the frequency of market clearing can induce large traders to trade faster, suppose that  $\Sigma_c^2$  is zero (i.e. large traders are not expected to initiate any trades) and that there are many trading dates. A large trader who happens to have a non-zero hedging demand will obviously trade a small (and decreasing) quantity at each date. Suppose now that the number of trading dates increases. In this case it will not be optimal for him to complete his trades in the same *calendar* time. Indeed, since his trades are very small, the reduction in the price impact caused by further spreading them over the larger number of trading dates is very small. However, the gain due to impatience in completing trades faster is not so small. The above intuition is supported by a heuristic analysis that shows that if  $\Sigma_c^2 = 0$ ,  $a^l$  varies in  $1/\sqrt{L}$ . Therefore the calendar time in which large traders complete almost all their trades converges to 0 as  $L$  gets large.

In other words, large traders will trade faster when the market clearing frequency increases, if the market does not strongly react to the fact that they spread their trades more. This will be the case if they initiate a small fraction of trades (which we assumed 0 in the above example).

Our heuristic analysis also shows that if  $\Sigma_c^2 \neq 0$ ,  $a^l$  varies in  $1/L$  and the calendar time in which large traders complete a given amount of their trades converges to a non-zero limit as  $L$  gets large. We hope to make this heuristic analysis rigorous in future work.

---

<sup>31</sup>Suppose, for instance, that the market believes that there is buying pressure. Suppose also that buyers instead of carrying all their trades at the current date postpone some of them. Then all traders will submit higher demands *for a given price* at the current date. (In terms of the model  $A_c$  will be positive instead of 0.) This will imply that the current price will not be very different even though buyers buy less.

## 1.7 The Case where Large Traders' Hedging Demands are Public Information

In this section we analyse large traders' welfare if information regarding their trading needs is publicly known. More precisely, we will study the market outcome when all initial endowments  $e_i$  are public information (we call this case the public information case) and compare it to the outcome studied in the rest of the chapter where initial endowments were private information (the private information case). We will be particularly interested in knowing whether large traders can ever benefit if the rest of the market has information about their hedging demands. Although this does not necessarily imply that large traders would individually prefer to reveal information about their demands (i.e. preannounce), it suggests that this might be the case.<sup>32</sup> For simplicity, we assume that there are only two trading dates. We characterize the equilibrium and discuss large traders' welfare.

If the  $e_i$ 's are publicly known, the date 1 demand of large trader  $i$  has the form

$$x_i^1(p^1) = A^1 - B^1 p^1 - a^1 e_i^1 - A_s^1 \frac{\sum_{j=1}^N e_j^1}{N} \quad (1.7.1)$$

where  $A^1$ ,  $B^1$ ,  $a^1$ , and  $A_s^1$  are, again, parameters to be determined. Note that in contrast to the private information case, demands contain the extra term  $A_s^1 \sum_{j=1}^N e_j^1 / N$  since hedging demands are known at date 1.

As in the private information case, we find that trader  $i$ 's asset holdings after date 1 trade can be decomposed into the "idiosyncratic" part

$$e_i^2 = (1 - a^1) e_i^1 \quad (1.7.2)$$

and the "common" part

$$\bar{e}^2 = \frac{a^1 \sum_{j=1}^N e_j^1 + u^1}{N}. \quad (1.7.3)$$

---

<sup>32</sup>We hope to study individual incentives to preannounce in future work.

Moreover, trader  $i$ 's date 2 demand will have the following form

$$x_i^2(p^2) = A^2 - B^2 p^2 - a^2 e_i^2 - A_s^2 \frac{\sum_{j=1}^N e_j^2}{N} - A_g^2 e^2 \quad (1.7.4)$$

where  $A^2$ ,  $B^2$ ,  $a^2$ ,  $A_s^2$  and  $A_g^2$  are parameters to be determined.

As the analysis of section 1.3 suggests, the date 2 equilibrium is exactly the same as in the private information case. Indeed, at date 2 no trader uses his information concerning others' hedging needs. We prove this result in appendix 1.C.

Proposition 1.7.1 (proven in appendix 1.C) characterizes equilibrium at date 1.

**Proposition 1.7.1** *There exists a unique date 1 equilibrium in which traders' demands have the form given by equation 1.7.1. In this equilibrium, the parameters  $A^1$ ,  $B^1$ ,  $a^1$  and  $A_s^1$  are given by*

$$\frac{A^1}{B^1} = d^1 \quad (1.7.5)$$

$$\frac{N-2}{N-1} \frac{1}{B^1} = 2\alpha\sigma^2 \quad (1.7.6)$$

$$\frac{A_s^1}{B^1} = (1 - a^1)Q_3 \quad (1.7.7)$$

and

$$\left(a^1 \frac{N-1}{N-2} - 1\right) 2\alpha\sigma^2 = -(1 - a^1)Q_3 \quad (1.7.8)$$

where  $Q_3$  is a positive constant determined in appendix 1.C. In particular  $0 < a^1 < \frac{N-2}{N-1}$ .

The important difference with the private information case is that demands are more price-sensitive (i.e. the market is more liquid). Indeed, a price which is lower than expected can be attributed only to small traders (and not to large traders as well, as in that case). However in the public information case, at date 1 traders have information about others' hedging needs and use this information when forming their demands (i.e.  $A_s^1 > 0$ ). If, for example, they know that there are some large sellers in the market, they decrease their demands because they know that they will also be

buying in the future.

In the appendix we show that in the public information case large traders may trade faster or more slowly than in the private information case (i.e.  $a^1$  may be higher or lower). The reason is that in the public information case an additional date 1 trade will affect prices less since it will not be attributed to large traders and will not be interpreted as a signal of more trades in the same direction. Large traders benefit thus less from postponing their trades. However, the cost of postponing trades is also smaller since the date 1 price perfectly reflects large traders' hedging needs and date 2 trades and is not more advantageous than the date 2 price. Not surprisingly, if there are few small traders in the market large traders trade faster in the public information case and vice versa. (This is proven in appendix 1.C.)

As equation 1.5.1 shows, total welfare is lower in the private information case if there is less trade at date 1 ( $a^1$  smaller).<sup>33</sup> In appendix 1.C we provide an example where large traders' welfare (both in the "interim" sense, for all values of their initial endowment and in the ex-ante sense) is also lower. This is the case because, although they obtain a better price at date 1 than in the public information case, they trade much less at this date facing a less liquid market.

## 1.8 Conclusion

In recent years there has been substantial progress in understanding the behavior of large agents in financial markets. Most of the papers in this area have focused on the strategic behavior of large agents who have inside information about assets' fundamental values and trade in order to benefit from that information. To do so in a simple manner they have considerably simplified the behavior of agents who trade because of non-informational reasons.<sup>34</sup> Our work shows that these agents may also have incentives to behave strategically and examines their dynamic behavior as well as the efficiency of the market allocation.

---

<sup>33</sup>Equation 1.5.1 is also valid for the public information case. See appendix 1.C.

<sup>34</sup>There are some exceptions. See for instance Admati and Pfleiderer (1988).

In a world with no insiders but with large traders whose trades are not informationally motivated, many conclusions concerning the design of trading systems may be different from the ones obtained in a world with insiders.

Consider for instance the transparency of the trading system, i.e. how much information about the trading process is revealed to the market participants. In a world with insiders it seems natural that as much information about the trading process as possible should be made public, so that the differences in information about assets' fundamental values are attenuated. This might no longer be the case in a world with large "informationless" traders. Indeed in a transparent market, information about their remaining trades would be quickly revealed and prices would adjust to reflect these trades. These traders would then find the liquidity of the market substantially reduced and would reduce their trades to slow down the revelation of information. In addition they would have a smaller ex-ante incentive to participate in this market. We believe that this as well as other market microstructure issues can be addressed in this model and we hope to do so in future research.

# 1.A Appendix: Proofs of Results of Sections 1.3 and 1.4

## 1.A.1 Notation and Preliminary Derivations

In this subsection we will motivate equation 1.4.1 which gives a trader's demand at date  $l$ . Doing so we will introduce some notation and derive some equations which will be useful later.

Suppose that until date  $l - 1$  all traders' demands have the form of equation 1.4.1 (written for these dates). Then, in equilibrium, before date  $l$  trade the asset holdings of trader  $i$ 's can be decomposed into two parts: the "idiosyncratic" part,  $e_i^l$ , and the "common" part,  $\bar{e}^l$ . In addition trader  $i$  forms expectations about others' "idiosyncratic" asset holdings at that date. When forming his expectations, he uses his information about his own hedging demand (i.e.  $e_i^l$  or equivalently  $e_i^l$ ) as well as his knowledge of the total quantities supplied at each date,  $a^{l'} \sum_{j=1}^N e_j^{l'} + u^{l'}$  for  $l' < l$ .

It is a matter of algebra to show that conditional on his information, trader  $i$  believes that  $\sum_{j \neq i} e_j^l / N$  is normally distributed with mean

$$E_i^l \left( \frac{\sum_{j \neq i} e_j^l}{N} \right) = s^l - (\beta^{l-1,1} + \dots + \beta^{l-1,l-1}) \frac{e_i^l}{N} \quad (1.A.1)$$

and variance

$$V^l = \frac{(1 - a^1)^2 \cdot \dots \cdot (1 - a^{l-1})^2 \Sigma_e^2 (N - 1) \sigma_u^2}{\sum_{l' < l} [(1 - a^1)^2 \cdot \dots \cdot (1 - a^{l'-1})^2 (a^{l'})^2] \Sigma_e^2 (N - 1) + \sigma_u^2} \frac{1}{N^2} \quad (1.A.2)$$

Equation 1.A.1 shows that in equilibrium a trader's expectations about others' "idiosyncratic" asset holdings at date  $l$  depend on his own "idiosyncratic" asset holdings at that date, as well as on a signal  $s^l$  which is common to all traders. It is thus natural to look for an equilibrium where at date  $l$  his demand has the form given by equation 1.4.1. Since we always refer to trader  $i$  we drop the subscript  $i$  from the  $E$ 's that denote expectations.

In equation 1.A.1  $s^l$  is given by

$$s^l = (1 - a^1) \dots (1 - a^{l-1}) [\beta^{l-1,1} \eta^1 + \dots + \beta^{l-1,l-1} \eta^{l-1}]. \quad (1.A.3)$$

The  $\beta^{l-1,l'}$ 's are given by

$$\beta^{l-1,l'} = \frac{(1 - a^1)^2 \dots (1 - a^{l'-1})^2 (a^{l'})^2 \Sigma_e^2 (N - 1)}{\sum_{l'' < l} [(1 - a^1)^2 \dots (1 - a^{l''-1})^2 (a^{l''})^2] \Sigma_e^2 (N - 1) + \sigma_u^2}. \quad (1.A.4)$$

Finally the  $\eta^{l'}$ 's are given by

$$\eta^{l'} = \frac{\sum_{j=1}^N e_j^1}{N} + \frac{u^{l'}}{(1 - a^1) \dots (1 - a^{l'-1}) a^{l'} N}. \quad (1.A.5)$$

$\eta^{l'}$  is the quantity supplied at date  $l'$ ,  $a^{l'} \sum_{j=1}^N e_j^{l'} + u^{l'}$ , divided by  $N$  as well as by  $(1 - a^1) \dots (1 - a^{l'-1}) a^{l'}$ , i.e. it is normalized in terms of the first period asset holdings.

It is also easy to show that  $s^{l+1}$  is obtained recursively as a function of  $s^l$  and the quantity traded at date  $l$  in the following simple manner

$$s^{l+1} = (1 - a^l) [(1 - \beta^{l,l}) s^l + \beta^{l,l} (\frac{\sum_{j=1}^N e_j^l}{N} + \frac{u^l}{a^l N})]. \quad (1.A.6)$$

We can similarly find the following recursive relation

$$E^{l+1}(\frac{\sum_{j \neq i} e_j^{l+1}}{N}) = (1 - a^l) [(1 - \beta^{l,l}) E^l(\frac{\sum_{j \neq i} e_j^l}{N}) + \beta^{l,l} (\frac{\sum_{j \neq i} e_j^l}{N} + \frac{u^l}{a^l N})]. \quad (1.A.7)$$

## 1.A.2 Derivation of Traders' Value Functions and Characterization of Equilibrium

We will now assume that all traders submit their equilibrium demands and will recursively derive their value functions. As a byproduct we will derive necessary and sufficient conditions for an equilibrium.

In order to evaluate a trader's incentives to deviate from equilibrium, we must compute his value function after any sequence of actions that he may take, including

actions that he is not supposed to take in equilibrium. After a deviation by this trader the other large traders will however still believe that everybody acted according to equilibrium but that the total hedging demand was different. At the following dates they will thus still submit their equilibrium demands given by equation 1.4.1 (written for these dates) but they will use different values of  $s$  and  $\bar{e}$ .

We will now prove the following result.

**Theorem 1.A.1** *Suppose that we are at date  $l$  before trade takes place. Suppose that trader  $i$  has deviated from equilibrium at the previous dates so that his asset holdings are higher by  $\Delta e_i^l$  than what they would be if he had never deviated. (Each other large trader's asset holdings are thus lower by  $\Delta e_i^l/(N-1)$  and they all believe that  $\bar{e}^l$  is lower by this quantity.) Suppose also that the other large traders believe that  $s^l$  is lower by  $\Delta s^l$ . Then the expected utility of trader  $i$  (i.e. his value function) is*

$$\begin{aligned}
& -\exp\left(-\alpha\left(M_i^l + d^l(e_i^l + \bar{e}^l + \Delta e_i^l) - \frac{1}{2}\alpha\sigma^2(L-l+1)(e_i^l + \bar{e}^l + \Delta e_i^l)^2\right.\right. \\
& + Q_1^l\left(E^l\left(\frac{\sum_{j \neq i} e_j^l}{N}\right) + \frac{e_i^l}{N} - e_i^l\right)^2 - (Q_2^l\Delta s^l + Q_3^l\Delta e_i^l)\left(E^l\left(\frac{\sum_{j \neq i} e_j^l}{N}\right) + \frac{e_i^l}{N} - e_i^l\right) \\
& \left. + Q_4^l(\Delta s^l)^2 + Q_5^l(\Delta e_i^l)^2 + Q_6^l\Delta s^l\Delta e_i^l + Q_7^l\right). \tag{1.A.8}
\end{aligned}$$

$M_i^l$  are trader  $i$ 's holdings of the numéraire.<sup>35</sup>  $Q_k^l$  ( $k \in 1 \dots 7$ ) are constants.  $e_i^l$  and  $E^l \sum_{j \neq i} e_j^l/N$  have their equilibrium values.

The rest of this subsection is devoted to the proof of Theorem 1.A.1. In proving this theorem we will also derive necessary and sufficient conditions for an equilibrium and prove Proposition 1.4.1. We do not state these conditions at this stage in order not to lengthen the appendix.

We will prove Theorem 1.A.1 by induction. The theorem is obviously true at date  $L+1$ . At that date all the uncertainty has been resolved and a large trader's utility

---

<sup>35</sup>Since we always refer to trader  $i$  we drop the subscript  $i$  from the  $M$ 's.



is

$$-exp(-\alpha(M^{L+1} + d^{L+1}(e_i^{L+1} + \bar{e}^{L+1} + \Delta e_i^{L+1})))$$

which is equation 1.A.8 with the  $Q_k$ 's equal to 0.

Suppose now that equation 1.A.8 holds for date  $l + 1$  and let us show that it holds for date  $l$ . We will proceed in three steps. First we will use the induction hypothesis to simplify trader  $i$ 's expected utility at date  $l$ . Second we will determine the solution to his maximization problem at that date. Finally we will plug this solution back into the expression for his expected utility and show that the induction hypothesis holds for date  $l$ .

### Step 1

We denote trader  $i$ 's purchases at date  $\hat{l} \geq l$  by  $x^{\hat{l}}$ . (Since we always refer to trader  $i$  we drop the subscript  $i$ .) His final consumption is

$$\tilde{C}(1) = M^l - \sum_{\hat{l} \geq l} p^{\hat{l}} x^{\hat{l}} + d^l(e_i^l + \bar{e}^l + \Delta e_i^l + \sum_{\hat{l} \geq l} x^{\hat{l}}).$$

$\tilde{C}(1)$  is of course random and we have to take expectations. Applying the law of iterative expectations we thus get for trader  $i$ 's expected utility before trade at date  $l$

$$\begin{aligned} E^l - exp(-\alpha \tilde{C}(1)) &= E^l - exp(-\alpha \left( M^l - \sum_{\hat{l} \geq l} p^{\hat{l}} x^{\hat{l}} + d^l(e_i^l + \bar{e}^l + \Delta e_i^l + \sum_{\hat{l} \geq l} x^{\hat{l}}) \right)) \\ &= E^l E^{l+1} - exp(-\alpha \left( M^l - \sum_{\hat{l} \geq l} p^{\hat{l}} x^{\hat{l}} + d^l(e_i^l + \bar{e}^l + \Delta e_i^l + \sum_{\hat{l} \geq l} x^{\hat{l}}) \right)). \end{aligned} \quad (1.A.9)$$

We now apply the induction hypothesis for date  $l + 1$ . At that date  $M^{l+1} = M^l - p^l x^l$ , and trader  $i$ 's asset holdings are:  $e_i^l + \bar{e}^l + \Delta e_i^l + x^l$ . We use the natural notation,  $\Delta e_i^{l+1}$ , for the difference between these asset holdings and his asset holdings if he had not deviated from equilibrium at the previous dates. We also denote by  $\Delta s^{l+1}$  the "misperceptions" of the other traders about  $s^{l+1}$ . Expression 1.A.9 together with the induction hypothesis give us then for trader  $i$ 's expected utility

$$\begin{aligned}
& E^l - \exp(-\alpha \left( M^l - p^l x^l + d^l (c_i^l + \bar{c}^l + \Delta c_i^l + x^l) \right. \\
& - \frac{\alpha \sigma^2}{2} (L - l) (c_i^l + \bar{e}^l + \Delta c_i^l + x^l)^2 + Q_1^{l+1} (E^{l+1} (\frac{\sum_{j \neq i} c_j^{l+1}}{N}) + \frac{c_i^{l+1}}{N} - e_i^{l+1})^2 \\
& - (Q_2^{l+1} \Delta s^{l+1} + Q_3^{l+1} \Delta c_i^{l+1}) (E^{l+1} (\frac{\sum_{j \neq i} c_j^{l+1}}{N}) + \frac{c_i^{l+1}}{N} - e_i^{l+1}) \\
& \left. + Q_4^{l+1} (\Delta s^{l+1})^2 + Q_5^{l+1} (\Delta c_i^{l+1})^2 + Q_6^{l+1} \Delta s^{l+1} \Delta c_i^{l+1} + Q_7^{l+1} \right). \quad (1.A.10)
\end{aligned}$$

$d^{l+1}$  and the information about the total hedging demand revealed when trading at date  $l$  (and contained in  $x^l$  and  $E^{l+1}(\sum_{j \neq i} e_j^{l+1}/N)$ ) are independent. Noting in addition that conditional on date  $l$  information  $d^{l+1}$  is normally distributed with mean  $d^l$  and variance  $\sigma^2$ , expression 1.A.10 can be transformed into

$$\begin{aligned}
& E^l - \exp(-\alpha \left( M^l - p^l x^l + d^l (e_i^l + \bar{e}^l + \Delta e_i^l + x^l) \right. \\
& - \frac{\alpha \sigma^2}{2} (L - l + 1) (e_i^l + \bar{e}^l + \Delta e_i^l + x^l)^2 + Q_1^{l+1} (E^{l+1} (\frac{\sum_{j \neq i} e_j^{l+1}}{N}) + \frac{e_i^{l+1}}{N} - e_i^{l+1})^2 \\
& - (Q_2^{l+1} \Delta s^{l+1} + Q_3^{l+1} \Delta e_i^{l+1}) (E^{l+1} (\frac{\sum_{j \neq i} e_j^{l+1}}{N}) + \frac{e_i^{l+1}}{N} - e_i^{l+1}) \\
& \left. + Q_4^{l+1} (\Delta s^{l+1})^2 + Q_5^{l+1} (\Delta e_i^{l+1})^2 + Q_6^{l+1} \Delta s^{l+1} \Delta e_i^{l+1} + Q_7^{l+1} \right). \quad (1.A.11)
\end{aligned}$$

## Step 2

Trader  $i$  maximizes expression 1.A.11 w.r.t.  $x^l$ . To simplify later notation we define  $\Delta x^l$  by

$$x^l = \frac{a^l \sum_{j \neq i} e_j^l + u^l}{N} + \frac{a^l c_i^l}{N} - a^l c_i^l + \Delta x^l. \quad (1.A.12)$$

$\Delta x^l$  represents thus trader  $i$ 's deviation from his equilibrium strategy, which is to buy  $(a^l \sum_{j \neq i} e_j^l + u^l)/N + a^l c_i^l/N - a^l c_i^l$ .

Before solving the optimization problem we express some variables present in

1.A.11 in terms of a smaller set of “independent” variables.

Since all other traders submit demand functions given by equation 1.4.1 and trader  $i$  buys  $x^l$ , the market clearing condition together with equation 1.A.12 imply that the price  $p^l$  is given by

$$p^l = \frac{A^l}{B^l} - \frac{a^l \sum_{j \neq i} e_j^l + u^l}{NB^l} - \frac{a^l e_i^l}{NB^l} - \frac{A_s^l}{B^l} (s^l - \Delta s^l) - \frac{A_z^l}{B^l} (\bar{e}^l - \frac{\Delta e_i^l}{N-1}) + \frac{\Delta x^l}{(N-1)B^l}. \quad (1.A.13)$$

The difference between trader  $i$  asset holdings at date  $l+1$  (and before trade takes place) and his asset holdings if he had not deviated from equilibrium at the previous dates, which was denoted by  $\Delta e_i^{l+1}$  is clearly equal to  $\Delta e_i^l + \Delta x^l$ .

We now come to  $\Delta s^{l+1}$  i.e. the “misperceptions” of the other traders about  $s^{l+1}$ . Since large trader  $i$  deviates from his equilibrium strategy at date  $l$  by buying  $\Delta x^l$  more, each other trader buys  $\Delta x^l / (N-1)$  less. They thus believe that the total quantity supplied at date  $l$  is lower by  $\Delta x^l N / (N-1)$  Equation 1.A.6 implies then that

$$\Delta s^{l+1} = (1 - a^l) \left[ (1 - \beta^{l,l}) \Delta s^l + \frac{\beta^{l,l}}{a^l} \frac{\Delta x^l}{N-1} \right]. \quad (1.A.14)$$

Using equations 1.A.1, 1.A.7, 1.A.13, 1.A.12 and 1.A.14 as well as the expression for  $\Delta e_i^{l+1}$  we can rewrite expression 1.A.11 as

$$\begin{aligned} & E^l - \exp(-\alpha(M^l - \left( \frac{A^l}{B^l} - \frac{a^l \sum_{j \neq i} e_j^l + u^l}{NB^l} - \frac{a^l e_i^l}{NB^l} - \frac{A_s^l}{B^l} (s^l - \Delta s^l) \right. \\ & \left. - \frac{A_z^l}{B^l} (\bar{e}^l - \frac{\Delta e_i^l}{N-1}) + \frac{\Delta x^l}{(N-1)B^l} \right) \left( \frac{a^l \sum_{j \neq i} e_j^l + u^l}{N} + \frac{a^l e_i^l}{N} - a^l e_i^l + \Delta x^l \right) \\ & + d^l \left( e_i^l + \bar{e}^l + \Delta e_i^l + \frac{a^l \sum_{j \neq i} e_j^l + u^l}{N} + \frac{a^l e_i^l}{N} - a^l e_i^l + \Delta x^l \right) \\ & - \frac{\alpha \sigma^2}{2} (L - l + 1) \left( e_i^l + \bar{e}^l + \Delta e_i^l + \frac{a^l \sum_{j \neq i} e_j^l + u^l}{N} + \frac{a^l e_i^l}{N} - a^l e_i^l + \Delta x^l \right)^2 \\ & + Q_1^{l+1} (1 - a^l)^2 \left( (1 - \beta^{l,l}) (s^l - (\beta^{l-1,l} + \dots + \beta^{l-1,l-1}) \frac{e_i^l}{N}) \right) \end{aligned}$$

$$\begin{aligned}
& +\beta^{l,l}\left(\frac{\sum_{j\neq i}e_j^l}{N}+\frac{u^l}{a^lN}\right)+\frac{e_i^l}{N}-e_i^l\Big)^2 \\
& -\left(Q_2^{l+1}(1-a^l)\left[(1-\beta^{l,l})\Delta s^l+\frac{\beta^{l,l}}{a^l}\frac{\Delta x^l}{N-1}\right]+Q_3^{l+1}(\Delta e_i^l+\Delta x^l)\right)(1-a^l) \\
& \left((1-\beta^{l,l})(s^l-(\beta^{l-1,1}+\dots+\beta^{l-1,l-1})\frac{e_i^l}{N})+\beta^{l,l}\left(\frac{\sum_{j\neq i}e_j^l}{N}+\frac{u^l}{a^lN}\right)+\frac{e_i^l}{N}-e_i^l\right) \\
& +Q_4^{l+1}(1-a^l)^2\left((1-\beta^{l,l})\Delta s^l+\frac{\beta^{l,l}}{a^l}\frac{\Delta x^l}{N-1}\right)^2+Q_5^{l+1}(\Delta e_i^l+\Delta x^l)^2 \\
& +Q_6^{l+1}(1-a^l)\left((1-\beta^{l,l})\Delta s^l+\frac{\beta^{l,l}}{a^l}\frac{\Delta x^l}{N-1}\right)(\Delta e_i^l+\Delta x^l)+Q_7^{l+1}). \tag{1.A.15}
\end{aligned}$$

It is well-known (see Kyle (1989)) that since trader  $i$  conditions his demand on price, he can choose  $\Delta x^l$  as if he knew the information about the total hedging demand revealed at date  $l$ . He thus maximizes expression 1.A.15 w.r.t.  $\Delta x^l$  for any value of the quantity that the other market participants supply at date  $l$ ,  $a^l \sum_{j\neq i} e_j^l + u^l$ . The first-order condition of the above problem can be derived in a straightforward manner and is omitted. It is also easy to show that the second-order condition is

$$\begin{aligned}
\mathcal{S} \equiv & \frac{2}{(N-1)B^l} + \alpha\sigma^2(L+1-l) - 2Q_4^{l+1}(1-a^l)^2 \left(\frac{\beta^{l,l}}{a^l} \frac{1}{(N-1)}\right)^2 \\
& - 2Q_5^{l+1} - 2Q_6^{l+1}(1-a^l) \frac{\beta^{l,l}}{a^l} \frac{1}{(N-1)} > 0. \tag{1.A.16}
\end{aligned}$$

To derive necessary conditions for an equilibrium we suppose that trader  $i$  has not deviated from equilibrium at the previous dates, i.e.  $\Delta s^l = 0$  and  $\Delta e_i^l = 0$ . Then, he should not have an incentive to deviate at date  $l$  (i.e.  $\Delta x^l = 0$ ) for any values of  $d^l$ ,  $a^l \sum_{j\neq i} e_j^l + u^l$ ,  $e_i^l$ ,  $s^l$  and  $\bar{e}^l$ . We thus set  $\Delta s^l = 0$ ,  $\Delta e_i^l = 0$  and  $\Delta x^l = 0$  in the first-order condition and identify terms in  $d^l$ ,  $a^l \sum_{j\neq i} e_j^l + u^l$ ,  $e_i^l$ ,  $s^l$  and  $\bar{e}^l$ . Identification of the terms in  $d^l$ ,  $a^l \sum_{j\neq i} e_j^l + u^l$ ,  $s^l$  and  $\bar{e}^l$  gives us equations 1.4.11 to 1.4.14. Identifying terms in  $e_i^l$  we get

$$\frac{2a^l}{NB^l} - \alpha\sigma^2(L+1-l)\left(1 - \frac{N-1}{N}a^l\right) + (1-a^l)\left(Q_2^{l+1}(1-a^l)\frac{\beta^{l,l}}{a^l}\frac{1}{N-1} + Q_3^{l+1}\right)$$

$$\left(\frac{N-1}{N} + \frac{1}{N}(1 - \beta^{l,l})(\beta^{l-1,1} + \dots + \beta^{l-1,l-1})\right) = 0. \quad (1.A.17)$$

Equation 1.A.17 together with equation 1.4.13 give us equation 1.4.15. We can also combine equation 1.A.17 with equation 1.4.14 to get the following equation that will be useful later

$$\begin{aligned} \frac{2a^l}{NB^l} + \frac{A_s^l}{NB^l}(\beta^{l-1,1} + \dots + \beta^{l-1,l-1}) - \alpha\sigma^2(L+1-l)\left(1 - \frac{N-1}{N}a^l\right) \\ + (1-a^l)\left(Q_2^{l+1}(1-a^l)\frac{\beta^{l,l}}{a^l}\frac{1}{N-1} + Q_3^{l+1}\right)\frac{N-1}{N} = 0. \end{aligned} \quad (1.A.18)$$

Finally, using the necessary conditions we can derive  $\Delta x^{l*}$  (i.e. the optimum value of  $\Delta x^l$ ) from the first-order condition (written now for any values of  $\Delta s^l$  and  $\Delta e_i^l$ ). We find

$$\begin{aligned} \Delta x^{l*} = -\frac{1}{S} \left( \frac{A_s^l}{B^l} \Delta s^l + \frac{A_e^l}{B^l} \frac{\Delta e_i^l}{N-1} + \alpha\sigma^2(L+1-l)\Delta e_i^l - 2Q_4^{l+1}(1-a^l)^2(1-\beta^{l,l}) \right. \\ \left. \frac{\beta^{l,l}}{a^l} \frac{1}{N-1} \Delta s^l - 2Q_5^{l+1} \Delta e_i^l - Q_6^{l+1}(1-a^l)[(1-\beta^{l,l})\Delta s^l + \frac{\beta^{l,l}}{a^l} \frac{1}{N-1} \Delta e_i^l] \right). \end{aligned} \quad (1.A.19)$$

Before proceeding to step 3 of the proof we briefly discuss the optimization problem of trader  $i$  when he has not deviated from equilibrium at the previous dates, i.e. expression 1.A.15 for  $\Delta s^l = 0$  and  $\Delta e_i^l = 0$ . For concreteness we suppose that this trader is a buyer in equilibrium and considers deviating from his equilibrium strategy reducing his purchases by a small quantity. If the market were not open in the future, his payoff would change because (i) the price at which the rest of his purchases are executed would fall, (ii) he would not pay for that quantity and (iii) he would not receive the hedging benefit of having that quantity as well as the final payoff it provides. These effects appear by differentiating the first four lines of expression 1.A.15. Since the market is open in the future, the value of buying less is larger. First, trader  $i$  can also hedge at a future date and buying more now has less hedging value. In addition other traders' future demands are lower for two reasons. First they are more willing to sell having sold less, and second they expect to be selling less in

the future too (since they believe that  $s^{l+1}$  is higher). Trader  $i$ 's future purchases are thus executed at lower expected prices. These effects appear by differentiating the seventh and eighth lines of expression 1.A.15<sup>36</sup> and their interpretation will become clear with the recursive relations for the  $Q_k$ 's derived at the end of this subsection.

### Step 3

We will now rewrite expression 1.A.15 for the value of  $\Delta x^l$  that maximizes it,  $\Delta x^{l*}$ . Since this expression is quadratic in  $\Delta x^l$ , we get<sup>37</sup>

$$\begin{aligned}
& - \left( \frac{A^l}{B^l} - \frac{a^l \sum_{j \neq i} e_j^l + u^l}{NB^l} - \frac{a^l e_i^l}{NB^l} - \frac{A_s^l}{B^l} s^l - \frac{A_{\bar{e}}^l}{B^l} \bar{e}^l \right) \left( \frac{a^l \sum_{j \neq i} e_j^l + u^l}{N} + \frac{a^l e_i^l}{N} - a^l e_i^l \right) \\
& \quad + d^l \left( e_i^l + \bar{e}^l + \Delta e_i^l + \frac{a^l \sum_{j \neq i} e_j^l + u^l}{N} + \frac{a^l e_i^l}{N} - a^l e_i^l \right) \\
& \quad - \frac{\alpha \sigma^2}{2} (L - l + 1) \left( e_i^l + \bar{e}^l + \Delta e_i^l + \frac{a^l \sum_{j \neq i} e_j^l + u^l}{N} + \frac{a^l e_i^l}{N} - a^l e_i^l \right)^2 \\
& \quad + Q_1^{l+1} (1 - a^l)^2 \left( (1 - \beta^{l,l}) E^l \left( \frac{\sum_{j \neq i} e_j^l}{N} \right) + \beta^{l,l} \left( \frac{\sum_{j \neq i} e_j^l}{N} + \frac{u^l}{a^l N} \right) + \frac{e_i^l}{N} - e_i^l \right)^2 \\
& \quad - \left( \frac{A_s^l}{B^l} \Delta s^l + \frac{A_{\bar{e}}^l}{B^l} \frac{\Delta e_i^l}{N - 1} \right) \left( \frac{a^l \sum_{j \neq i} e_j^l + u^l}{N} + \frac{a^l e_i^l}{N} - a^l e_i^l \right) \\
& \quad - \left( Q_2^{l+1} (1 - a^l) (1 - \beta^{l,l}) \Delta s^l + Q_3^{l+1} \Delta e_i^l \right) (1 - a^l) \\
& \quad \left( (1 - \beta^{l,l}) E^l \left( \frac{\sum_{j \neq i} e_j^l}{N} \right) + \beta^{l,l} \left( \frac{\sum_{j \neq i} e_j^l}{N} + \frac{u^l}{a^l N} \right) + \frac{e_i^l}{N} - e_i^l \right) \\
& \quad + Q_4^{l+1} (1 - a^l)^2 \left( (1 - \beta^{l,l}) \Delta s^l \right)^2 + Q_5^{l+1} (\Delta e_i^l)^2 + Q_6^{l+1} (1 - a^l) (1 - \beta^{l,l}) \Delta s^l \Delta e_i^l \\
& \quad + Q_7^{l+1} + \frac{1}{2} S (\Delta x^{l*})^2. \tag{1.A.20}
\end{aligned}$$

Using the first-order condition written for  $\Delta s^l = 0$  and  $\Delta e_i^l = 0$ , we now eliminate

<sup>36</sup>The fifth and sixth lines are constant w.r.t.  $\Delta x^l$  and the last two lines are in  $(\Delta x^l)^2$ .

<sup>37</sup>The expression  $A + Bx - (1/2)Cx^2$ ,  $C > 0$  is maximum for  $x^* = B/C$ . This maximum is equal to  $A + B^2/2C$  or, equivalently,  $A + (1/2)C(x^*)^2$ .

the price at date  $l$ ,

$$\frac{A^l}{B^l} - \frac{a^l \sum_{j \neq i} e_j^l + u^l}{NB^l} - \frac{a^l e_i^l}{NB^l} - \frac{A_s^l}{B^l} s^l - \frac{A_{\bar{e}}^l}{B^l} \bar{e}^l$$

in expression 1.A.20. We get

$$\begin{aligned} & d^l(e_i^l + \bar{e}^l + \Delta e_i^l) - \frac{\alpha \sigma^2}{2}(L - l + 1)(e_i^l + \bar{e}^l + \Delta e_i^l)^2 \\ & + \left( \frac{1}{(N-1)B^l} + \frac{1}{2}\alpha \sigma^2(L+1-l) \right) \left( \frac{a^l \sum_{j \neq i} e_j^l + u^l}{N} + \frac{a^l e_i^l}{N} - a^l e_i^l \right)^2 \\ & + Q_1^{l+1}(1-a^l)^2 \left( (1-\beta^{l,l})E^l\left(\frac{\sum_{j \neq i} e_j^l}{N}\right) + \beta^{l,l}\left(\frac{\sum_{j \neq i} e_j^l}{N} + \frac{u^l}{a^l N}\right) + \frac{e_i^l}{N} - e_i^l \right)^2 \\ & + \left( Q_2^{l+1}(1-a^l)\frac{\beta^{l,l}}{a^l} \frac{1}{N-1} + Q_3^{l+1} \right) (1-a^l) \left( (1-\beta^{l,l})E^l\left(\frac{\sum_{j \neq i} e_j^l}{N}\right) \right. \\ & \left. + \beta^{l,l}\left(\frac{\sum_{j \neq i} e_j^l}{N} + \frac{u^l}{a^l N}\right) + \frac{e_i^l}{N} - e_i^l \right) \left( \frac{a^l \sum_{j \neq i} e_j^l + u^l}{N} + \frac{a^l e_i^l}{N} - a^l e_i^l \right) \\ & - \left( \frac{A_s^l}{B^l} \Delta s^l + \frac{A_{\bar{e}}^l}{B^l} \frac{\Delta e_i^l}{N-1} + \alpha \sigma^2(L+1-l)\Delta e_i^l \right) \left( \frac{a^l \sum_{j \neq i} e_j^l + u^l}{N} + \frac{a^l e_i^l}{N} - a^l e_i^l \right) \\ & - \left( Q_2^{l+1}(1-a^l)(1-\beta^{l,l})\Delta s^l + Q_3^{l+1}\Delta e_i^l \right) (1-a^l) \\ & \left( (1-\beta^{l,l})E^l\left(\frac{\sum_{j \neq i} e_j^l}{N}\right) + \beta^{l,l}\left(\frac{\sum_{j \neq i} e_j^l}{N} + \frac{u^l}{a^l N}\right) + \frac{e_i^l}{N} - e_i^l \right) \\ & + Q_4^{l+1}(1-a^l)^2 \left( (1-\beta^{l,l})\Delta s^l \right)^2 + Q_5^{l+1}(\Delta e_i^l)^2 + Q_6^{l+1}(1-a^l)(1-\beta^{l,l})\Delta s^l \Delta e_i^l \\ & + Q_7^{l+1} + \frac{1}{2}S(\Delta x^l)^2. \end{aligned} \tag{1.A.21}$$

Finally we have to take expectations w.r.t. the quantity that market participants, except trader  $i$ , supply at date  $l$  i.e.  $a^l \sum_{j \neq i} e_j^l + u^l$ . Conditional on trader  $i$ 's information before date  $l$  trade,  $(a^l \sum_{j \neq i} e_j^l + u^l)/N$  is normally distributed with mean  $a^l E^l \sum_{j \neq i} e_j^l / N$  and variance  $(a^l)^2 V^l + \sigma_u^2 / N^2$ . To take expectations we define  $\epsilon$  by

$$\epsilon = \frac{a^l \sum_{j \neq i} e_j^l + u^l}{N} - \frac{a^l E^l \sum_{j \neq i} e_j^l}{N}$$

and rewrite expression 1.A.21 separating terms that do not depend on  $\epsilon$ , terms in  $\epsilon$  and terms in  $\epsilon^2$ . We then use the formula

$$\begin{aligned} E\left(\exp(-\alpha(\mathcal{N}\epsilon + \mathcal{D}\epsilon^2))\right) &= \frac{1}{\sqrt{2\pi}\sigma_\epsilon} \int_{-\infty}^{\infty} \exp(-\alpha(\mathcal{N}\epsilon + \mathcal{D}\epsilon^2)) \exp(-\frac{\epsilon^2}{2\sigma_\epsilon^2}) d\epsilon \\ &= \frac{1}{\sqrt{1 + 2\alpha\mathcal{D}\sigma_\epsilon^2}} \exp\left(\frac{\alpha^2\mathcal{N}^2\sigma_\epsilon^2}{2(1 + 2\alpha\mathcal{D}\sigma_\epsilon^2)}\right). \end{aligned} \quad (1.A.22)$$

Straightforward algebra shows that trader  $i$ 's expected utility has the form given in the statement of Theorem 1.A.1 and that the induction hypothesis holds. The  $Q_k$ 's are given by the following equations

$$\begin{aligned} Q_1^l &= \left(\frac{1}{(N-1)B^l} + \frac{1}{2}\alpha\sigma^2(L+1-l)\right) (a^l)^2 + Q_1^{l+1}(1-a^l)^2 \\ &+ \left(Q_2^{l+1}(1-a^l)\frac{\beta^{l,l}}{a^l} \frac{1}{N-1} + Q_3^{l+1}\right) (1-a^l)a^l - \frac{\alpha\mathcal{N}_1^2 \left((a^l)^2V^l + \frac{\sigma^2}{N^{\frac{1}{2}}}\right)}{2(1 + 2\alpha\mathcal{D} \left((a^l)^2V^l + \frac{\sigma^2}{N^{\frac{1}{2}}}\right))} \end{aligned} \quad (1.A.23)$$

$$Q_2^l = \frac{A_2^l}{B^l} a^l + Q_2^{l+1}(1-a^l)^2(1-\beta^{l,l}) - \frac{\alpha\mathcal{N}_1\mathcal{N}_2 \left((a^l)^2V^l + \frac{\sigma^2}{N^{\frac{1}{2}}}\right)}{(1 + 2\alpha\mathcal{D} \left((a^l)^2V^l + \frac{\sigma^2}{N^{\frac{1}{2}}}\right))} \quad (1.A.24)$$

$$Q_3^l = \left(\frac{A_3^l}{B^l} \frac{1}{N-1} + \alpha\sigma^2(L+1-l)\right) a^l + Q_3^{l+1}(1-a^l) - \frac{\alpha\mathcal{N}_1\mathcal{N}_3 \left((a^l)^2V^l + \frac{\sigma^2}{N^{\frac{1}{2}}}\right)}{(1 + 2\alpha\mathcal{D} \left((a^l)^2V^l + \frac{\sigma^2}{N^{\frac{1}{2}}}\right))} \quad (1.A.25)$$

$$\begin{aligned} Q_4^l(\Delta s^l)^2 + Q_5^l(\Delta e_i^l)^2 + Q_6^l\Delta s^l\Delta e_i^l &= Q_4^{l+1} \left((1-a^l)(1-\beta^{l,l})\Delta s^l\right)^2 + Q_5^{l+1}(\Delta e_i^l)^2 \\ &+ Q_6^{l+1}(1-a^l)(1-\beta^{l,l})\Delta s^l\Delta e_i^l + \frac{1}{2}\mathcal{S}(\Delta x^{l*})^2 - \frac{\alpha(\mathcal{N}_2\Delta s^l + \mathcal{N}_3\Delta e_i^l)^2 \left((a^l)^2V^l + \frac{\sigma^2}{N^{\frac{1}{2}}}\right)}{2(1 + 2\alpha\mathcal{D} \left((a^l)^2V^l + \frac{\sigma^2}{N^{\frac{1}{2}}}\right))} \end{aligned} \quad (1.A.26)$$

and

$$\exp(-\alpha Q_7^l) = \exp(-\alpha Q_7^{l+1}) \frac{1}{\sqrt{(1 + 2\alpha\mathcal{D} \left((a^l)^2V^l + \frac{\sigma^2}{N^{\frac{1}{2}}}\right))}}. \quad (1.A.27)$$

$\mathcal{N}_1$ ,  $\mathcal{N}_2$ ,  $\mathcal{N}_3$  and  $\mathcal{D}$  are given by



$$\begin{aligned} \mathcal{N}_1 = & 2 \left( \frac{1}{(N-1)B^l} + \frac{1}{2} \alpha \sigma^2 (L+1-l) \right) a^l + 2Q_1^{l+1} (1-a^l)^2 \frac{\beta^{l,l}}{a^l} \\ & + \left( Q_2^{l+1} (1-a^l) \frac{\beta^{l,l}}{a^l} \frac{1}{N-1} + Q_3^{l+1} \right) (1-a^l) (1+\beta^{l,l}) \end{aligned} \quad (1.A.28)$$

$$\mathcal{N}_2 = \left( \frac{A_s^l}{B^l} + Q_2^{l+1} (1-a^l)^2 (1-\beta^{l,l}) \frac{\beta^{l,l}}{a^l} \right) \quad (1.A.29)$$

$$\mathcal{N}_3 = \left( \frac{A_s^l}{B^l} \frac{1}{N-1} + \alpha \sigma^2 (L+1-l) + Q_3^{l+1} (1-a^l) \frac{\beta^{l,l}}{a^l} \right) \quad (1.A.30)$$

and

$$\begin{aligned} \mathcal{D} = & \left( \frac{1}{(N-1)B^l} + \frac{1}{2} \alpha \sigma^2 (L+1-l) \right) + Q_1^{l+1} (1-a^l)^2 \left( \frac{\beta^{l,l}}{a^l} \right)^2 \\ & + \left( Q_2^{l+1} (1-a^l) \frac{\beta^{l,l}}{a^l} \frac{1}{N-1} + Q_3^{l+1} \right) (1-a^l) \frac{\beta^{l,l}}{a^l}. \end{aligned} \quad (1.A.31)$$

We briefly discuss equations 1.A.24 and 1.A.25 in order to interpret the coefficients  $Q_2$  and  $Q_3$ . As it was stated before, in a dynamic model trader  $i$  who, for instance, is a buyer benefits more by buying less at a given date than in a static model. First, this trader can also hedge at a future date and buying more now has less hedging value. In addition other traders' future demands are lower for two reasons. First they are more willing to sell having sold less, and second they expect to be selling less in the future too (since they believe that  $s$  is higher). Trader  $i$ 's future purchases are thus executed at lower expected prices.

The term in  $Q_2$  in the value function captures this last benefit of buying less, i.e. to "manipulate" others' beliefs. This term is proportional to the other traders' "misperceptions" of  $s$  and to trader  $i$ 's expected future purchases. As equation 1.A.24 shows  $Q_2^l$  is higher if the market attributes a high weight on  $s^l$ , i.e. if  $A_s^l/B^l$  is high, and if the benefit of "manipulating" others' beliefs from date  $l+1$  on,  $Q_2^{l+1}$  is high. The last term in this equation implies that the benefit of buying less and manipulating others' beliefs is lower. The reason is that this benefit is risky; it is higher when there

is substantial (unexpected) hedging demand at future dates, the price is low and trader  $i$  buys more i.e. exactly when his payoff is high. This last term disappears if we assume that price movements coming from fundamentals are much larger than price movements coming from the trades of the different market participants.

The term in  $Q_3$  in the value function captures the first two benefits of buying less (relative to a static model). The discussion is very similar in spirit to the one given above and is omitted.

Theorem 1.A.1 is now proven. At the same time we derived necessary and sufficient conditions for an equilibrium and proved Proposition 1.4.1. Parameters  $A^l$ ,  $B^l$ ,  $a^l$ ,  $A'_l$  and  $A''_l$  (for  $1 \leq l \leq L$ ) define equilibrium demands only if there exist  $Q_2^{l+1}$  and  $Q_3^{l+1}$  so that equations 1.4.11 to 1.4.15 hold. In addition  $Q_1^l$ ,  $Q_2^l$  and  $Q_3^l$  are linked to their values at date  $l + 1$  by equations 1.A.23, 1.A.24 and 1.A.25. Moreover  $Q_1^{L+1} = 0$ ,  $Q_2^{L+1} = 0$  and  $Q_3^{L+1} = 0$ .

We have also derived sufficient conditions. If the second-order condition 1.A.16 holds, then parameters  $A^l$ ,  $B^l$ ,  $a^l$ ,  $A'_l$  and  $A''_l$  (that satisfy the necessary conditions) define equilibrium demands. The relations between  $Q_4^l$ ,  $Q_5^l$  and  $Q_6^l$  and their values at date  $l + 1$  as well as  $Q_1^{l+1}$ ,  $Q_2^{l+1}$  and  $Q_3^{l+1}$  can be derived from equation 1.A.26. In addition,  $Q_4^{L+1} = 0$ ,  $Q_5^{L+1} = 0$  and  $Q_6^{L+1} = 0$ .

In the multiperiod case we have neither shown that an equilibrium exists i.e. that the necessary conditions admit a solution which also satisfies the second-order conditions, nor that an equilibrium is unique. In all the numerical simulations that we have carried out the second-order condition was satisfied. We have shown existence and uniqueness in the two-period case (studied in subsection 1.A.5) and we hope to generalize these results to the multiperiod case.

### 1.A.3 A Numerical Algorithm for Solving the Model

In this subsection we present the numerical algorithm that we used to determine a solution to equations 1.4.11 to 1.4.15 and 1.A.23, 1.A.24 and 1.A.25. This algorithm is similar in spirit to the one followed in Kyle (1985).

We first set  $Q_1^{L+1} = 0$ ,  $Q_2^{L+1} = 0$  and  $Q_3^{L+1} = 0$ . We now have to go backwards.

First, knowing  $Q_2^{l+1}$  and  $Q_3^{l+1}$  we determine  $a^l$  from equation 1.4.15. We then use  $Q_2^{l+1}$ ,  $Q_3^{l+1}$  and  $a^l$  to determine  $A^l$ ,  $B^l$ ,  $A_s^l$  and  $A_z^l$  from equations 1.4.11 to 1.4.14. Finally, we determine  $Q_1^l$ ,  $Q_2^l$  and  $Q_3^l$  from equations 1.A.23, 1.A.24 and 1.A.25.

In order to do this, however, we need to know the  $\beta$ 's at date  $l$  as well as  $V^l$ . This is equivalent to knowing

$$F^l = (1 - a^1) \cdots (1 - a^{l-1}) \quad (1.A.32)$$

and

$$U^l = \sum_{l'' < l} [(1 - a^1)^2 \cdots (1 - a^{l''-1})^2 (a^{l''})^2]. \quad (1.A.33)$$

These can also be defined recursively but from date 1 (not date  $L + 1$ ).

Since we want to start from date  $L + 1$  we will take arbitrary values for  $F^{L+1}$  and  $U^{L+1}$ . The  $\beta$ 's at date  $l$  and  $V^l$  can easily be expressed as a function of  $F^{l+1}$ ,  $U^{l+1}$  and  $a^l$ . For instance

$$\beta^{l,l} = \frac{\left(\frac{F^{l+1}}{1-a^l}\right)^2 (a^l)^2 \Sigma_e^2 (N-1)}{U^{l+1} \Sigma_e^2 (N-1) + \sigma_u^2}.$$

Given thus  $Q_2^{l+1}$ ,  $Q_3^{l+1}$ ,  $F^{l+1}$  and  $U^{l+1}$ , we solve for  $a^l$ . Then we proceed as outlined above. We can also derive  $F^l$  and  $U^l$  from  $F^{l+1}$ ,  $U^{l+1}$  and  $a^l$ . Indeed, equations 1.A.32 and 1.A.33 imply that

$$F^l = \frac{F^{l+1}}{1 - a^l}$$

and

$$U^l = U^{l+1} - \left(\frac{F^{l+1}}{1 - a^l} a^l\right)^2.$$

We repeat this operation until  $F^1 = 1$  and  $U^1 = 0$ .

#### 1.A.4 An Alternative Characterization of Equilibrium

In this subsection we present an alternative way of characterizing equilibrium. In subsection 1.A.2 we wrote that the expected utility of trader  $i$  did not change when

he modified his current trades but kept future trades the same. Here we will use the results of that subsection and write that trader  $i$ 's expected utility remains the same when he modifies his trades at the current date by a given quantity and his trades at the next date by the *opposite* quantity. This is thus equivalent to combining two of the first-order conditions derived in subsection 1.A.2 at successive dates.

The equations that we will derive have not been used much in proofs. We believe however that they better illustrate agents' tradeoffs; this is why the discussion in sections 1.3 and 1.4 was based on these equations.

We assume that we are at date  $l$ , trader  $i$  has not deviated at the previous dates and considers modifying his trades at that date by  $\Delta x^l$  and his trades at date  $l + 1$  by  $-\Delta x^l$ . Proceeding as in the derivation of equation 1.A.11 (but going until date  $l + 2$ ) we find that his expected payoff is

$$\begin{aligned}
E^l - \exp(-\alpha (M^l - p^l x^l - p^{l+1} x^{l+1} + d^{l+1}(e_i^l + \bar{e}^l + x^l + x^{l+1})) \\
- \frac{1}{2} \alpha \sigma^2 (L - l)(e_i^l + \bar{e}^l + x^l + x^{l+1})^2 + Q_1^{l+2} (E^{l+2} (\frac{\sum_{j \neq i} e_j^{l+2}}{N}) + \frac{e_i^{l+2}}{N} - e_i^{l+2})^2 \\
- Q_2^{l+2} \Delta s^{l+2} + Q_4^{l+2} (\Delta s^{l+2})^2 + Q_7^{l+2})). \tag{1.A.34}
\end{aligned}$$

In contrast to equation 1.A.11 there are no terms in  $\Delta e_i^{l+2}$ . This happens because trader  $i$  conducts two trades in the opposite direction so that  $\Delta e_i^{l+2} = 0$ .

$x^l$  and  $x^{l+1}$  represent trader  $i$ 's purchases at date  $l$ . As before, we have

$$x^l = \frac{a^l \sum_{j \neq i} e_j^l + u^l}{N} + \frac{a^l e_i^l}{N} - a^l e_i^l + \Delta x^l \tag{1.A.35}$$

and

$$x^{l+1} = \frac{a^{l+1} \sum_{j \neq i} e_j^{l+1} + u^{l+1}}{N} + \frac{a^{l+1} e_i^{l+1}}{N} - a^{l+1} e_i^{l+1} - \Delta x^{l+1}. \tag{1.A.36}$$

The price,  $p^l$ , is given by equation 1.A.13, for  $\Delta e_i^l = 0$  and  $\Delta s^l = 0$ .  $p^{l+1}$  is given by the same equation with  $\Delta e_i^{l+1} = \Delta x^l / (N - 1)$ ,  $\Delta s^{l+1} = (1 - a^l)(\beta^{l,l} / a^l) \Delta x^l / (N - 1)$  and  $\Delta x^{l+1} = -\Delta x^l$ .

Finally, use of the recursive equation 1.A.6 for  $s$  shows that

$$\begin{aligned}\Delta s^{l+2} &= (1 - a^{l+1})\left((1 - a^l)(1 - \beta^{l+1,l+1})\frac{\beta^{l,l}}{a^l} - \frac{\beta^{l+1,l+1}}{a^{l+1}}\right)\frac{\Delta x^l}{N-1} \\ &= (1 - a^{l+1})\left((1 - a^l)\frac{\beta^{l+1,l}}{a^l} - \frac{\beta^{l+1,l+1}}{a^{l+1}}\right)\frac{\Delta x^l}{N-1}.\end{aligned}\quad (1.A.37)$$

Plugging  $x^l, x^{l+1}, p^l, p^{l+1}$  and  $\Delta s^{l+2}$  back in expression 1.A.34 and using equation 1.A.7, we get

$$\begin{aligned}E^l - \exp(-\alpha(M^l - \left(\frac{A^l}{B^l} - \frac{a^l \sum_{j \neq i} e_j^l + u^l}{NB^l} - \frac{a^l e_i^l}{NB^l} - \frac{A_s^l}{B^l} s^l - \frac{A_s^l}{B^l} \bar{e}^l + \frac{\Delta x^l}{(N-1)B^l}\right)) \\ \left(\frac{a^l \sum_{j \neq i} e_j^l + u^l}{N} + \frac{a^l e_i^l}{N} - a^l e_i^l + \Delta x^l\right) - \left(\frac{A^{l+1}}{B^{l+1}} - \frac{a^{l+1} \sum_{j \neq i} e_j^{l+1} + u^{l+1}}{NB^{l+1}} - \frac{a^{l+1} e_i^{l+1}}{NB^{l+1}}\right. \\ \left. - \frac{A_s^{l+1}}{B^{l+1}}(s^{l+1} - (1 - a^l)\frac{\beta^{l,l}}{a^l} \frac{\Delta x^l}{N-1}) - \frac{A_s^{l+1}}{B^{l+1}}(\bar{e}^{l+1} - \frac{\Delta x^l}{N-1}) - \frac{\Delta x^{l+1}}{(N-1)B^{l+1}}\right) \\ \left(\frac{a^{l+1} \sum_{j \neq i} e_j^{l+1} + u^{l+1}}{N} + \frac{a^{l+1} e_i^{l+1}}{N} - a^{l+1} e_i^{l+1} - \Delta x^{l+1}\right) \\ + d^{l+1} \left(e_i^{l+1} + \bar{e}^{l+1} + \frac{a^{l+1} \sum_{j \neq i} e_j^{l+1} + u^{l+1}}{N} + \frac{a^{l+1} e_i^{l+1}}{N} - a^{l+1} e_i^{l+1}\right) \\ - \frac{1}{2} \alpha \sigma^2 (L - l) \left(e_i^{l+1} + \bar{e}^{l+1} \frac{a^{l+1} \sum_{j \neq i} e_j^{l+1} + u^{l+1}}{N} + \frac{a^{l+1} e_i^{l+1}}{N} - a^{l+1} e_i^{l+1}\right)^2 \\ + Q_1^{l+2} (1 - a^{l+1})^2 \left((1 - \beta^{l+1,l+1}) E^l \left(\frac{\sum_{j \neq i} e_j^{l+1}}{N}\right) \right. \\ \left. + \beta^{l+1,l+1} \left(\frac{\sum_{j \neq i} e_j^{l+1}}{N} + \frac{u^{l+1}}{a^{l+1} N}\right) + \frac{e_i^{l+1}}{N} - e_i^{l+1}\right)^2 \\ - Q_2^{l+2} (1 - a^{l+1})^2 \left((1 - a^l) \frac{\beta^{l+1,l}}{a^l} - \frac{\beta^{l+1,l+1}}{a^{l+1}}\right) \frac{\Delta x^l}{N-1} \\ \left((1 - \beta^{l+1,l+1}) E^{l+1} \left(\frac{\sum_{j \neq i} e_j^{l+1}}{N}\right) + \beta^{l+1,l+1} \left(\frac{\sum_{j \neq i} e_j^{l+1}}{N} + \frac{u^{l+1}}{a^{l+1} N}\right) + \frac{e_i^{l+1}}{N} - e_i^{l+1}\right) \\ + Q_4^{l+2} (1 - a^{l+1})^2 \left((1 - a^l) \frac{\beta^{l+1,l}}{a^l} - \frac{\beta^{l+1,l+1}}{a^{l+1}}\right)^2 \left(\frac{\Delta x^l}{N-1}\right)^2 + Q_7^{l+1}),\end{aligned}\quad (1.A.38)$$

In equilibrium trader  $i$  does not have an incentive to modify his trades and this expression is maximum for  $\Delta x^l = 0$  for *any* value of the quantity that the other market participants supply at date  $l$ ,  $a^l \sum_{j \neq i} e_j^l + u^l$ . The first-order condition of the above problem is

$$\begin{aligned}
& E^{l+1} - \exp[-\alpha(-\left(\frac{A^{l+1}}{B^{l+1}} - \frac{a^{l+1} \sum_{j \neq i} e_j^{l+1} + u^{l+1}}{NB^{l+1}} - \frac{a^{l+1} e_i^{l+1}}{NB^{l+1}} - \frac{A_s^{l+1}}{B^{l+1}} s^{l+1} \right. \\
& \quad \left. - \frac{A_e^{l+1}}{B^{l+1}} \bar{e}^{l+1}\right) \left(\frac{a^{l+1} \sum_{j \neq i} e_j^{l+1} + u^{l+1}}{N} + \frac{a^{l+1} e_i^{l+1}}{N} - a^{l+1} e_i^{l+1}\right) \\
& \quad + d^{l+1} \left(e_i^{l+1} + \bar{e}^{l+1} + \frac{a^{l+1} \sum_{j \neq i} e_j^{l+1} + u^{l+1}}{N} + \frac{a^{l+1} e_i^{l+1}}{N} - a^{l+1} e_i^{l+1}\right) \\
& \quad - \frac{1}{2} \alpha \sigma^2 (L-l) \left(e_i^{l+1} + \bar{e}^{l+1} \frac{a^{l+1} \sum_{j \neq i} e_j^{l+1} + u^{l+1}}{N} + \frac{a^{l+1} e_i^{l+1}}{N} - a^{l+1} e_i^{l+1}\right)^2 \\
& \quad + Q_1^{l+2} (1-a^{l+1})^2 \left((1-\beta^{l+1,l+1}) E^l \left(\frac{\sum_{j \neq i} e_j^{l+1}}{N}\right) \right. \\
& \quad \left. + \beta^{l+1,l+1} \left(\frac{\sum_{j \neq i} e_j^{l+1}}{N} + \frac{u^{l+1}}{a^{l+1} N}\right) + \frac{e_i^{l+1}}{N} - e_i^{l+1}\right)^2 + Q_7^{l+1}] \\
& \quad \left[-\left(\frac{A^l}{B^l} - \frac{a^l \sum_{j \neq i} e_j^l + u^l}{NB^l} - \frac{a^l e_i^l}{NB^l} - \frac{A_s^l}{B^l} s^l - \frac{A_e^l}{B^l} \bar{e}^l\right) \right. \\
& \quad \left. - \frac{1}{(N-1)B^l} \left(\frac{a^l \sum_{j \neq i} e_j^l + u^l}{N} + \frac{a^l e_i^l}{N} - a^l e_i^l\right) \right. \\
& \quad \left. + \left(\frac{A^{l+1}}{B^{l+1}} - \frac{a^{l+1} \sum_{j \neq i} e_j^{l+1} + u^{l+1}}{NB^{l+1}} - \frac{a^{l+1} e_i^{l+1}}{NB^{l+1}} - \frac{A_s^{l+1}}{B^{l+1}} s^{l+1} - \frac{A_e^{l+1}}{B^{l+1}} \bar{e}^{l+1}\right) \right. \\
& \quad \left. + \left(\frac{1}{(N-1)B^{l+1}} - \frac{A_e^{l+1}}{B^{l+1}} \frac{1}{N-1} - \frac{A_s^{l+1}}{B^{l+1}} (1-a^l) \frac{\beta^{l,l}}{a^l} \frac{1}{N-1}\right) \right. \\
& \quad \left. \left(\frac{a^{l+1} \sum_{j \neq i} e_j^{l+1} + u^{l+1}}{N} + \frac{a^{l+1} e_i^{l+1}}{N} - a^{l+1} e_i^{l+1}\right) \right. \\
& \quad \left. - Q_2^{l+2} (1-a^{l+1})^2 \left((1-a^l) \frac{\beta^{l+1,l}}{a^l} - \frac{\beta^{l+1,l+1}}{a^{l+1}}\right) \frac{1}{N-1} \right]
\end{aligned}$$

$$\left[ (1 - \beta^{l+1, l+1}) E^{l+1} \left( \frac{\sum_{j \neq i} e_j^{l+1}}{N} \right) + \beta^{l+1, l+1} \left( \frac{\sum_{j \neq i} e_j^{l+1}}{N} + \frac{u^{l+1}}{a^{l+1} N} \right) + \frac{e_i^{l+1}}{N} - e_i^{l+1} \right] = 0. \quad (1.A.39)$$

In equation 1.A.39,  $E^{l+}$  denotes expectations taken after the information about the total hedging demand at date  $l$  has been revealed.

The second term in brackets (i.e. the last seven lines of equation 1.A.39) represents the change in trader  $i$ 's trading proceeds when he modifies his date  $l$  trades by  $\Delta x^l$  and his date  $l+1$  trades by  $-\Delta x^l$ . This change is evaluated for all realizations of the uncertainty which in turn can be decomposed into uncertainty about fundamentals (term  $A^{l+1}/B^{l+1} \equiv d^{l+1}$ ) and into uncertainty about the quantity that the other market participants supply at date  $l+1$ ,  $a^{l+1} \sum_{j \neq i} e_j^{l+1} + u^{l+1}$ . Equation 1.A.39 states that the product of trader  $i$ 's marginal utility (which is the term inside the exponential) times the change in his trading proceeds is, in expectation, zero.

To evaluate this expectation we first transform the term inside the exponential eliminating the price at date  $l+1$  exactly as we did to derive expression 1.A.21. We then define  $\epsilon$  by

$$\epsilon = \frac{a^{l+1} \sum_{j \neq i} e_j^{l+1} + u^{l+1}}{N} - \frac{a^{l+1} E^{l+1} \sum_{j \neq i} e_j^{l+1}}{N}$$

as in subsection A.2. We finally rewrite equation 1.A.39 separating terms which are known at  $l^+$ , terms in  $d^{l+1} - d^l$ , terms in  $\epsilon$  and terms in  $\epsilon^2$  and compute the various integrals. These computations are very simple and require use of formula 1.A.22 and of the following formula

$$\begin{aligned} E \left( \exp(-\alpha(\mathcal{N}\epsilon + \mathcal{D}\epsilon^2)) \epsilon \right) &= \frac{1}{\sqrt{2\pi\sigma_\epsilon}} \int_{-\infty}^{\infty} \exp(-\alpha(\mathcal{N}\epsilon + \mathcal{D}\epsilon^2)) \exp\left(-\frac{\epsilon^2}{2\sigma_\epsilon^2}\right) \epsilon d\epsilon \\ &= \frac{1}{\sqrt{1 + 2\alpha\mathcal{D}\sigma_\epsilon^2}} \frac{\alpha\mathcal{N}\sigma_\epsilon^2}{(1 + 2\alpha\mathcal{D}\sigma_\epsilon^2)} \exp\left(\frac{\alpha^2\mathcal{N}^2\sigma_\epsilon^2}{2(1 + 2\alpha\mathcal{D}\sigma_\epsilon^2)}\right). \end{aligned} \quad (1.A.40)$$

We thus get, rearranging terms

$$-\frac{1}{(N-1)B^l} \left( \frac{a^l \sum_{j \neq i} e_j^l + u^l}{N} + \frac{a^l e_i^l}{N} - a^l e_i^l \right)$$

$$\begin{aligned}
& + \left( \frac{1}{(N-1)B^{l+1}} - \frac{A_s^{l+1}}{B^{l+1}} \frac{1}{N-1} - \frac{A_s^{l+1}}{B^{l+1}} (1-a^l) \frac{\beta^{l,l}}{a^l} \frac{1}{N-1} \right) \\
& \quad \left( \frac{a^{l+1} E^{l+1} \sum_{j \neq i} e_j^{l+1}}{N} + \frac{a^{l+1} e_i^{l+1}}{N} - a^{l+1} e_i^{l+1} \right) \\
& \quad - \left( -\frac{a^l \sum_{j \neq i} e_j^l + u^l}{NB^l} - \frac{a^l e_i^l}{NB^l} - \frac{A_s^l}{B^l} s^l - \frac{A_s^l}{B^l} \bar{e}^l \right) \\
& + \left( -\frac{a^{l+1} E^{l+1} \sum_{j \neq i} e_j^{l+1}}{NB^{l+1}} - \frac{a^{l+1} e_i^{l+1}}{NB^{l+1}} - \frac{A_s^{l+1}}{B^{l+1}} s^{l+1} - \frac{A_s^{l+1}}{B^{l+1}} \bar{e}^{l+1} \right) \\
& \quad - \alpha \sigma^2 (e_i^{l+1} + \bar{e}^{l+1}) - Q_2^{l+2} (1-a^{l+1})^2 \\
& \left( (1-a^l) \frac{\beta^{l+1,l}}{a^l} - \frac{\beta^{l+1,l+1}}{a^{l+1}} \right) \frac{1}{N-1} \left( E^{l+1} \left( \frac{\sum_{j \neq i} e_j^{l+1}}{N} \right) + \frac{e_i^{l+1}}{N} - e_i^{l+1} \right) \\
& - \left[ \left( \frac{1}{(N-1)B^{l+1}} - \frac{A_s^{l+1}}{B^{l+1}} \frac{1}{N-1} - \frac{A_s^{l+1}}{B^{l+1}} (1-a^l) \frac{\beta^{l,l}}{a^l} \frac{1}{N-1} \right) \right. \\
& \quad \left. - \frac{1}{B^{l+1}} - Q_2^{l+2} (1-a^{l+1})^2 \left( (1-a^l) \frac{\beta^{l+1,l}}{a^l} - \frac{\beta^{l+1,l+1}}{a^{l+1}} \right) \frac{\beta^{l+1,l+1}}{a^{l+1}} \frac{1}{N-1} \right] \\
& \left( E^{l+1} \left( \frac{\sum_{j \neq i} e_j^{l+1}}{N} \right) + \frac{e_i^{l+1}}{N} - e_i^{l+1} \right) \frac{\alpha \mathcal{N}_1 \left( (a^{l+1})^2 V^{l+1} + \frac{\sigma^2}{N^2} \right)}{(1+2\alpha \mathcal{D} \left( (a^{l+1})^2 V^{l+1} + \frac{\sigma^2}{N^2} \right))} = 0. \quad (1.A.41)
\end{aligned}$$

$\mathcal{N}_1$  and  $\mathcal{D}$  are defined by equations 1.A.28 and 1.A.31, by substituting  $l+1$  for  $l$ .

We will slightly transform equation 1.A.41 in order to interpret the various terms. Equations 1.4.13 and 1.4.14, written for date  $l+1$  imply that

$$\frac{N-2}{N-1} \frac{1}{B^{l+1}} = (L-l) \alpha \sigma^2 + \frac{A_s^{l+1}}{B^{l+1}} \frac{1}{1-\beta^{l+1,l+1}} \frac{\beta^{l+1,l+1}}{a^{l+1}}. \quad (1.A.42)$$

Equation 1.A.42 together with equation 1.4.12 imply that

$$\begin{aligned}
& \frac{1}{(N-1)B^{l+1}} - \frac{A_s^{l+1}}{B^{l+1}} \frac{1}{N-1} - \frac{A_s^{l+1}}{B^{l+1}} (1-a^l) \frac{\beta^{l,l}}{a^l} \frac{1}{N-1} = \\
& \frac{1}{(N-1)^2 B^{l+1}} - \frac{A_s^{l+1}}{B^{l+1}} \frac{1}{1-\beta^{l+1,l+1}} \left( (1-a^l) \frac{\beta^{l+1,l}}{a^l} - \frac{\beta^{l+1,l+1}}{a^{l+1}} \right) \frac{1}{N-1}. \quad (1.A.43)
\end{aligned}$$

Using equations 1.A.24, 1.A.29 and 1.A.43, equation 1.A.41 can be transformed into



$$\begin{aligned}
& -\frac{1}{(N-1)B^l} \left( \frac{a^l \sum_{j \neq i} e_j^l + u^l}{N} + \frac{a^l e_i^l}{N} - a^l e_i^l \right) \\
& + \frac{1}{(N-1)^2 B^{l+1}} \left( \frac{a^{l+1} E^{l+1} \sum_{j \neq i} e_j^{l+1}}{N} + \frac{a^{l+1} e_i^{l+1}}{N} - a^{l+1} e_i^{l+1} \right) \\
& - \left( -\frac{a^l \sum_{j \neq i} e_j^l + u^l}{NB^l} - \frac{a^l e_i^l}{NB^l} - \frac{A_s^l s^l}{B^l} - \frac{A_{\bar{e}}^l \bar{e}^l}{B^l} \right) \\
& + \left( -\frac{a^{l+1} E^{l+1} \sum_{j \neq i} e_j^{l+1}}{NB^{l+1}} - \frac{a^{l+1} e_i^{l+1}}{NB^{l+1}} - \frac{A_s^{l+1} s^{l+1}}{B^{l+1}} - \frac{A_{\bar{e}}^{l+1} \bar{e}^{l+1}}{B^{l+1}} \right) \\
& - \frac{Q_2^{l+1}}{(1 - \beta^{l+1, l+1})} \left( (1 - a^l) \frac{\beta^{l+1, l}}{a^l} - \frac{\beta^{l+1, l+1}}{a^{l+1}} \right) \frac{1}{N-1} \left( E^{l+1} \left( \frac{\sum_{j \neq i} e_j^{l+1}}{N} \right) + \frac{e_i^{l+1}}{N} - e_i^{l+1} \right) \\
& - \alpha \sigma^2 (e_i^{l+1} + \bar{e}^{l+1}) + \frac{N(N-2)}{(N-1)^2} \frac{1}{B^{l+1}} \\
& \left( E^{l+1} \left( \frac{\sum_{j \neq i} e_j^{l+1}}{N} \right) + \frac{e_i^{l+1}}{N} - e_i^{l+1} \right) \frac{\alpha \mathcal{N}_1 \left( (a^{l+1})^2 V^{l+1} + \frac{\sigma^2}{N^2} \right)}{(1 + 2\alpha \mathcal{D} \left( (a^{l+1})^2 V^{l+1} + \frac{\sigma^2}{N^2} \right))} = 0. \quad (1.A.44)
\end{aligned}$$

We now use equation 1.A.44 to illustrate the discussion in sections 1.3 and 1.4. Suppose that trader  $i$  who for concreteness is a buyer reduces his date  $l$  purchases by a small quantity and increases his date  $l+1$  purchases by the same quantity. Then the following will happen:

- He will benefit since the price at which the rest of his date  $l$  purchases are executed will decrease. This effect corresponds to the first term in parentheses in equation 1.A.44. At the same time the date  $l+1$  price will increase because of the imperfectly competitive behavior of the other traders. This latter effect corresponds to the second term in parentheses and is  $N-1$  times smaller than the first effect.
- He will bear the cost of holding a larger unhedged position between dates  $l$  and  $l+1$ . Prices at date  $l+1$  may change because of fundamentals and also because of agents' trades. The cost of bearing more fundamental risk corresponds to the

term  $-\alpha\sigma^2(e_i^{l+1} + \bar{e}^{l+1})$  in equation 1.A.44. The cost of bearing more risk of a price change due to unexpected hedging demand corresponds to the last term in equation 1.A.44. If we assume that price movements coming from fundamentals are much larger than price movements coming from the trades of the different market participants, this last term naturally disappears.

- He will buy the extra unit at the date  $l + 1$  price instead from the date  $l$  price. The third and fourth terms in parentheses in equation 1.A.44 correspond to the difference between the expected date  $l + 1$  price and the date  $l$  price. The date  $l$  price does not fully reflect the impact of trader  $i$ 's date  $l + 1$  trades on the date  $l + 1$  price,  $a^{l+1}e_i^{l+1}/NB^{l+1}$ .
- By “manipulating” the market beliefs about the total hedging demand he will affect the prices at which his future purchases are executed. Not surprisingly this effect corresponds to the term in  $Q_2^{l+1}$  in equation 1.A.44. It is also proportional to the market “misperceptions” which will be important if large traders are expected to trade substantially different quantities at dates  $l$  and  $l + 1$  so that  $(1 - a^l)\beta^{l+1,l}/a^l - \beta^{l+1,l+1}/a^{l+1}$  is large in absolute value.

If we express  $E^{l+1} \sum_{j \neq i} e_j^{l+1}$ ,  $e_i^{l+1}$ ,  $s^{l+1}$  and  $\bar{e}^{l+1}$  as a function of date  $l$  variables and identify terms, we can get equations similar to 1.4.11 to 1.4.15.

### 1.A.5 The Two-Period Case

In this subsection we will prove Lemma 1.3.1 and Proposition 1.3.1. Having characterized equilibrium in Theorem 1.A.1, we only need to prove that an equilibrium indeed exists (i.e. the necessary conditions admit a solution which also satisfies the second-order condition) and that it is unique.

Since the  $Q_k^3$ 's are all zero, equation 1.4.15 implies that  $a^2 = (N - 2)/(N - 1)$ .  $A^2$ ,  $B^2$ ,  $A_s^2$  and  $A_z^2$  are also clearly uniquely defined by equations 1.3.11 to 1.3.14, and the second-order condition at date 2 obviously holds. Simple algebra shows that

$$Q_1^2 = \frac{Q_3^2}{2} = Q_6^2 = \alpha\sigma^2 \frac{N(N-2)}{2(N-1)^2}$$

$$\left(1 + \alpha^2\sigma^2 \frac{1}{N(N-2)} \left[ \left(\frac{N-2}{N-1}\right)^2 \frac{(1-a^1)^2 \Sigma_e^2(N-1)\sigma_u^2}{(a^1)^2 \Sigma_e^2(N-1) + \sigma_u^2} + \sigma_u^2 \right] \right)^{-1} \quad (1.A.45)$$

and that  $Q_2^2 = Q_4^2 = 0$ .

Going back to date 1, equation 1.3.19 can be written as

$$\frac{1 - \frac{N-1}{N-2}a^1}{1 - a^1} = \frac{N(N-2)}{2(N-1)^2} \left( \frac{N-1}{N} + \frac{2(N-1)}{N(N-2)} \frac{(a^1)^2 \Sigma_e^2(N-1)}{(a^1)^2 \Sigma_e^2(N-1) + \sigma_u^2} \right)$$

$$\left(1 + \alpha^2\sigma^2 \frac{1}{N(N-2)} \left[ \left(\frac{N-2}{N-1}\right)^2 \frac{(1-a^1)^2 \Sigma_e^2(N-1)\sigma_u^2}{(a^1)^2 \Sigma_e^2(N-1) + \sigma_u^2} + \sigma_u^2 \right] \right)^{-1}. \quad (1.A.46)$$

The LHS function is continuously decreasing from  $(N-1)/(N-2)$  to  $-\infty$  in  $(-\infty, 1)$ , its value is 1 at 0 and 0 at  $(N-2)/(N-1)$ . In  $(1, \infty)$  it continuously decreases from  $\infty$  to  $(N-1)/(N-2)$ . The RHS function is continuous, positive, bounded above by  $N/2(N-1)$  and increases in  $[0, 1]$ . Since for  $N > 3$  we have  $N/2(N-1) < 1$ , equation 1.3.19 has a unique real solution,  $a^1 \in (0, (N-2)/(N-1))$ . Given  $Q_2^2, Q_3^2$  and  $a^1$ ,  $A^1$  and  $B^1$  are clearly uniquely defined by equations 1.3.17 and 1.3.18. The second-order condition at date 1 also holds since  $Q_6^2 \leq \alpha\sigma^2/2$ .

Finally, if  $\sigma_u^2$  decreases the RHS function increases uniformly. As a result  $a^1$  decreases.

## 1.B Appendix: Determination of Total Welfare

In this section we will determine total welfare. In subsection 1.B.1 we study large traders' welfare and in subsection 1.B.2 we study small traders' welfare and derive equation 1.5.1.

## 1.B.1 Large traders' welfare

To derive a large trader's expected utility before trade at date 1 we use Theorem 1. Omitting numéraire holdings we get

$$- \exp(-\alpha(d^1 e_i^1 - \frac{1}{2}\alpha\sigma^2 L(e_i^1)^2 + Q_1^1 \left(\frac{N-1}{N}\right)^2 (e_i^1)^2 + Q_7^1)). \quad (1.B.1)$$

To derive the trader's expected utility at time 0,<sup>38</sup> in the "interim" sense (i.e. for a given value of his initial endowment) we take expectations w.r.t.  $d^1$ . Since as of time 0,  $d^1$  has zero mean, we get

$$- \exp(-\alpha(-\frac{1}{2}\alpha\sigma^2(L+1)(e_i^1)^2 + Q_1^1 \left(\frac{N-1}{N}\right)^2 (e_i^1)^2 + Q_7^1)). \quad (1.B.2)$$

To derive expected utility in the "ex-ante" sense, we take expectations w.r.t.  $e_i^1$ . Using equation 1.A.22 we get

$$- \exp(-\alpha Q_7^1) \frac{1}{\sqrt{(1 - \alpha(\alpha\sigma^2(L+1) - 2Q_1^1 \left(\frac{N-1}{N}\right)^2) \Sigma_e^2)}}. \quad (1.B.3)$$

A large trader's welfare in the "interim" sense is thus

$$- \frac{1}{2}\alpha\sigma^2(L+1)(e_i^1)^2 + Q_1^1 \left(\frac{N-1}{N}\right)^2 (e_i^1)^2 + Q_7^1 \quad (1.B.4)$$

and in the "ex-ante" sense

$$\frac{1}{2\alpha} \log(1 - \alpha(\alpha\sigma^2(L+1) - 2Q_1^1 \left(\frac{N-1}{N}\right)^2) \Sigma_e^2) + Q_7^1. \quad (1.B.5)$$

We will now assume that price movements coming from fundamentals dominate price movements due to the trades of the different market participants, i.e. that terms in  $\alpha^2\sigma^2\Sigma_e^2$  and terms in  $\alpha^2\sigma^2\sigma_u^2$  are much smaller than 1, and derive expressions for  $Q_1^1$  (step 1),  $Q_7^1$  (step 2) and "ex-ante" welfare (step 3).

### Step 1 $Q_1^1$

---

<sup>38</sup>Remember that date 1 is time  $\frac{1}{L+1}$ .

Under the above assumption, equation 1.A.23 simplifies to

$$Q_1^l = \left( \frac{1}{(N-1)B^l} + \frac{1}{2}\alpha\sigma^2(L+1-l) \right) (a^l)^2 + Q_1^{l+1}(1-a^l)^2 \\ + \left( Q_2^{l+1}(1-a^l)\frac{\beta^{l,l}}{a^l}\frac{1}{N-1} + Q_3^{l+1} \right) (1-a^l)a^l. \quad (1.B.6)$$

We will now prove the following result

**Proposition 1.B.1**  $Q_1^l$  is given by

$$\left( \frac{N-1}{N} \right)^2 Q_1^l = - \sum_{L \geq l' \geq l} \frac{(1-a^l)^2 \dots (1-a^{l'-1})^2 (a^l)^2}{NB^{l'}} \frac{N-1}{N} \\ - \sum_{L \geq l' \geq l} \frac{A_s^{l'}}{NB^{l'}} (\beta^{l'-1,1} + \dots + \beta^{l'-1,l'-1}) (1-a^l)^2 \dots (1-a^{l'-1})^2 a^{l'} \frac{N-1}{N} - \\ \sum_{L \geq l' \geq l+1} \alpha\sigma^2(L+1-l') \left( \sum_{l''-1 \geq l'' \geq l} (1-a^l) \dots (1-a^{l''-1}) a^{l''} \right) (1-a^l) \dots (1-a^{l'-1}) a^{l'} \frac{N-1}{N^2} \\ - \frac{1}{2}\alpha\sigma^2 \sum_{L \geq l' \geq l} \left( \frac{1}{N} + (1-a^l) \dots (1-a^{l'}) \frac{N-1}{N} \right)^2 + \frac{1}{2}\alpha\sigma^2(L+1-l). \quad (1.B.7)$$

We will prove Proposition 1.B.1 by induction. The proposition is obviously true at date  $L+1$  since  $Q_1^{L+1} = 0$ . Assuming that the proposition is true at  $l+1$  we will show that it is true at  $l$ . To do this we will use the recursive equation 1.B.6 for  $Q_1^l$  and equation 1.A.18.

Multiplying both sides of equation 1.B.6 by  $((N-1)/N)^2$  and eliminating the term in  $Q_2^{l+1}(1-a^l)\beta^{l,l}/a^l(N-1) + Q_3^{l+1}$ , using equation 1.A.18, we get

$$\left( \frac{N-1}{N} \right)^2 Q_1^l = - \frac{(a^l)^2 N-1}{NB^l} \frac{N-1}{N} - \frac{A_s^{l-1}}{NB^{l-1}} (\beta^{l-1,1} + \dots + \beta^{l-1,l-1}) a^l \frac{N-1}{N} + \\ \alpha\sigma^2(L+1-l) a^l \frac{N-1}{N} - \frac{1}{2}\alpha\sigma^2(L+1-l) (a^l)^2 \left( \frac{N-1}{N} \right)^2 + Q_1^{l+1} (1-a^l)^2 \left( \frac{N-1}{N} \right)^2. \quad (1.B.8)$$

Using the induction hypothesis and adding and subtracting the term

$$-\alpha\sigma^2(L-l) a^l [(1-a^l)a^{l+1} + \dots + (1-a^l) \dots (1-a^{L-1})a^L] \frac{N-1}{N^2}$$

we get

$$\begin{aligned}
\left(\frac{N-1}{N}\right)^2 Q_1^l &= - \sum_{L \geq l' \geq l} \frac{(1-a^l)^2 \cdots (1-a^{l'-1})^2 (a^l)^2}{NB^{l'}} \frac{N-1}{N} \\
&- \sum_{L \geq l' \geq l} \frac{A_s^{l'}}{NB^{l'}} (\beta^{l'-1,1} + \cdots + \beta^{l'-1,l'-1}) (1-a^l)^2 \cdots (1-a^{l'-1})^2 a^{l'} \frac{N-1}{N} - \\
&\sum_{L \geq l' \geq l+1} \alpha \sigma^2 (L+1-l') \left( \sum_{l''-1 \geq l'' \geq l} (1-a^l) \cdots (1-a^{l''-1}) a^{l''} \right) (1-a^l) \cdots (1-a^{l'-1}) a^{l'} \frac{N-1}{N^2} \\
&+ \alpha \sigma^2 (L-l) a^l [(1-a^l) a^{l+1} + \cdots + (1-a^l) \cdots (1-a^{L-1}) a^L] \frac{N-1}{N^2} - \\
&\frac{1}{2} \alpha \sigma^2 (1-a^l)^2 \sum_{L \geq l' \geq l+1} \left( \frac{1}{N} + (1-a^{l+1}) \cdots (1-a^{l'}) \frac{N-1}{N} \right)^2 + \alpha \sigma^2 (L+1-l) a^l \frac{N-1}{N} \\
&- \frac{1}{2} \alpha \sigma^2 (L+1-l) (a^l)^2 \left( \frac{N-1}{N} \right)^2 + \frac{1}{2} \alpha \sigma^2 (L-l) (1-a^l)^2. \tag{1.B.9}
\end{aligned}$$

It is then a matter of simple algebra to transform the last three lines the above of equation so that it becomes exactly equation 1.B.7.  $\square$

## Step 2 $Q_7^l$

Equation 1.A.27 similarly simplifies to

$$Q_7^l = Q_7^{l+1} + \mathcal{D} \left( (a^l)^2 V^l + \frac{\sigma_u^2}{N^2} \right). \tag{1.B.10}$$

Before deriving an expression for  $Q_7^l$ , we will derive an alternative expression for  $Q_1^l$  in Proposition 1.B.2.

**Proposition 1.B.2**  $Q_1^l$  is given by

$$\begin{aligned}
Q_1^l &= \sum_{L \geq l' \geq l} \frac{(1-a^l)^2 \cdots (1-a^{l'-1})^2 (a^l)^2}{B^{l'}} + \sum_{L \geq l' \geq l} \frac{A_s^{l'}}{B^{l'}} (1-a^l)^2 \cdots (1-a^{l'-1})^2 a^{l'} + \\
&\sum_{L \geq l' \geq l+1} \alpha \sigma^2 (L+1-l') \left( \sum_{l''-1 \geq l'' \geq l} (1-a^l) \cdots (1-a^{l''-1}) a^{l''} \right) (1-a^l) \cdots (1-a^{l'-1}) a^{l'} \\
&- \frac{1}{2} \alpha \sigma^2 \sum_{L \geq l' \geq l} \left( 1 - (1-a^l) \cdots (1-a^{l'}) \right)^2. \tag{1.B.11}
\end{aligned}$$

We also prove Proposition 1.B.2 by induction. The proposition is obviously true at date  $L+1$ . Assuming that the proposition is true at  $l+1$  we will show that it is true at  $l$ . To do this we will use the recursive equation 1.B.6 for  $Q_1^l$  as before, and equations 1.4.13 and 1.4.14.

Multiplying equation 1.4.13 by  $a^l$  and adding it to equation 1.4.14 we get

$$\frac{N-2}{N-1} \frac{a^l}{B^l} - \alpha \sigma^2 (L+1-l) a^l + \frac{A_2^l}{B^l} = (Q_2^{l+1}(1-a^l)\beta^{l,l}/a^l(N-1) + Q_3^{l+1})(1-a^l). \quad (1.B.12)$$

Using the above equation to eliminate the term in  $Q_2^{l+1}(1-a^l)\beta^{l,l}/a^l(N-1) + Q_3^{l+1}$  in the recursive equation for  $Q_1^l$ , 1.B.6, we obtain

$$Q_1^l = \frac{(a^l)^2}{B^l} + \frac{A_2^{l-1}}{B^{l-1}} a^l - \frac{1}{2} \alpha \sigma^2 (L+1-l) (a^l)^2 + Q_1^{l+1} (1-a^l)^2. \quad (1.B.13)$$

We now use the induction hypothesis and add and subtract the term

$$\alpha \sigma^2 (L-l) a^l [(1-a^l)a^{l+1} + \dots + (1-a^l) \dots (1-a^{L-1})a^L].$$

Equation 1.B.13 then becomes

$$\begin{aligned} Q_1^l &= \sum_{L \geq l' \geq l} \frac{(1-a^l)^2 \dots (1-a^{l'-1})^2 (a^{l'})^2}{B^{l'}} + \sum_{L \geq l' \geq l} \frac{A_2^{l'}}{B^{l'}} (1-a^l)^2 \dots (1-a^{l'-1})^2 a^{l'} \\ &+ \sum_{L \geq l'' \geq l+1} \alpha \sigma^2 (L+1-l'') \left( \sum_{l''-1 \geq l''' \geq l} (1-a^l) \dots (1-a^{l'''-1}) a^{l'''} \right) (1-a^l) \dots (1-a^{l''-1}) a^{l''} \\ &\quad - \alpha \sigma^2 (L-l) a^l [(1-a^l)a^{l+1} + \dots + (1-a^l) \dots (1-a^{L-1})a^L] \\ &- \frac{1}{2} \alpha \sigma^2 (1-a^l)^2 \sum_{L \geq l'' \geq l+1} \left( 1 - (1-a^{l+1}) \dots (1-a^{l''}) \right)^2 - \frac{1}{2} \alpha \sigma^2 (L+1-l) (a^l)^2. \quad (1.B.14) \end{aligned}$$

It is again very simple to transform the last two lines of the above equation so that it becomes exactly equation 1.B.11.  $\square$

We finally establish the following result

**Proposition 1.B.3**  $Q_7^l$  is given by

$$\begin{aligned}
Q_7^l &= \sum_{L \geq l' \geq l} \left( \frac{1}{B^{l'}} - \frac{1}{2} \alpha \sigma^2 (L+1-l') \right) \frac{\sigma_u^2}{N^2} + \\
&\left( \sum_{L \geq l' \geq l} \frac{(1-a^l)^2 \cdots (1-a^{l'-1})^2 (a^{l'})^2}{B^{l'}} - \frac{1}{2} \alpha \sigma^2 \sum_{L \geq l' \geq l} \left( 1 - (1-a^l) \cdots (1-a^{l'}) \right)^2 + \right. \\
&\sum_{L \geq l' \geq l+1} \alpha \sigma^2 (L+1-l') \left( \sum_{l'-1 \geq l'' \geq l} (1-a^l) \cdots (1-a^{l''-1}) a^{l''} \right) (1-a^l) \cdots (1-a^{l'-1}) a^{l'} \Big) V^l \\
&+ \sum_{L \geq l' \geq l+1} \frac{A_2^{l'}}{B^{l'}} \left( (1-a^l)^2 \cdots (1-a^{l'-1})^2 a^{l'} V^l - a^{l'} V^{l'} \right). \tag{1.B.15}
\end{aligned}$$

Proposition 1.B.3 will also be proven by induction. The proposition is obviously true at date  $L+1$ . Assuming that the proposition is true at  $l+1$  we will show that it is true at  $l$ . For this we will use the recursive equation 1.B.10 for  $Q_7^l$ , and equation 1.4.13.

We first eliminate the term in  $Q_2^{l+1}(1-a^l)\beta^{l,l}/a^l(N-1) + Q_3^{l+1}$  in equation 1.B.10 using equation 1.4.13. We get

$$Q_7^l = \left( \frac{1}{B^{l'}} - \frac{1}{2} \alpha \sigma^2 (L+1-l') \right) \left( (a^l)^2 V^l + \frac{\sigma_u^2}{N^2} \right) + Q_1^{l+1} (1-a^l)^2 \left( \frac{\beta^{l,l}}{a^l} \right)^2 + Q_7^{l+1}. \tag{1.B.16}$$

Using the induction hypothesis as well as the expression for  $Q_1^{l+1}$  derived in Proposition 1.B.2. we obtain

$$\begin{aligned}
Q_7^l &= \sum_{L \geq l' \geq l} \left( \frac{1}{B^{l'}} - \frac{1}{2} \alpha \sigma^2 (L+1-l') \right) \frac{\sigma_u^2}{N^2} + \left( \frac{(a^l)^2}{B^l} - \frac{1}{2} \alpha \sigma^2 (L+1-l) (a^l)^2 \right) V^l \\
&+ \left( (1-a^l)^2 \left( \frac{\beta^{l,l}}{a^l} \right)^2 \left( (a^l)^2 V^l + \frac{\sigma_u^2}{N^2} \right) + V^{l+1} \right) \\
&\left( \sum_{L \geq l' \geq l+1} \frac{(1-a^{l+1})^2 \cdot (1-a^{l'-1})^2 (a^{l'})^2}{B^{l'}} - \frac{\alpha \sigma^2}{2} \sum_{L \geq l' \geq l+1} \left( 1 - (1-a^{l+1}) \cdot (1-a^{l'}) \right)^2 + \right.
\end{aligned}$$



$$\begin{aligned}
& \sum_{L \geq l' \geq l+2} \alpha \sigma^2 (L+1-l') \left( \sum_{l''-1 \geq l''' \geq l+1} (1-a^{l'+1}) \cdot (1-a^{l'''-1}) a^{l'''} \right) (1-a^{l'+1}) \cdot (1-a^{l''-1}) a^{l''} \\
& \quad + \frac{A_s^{l'+1}}{B^{l'+1}} a^{l'+1} (1-a^l)^2 \left( \frac{\beta^{l,l}}{a^l} \right)^2 \left( (a^l)^2 V^l + \frac{\sigma_u^2}{N^2} \right) + \sum_{L \geq l' \geq l+2} \frac{A_s^{l'}}{B^{l'}} \\
& \left( (1-a^{l'+1})^2 \cdot (1-a^{l''-1})^2 a^{l''} \left( (1-a^l)^2 \left( \frac{\beta^{l,l}}{a^l} \right)^2 \left( (a^l)^2 V^l + \frac{\sigma_u^2}{N^2} \right) + V^{l+1} \right) - a^{l''} V^{l''} \right). \tag{1.B.17}
\end{aligned}$$

Equation 1.B.17 can be simplified noting that

$$\left( (1-a^l)^2 \left( \frac{\beta^{l,l}}{a^l} \right)^2 \left( (a^l)^2 V^l + \frac{\sigma_u^2}{N^2} \right) + V^{l+1} \right) = (1-a^l)^2 V^l. \tag{1.B.18}$$

We can derive this equation by using the expressions for the  $\beta$ 's and the  $V$ 's obtained in subsection 1.A.1 or, more directly, by writing that

$$\text{var}^l \left( \frac{\sum_{j \neq i} e_j^{l+1}}{N} \right) = \text{var}^l \left( E^{l+1} \left( \frac{\sum_{j \neq i} e_j^{l+1}}{N} \right) \right) + E^l \left( \text{var}^{l+1} \left( \frac{\sum_{j \neq i} e_j^{l+1}}{N} \right) \right).$$

Using equation 1.B.18 and proceeding as in the proof of proposition 1.B.2 we can easily transform equation 1.B.17 so that it becomes equation 1.B.15.  $\square$

Before deriving “ex-ante” welfare we write equation 1.B.15 for  $l = 1$ . Noting that  $V^1 = \Sigma_e^2(N-1)/N^2$  we get

$$\begin{aligned}
Q_7^1 &= \sum_{L \geq l' \geq 1} \left( \frac{1}{B^{l'}} - \frac{1}{2} \alpha \sigma^2 (L+1-l') \right) \frac{\sigma_u^2}{N^2} + \\
& \left( \sum_{L \geq l' \geq 1} \frac{(1-a^1)^2 \cdot \dots \cdot (1-a^{l'-1})^2 (a^{l'})^2}{B^{l'}} - \frac{1}{2} \alpha \sigma^2 \sum_{L \geq l' \geq 1} \left( 1 - (1-a^1) \cdot \dots \cdot (1-a^{l'}) \right)^2 + \right. \\
& \sum_{L \geq l' \geq 2} \alpha \sigma^2 (L+1-l') \left( \sum_{l''-1 \geq l''' \geq 1} (1-a^1) \cdot \dots \cdot (1-a^{l'''-1}) a^{l'''} \right) (1-a^1) \cdot \dots \cdot (1-a^{l''-1}) a^{l''} \\
& \left. + \sum_{L \geq l' \geq 2} \frac{A_s^{l'}}{B^{l'}} (\beta^{l''-1,1} + \dots + \beta^{l''-1,l''-1}) (1-a^l)^2 \cdot \dots \cdot (1-a^{l''-1})^2 a^{l''} \right) \frac{\Sigma_e^2(N-1)}{N^2}. \tag{1.B.19}
\end{aligned}$$

### Step 3 “Ex-ante” welfare

Assuming that terms in  $\alpha^2\sigma^2\Sigma_c^2$  and terms in  $\alpha^2\sigma^2\sigma_u^2$  are much smaller than 1, expression 1.B.5 simplifies to

$$-\frac{1}{2}\alpha\sigma^2(L+1)\Sigma_c^2 + Q_1^1\left(\frac{N-1}{N}\right)^2\Sigma_c^2 + Q_1^1. \quad (1.B.20)$$

Using equation 1.B.7 for  $l=1$  as well as equation 1.B.19 we can rewrite expression 1.B.20 as

$$\sum_{L \geq l' \geq 1} \left( \frac{1}{B^{l'}} - \frac{1}{2}\alpha\sigma^2(L+1-l') \right) \frac{\sigma_u^2}{N^2} - \frac{1}{2}\alpha\sigma^2\Sigma_c^2 \left( 1 + \sum_{L \geq l' \geq 1} \left[ \left( \frac{1}{N} + (1-a^1) \cdots (1-a^{l'}) \frac{N-1}{N} \right)^2 + \frac{N-1}{N^2} \left( 1 - (1-a^1) \cdots (1-a^{l'}) \right)^2 \right] \right). \quad (1.B.21)$$

## 1.B.2 Small Traders' Welfare and Total Welfare

The welfare of small traders is

$$E^1(p^1u^1 + \cdots + p^Lu^L). \quad (1.B.22)$$

Since  $u^l$  is independent of all information before date  $l$ , this expression simplifies to

$$-\sum_{L \geq l' \geq 1} \frac{1}{B^{l'}} \frac{\sigma_u^2}{N}. \quad (1.B.23)$$

Multiplying expression 1.B.21 by  $N$  (since there are  $N$  large traders) and adding expression 1.B.23 we get the RHS of equation 1.5.1. This equation shows that total welfare decreases if trade takes place more slowly. The speed of trade is the *only* factor that affects welfare; the riskiness of transfers among the various agents due to hedging demand uncertainty does not play a role. Our assumptions that price movements due to fundamentals are much more important than price movements due to the trades of the different market participants and that small traders care only

about the expectation of their trading proceeds are crucial to this result.

## 1.C Appendix: The Public Information Case

### 1.C.1 Proof of Proposition 1.7.1

In order to prove Proposition 1.7.1, we will first derive traders' value functions as well as necessary and sufficient conditions for an equilibrium. For generality we will consider the case where there is any number of trading dates (i.e. not just two). The analysis parallels that of subsection A.2 and most of it will be omitted. We will then show that in the two-period case there exists a unique equilibrium.

#### Derivation of Traders' Value Functions and Characterization of Equilibrium

It seems natural to look for an equilibrium where trader  $i$ 's date  $l$  demand has the form

$$x_i^l(p^l) = A^l - B^l p^l - a^l e_i^l - A_s^l \frac{\sum_{j=1}^N e_j^l}{N} - A_{\bar{e}}^l \bar{e}^l. \quad (1.C.1)$$

This equation is similar to equation 1.4.1 which gives trader  $i$ 's demand in the private information case. The difference is that  $s^l$  is replaced by its "true" value,  $\sum_{j=1}^N e_j^l$ .

Theorem 1.C.1 is the counterpart of Theorem 1.A.1.

**Theorem 1.C.1** *Suppose that we are at date  $l$  before trade takes place. Suppose that trader  $i$  has deviated from equilibrium at the previous dates so that his asset holdings are higher by  $\Delta e_i^l$  than what they would be if he had never deviated. (Each other large trader's asset holdings are thus lower by  $\Delta e_i^l / (N - 1)$  and they all believe that  $\bar{e}^l$  is lower by this quantity.) Then the expected utility of trader  $i$  (i.e. his value function) is*

$$-\exp\left(-\alpha \left( M^l + d^l (e_i^l + \bar{e}^l + \Delta e_i^l) - \frac{\alpha \sigma^2}{2} (L - l + 1) (e_i^l + \bar{e}^l + \Delta e_i^l)^2 \right)\right)$$

$$+Q_1^l \left( \frac{\sum_{j=1}^N e_j^l}{N} - e_i^l \right)^2 - Q_3^l \Delta e_i^l \left( \frac{\sum_{j=1}^N e_j^l}{N} - e_i^l \right) + Q_5^l (\Delta e_i^l)^2 + Q_7^l \Big). \quad (1.C.2)$$

$Q_k^l$  ( $k \in 1, 3, 5, 7$ ) are constants. The  $e_j^l$ 's have their equilibrium values.

Theorem 1.C.1 is again proven by induction. The proof is very similar to that of Theorem 1.A.1 and is thus omitted. We only give the necessary and sufficient conditions for an equilibrium.

The parameters  $A^l$ ,  $B^l$ ,  $a^l$ ,  $A_s^l$ , and  $A_e^l$  are given by

$$\frac{A^l}{B^l} = d^l \quad (1.C.3)$$

$$\frac{A_e^l}{B^l} = (L + 1 - l)\alpha\sigma^2 \quad (1.C.4)$$

$$\frac{N-2}{N-1} \frac{1}{B^l} = (L + 1 - l)\alpha\sigma^2 \quad (1.C.5)$$

$$\frac{A_s^l}{B^l} = (l - a^l)Q_3 \quad (1.C.6)$$

and

$$\left( a^l \frac{N-1}{N-2} - 1 \right) (L + 1 - l)\alpha\sigma^2 = -(1 - a^l)Q_3. \quad (1.C.7)$$

The second-order condition is

$$S \equiv \frac{2}{(N-1)B^l} + \alpha\sigma^2(L + 1 - l) - 2Q_5^{l+1} > 0. \quad (1.C.8)$$

Finally, the  $Q_k$ 's are given by the following equations

$$Q_1^l = \left( \frac{1}{(N-1)B^l} + \frac{1}{2}\alpha\sigma^2(L + 1 - l) \right) (a^l)^2 + Q_1^{l+1}(1 - a^l)^2 \\ + Q_3^{l+1}(1 - a^l)a^l - \frac{\alpha\mathcal{N}_1^2 \frac{\sigma^2}{N^{\frac{1}{2}}}}{2(1 + 2\alpha\mathcal{D} \frac{\sigma^2}{N^{\frac{1}{2}}})} \quad (1.C.9)$$

$$Q_3^l = \left( \frac{A_e^l}{B^l} \frac{1}{N-1} + \alpha\sigma^2(L + 1 - l) \right) a^l + Q_3^{l+1}(1 - a^l) - \frac{\alpha\mathcal{N}_1\mathcal{N}_3 \frac{\sigma^2}{N^{\frac{1}{2}}}}{(1 + 2\alpha\mathcal{D} \frac{\sigma^2}{N^{\frac{1}{2}}})} \quad (1.C.10)$$

$$Q_5^l = Q_5^{l+1} + \frac{\left(\frac{A_5^l}{B^l} \frac{1}{N-1} + \alpha\sigma^2(L+1-l) - 2Q_5^{l+1}\right)^2}{4\left(\frac{2}{(N-1)B^l} + \alpha\sigma^2(L+1-l) - 2Q_5^{l+1}\right)} - \frac{\alpha\mathcal{N}_3^2 \frac{\sigma_u^2}{N^{\frac{1}{2}}}}{(1 + 2\alpha\mathcal{D} \frac{\sigma_u^2}{N^{\frac{1}{2}}})} \quad (1.C.11)$$

and

$$\exp(-\alpha Q_7^l) = \exp(-\alpha Q_7^{l+1}) \frac{1}{\sqrt{(1 + 2\alpha\mathcal{D} \frac{\sigma_u^2}{N^{\frac{1}{2}}})}}. \quad (1.C.12)$$

$\mathcal{N}_1, \mathcal{N}_3$  and  $\mathcal{D}$  are given by

$$\mathcal{N}_1 = 2 \left( \frac{1}{(N-1)B^l} + \frac{1}{2}\alpha\sigma^2(L+1-l) \right) a^l + Q_3^{l+1}(1-a^l) \quad (1.C.13)$$

$$\mathcal{N}_3 = \frac{A_3^l}{B^l} \frac{1}{N-1} + \alpha\sigma^2(L+1-l) \quad (1.C.14)$$

and

$$\mathcal{D} = \frac{1}{(N-1)B^l} + \frac{1}{2}\alpha\sigma^2(L+1-l). \quad (1.C.15)$$

## The Two-Period Case

Equations 1.C.3 to 1.C.7 and 1.C.8 imply that the parameters characterizing equilibrium at date 2 are exactly the same as in the private information case and that the second-order condition holds. Simple algebra shows that

$$Q_1^2 = \frac{Q_3^2}{2} = Q_5^2 = \alpha\sigma^2 \frac{N(N-2)}{2(N-1)^2} \frac{1}{1 + \alpha^2\sigma^2 \frac{\sigma_u^2}{N(N-2)}}. \quad (1.C.16)$$

Going back to date 1, equation 1.7.8 can be written as

$$\frac{1 - \frac{N-1}{N-2}a^1}{1 - a^1} = \frac{N(N-2)}{2(N-1)^2} \frac{1}{1 + \alpha^2\sigma^2 \frac{\sigma_u^2}{N(N-2)}}. \quad (1.C.17)$$

A similar argument to the one given in subsection A.5 shows that equation 1.C.17 has a unique real solution,  $a^1 \in (0, (N-2)/(N-1))$ , that  $A^1, B^1$  and  $A_5^1$  are uniquely defined and that the second-order condition at date 1 also holds.

We finally compare  $a^1$  with its counterpart in the private information case. Suppose first that  $\sigma_u^2$  is high. Since  $a^1 \in (0, (N-2)/(N-1))$  and is thus bounded, the RHS of equation 1.A.46 is smaller than the RHS of equation 1.C.17. Given that the LHS of both equations is the same and decreasing in  $a^1$ ,  $a^1$  is higher in the private information case.

Suppose now that  $\sigma_u^2$  is small. We first “extend” equation 1.A.46 to the case  $\sigma_u^2 = 0$  in the natural way

$$\frac{1 - \frac{N-1}{N-2}a^1}{1 - a^1} = \frac{N}{2(N-1)}. \quad (1.C.18)$$

This equation admits a unique solution  $a_0^1 > 0$ . The implicit function theorem guarantees that this solution can be continuously extended for  $\sigma_u^2 > 0$ .<sup>39</sup> Of course, by uniqueness, this extension coincides with the solution  $a^1$  that we already have found. Since  $N/2(N-1) > N(N-2)/2(N-1)^2$ ,  $a_0^1$  and the solution of equation 1.A.46 for  $\sigma_u^2$  small, is smaller than the solution of equation 1.C.17 for  $\sigma_u^2$  small.

## 1.C.2 Welfare

In this subsection we will determine large traders’ welfare and provide an example where where their welfare in both the “interim” and the “ex-ante” sense is lower in the private information case than in the public information case. We, again, consider the multiperiod case.

Using Theorem 1.C.1 and proceeding as in subsection 1.B.1 we find that a large trader’s expected utility at time 0 is

$$- \exp\left(-\alpha\left(-\frac{1}{2}\alpha\sigma^2(L+1)(e_i^1)^2 + Q_1^1 \left(\frac{\sum_{j=1}^N e_j^1}{N} - e_i^1\right)^2 + Q_7^1\right)\right). \quad (1.C.19)$$

To derive expected utility in the “interim” sense, we take expectations w.r.t

---

<sup>39</sup>The difference between the RHS and the LHS of equation 1.A.46 viewed as a function of  $a^1$  and  $\sigma_u^2$  is continuously differentiable in a neighborhood of  $(a_0^1, 0)$  and has a non-zero partial derivative w.r.t.  $a^1$ .

$\sum_{j \neq i} e_j^1/N$ . Using equation 1.A.22 we get

$$- \exp\left(-\alpha\left(-\frac{1}{2}\alpha\sigma^2(L+1)(e_i^1)^2 + \frac{Q_1^1 \left(\frac{N-1}{N}\right)^2 (e_i^1)^2}{1 + 2\alpha Q_1^1 \frac{N-1}{N^2} \Sigma_e^2} + Q_7^1\right)\right) \frac{1}{\sqrt{1 + 2\alpha Q_1^1 \frac{N-1}{N^2} \Sigma_e^2}}. \quad (1.C.20)$$

To derive expected utility in the ex-ante sense, we take expectations w.r.t.  $e_i^1$ . Using equation 1.A.22 we get

$$- \exp(-\alpha Q_7^1) \frac{1}{\sqrt{1 + 2\alpha Q_1^1 \frac{N-1}{N^2} \Sigma_e^2}} \frac{1}{\sqrt{\left(1 - \alpha\left(\alpha\sigma^2(L+1) - 2\frac{Q_1^1 \left(\frac{N-1}{N}\right)^2}{1 + 2\alpha Q_1^1 \frac{N-1}{N^2} \Sigma_e^2}\right)\Sigma_e^2\right)}}. \quad (1.C.21)$$

A large trader's welfare in the "interim" sense is thus

$$-\frac{1}{2}\alpha\sigma^2(L+1)(e_i^1)^2 + \frac{Q_1^1 \left(\frac{N-1}{N}\right)^2 (e_i^1)^2}{1 + 2\alpha Q_1^1 \frac{N-1}{N^2} \Sigma_e^2} + \frac{1}{2\alpha} \log\left(1 + 2\alpha Q_1^1 \frac{N-1}{N^2} \Sigma_e^2\right) + Q_7^1 \quad (1.C.22)$$

and in the "ex-ante" sense

$$\frac{1}{2\alpha} \log\left(1 + 2\alpha Q_1^1 \frac{N-1}{N^2} \Sigma_e^2\right) + \frac{1}{2\alpha} \log\left(1 - \alpha\left(\alpha\sigma^2(L+1) - 2\frac{Q_1^1 \left(\frac{N-1}{N}\right)^2}{1 + 2\alpha Q_1^1 \frac{N-1}{N^2} \Sigma_e^2}\right)\Sigma_e^2\right) + Q_7^1. \quad (1.C.23)$$

Having derived expressions 1.C.22 and 1.C.23 we now give an example where the welfare of a large trader in both the "interim" and the "ex-ante" sense is lower in the private information case than in the public information case. We assume that  $N = 5$ ,  $L = 2$ ,  $\Sigma_e^2/N\Sigma_u^2 = 0.01$  and that terms in  $\alpha^2\sigma^2\Sigma_e^2$  and  $\alpha^2\sigma^2\sigma_u^2$  are much smaller than 1. We then find that in the private information case "interim" welfare is given by

$$\alpha\Sigma^2(-0.311(e_i^1)^2 + 0.0478\Sigma_e^2)$$

while in the public information case it is given by

$$\alpha \Sigma^2 (-0.304(e_i^1)^2 + 0.0547 \Sigma_e^2).$$

Ex-ante welfare is also clearly lower in the private information case.

The derivation of equation 1.5.1 in the public information case is similar in spirit to the one in the private information case and is thus omitted.



## References

- Admati, Anat R. and Paul Pfleiderer (1988), "*A Theory of Intraday Patterns: Volume and Price Variability*", *The Review of Financial Studies*, 1:3-40.
- Admati, Anat R. and Paul Pfleiderer (1991), "*Sunshine Trading and Financial Market Equilibrium*", *The Review of Financial Studies*, 4:443-481.
- Chan, Louis K. C. and Josef Lakonishok (1992), "*Institutional Trades and Intra-Day Stock Price Behavior*", *Journal of Financial Economics*, forthcoming.
- Chan, Louis K. C. and Josef Lakonishok (1993), "*The Behavior of Stock Prices Around Institutional Trades*", mimeo, University of Illinois at Urbana-Champaign.
- Cohen, Kalman J. et al (1986), "*The Microstructure of Securities Markets*", Prentice Hall, Englewood Cliffs, New Jersey.
- Foster, Douglas F. and S. Viswanathan (1992), "*The Effect of Public Information and Competition on Trading Volume and Price Volatility*", *Review of Financial Studies*, forthcoming.
- Gresik, Thomas A. and Mark A. Satterthwaite (1989), "*The Rate at which a Simple Market Converges to Efficiency as the Number of Traders Increases. An Asymptotic Result for Optimal Trading Mechanisms*", *Journal of Economic Theory*, 48:304-332.
- Grossman, Sanford J. and Merton H. Miller (1988), "*Liquidity and Market Structure*", *The Journal of Finance*, 43:617-637.
- Holden, Craig W. and Avanidhar Subrahmanyam (1992), "*Long-Lived Private Information and Imperfect Competition*", *Journal of Finance*, 47:247-270.
- Holthausen, Robert, Richard Leftwich and David Mayers (1987), "*The Effect of Large Block Transactions on Security Prices: A Cross-Sectional Analysis*", *Journal of Financial Economics*, 19:237-268.

- Kraus, Alan and Hans Stoll (1972), "*Large-Block Transactions, the Speed of Response, and Temporary and Permanent Stock-Price Effects*", *Journal of Financial Economics*, 26:71-95.
- Kyle, Albert S. (1985), "*Continuous Auctions and Insider Trading*", *Econometrica*, 53:1315-1336.
- Kyle, Albert S. (1989), "*Informed Speculation and Imperfect Competition*", *Review of Economic Studies*, 56:517-556.
- Madhavan, Ananth (1990), "*Security Prices and Market Transparency*", mimeo, Wharton School.
- McAfee, Preston R. and John McMillan (1987), "*Auctions and Bidding*", *Journal of Economic Literature*, 25:699-738.
- Pagano, Marco and Ailsa Roëll (1990), "*Stock Markets*", *Economic Policy*, April 1990:63-115.
- Satterthwaite, Mark A. and Steven R. Williams (1989), "*The Rate of Convergence to Efficiency in the Buyer's Bid Double Auction as the Market Becomes Large*", *Review of Economic Studies*, 56:477-498.
- Vives, Xavier (1992), "*The Speed of Information Revelation in a Financial Market Mechanism*", ESF, Financial Market Network, W.P. No16.
- Wang, Jiang (1989), "*Asset Prices, Stock Returns, Price Volatility, Risk Premium and Trading Strategies Under Asymmetric Information*", forthcoming, *Review of Economic Studies*.
- Wang, Jiang (1991), "*A Model of Competitive Stock Trading Volume*", mimeo, Sloan School of Management.
- Wilson, Robert (1985), "*Incentive Efficiency of Double Auctions*", *Econometrica*, 59:1101-1115.

Wilson, Robert (1986), "*Equilibria in Bid-Ask Markets*", in G.Feiwel(ed.) *Arrow and the Ascent of Economic Theory: Essays in the Honor of Kenneth J. Arrow*, London: The Macmillan Press Ltd.

Wilson, Robert (1987), "*Game-Theoretic Analysis of Trading Processes*", in T.Bewley (ed.) *Advances in Economic Theory: Fifth World Congress*, Cambridge: Cambridge University Press.

Wilson, Robert (1990), "*Two Surveys: Auctions and Entry Deterrence*", Technical Report No. 3, Stanford University.

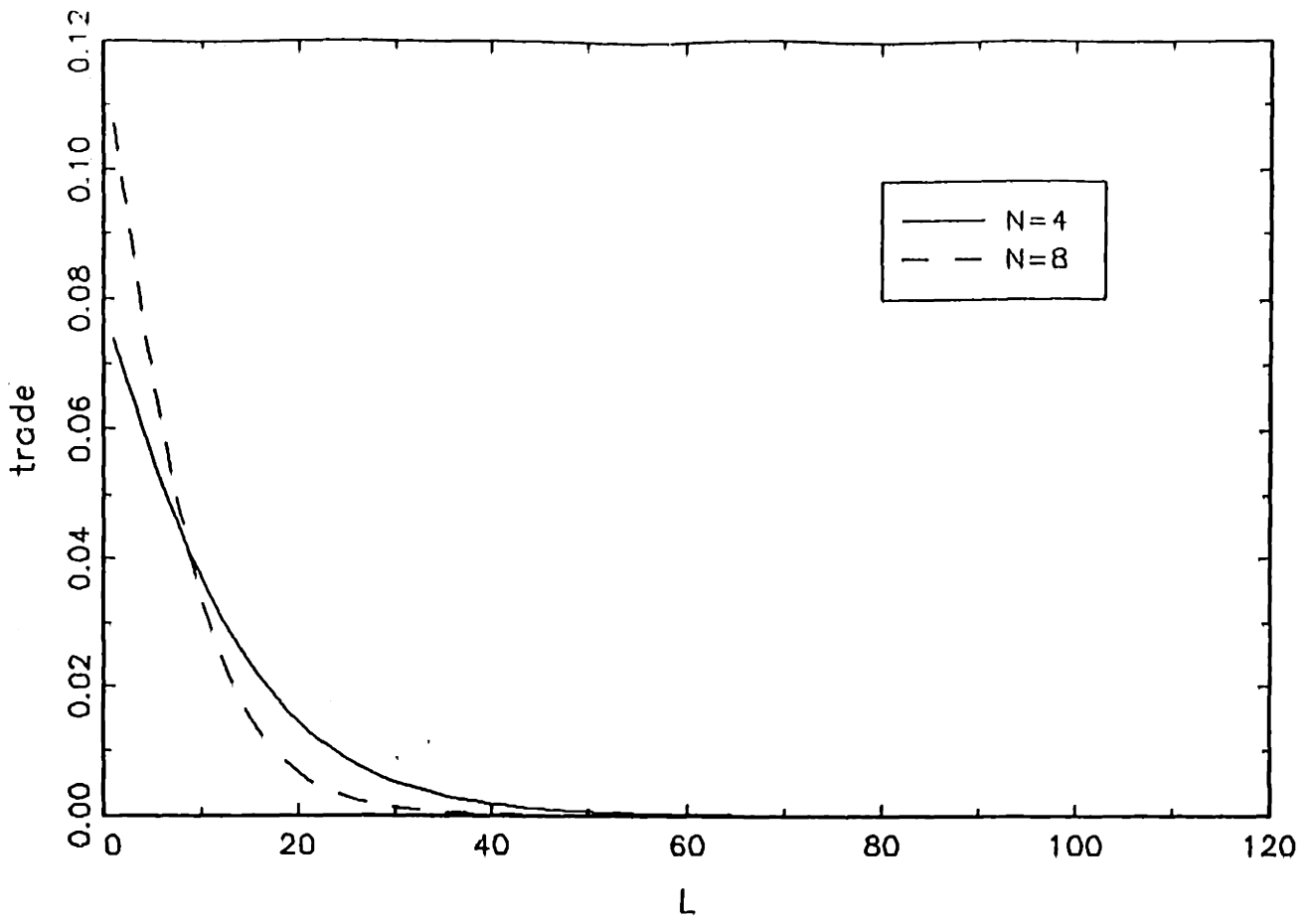


Figure 1-1: Trading patterns for noise = 30

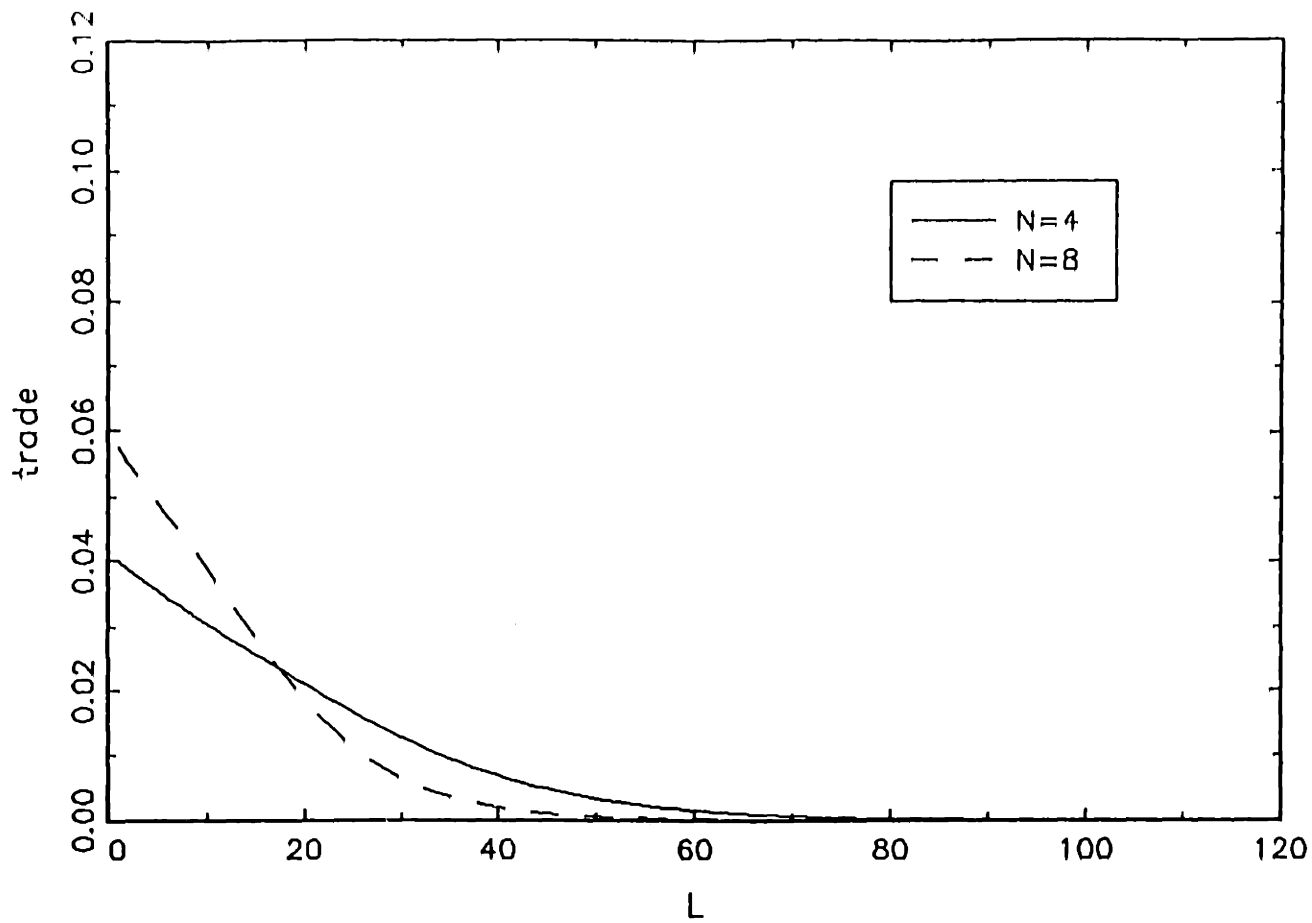


Figure 1-2: Trading patterns for noise = 4

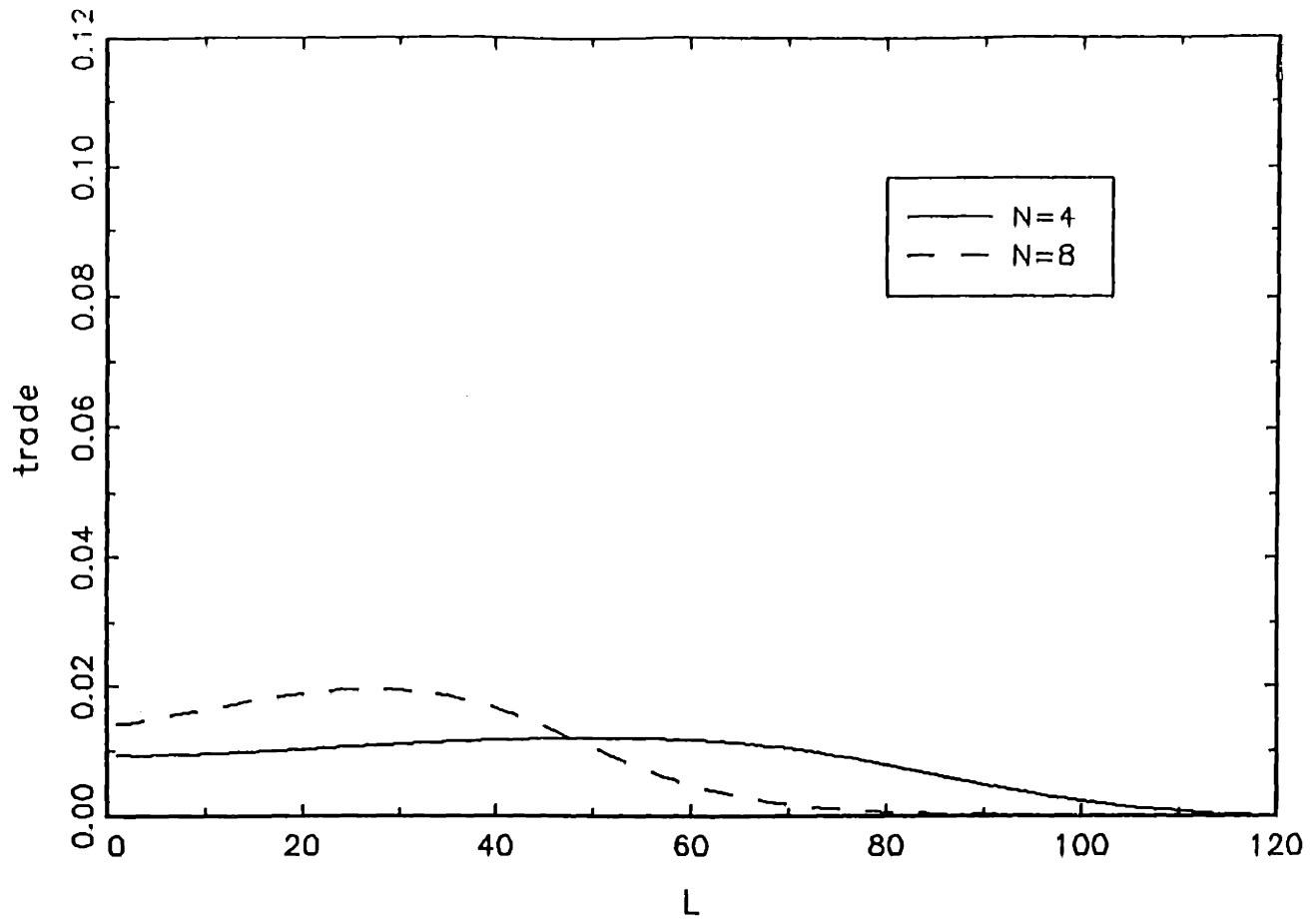


Figure 1-3: Trading patterns for noise = 0.1

## Chapter 2

# Equilibrium Interest Rate and Liquidity Premium Under Proportional Transactions Costs

### 2.1 Introduction

Although most of asset pricing theory assumes frictionless markets, transactions costs are ubiquitous in financial markets. Transactions costs can be decomposed into (i) direct transactions costs such as brokerage commissions, exchange fees and transactions taxes, (ii) bid-ask spread, (iii) market impact costs and (iv) delay and search costs.<sup>1</sup> Aiyagari and Gertler (1991), report that typical (retail) brokerage costs for common stocks average 2% of the dollar amount of the trade while the bid-ask spread for actively traded stocks averages around .5%. Moreover, transactions costs vary across assets and over time. Money market accounts are clearly more liquid than stocks. In addition, deregulation as well as changes in information technology have reduced (but not eliminated) transactions costs.

Empirical work on transactions costs documents not only their magnitude but their important effect on rates of return. Amihud and Mendelson (1986) show that

---

<sup>1</sup>See Amihud and Mendelson (1991a).

the risk-adjusted average return on stocks is positively related to their bid-ask spread. Even more direct evidence can be found by comparing two assets with exactly the same cash flows but different liquidity: (i) restricted ("letter") stocks which cannot be publicly traded for 2 years sell at a 35% discount below regular stocks<sup>2</sup> and (ii) the average yield differential between Treasury Notes close to maturity and more liquid Treasury Bills is about .43%.<sup>3 4</sup>

The evidence above shows that liquidity is an important determinant of assets' returns and should be incorporated into asset pricing theory. Understanding the impact of transactions costs on assets' returns will shed some light on some policy issues as well. Transactions taxes and differential taxation of long and short-term capital gains both reduce liquidity and therefore affect assets' returns. As a result, investment decisions will change with additional welfare implications.

In this chapter we analyze the impact of transactions costs on the rates of return on liquid and illiquid assets, in a general equilibrium framework. We are interested in questions such as: On what characteristics of the economy does the liquidity premium (the difference between the rates of return on illiquid and liquid assets) depend? How do transactions costs affect the rates of return on liquid and illiquid assets? Are these effects first or second order effects?

Despite their importance for asset pricing, these questions have so far not been satisfactorily addressed in the theoretical literature. A major reason is that to answer them, one has to move away from the basic model of asset pricing, namely the representative agent model. (One cannot understand the impact of trade frictions in a model where there is no trade.) Unfortunately, models with heterogeneous agents (and trade), tend to be quite intractable analytically.

Since our objective is to understand the effects of asset liquidity on asset pricing, we take risk out of the picture: All the assets that we consider (liquid or illiquid) pay a constant stream of dividends. The analysis of the joint effects of risk and liquidity is an interesting question that we leave for future research.

---

<sup>2</sup>See Silber (1991).

<sup>3</sup>See Amihud and Mendelson (1991a).

<sup>4</sup>More evidence is also presented in Boudoukh and Whitelaw (1991).



There are many ways to construct a deterministic economy with heterogeneous agents. Agents may trade because of differences in preferences or endowments, that is, they may have different preferences for current versus future consumption, or they may have different labor income paths.<sup>5</sup> In our economy both motives exist. More precisely, our economy is a tractable version of a multiperiod overlapping generations economy, the perpetual youth economy, first studied by Blanchard (1985). We believe, however, that our results on the effects of transactions costs on asset returns could appear in other contexts as well.

In our model, agents face a constant probability of death (this is the key assumption that makes things tractable), and the population is kept constant by an inflow of new arrivals. Agents start with no financial wealth and receive a decreasing stream of labor income over their lifetimes. In addition they can invest in long-term assets which pay a constant stream of dividends. There are two such assets, the liquid asset and the illiquid asset. The liquid asset is traded without transactions costs, while trading the illiquid asset entails proportional transactions costs. Neither asset can be sold short. In this economy agents buy and sell assets for life-cycle motives. In fact, they accumulate the higher yielding illiquid asset for long-term investment purposes and the liquid asset for short-term investment needs.

We find that when transactions costs increase, the rate of return on the liquid asset decreases while the rate of return on the illiquid asset may increase or decrease. We also find, quite naturally, that the liquidity premium increases. The effects of transactions costs on both the rate of return on the liquid asset and the liquidity premium, are stronger the higher the fraction of the illiquid asset in the economy. Finally, transactions costs have first order effects on asset returns and on the liquidity premium.

The reason why the rate of return on the liquid asset falls in response to an increase in transactions costs, can be briefly summarized as follows: Suppose that transactions costs increase from 0 to  $\epsilon$  and that the rate of return on the liquid asset

---

<sup>5</sup>In a stochastic economy, differential information together with liquidity shocks may also generate trade (see Wang (1992)).

stays the same in equilibrium. Then, in equilibrium, the rate of return on the illiquid asset must increase (by the liquidity premium) which implies that this asset becomes cheaper. Agents now consume more since they face better investment opportunities (they have the liquid asset at the same rate as before, and an additional investment opportunity). Moreover, they substitute consumption over time so that they buy more of the cheaper illiquid asset and hold it for a longer period. Thus, they will demand more securities for two reasons. The first reason is that they have to finance higher future consumption, selling the cheaper illiquid asset and paying transactions costs. The second reason is that, by substitution, they want to buy more of the cheaper illiquid asset and hold it for a longer period. As a result, total asset demand goes up. The rate of return on the liquid asset has to fall to restore equilibrium. In addition, if there are more illiquid assets in the economy, total asset demand will increase more, and the rate of return on the liquid asset will have to fall by more.

We cannot infer whether the rate of return on the illiquid asset will increase or decrease, by a similar reasoning. Indeed, suppose that transactions costs increase from 0 to  $\epsilon$  and that the rate of return on the illiquid asset is unaffected in equilibrium. The rate of return on the liquid asset then has to fall by an amount equal to the liquidity premium. This time agents face worse investment opportunities since (i) the price of the liquid asset increases and (ii) trading in the illiquid asset is subject to transactions costs. Agents' consumption shifts down uniformly. Furthermore, by substitution they accumulate less of the illiquid asset, but hold it for a longer period. They also accumulate less of the liquid asset. The effect on the total demand<sup>6</sup> for securities is ambiguous. First, agents have to finance lower future consumption selling the more expensive liquid asset but they pay transactions costs when selling the illiquid asset. Second, agents buy less of the liquid and illiquid assets, but hold the illiquid asset for a longer period.

Finally, the liquidity premium depends on the minimum holding period of the illiquid asset. It increases with the fraction of illiquid assets in the economy, since this period gets shorter.

---

<sup>6</sup>In number of shares.

There is a growing literature studying asset market frictions such as transactions costs, short sale constraints or borrowing constraints. This literature addresses three basic questions. The first question is to find the optimal consumption/investment policy given price processes and imperfections. The objectives of the body of literature addressing this question are: (i) to derive the asset demand for a particular price process and (ii) to evaluate the cost that market imperfections impose upon market participants (given the price process). The answer to (ii) sheds some light upon the “equilibrium implications” of market frictions.<sup>7</sup> The equilibrium determination of prices in markets with frictions, *taking the financial structure* (and the imperfections, in particular) *as given*, is the second question raised in the literature on market frictions. It is also the question addressed here. Finally, the third question addressed by the literature on market frictions is to endogenize the financial structure<sup>8</sup>. While we consider this question to be a fundamental one we do not address it here i.e. we take the financial structure as given.

Most of the work on the equilibrium implications of market frictions, considers either a static framework along the lines of the Capital Asset Pricing Model (see among others, Brennan (1975), Goldsmith (1976), Levy (1978) and Mayshar (1979) and (1981) for a partial equilibrium analysis, and Fremault (1991) and Michaely and Vila (1992) for a general equilibrium treatment) or an overlapping generations economy where agents live for only two periods (Pagano (1989)). Although these models give us useful insights, they are not adequate for answering several of the questions we are interested in. In static models, assets are not sold but only liquidated. Moreover, in a static model (as well as in a two period overlapping generations model), agents cannot choose when to buy or sell assets, and the holding period is the same for all assets. It is thus clear that many of our results would not appear in that simplified framework.

---

<sup>7</sup>See for instance, Constantinides (1986), Davis and Norman (1990), Duffie and Sun (1989), Dumas and Luciano (1991), Fleming et al. (1992), Grossman and Vila (1992), Tuckman and Vila (1992), Vila and Zariphopoulou (1990), among many others.

<sup>8</sup>See for instance Allen and Gale (1988), Boudoukh and Whitelaw (1992), Duffie and Jackson (1989) and Ohashi (1992).

In a context directly related to our work, Amihud and Mendelson (1986) consider a dynamic model where investors have different horizons. They argue that investors with, say, an investment horizon of 4 years who face a 2% roundtrip transactions cost when buying and selling assets, will lose approximately 2/4% (.5%) per year because of the transactions cost. Hence, they will require a rate of return of .5% higher on illiquid assets than on liquid assets. Consequently, the liquidity premium on assets which appeal to investors with a 4 year horizon must be approximately .5%. The above reasoning implies that investors with longer horizons are less affected by transactions costs and would select higher yielding illiquid assets. By contrast, investors with shorter horizons select low yield liquid assets. This clientele effect explains the empirical fact that the cross-sectional relation between transactions costs and asset returns is concave.<sup>9</sup> The analysis above, while insightful, takes investors' horizons as given and does not explain how they change in response to an increase in transactions costs. Moreover since, as in the previous papers, the rate of return on the liquid asset is assumed for simplicity to be fixed, only the effect of transactions costs on the differentials of rates of return and not on their levels can be examined.

Two recent papers, one by Aiyagari and Gertler (1991), and one by Heaton and Lucas (1992) consider dynamic models where investors' horizons are endogenous. In their models, agents are infinitely lived, face labor income uncertainty, and trade assets for consumption-smoothing purposes. These papers seek to solve the equity premium puzzle (see Mehra and Prescott (1985)), i.e. to explain the differential rates of return between the stock and the bond market. Aiyagari and Gertler (1991) argue that differential transactions costs between these two markets account for part of the equity premium. In their model (as in ours), the 'stock' is riskless and therefore the equity premium is due to transactions costs and not to risk. Hence their model explains the fraction of the equity premium which is in fact a liquidity premium. They do not however analyze the effect of transactions costs on the level of rates of return as they take the rate of return on the liquid asset as given. By contrast with Aiyagari and Gertler (1991), Heaton and Lucas (1992) allow for a truly risky asset

---

<sup>9</sup>By contrast, if all investors had the same horizon this relation would be linear.

as well as for aggregate labor income uncertainty. They argue that transactions costs prevent investors from reducing the variability of their consumption by intertemporal smoothing thereby raising the equity premium. In addition, they endogenize the rate of return on the liquid asset and find that it falls in response to increased transactions costs.

While in our model agents save for life-cycle purposes rather than because of labor income uncertainty, our results are consistent with the numerical simulation results of the above two papers. We find in particular that transactions costs create a liquidity premium, as in Aiyagari and Gertler (1991), and that they cause the rate of return on the liquid asset to fall, as in Heaton and Lucas (1992). The contribution of our work is twofold: First, our closed form analysis allows us to precisely identify the different effects of transactions costs on asset demands and on rates of return. Second, we are able to easily perform and interpret various comparative statics.

The remainder of the chapter is structured as follows: In section 2.2, we describe the model. We determine asset returns when there are no transactions costs in section 2.3. In section 2.4, we consider the case where there are transactions costs. In section 2.5, we illustrate our general results with some numerical examples. Section 2.6 contains concluding remarks and all proofs appear in the appendix.

## 2.2 The Model

To analyze the impact of transactions costs on the return on assets and on the liquidity premium, we have adapted Blanchard's (1985) model of perpetual youth. A simplified exposition of the original model can be found in Blanchard and Fisher (1989).

We consider a continuous time overlapping generations economy with a continuum of agents with total mass equal to 1. An agent in this economy faces a constant probability of death per unit time,  $\lambda$ . In addition, we assume that death is independent across agents and that agents are born at a rate equal to  $\lambda$ . Therefore the population is stationary, with total mass equal to one and the distribution of age,  $t$ , has a density function equal to  $\lambda \exp(-\lambda t)$ . Although agents can live arbitrary long

lives in this economy, their life expectancy is bounded and equal to  $1/\lambda$ .

Agents are born with zero financial wealth and receive an exogenous labor income  $y_t$  over their lifetimes. We assume that  $y_t$  declines exponentially with age  $t$

$$y_t = \bar{y}e^{-\delta t}; \quad \delta \geq 0 \quad (2.2.1)$$

The aggregate labor income  $Y$  is constant and equal to

$$Y = \int_0^{\infty} \lambda e^{-\lambda t} y_t dt = \frac{\lambda}{\lambda + \delta} \bar{y} \quad (2.2.2)$$

The financial structure in this economy is given as follows. All assets in this economy are real perpetuities which pay a constant flow of dividends  $D$  per unit time. The total supply of perpetuities is normalized to one so that  $D$  is also the aggregate dividend. There are two such perpetuities. The *liquid* asset, in total supply  $1 - k$  ( $0 \leq k \leq 1$ ), can be exchanged without transactions costs. The price of the liquid asset is denoted by  $p$  and the rate of return on liquid assets is denoted by  $r = D/p$ . The *illiquid* asset, in total supply of  $k$ , has a price equal to  $P$  and a rate of return equal to  $R = D/P$ <sup>10</sup>. Trading in the illiquid asset is subject to proportional transactions costs: when buying (or selling)  $x$  shares of the illiquid asset the agent must pay  $\epsilon x P$  transactions costs. Because of transactions costs, the rate of return on the illiquid asset and on the liquid asset will be different. The *liquidity premium*  $\mu$  is defined as

$$\mu = R - r. \quad (2.2.3)$$

Finally, none of these assets can be sold short.<sup>11</sup>

Over the course of their lives, agents accumulate both assets. Since death is stochastic it imposes a financial risk upon the agents namely that of losing their

---

<sup>10</sup>Note that we have defined  $R$  as the rate of return *before transactions costs*. The rate of return *net of transactions costs* depends upon the holding period and is therefore investor specific.

<sup>11</sup>If short sales were costless agents would not sell the illiquid asset but would short the liquid asset instead. Our results do not change if we assume that the cost of short selling is higher than  $\epsilon$ , which is a reasonable assumption (see for instance Boudoukh and Whitelaw (1991) and Tuckman and Vila (1992) for evidence on short sale costs).

accumulated holdings.<sup>12</sup> We assume that this risk is fully and costlessly insurable in the following way: there exist insurance companies which pay shares of assets to the living participants in exchange for a claim on their estate. For example an insurance company that insures one share of, say, the liquid asset will pay a premium of  $\pi dt$  additional shares of the liquid asset per unit time  $dt$  to a living participant. Its compensation is to collect the share in the event of death. We assume that (i) the insurance market is perfectly competitive, (ii) insurance companies transfer assets costlessly<sup>13</sup> and (iii) death is an idiosyncratic risk. As a result, the premium  $\pi dt$  must be equal to the probability of death  $\lambda dt$ , for both the liquid and illiquid asset. Finally since, as previously indicated, agents do not derive any utility from their estate they will purchase full insurance.

We assume that agents maximize at time 0 the expected value of a time separable utility function of their consumption i.e.

$$E \left[ \int_0^{\infty} u(c_t) e^{-\beta t} dt \right]. \quad (2.2.4)$$

Since the only uncertainty comes from the possibility of death we can write equation 2.2.4 as<sup>14</sup>

$$\int_0^{\infty} u(c_t) e^{-(\beta+\lambda)t} dt. \quad (2.2.5)$$

We also assume that the utility function exhibits a constant elasticity of substitution equal to  $1/A$  i.e.

$$u(c) = \frac{1}{1-A} c^{1-A}.^{15} \quad (2.2.6)$$

We focus on the stationary equilibria of this economy. In a stationary equilibrium, the rates of return  $r$  and  $R$  are constant. We seek to understand the determination

---

<sup>12</sup>Agents do not leave any heir behind and care only about themselves.

<sup>13</sup>The introduction of insurance companies is a convenient way to close the model. Our discussion in the introduction suggests that our results would carry through in a multi-period overlapping generations model with deterministic death. The latter is much more difficult to solve analytically.

<sup>14</sup>See Blanchard and Fisher (1989) for details.

<sup>15</sup>The case  $A=1$  corresponds to  $u(c) = \log c$ .

of  $r$ ,  $R$  and  $\mu$  as functions of the parameters of the model:  $\epsilon$ ,  $k$ ,  $\lambda$ ,  $\delta$ ,  $\beta$ ,  $A$ ,  $Y$  and  $D$ .

## 2.3 The No Transactions Costs Case

In this section, we analyze the determination of the interest rate in the benchmark case when transactions costs are equal to zero. In this case there is no difference between the liquid asset and the illiquid asset:  $r = R$  and  $\mu = 0$ .

### 2.3.1 The consumer's problem

The *financial wealth*  $w_t$  of the consumer at date  $t$  is defined as the value of the consumer's assets. That is if  $x_t$  is the number of shares that the consumer owns at date  $t$

$$w_t = px_t. \quad (2.3.1)$$

At date  $t$ , the consumer receives a labor income  $y_t$  per unit time. His financial income (per unit time) entails  $Dx_t$  in dividend income plus  $\lambda x_t$  shares worth  $\lambda px_t$ . Since he consumes  $c_t$  per unit time the dynamics of his wealth are

$$dw_t = Dx_t dt + \lambda px_t dt + (y_t - c_t) dt = (r + \lambda)w_t dt + (y_t - c_t) dt$$

$$c_t \geq 0; \quad w_0 = 0; \quad w_t \geq 0. \quad (2.3.2)$$

From equations 2.2.5 and 2.3.2, the consumer's problem is the optimization problem of an infinitely lived consumer with discount factor  $\beta + \lambda$  who faces a constant interest rate. This constant interest rate equals the rate of return on the perpetuity,  $r$ , plus the premium paid by the insurance company,  $\lambda$ . Hence the consumer's problem can be written as

$$\max \int_0^{\infty} u(c_t) e^{-(\beta+\lambda)t} dt = \max \int_0^{\infty} \frac{1}{1-A} c^{1-A} e^{-(\beta+\lambda)t} dt$$



$$s.t. \int_0^{\infty} c_t e^{-(r+\lambda)t} dt = \max \int_0^{\infty} y_t e^{-(r+\lambda)t} dt; \quad w_t \geq 0. \quad (2.3.3)$$

The problem 2.3.3 above admits the following solution<sup>16</sup>

$$c_t = \bar{y} \frac{\psi}{\phi} e^{-\omega t} \quad (2.3.4)$$

with

$$\omega = \frac{\beta - r}{A}$$

$$\phi = r + \lambda + \delta$$

and

$$\psi = r + \lambda + \omega.$$

Finally, the consumer's financial wealth at date  $t$  equals

$$w_t = \bar{y} \frac{e^{-\omega t} - e^{-\delta t}}{\phi}. \quad (2.3.5)$$

## 2.3.2 Equilibrium

In equilibrium the aggregate financial wealth

$$\int_0^{\infty} \lambda e^{-\lambda t} w_t dt \quad (2.3.6)$$

equals the market value of the perpetuities i.e.  $p = D/r$ .

Using equation 2.3.5 we can show that the equilibrium interest rate solves the equation

$$\frac{r^*(\delta - \omega^*)}{\phi^*(\lambda + \omega^*)} = \frac{D}{Y} \quad (2.3.7)$$

---

<sup>16</sup>To calculate this solution we have assumed that the borrowing constraint is not binding, i.e.  $\delta > \omega \equiv (\beta - r)/A$ , and that the maximum in 2.3.3 is finite, i.e.  $\psi \equiv r + \lambda + \omega > 0$ . Both restrictions hold in equilibrium. (See appendix 2.A for details.)

where  $r^*$ ,  $\omega^*$ ,  $\phi^*$  and  $\psi^*$  denote the equilibrium values of  $r$ ,  $\omega$ ,  $\phi$  and  $\psi$ , respectively.

Equation 2.3.7 determines the interest rate  $r^*$  uniquely. As expected the interest rate goes up with the discount factor  $\beta$ , the probability of death  $\lambda$  and the ratio of aggregate financial income over aggregate labor income  $D/Y$ . The interest rate goes down with the rate of decline of labor income,  $\delta$ , since an increase in  $\delta$  leads to greater incentives to save. Finally if the coefficient of elasticity of substitution,  $1/A$ , goes up the interest rate goes down provided that  $r^*$  be greater than  $\beta$ . Otherwise the interest rate goes up.

In equilibrium, agents use the financial markets to smooth their consumption over their lifetime: they buy assets when they are young and begin selling assets at age  $\tau^*$  where  $\tau^*$  solves

$$r^*w_{\tau^*} + y_{\tau^*} - c_{\tau^*} = 0.^{17} \quad (2.3.8)$$

From 2.3.5,  $\tau^*$  is given by

$$\tau^* = \frac{1}{\omega^* - \delta} \log \left( \frac{\omega^* + \lambda}{\delta + \lambda} \right). \quad (2.3.9)$$

The aggregate dollar volume in this economy equals

$$T = \int_0^{\infty} \lambda e^{-\lambda t} |rw_t + y_t - c_t| dt = 2\lambda w_{\tau^*} e^{-\lambda \tau^*} = 2Y \frac{\delta - \omega^*}{\phi^*} e^{-(\lambda + \omega^*)\tau^*}. \quad (2.3.10)$$

## 2.4 Transactions Costs and Assets' Returns

In this section, we determine the rate of return on the liquid asset,  $r$ , the rate of return on the illiquid asset,  $R$ , and the liquidity premium  $\mu$  in the presence of transactions

---

<sup>17</sup>Note that we do not consider the payment of shares by insurance companies to be a trade. Hence although the agent's portfolio may still be growing ( $dW_t > 0$ ), the agent is considered a seller if his portfolio grows at a rate lower than  $\lambda$ . In the absence of transactions costs, this assumption simply amounts to defining who is called a seller and who is not. With transactions costs, however, matters are different. Since we have assumed that insurance companies pay living participants *shares* of assets as opposed to cash and that this transfer is costless, our definition of a seller is the correct one.

costs. We will consider the case of small transactions costs and focus on their first order effect on equilibrium variables. For this purpose we write

$$r(\epsilon) = r^* + (b - m^*)\epsilon + o(\epsilon)$$

$$R(\epsilon) = r^* + b\epsilon + o(\epsilon)$$

$$\mu(\epsilon) = m^*\epsilon + o(\epsilon)$$

where  $b$  and  $m^*$  are the first order equilibrium effects that we seek to calculate. We consider the case where the supply of the illiquid asset,  $k$ , is less than one so both assets are available to consumers. The case where all assets are illiquid, i.e.  $k = 1$ , is somewhat different and is studied in appendix 2.E.

Before proceeding with the formal derivations, it is useful to show that in equilibrium the liquidity premium per unit of transactions costs  $\mu/\epsilon$  must be greater than the rate of return on the liquid asset,  $r$ , or equivalently  $R > r(1 + \epsilon)$ . This is because, since agents are born without any financial assets, in equilibrium they must buy the illiquid asset at some point in their lives. Now consider an agent who buys for one dollar worth of asset *inclusive of transactions costs* at date  $t$  and sells it  $\Delta t$  periods later. If he buys the liquid asset his cash flows are

-1 at date  $t$

$rds$  for  $s$  between  $t$  and  $t + \Delta t$  and

+1 at date  $t + \Delta t$ .

If he buys the illiquid asset, given the transactions costs he will get  $1/[P(1 + \epsilon)]$  shares. Hence his cash flows are

-1 at date  $t$

$R/(1 + \epsilon)ds$  for  $s$  between  $t$  and  $t + \Delta t$  and

$(1 - \epsilon)/(1 + \epsilon)$  at date  $t + \Delta t$ .

If  $R \leq r(1 + \epsilon)$ , i.e. if  $\mu \leq r\epsilon$ , then buying the liquid asset always dominates buying the illiquid asset. Hence

$$\mu > r\epsilon. \tag{2.4.1}^{18}$$

In particular, this means that the effect of transactions costs on the liquidity premium is at least a first order effect. With this a priori information about equilibrium prices, we next characterize the investor's demand for liquid and illiquid assets when  $\mu > r\epsilon$ .

### 2.4.1 The consumer's problem

With transactions costs, the consumer's financial wealth  $w_t$  is the sum of the value of his liquid portfolio, denoted by  $a_t$ , and of the value of his illiquid portfolio, denoted by  $A_t$ . Denoting by  $i_t$  (respectively  $I_t$ ) the incremental dollar investment in the liquid asset (respectively illiquid asset), the dynamics of  $a_t$  and  $A_t$  are given by

$$\begin{aligned} da_t &= \lambda a_t dt + i_t dt; & a_0 &= 0; & a_t &\geq 0 \\ dA_t &= \lambda A_t dt + I_t dt; & A_0 &= 0; & A_t &\geq 0 \\ c_t &= y_t + ra_t + RA_t - i_t - I_t - \epsilon|I_t|; & c_t &\geq 0. \end{aligned} \tag{2.4.2}$$

From 2.4.2 above, we can see that the agent's consumption equals the labor income  $y_t$ , plus the dividend income  $ra_t + RA_t$ , minus purchases of liquid assets  $i_t$ , minus purchases of illiquid assets  $I_t$ , minus transactions costs  $\epsilon|I_t|$ .

With transactions costs, the consumer's problem becomes far more complex. Proposition 2.4.1 (proven in appendix 2.B) describes the optimal policy of the consumer for small transactions costs and for a subset of values of  $r$  and  $R$  that are of interest, i.e. such that their equilibrium values belong to this subset.

**Proposition 2.4.1** *For  $\epsilon$  small and for  $r$  and  $R$  belonging to a subset of their possible values, the optimal policy has the following form: The consumer buys the*

---

<sup>18</sup>This lower bound is reached asymptotically when the holding period of illiquid assets goes to infinity, that is when the fraction of illiquid assets,  $k$ , goes to zero.

*illiquid asset until an age  $\tau_1$ . He then buys the liquid asset. He next sells the liquid asset until an age  $\tau_1 + \Delta$ . At age  $\tau_1 + \Delta$ , he does not own any share of the liquid asset. He then start selling the illiquid asset until he dies.*

We find that in equilibrium, agents will buy high yield illiquid assets for long-term investment and low yield liquid assets for short-term investment. This fairly intuitive result is consistent with the analysis of Amihud and Mendelson (1986). The clientele for the illiquid asset are the agents of age less than  $\tau_1$  while the clientele for the liquid asset are the agents of age between  $\tau_1$  and the age at which they begin to sell it. The marginal investor is the investor who buys the illiquid asset at date  $\tau_1$  and sells it at date  $\tau_1 + \Delta$ . As in Amihud and Mendelson (1986), this investor determines the liquidity premium (see below).

The liquid and illiquid portfolios as function of age  $t$  are plotted in figure 2.1.

Proposition 2.4.1 presents the qualitative properties of the optimal consumption/investment policy. In what follows, these qualitative properties will allow us to calculate consumption at date  $t$ ,  $c_t$ , as a function of the initial consumption,  $c_0$ . We will also show how the intertemporal budget equation, properly modified to account for transactions costs, leads to the determination of the initial consumption  $c_0$ . Finally, we will show how the parameter  $\Delta$  can be easily calculated as function of the rate of return on the liquid asset,  $r$ , the liquidity premium,  $\mu$ , and the level of transactions costs,  $\epsilon$ . For the sake of the presentation, all technical details have been sent to appendix 2.B.

Over the course of his life the agent faces three interest rates.

First until age  $\tau_1$ , the interest rate which is relevant for the consumption-savings decision is

$$R_L + \lambda = \frac{R}{1 + \epsilon} + \lambda.$$

Indeed, consider a consumer who at  $t \in [0, \tau_1]$  decides to consume \$1 less, but wants to have the same wealth after  $t + dt$ . He then buys  $1/P(1 + \epsilon)$  illiquid securities. At  $s$  between  $t$  and  $t + dt$ , he consumes the extra dividend flows  $(D/P(1 + \epsilon))e^{\lambda(s-t)}$  and

at  $t + dt$ , he consumes the proceeds *from avoiding to buy*  $(1/P(1 + \epsilon))e^{\lambda dt}$  securities, ie  $e^{\lambda dt}$ . Hence by foregoing \$1 at  $t$  he gets  $\$1 + \lambda dt + (R/(1 + \epsilon))dt + o(dt)$  between  $t$  and  $t + dt$ . Therefore, for  $R$  given, higher transactions costs increase the desire to consume earlier rather than later. The reason is that the consumer has to buy an asset which is more expensive, but pays the same dividend.

Second, between ages  $\tau_1$  and  $\tau_1 + \Delta$ , the consumer invests in the liquid assets and therefore faces the interest rate  $r + \lambda$ .

Third and finally, after age  $\tau_1 + \Delta$ , when the consumer is divesting out of the illiquid assets, he faces a higher rate

$$R_B + \lambda = \frac{R}{1 - \epsilon} + \lambda.$$

Indeed, suppose that at  $t \in [\tau_1 + \Delta, \infty)$  he decides to consume \$1 less but wants to have the same wealth after  $t + dt$ . He sells  $1/P(1 - \epsilon)$  less illiquid securities. At  $s$  between  $t$  and  $t + dt$ , he consumes the extra dividend flow  $(D/P(1 - \epsilon))e^{\lambda(s-t)}$  and at  $t + dt$  he consumes the proceeds from selling  $(1/P(1 - \epsilon))e^{\lambda dt}$  securities i.e.  $e^{\lambda dt}$ . Hence by foregoing \$1 at  $t$  he gets  $\$1 + \lambda dt + (R/(1 - \epsilon))dt + o(dt)$  between  $t$  and  $t + dt$ . For  $R$  given, higher transactions costs increase the desire to consume later rather than earlier. The reason is that the consumer has to sell a cheaper asset that pays the same dividend.

We denote by  $\hat{\rho}(t)$  the interest rate relevant for date  $t$  i.e.

$$\begin{aligned}\hat{\rho}(t) &= R_L + \lambda \text{ for } t < \tau_1 \\ \hat{\rho}(t) &= r + \lambda \text{ for } \tau_1 \leq t < \tau_1 + \Delta \\ \hat{\rho}(t) &= R_B + \lambda \text{ for } \tau_1 + \Delta \leq t\end{aligned}$$

and by  $\rho(t)$  the *discount rate between date 0 and date t* i.e.

$$\rho(t) = \int_0^t \hat{\rho}(s) ds.$$

In appendix 2.B, we indeed show that the optimal consumption must satisfy

$$c_t = c_0 e^{-\omega(t)} \quad (2.4.3)$$

with

$$\omega(t) = \frac{(\beta + \lambda)t - \rho(t)}{A}$$

where the consumption at birth  $c_0$  is derived from the intertemporal budget constraint presented below.

Given proposition 2.4.1 and equation 2.4.2, it can easily be shown that the consumption path  $c_t$  must satisfy the *intertemporal budget equation*

$$\int_0^{\tau_1} \frac{y_t - c_t}{1 + \epsilon} e^{-\rho(t)} dt + \int_{\tau_1}^{\tau_1 + \Delta} (y_t + (R - r)A_t - c_t) e^{-\rho(t)} dt + \int_{\tau_1 + \Delta}^{\infty} \frac{y_t - c_t}{1 - \epsilon} e^{-\rho(t)} dt = 0 \quad (2.4.4)$$

with

$$A_t = A_{\tau_1} e^{\lambda(t - \tau_1)} = e^{\lambda(t - \tau_1)} \int_0^{\tau_1} \frac{y_s - c_s}{1 + \epsilon} e^{\rho(\tau_1) - \rho(s)} ds.$$

Equation 2.4.4 says that the Net Present Value of lifetime savings net of transactions costs must equal zero, where we define savings as total income minus consumption *minus* what must be reinvested in order for financial wealth to grow at the rate  $\hat{\rho}(t)$ . Between periods  $[0, \tau_1[$  and  $[\tau_1 + \Delta, \infty[$ , this latter quantity equals the dividend income and therefore savings equal  $y_t - c_t$ . Between  $\tau_1$  and  $\tau_1 + \Delta$ , only a fraction  $rA_t$  of the dividends from the illiquid portfolio must be reinvested and thus savings equal labor income,  $y_t$ , minus consumption  $c_t$  plus *excess* dividends  $(R - r)A_t$ . Finally savings are adjusted for transactions costs.

We now show how the minimum holding period of the illiquid asset,  $\Delta$ , can be calculated as a function of  $r$ ,  $\mu$  and  $\epsilon$ . Consider a consumer at age  $\tau_1$ . Since  $\tau_1$  and  $\Delta$  are optimally chosen, this consumer must be indifferent between investing in the illiquid asset and not doing so. Given that he starts selling the illiquid asset at  $\tau_1 + \Delta$  his change in utility if he buys one unit of the illiquid asset at  $\tau_1$  is

$$-u'(c_{\tau_1})(1 + \epsilon)P + u'(c_{\tau_1 + \Delta})(1 - \epsilon)Pe^{-\beta\Delta} + \int_{\tau_1}^{\tau_1 + \Delta} u'(c_t)De^{-\beta(t - \tau_1)}dt = 0. \quad (2.4.5)$$

Use of equation 2.4.3 and simple algebra show that the above equation can be rewritten as

$$\frac{\mu}{\epsilon} = r \frac{1 + e^{-r\Delta}}{1 - e^{-r\Delta}}. \quad (2.4.6)$$

Equation 2.4.6 shows that the minimum holding period of the illiquid asset,  $\Delta$ , is decreasing in its excess rate of return over the liquid asset,  $\mu$ , and increasing in transactions costs,  $\epsilon$ .<sup>19</sup>

From equation 2.4.6 it follows that the intertemporal budget constraint, for an optimal choice of  $\tau_1$  and  $\Delta$ , states that the Net Present Value of consumption equals the Net Present Value of income where the discount factor is  $\rho(t)$ , i.e.

$$\int_0^{\infty} (y_t - c_t)e^{-\rho(t)}dt = 0. \quad (2.4.7)$$

The initial consumption,  $c_0$  can be derived from equations 2.4.3 and 2.4.7.

Having characterized the solution to the consumer's problem we turn to the equilibrium determination of  $r$  and  $R$ .

## 2.4.2 Equilibrium

In equilibrium, the dollar demands for liquid and illiquid assets

$$\int_0^{\infty} \lambda e^{-\lambda t} a_t dt$$

and:

$$\int_0^{\infty} \lambda e^{-\lambda t} A_t dt$$

---

<sup>19</sup>We can also derive equation 2.4.6 by noting that between  $\tau_1$  and  $\tau_1 + \Delta$  the consumer invests in the liquid asset. Therefore the Net Present Value rule applies, and the Net Present Value of investing in the illiquid asset between these two dates is zero.



equal the assets' market value,  $(1 - k)D/r$  and  $kD/R$  respectively.

As we stated in the beginning of this section, we will consider small transactions costs (small values of  $\epsilon$ ), and find their first order effects on  $r$ ,  $R$  and  $\mu$ . Recall that we have defined  $b$  and  $m^*$  by

$$r(\epsilon) = r^* + (b - m^*)\epsilon + o(\epsilon) \quad (2.4.8)$$

$$R(\epsilon) = r^* + b\epsilon + o(\epsilon) \quad (2.4.9)$$

$$\mu(\epsilon) = m^*\epsilon + o(\epsilon) \quad (2.4.10)$$

and that  $r^*$ ,  $\omega^*$ ,  $\phi^*$  and  $\psi^*$  are the equilibrium values of  $r$ ,  $\omega$ ,  $\phi$  and  $\psi$  for  $\epsilon = 0$ .

We also define by  $\tau_1(\epsilon)$  and  $\Delta(\epsilon)$  as the equilibrium values of  $\tau_1$  and  $\Delta$  as a function of  $\epsilon$ , and by  $\tau_1^*$  and  $\Delta^*$  the respective limits of  $\tau_1(\epsilon)$  and  $\Delta(\epsilon)$  as  $\epsilon$  goes to zero. (Note that when  $\epsilon$  equals zero, the liquid asset and the illiquid asset are the same asset and therefore the holdings  $a_t$  and  $A_t$  are not well defined.) In other terms,  $\tau_1^*$  and  $\Delta^*$  are the zero-th order effects of transactions costs on holding periods.

The next proposition characterizes the equilibrium values of  $m$  and  $b$ .

**Proposition 2.4.2** *There exists an equilibrium where  $r$ ,  $R$  and  $\mu$  have the form of equations 2.4.8, 2.4.9 and 2.4.10. In equilibrium the first order effect of transactions costs on the liquidity premium,  $m^*$ , is positive while the first order effect on the rate of return on the liquid asset,  $b - m^*$ , is negative. The first order effect on the rate of return on the illiquid asset,  $b$ , has an ambiguous sign.*

The rigorous derivations of  $b - m^*$ ,  $b$  and  $m^*$ , as well as explicit formulae are presented in appendix 2.C.

Proposition 2.4.2, that we discuss in detail next, states that transactions costs decrease the rate of return on the liquid asset but have an ambiguous effect on the rate of return on the illiquid asset.

We discuss the results of proposition 2.4.2 in subsections 2.4.2.1., 2.4.2.2. and

2.4.2.3. In subsection 2.4.2.1, we characterize the parameters  $\tau_1^*$  and  $\Delta^*$  i.e. the zero-th order effect of transactions costs on optimal consumption/savings policies. In subsection 2.4.2.2 we go over the determination of the liquidity premium, and in the rather long subsection 2.4.2.3 we discuss the determination of the rates of return (or, more accurately, the first order effects of transactions costs on these variables.)

### Optimal Switching Times

The age at which agents switch from the illiquid asset to the liquid asset,  $\tau_1^*$  and the age at which they start selling the illiquid asset,  $\tau_1^* + \Delta^*$ , can be easily interpreted from the limit case as  $\epsilon$  goes to zero. Indeed consider the accumulation equations 2.4.2 with  $\epsilon = 0$  and  $r = R = r^*$ . Given the investor's total wealth  $w_t = a_t + A_t$ , the values of the liquid and illiquid portfolios over time are given by

$$\begin{aligned} A_t &= w_t \text{ and } a_t = 0 \text{ for } t < \tau_1^* \\ A_t &= w_{\tau_1^*} e^{\lambda(t-\tau_1^*)} \text{ and } a_t = w_t - w_{\tau_1^*} e^{\lambda(t-\tau_1^*)} \text{ for } \tau_1^* \leq t < \tau_1^* + \Delta^* \\ A_t &= w_t \text{ and } a_t = 0 \text{ for } \tau_1^* + \Delta^* \leq t. \end{aligned}$$

It follows that the values of  $\tau_1^*$  and  $\Delta^*$  can be calculated by noting (i) that total financial wealth  $w_t$  grows by  $\exp(\lambda\Delta^*)$  between  $\tau_1^*$  and  $\tau_1^* + \Delta^*$  i.e.

$$w_{\tau_1^* + \Delta^*} = w_{\tau_1^*} e^{\lambda\Delta^*} \quad (2.4.11)$$

and (ii) that aggregate liquid financial wealth must equal the supply of liquid assets i.e.

$$\int_{\tau_1^*}^{\tau_1^* + \Delta^*} \lambda e^{-\lambda t} a_t dt = \int_{\tau_1^*}^{\tau_1^* + \Delta^*} \lambda e^{-\lambda t} (w_t - w_{\tau_1^*} e^{\lambda(t-\tau_1^*)}) dt = (1 - k) \frac{D}{r^*}. \quad (2.4.12)$$

Using the expression for  $w_t$  from equation 2.3.5 as well as equations 2.4.11 and 2.4.12 above we obtain the values of  $\tau_1^*$  and  $\Delta^*$ .

## Liquidity Premium

In appendix 2.C we show that the first order effect of transactions costs on the liquidity premium,  $m$ , is given by

$$m^* = r^* \frac{1 + e^{-r^* \Delta^*}}{1 - e^{-r^* \Delta^*}}. \quad (2.4.13)$$

It is fairly easy to understand the determination of  $m^*$ . Equation 2.4.6 implies that if  $m^*$  were different from its value in equation 2.4.13, (i.e. if  $\mu$  were different in the first order), there would be a zero-th order change in  $\Delta$ . Therefore, there would a zero-th order change in the demand for liquid versus illiquid assets. (Although there would only be a first order change in total asset demand.)

## Rates of Return

The reasoning for the rates of return is more involved. To determine the parameter  $b$  (and  $b - m^*$ ) we will make the following exercise: We will assume that for fixed  $r$  transactions costs increase. In order to preserve equilibrium,  $R$  has to increase by  $m^* \epsilon$ , for  $m^*$  given by equation 2.4.13. We will then find by how much total asset demand and supply change in the first order and infer  $b - m^*$  by the equation that states that total asset demand equals total asset supply:

$$\int_0^{\infty} \lambda e^{-\lambda t} (a_t + A_t) dt = \int_0^{\infty} \lambda e^{-\lambda t} w_t dt = (1 - k)D/r + kD/R. \quad (2.4.14)$$

Since this exercise is useful for understanding why the rate of return on the liquid asset decreases when transactions costs increase, we go over it in some detail.

To determine total asset demand, we must first understand how the consumption path of the agent is modified by the change in transactions costs and asset returns that we are considering. Lemma 2.4.3 (proven in appendix 2.C) gives us the consumption of the agent at age 0.

**Lemma 2.4.3** *The consumption at date zero,  $c_0$  is given by*

$$c_0 = \bar{y} \frac{\psi^*}{\phi^*} + \epsilon C_W + \epsilon C_S + o(\epsilon) \quad (2.4.15)$$

where  $C_W$  is given by

$$C_W = \bar{y} \frac{\psi^*}{\phi^*} \left( \left( \frac{m^*}{r^*} - 1 \right) \int_0^{\tau_1^*} [(\lambda + \delta)e^{-\phi^* t} - (\lambda + \omega^*)e^{-\psi^* t}] dt \right. \\ \left. + \left( \frac{m^*}{r^*} + 1 \right) \int_{\tau_1^* + \Delta^*}^{\infty} [(\lambda + \delta)e^{-\phi^* t} - (\lambda + \omega^*)e^{-\psi^* t}] dt \right) \quad (2.4.16)$$

or alternatively

$$\bar{y} \frac{\psi^*}{\phi^*} \left( (m^* - r^*) \int_0^{\tau_1^*} (e^{-\psi^* t} - e^{-\phi^* t}) dt + (m^* + r^*) \int_{\tau_1^* + \Delta^*}^{\infty} (e^{-\psi^* t} - e^{-\phi^* t}) dt \right) \quad (2.4.17)$$

and  $C_S$  is given by

$$C_S = -\frac{1}{A} \bar{y} \frac{\psi^*}{\phi^*} \left( (m^* - r^*) \int_0^{\tau_1^*} e^{-\psi^* t} dt + (m^* + r^*) \int_{\tau_1^* + \Delta^*}^{\infty} e^{-\psi^* t} dt \right). \quad (2.4.18)$$

The terms  $C_W$  and  $C_S$  have a very intuitive interpretation. First,  $C_W$  can be interpreted as a wealth effect. For this we need to note that

$$\bar{y} \frac{|(\lambda + \delta)e^{-\phi^* t} - (\lambda + \omega^*)e^{-\psi^* t}|}{\phi^*}$$

is the present value of the dollar amount of transactions between  $\tau$  and  $\tau + d\tau$  in the case  $\epsilon = 0$ . The consumer when buying the illiquid asset between 0 and  $\tau_1$  pays the transactions costs but pays a lower price. When selling the illiquid asset (from  $\tau_1 + \Delta$  until he dies), he pays the transactions costs and receives a lower price. The term  $C_W$  is equal to the present discounted value of these “extra”<sup>20</sup> cash flows, times  $\psi^*$ , as expression 2.4.16 shows. Clearly, since the rate of return in the liquid asset is kept

---

<sup>20</sup>Compared to the case  $\epsilon = 0$ .

constant, the consumer can only be better off compared to the case  $\epsilon = 0$ , and this term is positive as we can see in expression 2.4.17.

The term  $C_S$  can be interpreted as a substitution effect. Since  $R/(1+\epsilon) > r$ , saving is more attractive from 0 to  $\tau_1$ . (Agents buy the illiquid asset paying transaction costs but at a lower price which more than compensates them.) It is also clear that  $R/(1-\epsilon) > r$ , therefore deferring consumption for later is also more attractive from  $\tau_1 + \Delta$  until death. Thus this term is negative.

Having interpreted the expression for  $c_0$ , we can briefly describe how the consumption path changes compared to the case  $\epsilon = 0$ . Because the consumer has better investment opportunities, (i.e. he has the liquid asset at the same price as before, and the illiquid asset), his consumption path goes up uniformly. (This is the wealth effect.) Because the illiquid asset is available at a lower price (which more than compensates transactions costs), and because the proceeds from selling it are lower (lower price plus transactions costs), the consumer changes the slope of the consumption path so that he buys more of the illiquid asset and holds it for a longer period. In other words, he buys more in the beginning of his life (he saves more) and he sells it at a lower rate (he defers consumption for later). This is the substitution effect.

In lemma 2.4.4 (proven in appendix 2.C) we determine total asset demand.

**Lemma 2.4.4** *Total asset demand is given by*

$$\frac{\delta - \omega^*}{(\lambda + \omega^*)\phi^*} Y + \epsilon W_W + \epsilon W_S + o(\epsilon) \quad (2.4.19)$$

where  $W_W$  and  $W_S$  are given by

$$W_W = 2 \frac{\lambda \bar{y}}{\phi^* r^*} (e^{-(\omega^* + \lambda)\tau_1^*} - e^{-(\delta + \lambda)\tau_1^*}) - \frac{m^*}{r^*} \int_0^\infty \lambda e^{-\lambda t} A_t dt +$$

$$\frac{\lambda}{(\lambda + \omega^*)} \frac{\bar{y} \psi^*}{\phi^* r^*} \left( (m^* - r^*) \int_0^{\tau_1^*} (e^{-\psi^* t} - e^{-\phi^* t}) dt + (m^* + r^*) \int_{\tau_1^* + \Delta^*}^\infty (e^{-\psi^* t} - e^{-\phi^* t}) dt \right) \quad (2.4.20)$$

and

$$W_S = \frac{1}{A} \frac{\lambda}{(\lambda + \omega^*)} \frac{\bar{y}\psi^*}{\phi^* r^*}$$

$$\left( (m^* - r^*) \int_0^{\tau_1^*} [e^{-(\lambda + \omega^*)t} - e^{-\psi^*t}] dt + (m^* + r^*) \int_{\tau_1^* + \Delta^*}^{\infty} (e^{-(\lambda + \omega^*)t} - e^{-\psi^*t}) dt \right) \quad (2.4.21)$$

The term  $W_W$  represents the additional demand for wealth (in the first order) of the consumer if the latter changes only the level but not the slope of his consumption path (i.e. if the wealth effect is present, but not the substitution effect) in response to the change in transaction costs and asset returns that we are studying. This additional demand for (dollar) wealth has an ambiguous sign because, on the one hand, higher future consumption to be financed and transactions costs to be paid when selling assets require more wealth, but on the other hand illiquid assets are cheaper.

The term  $W_S$  corresponds to the substitution effect: Indeed as it was said before the consumer changes the slope of his consumption path so that he buys more of the illiquid asset and holds it for a longer period. This implies more wealth accumulation. This term is positive and its magnitude depends on the elasticity of intertemporal substitution.

Finally total asset supply (in the first order) is:

$$\frac{D}{r^*} - \epsilon \frac{m^*}{r^*} \int_0^{\infty} \lambda e^{-\lambda t} A_t dt = \frac{D}{r^*} - \epsilon W_{supply} \quad (2.4.22)$$

It decreases since the illiquid asset is cheaper.

The difference between total asset demand and supply is  $W_W + W_S + W_{supply}$  and is *always* positive. It is easy to understand why this is so, based on our earlier discussion. Higher future consumption to be financed by selling the cheaper illiquid asset and paying transactions costs requires a larger *number* of securities to be held. (Although the dollar amount may be lower.) In addition, the agents change the slope of their consumption paths in order to buy more of the illiquid asset and hold it for a

longer period, making the imbalance between asset demand and supply even higher.

The value of  $b - m^*$  is then easily deduced, and is negative.

The above discussion which explained why the rate of return on the liquid asset falls, can be summarized as follows: Suppose that transactions costs increase from 0 to  $\epsilon$  and that the rate of return on the liquid asset stays the same in equilibrium. Then, (in equilibrium) the rate of return on the illiquid asset must increase (in the first order) by  $m^*\epsilon$ . Agents' consumption paths will shift uniformly up because there are more investment opportunities (wealth effect), and their slope will change so that they buy more of the illiquid asset and hold it for a longer period (substitution effect). Agents will thus demand more securities for two reasons. First, because they have to finance higher future consumption by selling the cheaper illiquid asset and paying transaction costs. Second, because they want to buy more of the illiquid asset and hold it for a longer period.

Although the first order effect of transactions costs on the rate of return on the liquid asset is unambiguous ( $b - m^*$  is negative), the first order effect on the rate of return on the illiquid asset (i.e. the sign of  $b$ ) is ambiguous. In what follows, we replicate (more briefly) the above exercise, assuming that this time, as transactions costs increase,  $R$  stays the same and  $r$  decreases in the first order, as determined above ( $r$  decreases by  $m^*\epsilon$ ).

This time, agents face worse investment opportunities. The price of the liquid asset increases while trading the illiquid asset entails transactions costs. This (wealth) effect implies then that their consumption paths shift down uniformly. On the other hand, by substitution, agents accumulate less of both assets but hold the illiquid asset for a longer period. The effect on the total demand for securities is ambiguous. Indeed, the future consumption to be financed is lower and the liquid asset is more expensive, but on the other hand transactions costs have to be paid. In addition, agents buy less of the liquid and illiquid assets, but hold the illiquid asset for a longer period.

### 2.4.3 Comparative Statics

In this subsection we study how the effects of transactions costs on assets' returns depend on the parameters of the model. (More precisely, we find how  $m^*$ ,  $b$  and  $b - m^*$  depend on these parameters.) The parameter that is of greatest interest is  $k$ , the fraction of illiquid assets to the total stock of assets. In lemma 2.4.5 (proven in appendix 2.D) we examine how  $m^*$ ,  $b$  and  $b - m^*$  depend on  $k$ .

**Lemma 2.4.5**  *$m^*$  increases in  $k$ ,  $b - m^*$  decreases in  $k$  while the dependence of  $b$  in  $k$  is ambiguous.*

We briefly discuss the results of this Lemma.

The dependence of  $m^*$  on  $k$  is relatively simple to understand. More illiquid assets in the economy imply that the minimum holding period of an illiquid asset becomes shorter. The liquidity premium must increase so that consumers are willing to hold illiquid assets for shorter periods.

The dependence of  $b - m^*$  on  $k$  can be explained in the light of the analysis of the previous subsection. There it was argued that to understand why the rate of return on the liquid asset falls in response to increased transactions costs, we could make the following experiment: We could suppose that transactions costs increased from 0 to  $\epsilon$  and that the rate of return on the illiquid asset had to increase (in the first order) by  $m^*\epsilon$ . We could then study the difference between demand and supply of total wealth and infer the direction of change of the rate of return on the liquid asset. In fact, we can also infer the *magnitude* of change of the rate of return on the liquid asset, studying the magnitude of the difference between total asset demand and total asset supply.

As  $k$  increases,  $m^*$  increases, therefore both the wealth effect and the substitution effect are stronger. This implies that the difference between asset demand and supply is greater, and the first order effect on  $r$  (i.e.  $b - m^*$ ) is bigger (in absolute value).

The effects of the other parameters on  $m^*$ ,  $b$  and  $b - m^*$  are of less interest and are not reported here.



## 2.5 Numerical Examples

In this section we present some numerical examples to illustrate the results of the previous sections. In all these examples we assume that  $A = 1$  ( $u(c) = \log c$ ) and that the level of transactions costs  $\epsilon$  equals 3% which is consistent with empirical evidence (see Aiyagari and Gertler (1991) for instance). In all figures 2.2 to 2.4, we plot various rates of return as a function of  $k$ , the supply of the illiquid asset. These figures are consistent with the results in proposition 2.4.2 and lemma 2.4.5, namely that (i) the liquidity premium is positive, (ii) the rate of return on the liquid asset goes down, (iii) the rate of return on the illiquid asset can go up or down, (iv) the effect of transactions costs on the liquidity premium and the rate of return on the liquid asset is large when  $k$  is close to 1.

The main quantitative observations are as follows: (i) When  $k$  is close to 1, the liquidity premium is significant (about 10% of the level of the rates of return). (ii) When  $k$  is close to 1, transactions costs cause a non-trivial fall in the rate of return on the liquid asset while the rate of return on the illiquid asset remains almost constant.

These quantitative results have important practical applications. To understand the impact of a change in transactions costs in the economy, it is important to understand how assets are differently affected by this change. A technological change, such as a reduction in computer cost, can be assumed to reduce transaction costs for all assets and in our model corresponds to the case  $k=1$ . Our results suggest that rates of return will not change much. By contrast, a reduction of transactions costs on one single asset (e.g. by the introduction of a derivative security) will increase the price of this asset without any significant impact on the other assets. Finally, a transaction tax on a significant subset of existing assets (stocks, real estate ..) will lower their value by an amount less than suggested by a simple partial equilibrium analysis which takes the rates of return on the other assets as given.

## 2.6 Conclusion

In this work we have constructed a fairly simple general equilibrium model of an imperfect capital market. Our main result is that while transactions costs tend to push the rate of return on illiquid assets upward, there is a general equilibrium effect which tends to lower rates of return. The net result is that the rate of return on liquid assets goes down while the rate of return on illiquid assets may go up or down. We believe that these results are robust to the specification of (i) the trading motives: life cycle, labor income shocks<sup>21</sup> or taste shocks and (ii) the preferences.<sup>22</sup>

Our model endogenously generates clienteles for assets with differential liquidity. This clientele effect is consistent with previous work by Amihud and Mendelson (1986). In fact, if we generalized our model to allow for many assets with different transactions costs, we would obtain the concave relationship between rates of return and transactions costs derived by these authors.

If transactions costs instead of being a pure destruction of resources, as we assumed here, were due to a transaction tax whose proceeds were redistributed to the agents, the results would be similar. Amihud and Mendelson (1990b) argue that, holding the risk free rate constant, a .5% transaction tax would lower the market value of the NYSE stocks by 13.8%. While we do not dispute the fact that a small transaction tax will increase the liquidity premium significantly, our results suggest that the risk free rate will fall so that the stock price fall is likely to be somewhat smaller.

This line of research can be pursued in (at least) two directions. First, the interaction between risk and liquidity is not fully understood. It would be interesting to construct tractable models to analyze the interaction between transactions costs and risk and examine in particular whether, as it has been argued, illiquid markets are more volatile since investors find it more costly to absorb liquidity shocks. Second and

---

<sup>21</sup>See for instance Amihud et al. (1992).

<sup>22</sup>The treatment of the perpetual youth model for a general utility function seems to us analytically intractable. In a companion note, we consider a two-period overlapping generations model similar in spirit to the model herein. This model is simpler but also much less rich. In particular, the holding period is the same for all assets and as a result the liquidity premium is fixed. In this simple model the results are independent of the functional form of preferences.

more importantly, very little is known about the determination of the level of transactions costs as well as the financial structure created to deal with these transactions costs.

## 2.A Appendix: The No Transactions Costs Case

This appendix considers the case when transactions costs are zero. We will first prove that in equilibrium  $\delta > \omega$  and  $\psi = r + \lambda + \omega > 0$ . We will then prove that the equilibrium is unique.

We must first calculate the optimal policy which entails solving

$$\begin{aligned} \max \int_0^{\infty} u(c_t) e^{-(\beta+\lambda)t} dt &= \max \int_0^{\infty} \frac{1}{1-A} c^{1-A} e^{-(\beta+\lambda)t} dt \\ \text{s.t. } \int_0^{\infty} c_t e^{-(r+\lambda)t} dt &= \max \int_0^{\infty} y_t e^{-(r+\lambda)t} dt; \quad w_t \geq 0. \end{aligned} \quad (2.A.1)$$

From He and Pages (1991), we know that a bounded value for 2.3.3 above exists provided that

$$\psi = r + \lambda + \omega > 0. \quad (2.A.2)$$

In that case, He and Pages show that

$$\begin{aligned} c_t &= y_t \text{ for every } t \text{ if } \omega = (\beta - r)/A \geq \delta \text{ and} \\ c_t &= (\bar{y}\psi/\phi) e^{-\omega t} \text{ for every } t \text{ if } \omega < \delta, \text{ with } \phi = r + \lambda + \delta. \end{aligned}$$

If  $\psi \leq 0$ , then the value of 2.3.3 is  $\infty$ . We show below that this cannot be the case in equilibrium.

In equilibrium, the resource constraint implies that

$$\int_0^{\infty} \lambda e^{-\lambda t} \leq D + Y. \quad (2.A.3)$$

Hence the equilibrium utility of the representative agent is bounded by the solution to the program below

$$\max \int_0^{\infty} \frac{1}{1-A} c^{1-A} e^{-(\beta+\lambda)t} dt$$

$$s.t. \int_0^{\infty} \lambda e^{-\lambda t} c_t dt = D + Y. \quad (2.A.4)$$

It is easy to show that 2.A.4 has a finite value.

If  $\omega \geq \delta$  then the agent does not buy any asset and therefore  $w_t \equiv 0$ . Obviously, this is not consistent with the equilibrium condition

$$\int_0^{\infty} \lambda e^{-\lambda t} w_t dt = \frac{D}{r}. \quad (2.A.5)$$

Hence  $\delta$  must be larger than  $\omega$  in equilibrium.

From the expression for  $c_t$  and 2.3.2 it follows easily that

$$w_t = \bar{y} \frac{e^{-\omega t} - e^{-\delta t}}{\phi}. \quad (2.A.6)$$

Combining 2.A.5 and 2.A.6 yields the equilibrium condition

$$\frac{r^*(\delta - \omega^*)}{\phi^*(\lambda + \omega^*)} = \frac{r^*(\delta - \frac{\beta - r^*}{A})}{(r^* + \lambda + \delta)(\lambda + \frac{\beta - r^*}{A})} = \frac{D}{Y} \quad (2.A.7)$$

which must be solved for  $r^*$ .

Simple algebraic manipulations show that 2.A.7 above has a unique positive solution  $r^*$ , that this solution is increasing in  $\beta$ ,  $\lambda$  and  $D/Y$  and that it is decreasing in  $\delta$ . It is increasing in  $A$  if  $r^*$  is greater than  $\beta$  and decreasing with  $A$  otherwise.

## 2.B Appendix: Proof of Proposition 2.4.1

The method of proof is as follows: We first define the control variables. We then derive heuristic conditions for an optimal control. Next we construct a candidate optimal control and show that is indeed optimal.

### Step 1 The control problem

Recall that  $i_t$  is the per unit time value of the liquid assets purchased at date  $t$  and that  $I_t$  is the per unit time value of the illiquid assets purchased at date  $t$ . Recall also equations 2.4.2

$$da_t = \lambda a_t dt + i_t dt; \quad a_0 = 0; \quad a_t \geq 0$$

$$dA_t = \lambda A_t dt + I_t dt; \quad A_0 = 0; \quad A_t \geq 0 \quad (2.B.1)$$

$$c_t = y_t + r a_t + R A_t - i_t - I_t - \epsilon |I_t|; \quad c_t \geq 0.$$

The control problem faced by the consumer is to maximize 2.2.5 with respect to the controls  $i_t$  and  $I_t$  subject to the above dynamics of  $a_t$  and  $A_t$ .

Formally, we say that a control  $(i(), I())$  is *admissible* if it is (i) piecewise continuous (ii) it satisfies the no short sale constraints  $a_t \geq 0$  and  $A_t \geq 0$  as well as the constraint  $c_t \geq 0$ . We denote by  $\mathcal{C}$  the set of admissible controls and by  $J(i(), I())$  the payoff function, i.e. the utility that the consumer enjoys if he follows the controls  $i_t = i(t)$  and  $I_t = I(t)$ . Using the fact that

$$a_t = \int_0^t i_s e^{\lambda(t-s)} ds \quad (2.B.2)$$

and

$$A_t = \int_0^t I_s e^{\lambda(t-s)} ds \quad (2.B.3)$$

the payoff function can be written as

$$J(i(), I()) = \int_0^\infty u(y_t + r \int_0^t i_s e^{\lambda(t-s)} ds + R \int_0^t I_s e^{\lambda(t-s)} ds - i_t - I_t - \epsilon |I_t|) e^{-(\beta+\lambda)t} dt \quad (2.B.4)$$

Hence the control problem is to maximize  $J(i(), I())$  with respect to  $(i(), I())$  belonging to  $\mathcal{C}$ .

## Step 2 Heuristic optimality conditions

To derive heuristic conditions for an optimal control, we take a candidate optimal control  $(i(), I())$ , denote by  $c()$  the resulting consumption and consider the possible perturbations:

(i) Suppose first that  $a_t > 0$ . Suppose that at  $t$ , the consumer changes his consumption by  $\alpha$  dollars (per unit time) and invests  $\alpha$  more dollars in the liquid asset. At  $s$  between  $t$  and  $t + dt$ , he consumes the extra dividend  $\alpha r e^{\lambda(s-t)}$  and at  $t + dt$  he sells  $\alpha e^{\lambda dt}$ . He does not change his investment decision thereafter. His payoff will change by

$$\alpha[-u'(c_t) + e^{-(\beta+\lambda)dt}(e^{\lambda dt} + r dt)u'(c_{t+dt})].$$

Since  $(i(), I())$  is optimal, the change in payoff must be non-positive for every  $\alpha$  and  $dt$  going to zero which yields the standard Euler equation

$$\frac{du'(c_t)}{dt} = (\beta - r)u'(c_t). \quad (2.B.5)$$

(ii) Suppose now that  $a_t = 0$ . Then the consumer can only *increase* his investment in the liquid asset and therefore condition 2.B.5 must be replaced by

$$\frac{du'(c_t)}{dt} \leq (\beta - r)u'(c_t). \quad (2.B.6)$$

(iii) Consider a policy where  $A_t > 0$  for every  $t$ , which will be the case in our candidate policy.

(iiia) Suppose  $I_t^* > 0$ . If between  $t$  and  $t + dt$ , the consumer changes his consumption by  $\alpha(1 + \epsilon)$  dollars (per unit time), invests  $\alpha$  (with  $\alpha + I_t^* > 0$ ) more dollars in the illiquid asset and does not change his investment decision thereafter, his payoff will change by

$$\alpha[-u'(c_t)(1 + \epsilon) + \int_t^\infty Ru'(c_s)e^{-\beta(s-t)} ds]dt$$

which must be non-positive for any  $\alpha$  and so

$$u'(c_t)(1 + \epsilon) = \int_t^\infty Ru'(c_s)e^{-\beta(s-t)} ds. \quad (2.B.7)$$

(iiib) Suppose now that  $I_t^* < 0$ . If between  $t$  and  $t + dt$  the consumer changes his

consumption by  $(1 - \epsilon)\alpha$  dollars (per unit time), invests  $\alpha$  (with  $\alpha + I_t^* < 0$ ) more dollars in the illiquid asset and does not change his investment decision thereafter<sup>23</sup> his payoff will change by

$$\alpha[-u'(c_t)(1 - \epsilon) + \int_t^\infty Ru'(c_s)e^{-\beta(s-t)}ds]dt$$

which must be non-positive for any  $\alpha$  and so

$$u'(c_t)(1 - \epsilon) = \int_t^\infty Ru'(c_s)e^{-\beta(s-t)}ds. \quad (2.B.8)$$

(iiic) If  $I_t = 0$ , then the first order condition reads as

$$u'(c_t)(1 - \epsilon) \leq \int_t^\infty Ru'(c_s)e^{-\beta(s-t)}ds \leq u'(c_t)(1 + \epsilon). \quad (2.B.9)$$

### Step 3 Construction of the candidate policy

Given  $c_0$ ,  $\tau_1$  and  $\Delta$  we define the candidate optimal consumption plan as in equation 2.4.3

$$c_t = c_0e^{-\omega(t)}$$

with

$$\omega(t) = \frac{(\beta + \lambda)t - \rho(t)}{A}.$$

This consumption plan will be completely defined once we specify  $c_0$ ,  $\tau_1$  and  $\Delta$ . We now motivate our definition of these parameters. To finance this consumption plan, the agent will buy shares of the illiquid asset from 0 to  $\tau_1$ . He will buy and then sell shares of the liquid asset between  $\tau_1$  and  $\tau_1 + \Delta$ . Finally he will sell the illiquid assets from  $\tau_1 + \Delta$  to  $\infty$ . From equations 2.4.2 it follows that

---

<sup>23</sup>Note that this perturbation is feasible only if  $A_t > 0$  for every  $t$ . In the optimal policy,  $A_t$  will indeed be positive for every  $t$ . However  $a_t = 0$  for  $t \leq \tau_1$  and  $t \geq \tau_1 + \Delta$ . For this reason, the optimality condition with respect to  $i_t$  is written as an Euler equation (2.B.5 and 2.B.6) while the optimality condition with respect to  $I_t$  is an integral condition.



$$A_t = \int_0^t \frac{y_s - c_s}{1 + \epsilon} e^{\rho(t) - \rho(s)} ds; \quad a_t = 0$$

for  $t < \tau_1$

$$A_t = A_{\tau_1} e^{\lambda(t - \tau_1)}; \quad a_t = \int_{\tau_1}^t (y_s - c_s + RA_s) e^{\rho(t) - \rho(s)} ds$$

for  $\tau_1 \leq t < \tau_1 + \Delta$  and

$$A_t = A_{\tau_1 + \Delta} e^{\rho(t) - \rho(\tau_1 + \Delta)} + \int_{\tau_1 + \Delta}^t \frac{y_s - c_s}{1 - \epsilon} e^{\rho(t) - \rho(s)} ds; \quad a_t = 0$$

for  $\tau_1 + \Delta \leq t$ .

From the above equations together with the transversality condition

$$\lim_{t \rightarrow \infty} A_t e^{-\rho(t)} = 0$$

we get

$$\int_0^{\tau_1} \frac{y_t - c_t}{1 + \epsilon} e^{-\rho(t)} dt = e^{r\Delta} \int_{\tau_1 + \Delta}^{\infty} \frac{c_t - y_t}{1 - \epsilon} e^{-\rho(t)} dt. \quad (2.B.10)$$

Finally

$$a_{\tau_1 + \Delta} = 0 = e^{\rho(\tau_1 + \Delta)} \int_{\tau_1}^{\tau_1 + \Delta} (y_t - c_t + RA_t) e^{-\rho(t)} dt. \quad (2.B.11)$$

We first define  $\Delta$  by 2.4.6 that is

$$\frac{\mu}{\epsilon} = \frac{R - r}{\epsilon} = r \frac{1 + e^{-r\Delta}}{1 - e^{-r\Delta}}$$

or equivalently

$$e^{-r\Delta} = \frac{\frac{R}{r} - 1 - \epsilon}{\frac{R}{r} - 1 + \epsilon}. \quad (2.B.12)$$

Adding equations 2.B.10 and 2.B.11 gives the intertemporal budget equation 2.4.4. We define  $c_0$  (as a function of  $\tau_1$ ,  $\epsilon$ ,  $r$  and  $R$ ) from the intertemporal budget equation 2.4.4 which together with 2.4.6 yields the simpler equation 2.4.7 below

$$\int_0^{\infty} (y_t - c_t)e^{-\rho(t)} dt = 0.$$

Finally we define  $\tau_1$  by equation 2.B.11.

We now show that this policy is well defined and admissible for  $\epsilon$  small and for  $r$  and  $R$  belonging to a subset of their equilibrium values. We will also show that it varies smoothly with  $\epsilon$ ,  $r$  and  $m$  where  $m$  is defined by

$$R \equiv r + m\epsilon.$$

We will show the following lemma

**Lemma 2.B.1** *Consider  $r$ ,  $m$  and  $m^*$  such that  $m^*/r^* > 1$  (recall that  $r^*$  is the equilibrium interest rate when transactions costs are zero). There exists  $\epsilon_0$  such that if*

$$(i) \quad 0 < \epsilon < \epsilon_0$$

$$(ii) \quad |r^* - r| \leq \epsilon_0, |m^* - m| \leq \epsilon_0$$

*the consumption plan defined above, where  $R = r + m\epsilon$  is well-defined and admissible. Moreover,  $\Delta$ ,  $c_0$  and  $\tau_1$  are infinitely differentiable ( $C^\infty$ ) in  $(\epsilon, r, m)$ .*

**Proof:** The intuition for the proof is simple. If  $\epsilon = 0$ , this consumption plan collapses to the optimal one for  $\epsilon = 0$ . That plan is admissible. The admissibility of the consumption plan for  $\epsilon > 0$  will follow by continuity.

We first note that the equation defining  $\Delta$  can be written as

$$e^{-r\Delta} = \frac{m - r}{m + r}. \tag{2.B.13}$$

Therefore  $\Delta$  is uniquely defined, is  $C^\infty$  in  $(\epsilon, r, m)$  and verifies  $0 < \underline{\Delta} \leq \Delta \leq \overline{\Delta} < \infty$  for  $\epsilon_0$  sufficiently small.

The equation giving  $c_0$  can be written as

$$\bar{y} \left( \frac{1 - e^{-(R_L + \lambda + \delta)\tau_1}}{R_L + \lambda + \delta} + e^{-(R_L + \lambda + \delta)\tau_1} \frac{1 - e^{-(r + \lambda + \delta)\Delta}}{r + \lambda + \delta} + e^{-(R_L + \lambda + \delta)\tau_1} \frac{e^{-(r + \lambda + \delta)\Delta}}{R_B + \lambda + \delta} \right) =$$

$$c_0 \left( \frac{1 - e^{-(R_L + \lambda + \omega_L)\tau_1}}{R_L + \lambda + \omega_L} + e^{-(R_L + \lambda + \omega_L)\tau_1} \frac{1 - e^{-(r + \lambda + \omega)\Delta}}{r + \lambda + \omega} + e^{-(R_L + \lambda + \omega_L)\tau_1} \frac{e^{-(r + \lambda + \omega)\Delta}}{R_B + \lambda + \omega_B} \right) \quad (2.B.14)$$

with

$$\omega_L = \frac{\beta - R_L}{A} \text{ and } \omega_B = \frac{\beta - R_B}{A},$$

From equation 2.B.14, it is obvious that  $c_0$  is uniquely defined and is  $C^\infty$  in  $(\epsilon, r, m, \tau_1)$ .

It is fairly straightforward to show that we can restrict  $\epsilon_0 > 0$  such that

$$\left| \frac{\partial c_0}{\partial \epsilon} \right| \leq \bar{y} K_c \text{ in } [0, \epsilon_0]$$

for all values of  $r$  and  $m$  in the set defined before, and for all  $\tau_1 \in [0, \infty[$ , where  $K_c$  is a positive constant.

We now turn to the equation defining  $\tau_1$ . This equation can be written as

$$\frac{1}{1 + \epsilon} \left( \bar{y} \frac{1 - e^{-(R_L + \lambda + \delta)\tau_1}}{R_L + \lambda + \delta} - c_0 \frac{1 - e^{-(R_L + \lambda + \omega_L)\tau_1}}{R_L + \lambda + \omega_L} \right) =$$

$$\frac{1}{1 - \epsilon} \left( c_0 e^{-(R_L + \lambda + \omega_L)\tau_1} \frac{e^{-(\lambda + \omega)\Delta}}{R_B + \lambda + \omega_B} - \bar{y} e^{-(R_L + \lambda + \delta)\tau_1} \frac{e^{-(\lambda + \delta)\Delta}}{R_B + \lambda + \delta} \right). \quad (2.B.15)$$

Straightforward algebra shows that we can rewrite this equation as

$$e^{-(r + \lambda + \delta)\tau_1} - e^{-(r + \lambda + \omega)\tau_1} = e^{-(r + \lambda + \delta)\tau_1 - (\lambda + \delta)\Delta} - e^{-(r + \lambda + \omega)\tau_1 - (\lambda + \omega)\Delta} + f\left(\frac{c_0}{\bar{y}}, r, m, \epsilon, \tau_1\right) \quad (2.B.16)$$

where  $f(., ., ., ., .)$  is a  $C^\infty$  function such that  $f = 0$  for  $\epsilon = 0$  and

$$\left| \frac{\partial f}{\partial \epsilon} \right| \leq K_f \text{ in } [0, \epsilon_0]$$

for all values of  $r$ ,  $m$  and  $\tau_1$  (restricting  $\epsilon_0$  if necessary).  $K_f$  is a positive constant.

Consider equation 2.B.16 for  $\epsilon = 0$

$$e^{-(\lambda+\delta)\tau_1} - e^{-(\lambda+\omega)\tau_1} = e^{-(r+\lambda+\delta)(\tau_1+\Delta)} - e^{-(r+\lambda+\omega)(\tau_1+\Delta)}. \quad (2.B.17)$$

The function  $g(\cdot) : t \rightarrow e^{-(\delta+\lambda)t} - e^{-(\omega+\lambda)t}$  has the following graph (see figure 2.5)

Therefore this equation has a unique solution  $\tau_1$  in  $(0, \tau^*)$ . It is easy to show (given  $\Delta \geq \underline{\Delta} > 0$ ) that there exists  $\zeta > 0$  such that  $\tau_1 \leq \tau^* - \zeta$  and  $\tau_1 + \Delta \geq \tau^* + \zeta$ . Using the implicit function theorem, the fact that  $0 < \underline{\Delta} \leq \Delta \leq \bar{\Delta} < \infty$  and the fact that

$$\left| \frac{\partial f}{\partial \epsilon} \right| \leq K_f \text{ in } [0, \epsilon_0]$$

for all values of  $r$ ,  $m$  and  $\tau_1$ , we can show that we can define a  $C^\infty$  function  $\tau_1(\epsilon, r, m)$  for  $\epsilon \leq \epsilon_0$  and for all values of  $r$  and  $m$ . Moreover, it is easy to show that

$$\left| \frac{\partial \tau_1}{\partial \epsilon} \right| \leq K_\tau \text{ in } [0, \epsilon_0]$$

uniformly in  $r$  and  $m$  (restricting  $\epsilon_0$  if necessary). This implies in particular that  $\tau_1 \leq \tau^* - \zeta/2$  and  $\tau_1 + \Delta \geq \tau^* + \zeta/2$ , for  $\epsilon$  small.

Summarizing our discussion above, it can be seen that  $c_0(\epsilon, r, m)$  is close to its  $\epsilon = 0$  value. It is less straightforward to interpret the values of  $\tau_1(0, r, m)$  and  $\Delta(0, r, m)$ . Indeed when  $\epsilon = 0$ , all assets are liquid and the switching times do not have any particular meaning. However, as seen above  $\tau_1(\epsilon, r, m)$  and  $\Delta(\epsilon, r, m)$  have well-defined positive limits as  $\epsilon$  goes to zero. As seen from 2.B.13, given  $m$  is easy to calculate  $\Delta(0, r, m) = \Delta(\epsilon, r, m)$ . Given  $m$  (or  $\Delta$ ), one can calculate  $\tau_1(0, r, m)$  directly from the no transactions costs case. This value  $\tau_1(0, r, m)$  is the time such that accumulated wealth (when  $\epsilon=0$ ) grows at a rate  $\lambda$  between  $\tau_1(0, r, m)$  and  $\tau_1(0, r, m) + \Delta$ .

Therefore, the consumption plan is well-defined and  $\Delta$ ,  $c_0$ ,  $\tau_1$  vary smoothly with  $\epsilon$ ,  $r$ ,  $m$ .

To show that it is admissible, we have to prove that  $I_t \geq 0$  in  $[0, \tau_1]$ ,  $I_t \leq 0$  in  $[\tau_1 + \Delta, \infty[$  and  $a_t \geq 0$ ,  $A_t \geq 0$ . We recall that

$$I_t = \frac{dA_t}{dt} - \lambda A_t.$$

We briefly sketch proofs of the above statements.

In  $[0, \tau_1]$ , we have

$$I_t = \frac{RA_t + y_t - c_t}{1 + \epsilon}.$$

It is easy to see that  $I_t = I_{t|\epsilon=0} + g(\epsilon, r, m, t)$  where  $g=0$  for  $\epsilon=0$  and

$$\left| \frac{\partial f}{\partial \epsilon} \right| \leq K_g$$

uniformly. Since  $\tau_1 \leq \tau^* + \zeta/2$  and  $I_{t|\epsilon=0} \geq \theta > 0$  in  $[0, \tau^* - \zeta/2]$ , it follows that for  $\epsilon \leq \epsilon_0$  (restricting again  $\epsilon_0$ ),  $I_t > 0$ , which implies that  $A_t \geq 0$ .

In  $[\tau_1, \tau_1 + \Delta]$  again,  $a_t$  and  $i_t$  are very close to their  $\epsilon = 0$  counterparts. We can easily show by continuity that

$$a_t > 0 \text{ in } [\tau^* - \zeta/4, \tau^* + \zeta/4]$$

$$i_t > 0 \text{ in } [\tau_1, \tau^* - \zeta/4]$$

$$i_t < 0 \text{ in } [\tau^* + \zeta/4, \tau_1 + \Delta].$$

This implies that  $a_t \geq 0$  in  $[\tau_1, \tau_1 + \Delta]$ .

In  $[\tau_1 + \Delta, \infty[$ , simple calculations using

$$A_t = \int_t^\infty \frac{c_s - y_s}{1 - \epsilon} e^{\rho(t) - \rho(s)} ds$$

show that

$$I_t = \frac{1}{1 - \epsilon} \left( \frac{\lambda + \delta}{R_B + \lambda + \delta} \bar{y} e^{-\delta t} - \frac{\lambda + \omega_B}{R_B + \lambda + \omega_B} c_0 e^{-\omega_L \tau_1} e^{-\omega \Delta} e^{-\omega_B(t - (\tau_1 + \Delta))} \right).$$

The continuity argument can be applied in a compact set  $[\tau^* + \zeta/2, T]$  to show

that  $I_t < 0$ , while since  $\delta - \omega_B \geq \eta > 0$  for  $\epsilon_0$  small,  $I_t$  will be negative if  $T$  is large enough.

Finally

$$A_t = \frac{1}{1 - \epsilon} \left( \frac{c_0}{R_B + \lambda + \omega_B} e^{-\omega_L \tau_1} e^{-\omega \Delta} e^{-\omega_B(t - (\tau_1 + \Delta))} - \frac{\bar{y}}{R_B + \lambda + \delta} e^{-\delta t} \right)$$

and similar arguments show that  $A_t \geq 0$ . Therefore the consumption plan is admissible. This ends the proof of lemma 2.B.1.  $\square$

#### Step 4 Optimality of the control

Having shown that our candidate optimal control is well defined and admissible, we will show that it is indeed optimal. For this purpose, we first show that it satisfies 2.B.5-2.B.9.

**Lemma 2.B.2** *Our candidate control satisfies conditions 2.B.5-2.B.9.*

**Proof:** It is obvious that

in  $[0, \tau_1]$

$$\frac{d \log u'(c_t)}{dt} = \beta - R_L \leq \beta - r$$

in  $[\tau_1, \tau_1 + \Delta]$

$$\frac{d \log u'(c_t)}{dt} = \beta - r$$

and that in  $[\tau_1 + \Delta, \infty[$

$$\frac{d \log u'(c_t)}{dt} = \beta - R_B \leq \beta - r.$$

Hence 2.B.5-2.B.6 are satisfied.

For  $t \in [\tau_1 + \Delta, \infty[$

$$\int_t^\infty u'(c_t) e^{-\beta(s-t)} ds = u'(c_t) \int_t^\infty e^{-R_B(s-t)} ds = \frac{u'(c_t)}{R_B} = \frac{u'(c_t)(1-\epsilon)}{R}.$$

It follows that 2.B.8 is satisfied.

For  $t \in [\tau_1, \tau_1 + \Delta]$

$$\begin{aligned} \int_t^\infty u'(c_t) e^{-\beta(s-t)} ds &= u'(c_t) \left( \int_t^{\tau_1+\Delta} e^{-r(s-t)} ds + \int_{\tau_1+\Delta}^\infty e^{-r(\tau_1+\Delta-t)} e^{-R_B(s-(\tau_1+\Delta))} ds \right) \\ &= u'(c_t) \left( \frac{1 - e^{-r(\tau_1+\Delta-t)}}{r} + (1-\epsilon) \frac{e^{-r(\tau_1+\Delta-t)}}{R} \right). \end{aligned}$$

In addition

$$\frac{1-\epsilon}{R} \leq \frac{1 - e^{-r(\tau_1+\Delta-t)}}{r} + (1-\epsilon) \frac{e^{-r(\tau_1+\Delta-t)}}{R} \leq \frac{1 - e^{-r\Delta}}{r} + (1-\epsilon) \frac{e^{-r\Delta}}{R}.$$

From equation 2.B.12 we get that

$$\frac{1 - e^{-r\Delta}}{r} + (1-\epsilon) \frac{e^{-r\Delta}}{R} = \frac{1+\epsilon}{R}. \quad (2.B.18)$$

The inequalities 2.B.9 follow.

Finally, for  $t \in [0, \tau_1]$ , we have

$$\begin{aligned} \int_t^\infty u'(c_t) e^{-\beta(s-t)} ds &= u'(c_t) \left( (1+\epsilon) \frac{1 - e^{-\frac{R}{1+\epsilon}(\tau_1-t)}}{R} + e^{-\frac{R}{1+\epsilon}(\tau_1-t)} \frac{1 - e^{-r\Delta}}{r} \right. \\ &\quad \left. + (1-\epsilon) e^{-\frac{R}{1+\epsilon}(\tau_1-t)} \frac{e^{-r\Delta}}{R} \right) = \frac{1+\epsilon}{R} \end{aligned}$$

because of equation 2.B.18.

This proves equation 2.B.7 and ends our proof of lemma 2.B.2.  $\square$

We now show that the candidate policy is indeed optimal.

**Lemma 2.B.3** *The candidate policy is optimal.*

**Proof:** To derive this lemma we shall show that no perturbation of the candidate

policy can be utility enhancing.

Consider an alternative policy  $(i_t + \delta i_t, I_t + \delta I_t)$  which induces consumption  $c_t + \delta c_t$ . We know from equations 2.4.2, 2.B.2 and 2.B.3 that

$$\delta c_t = r \int_0^t \delta i_s e^{\lambda(t-s)} ds + R \int_0^t \delta I_s e^{\lambda(t-s)} ds - \delta i_t - \delta I_t - \epsilon(|I_t + \delta I_t| - |I_t|).$$

Using the concavity of  $u(c_t)$ , it follows that

$$u(c_t + \delta c_t) \leq u(c_t) + u'(c_t)$$

$$\left( r \int_0^t \delta i_s e^{\lambda(t-s)} ds + R \int_0^t \delta I_s e^{\lambda(t-s)} ds - \delta i_t - \delta I_t - \epsilon(|I_t + \delta I_t| - |I_t|) \right). \quad (2.B.19)$$

We next multiply equation 2.B.19 by  $e^{-(\beta+\lambda)t}$ , integrate from 0 to  $t$ , and get

$$\int_0^t u(c_s + \delta c_s) e^{-(\beta+\lambda)s} ds \leq \int_0^t u(c_s) e^{-(\beta+\lambda)s} ds + K_i(t) + K_I(t)$$

with

$$K_i(t) = \int_0^t u'(c_s) \left( r \int_0^s \delta i_h e^{\lambda(s-h)} dh - \delta i_s \right) e^{-(\beta+\lambda)s} ds$$

and

$$K_I(t) = \int_0^t u'(c_s) \left( R \int_0^s \delta I_h e^{\lambda(s-h)} dh - \delta I_s - \epsilon(|I_t + \delta I_t| - |I_t|) \right) e^{-(\beta+\lambda)s} ds.$$

We will show that when  $t$  goes to infinity,  $K_i(t)$  and  $K_I(t)$  are asymptotically non-positive.

Integrating the second term of  $K_i(t)$  by parts, we get

$$\begin{aligned} & \int_0^t u'(c_s) \delta i_s e^{-(\beta+\lambda)s} ds = \\ & \left[ \left( \int_0^s \delta i_h e^{-\lambda h} dh \right) u'(c_s) e^{-\beta s} \right]_0^t - \int_0^t \left( \int_0^s \delta i_h e^{-\lambda h} dh \right) \frac{d(u'(c_s) e^{-\beta s})}{ds} ds. \end{aligned}$$



Therefore,  $K_i(t)$  equals

$$\int_0^t \left( u'(c_s)(r - \beta) + \frac{du'(c_s)}{ds} \right) \left( \int_0^s \delta i_h e^{-\lambda h} dh \right) e^{-\beta s} ds - \left( \int_0^t \delta i_s e^{-\lambda s} ds \right) u'(c_t) e^{-\beta t}. \quad (2.B.20)$$

We note that

$$\delta a_s e^{-\lambda s} = \int_0^s \delta i_h e^{-\lambda h} dh.$$

If  $a_s > 0$ , then condition 2.B.5 holds and the integrand in the first term of 2.B.20 above must be zero. If  $a_s = 0$ , then the short sale constraint  $\delta a_s \geq 0$  and condition 2.B.6 ensure that this integrand is non-positive. For  $t \geq \tau_1 + \Delta$ ,  $a_t = 0$  and thus  $\delta a_t$  must also be greater or equal than 0. This implies that the second term is non-positive, for  $t$  large enough. We have therefore proven that  $K_i(t)$  is non-positive for large  $t$ .

Consider now  $K_I(t)$ . Integrating by parts the first term, we get

$$\begin{aligned} & \int_0^t u'(c_s) \left( \int_0^s \delta I_h e^{\lambda(s-h)} dh \right) e^{-(\beta+\lambda)s} ds = \\ & \left[ - \left( \int_0^s \delta I_h e^{-\lambda h} dh \right) \left( \int_s^\infty u'(c_h) e^{-\beta h} dh \right) \right]_0^t + \int_0^t \left( \int_s^\infty u'(c_h) e^{-\beta h} dh \right) \delta I_s e^{-\lambda s} ds. \end{aligned}$$

Therefore  $K_I(t)$  is equal to

$$\begin{aligned} & \int_0^t \left( (-u'(c_s) + R \int_s^\infty u'(c_h) e^{-\beta(h-s)} dh) \delta I_s - u'(c_s) \epsilon (|I_s + \delta I_s| - |I_s|) \right) e^{-(\beta+\lambda)s} ds \\ & - R \left( \int_t^\infty u'(c_s) e^{-\beta s} ds \right) \left( \int_0^t \delta I_s e^{-\lambda s} ds \right). \quad (2.B.21) \end{aligned}$$

If  $I_s > 0$ ,  $|I_s + \delta I_s| - |I_s| \geq \delta I_s$ , the integrand in 2.B.21 is less or equal to

$$\left( -u'(c_s)(1 + \epsilon) + \int_s^\infty R u'(c_h) e^{-\beta(h-s)} dh \right) \delta I_s = 0$$

by condition 2.B.7.

If  $I_s < 0$ ,  $|I_s + \delta I_s| - |I_s| \geq -\delta I_s$ , the integrand is less or equal to

$$\left( -u'(c_s)(1 - \epsilon) + \int_s^\infty Ru'(c_h)e^{-\beta(h-s)}dh \right) \delta I_s = 0$$

by condition 2.B.8.

If  $I_s = 0$  and  $\delta I_s > 0$ , the integrand is

$$\left( -u'(c_s)(1 + \epsilon) + \int_s^\infty Ru'(c_h)e^{-\beta(h-s)}dh \right) \delta I_s \leq 0$$

by condition 2.B.9.

If  $I_s = 0$  and  $\delta I_s < 0$ , the integrand is

$$\left( -u'(c_s)(1 - \epsilon) + \int_s^\infty Ru'(c_h)e^{-\beta(h-s)}dh \right) \delta I_s \leq 0$$

by condition 2.B.9.

Therefore the first term in  $K_I(t)$  is less or equal to 0.

For the second term, note that

$$\delta A_s e^{-\lambda s} = \int_0^s \delta I_h e^{-\lambda h} dh$$

and that

$$\int_t^\infty u'(c_s) e^{-\beta s} ds = u'(c_t) e^{-\beta t} \frac{1 - \epsilon}{R} = e^{-\rho(t)} \frac{1 - \epsilon}{R}$$

for  $t \geq \tau_1 + \Delta$ .

Since  $\delta A_t \geq -A_t$ ,  $(\delta A_t + A_t \geq 0)$ ,  $A_t \sim \exp(-\omega_B t)$  and  $\lambda + \omega_B > \eta > 0$  for  $\epsilon$  small, it is clear that the term:

$$-R \left( \int_t^\infty u'(c_s) e^{-\beta s} ds \right) \left( \int_0^t \delta I_s e^{-\lambda s} ds \right)$$

is smaller than an arbitrary  $\xi > 0$  for  $t$  large enough. It follows that  $K_I(t)$  is asymp-

totally non positive. This concludes the proof of lemma 2.B.3.  $\square$

## 2.C Appendix: Proofs of Proposition 2.4.2 and Lemmas 2.4.3 and 2.4.4

### 2.C.1 Proof of Lemma 2.4.3

Suppose that  $r = r^*$  and that  $R = r^* + m^*\epsilon$ . From appendix 2.B, we know that for  $\epsilon$  small, the optimal consumption at time 0 is given by equation 2.4.7, which as we have seen can be rewritten as

$$\bar{y} \left( \frac{1 - e^{-(R_L + \lambda + \delta)\tau_1}}{R_L + \lambda + \delta} + e^{-(R_L + \lambda + \delta)\tau_1} \frac{1 - e^{-(r^* + \lambda + \delta)\Delta}}{r^* + \lambda + \delta} + e^{-(R_L + \lambda + \delta)\tau_1} \frac{e^{-(r^* + \lambda + \delta)\Delta}}{R_B + \lambda + \delta} \right) =$$

$$c_0 \left( \frac{1 - e^{-(R_L + \lambda + \omega_L)\tau_1}}{R_L + \lambda + \omega_L} + e^{-(R_L + \lambda + \omega_L)\tau_1} \frac{1 - e^{-(r^* + \lambda + \omega)\Delta}}{r^* + \lambda + \omega} + e^{-(R_L + \lambda + \omega_L)\tau_1} \frac{e^{-(r^* + \lambda + \omega)\Delta}}{R_B + \lambda + \omega_B} \right).$$

The first term in brackets (divided by  $\bar{y}$ ) can be written as

$$\frac{1}{R_L + \lambda + \delta} + e^{-(R_L + \lambda + \delta)\tau_1} \left( \frac{1}{r^* + \lambda + \delta} - \frac{1}{R_L + \lambda + \delta} \right)$$

$$+ e^{-(R_L + \lambda + \delta)\tau_1} e^{-(r^* + \lambda + \delta)\Delta} \left( \frac{1}{R_B + \lambda + \delta} - \frac{1}{r^* + \lambda + \delta} \right). \quad (2.C.1)$$

Straightforward algebra shows that this term can be written as

$$\frac{1}{\phi^*} \left( 1 - \frac{m^* - r^*}{\phi^*} \epsilon + \frac{m^* - r^*}{\phi^*} \epsilon e^{-\phi^* \tau_1^*} + \frac{m^* + r^*}{\phi^*} \epsilon e^{-\phi^* (\tau_1^* + \Delta^*)} + o(\epsilon) \right) \quad (2.C.2)$$

or equivalently, as

$$\frac{1}{\phi^*} \left( 1 - \epsilon(m^* - r^*) \int_0^{\tau_1^*} e^{-\phi^* t} dt - \epsilon(m^* + r^*) \int_{\tau_1^* + \Delta^*}^{\infty} e^{-\phi^* t} dt + o(\epsilon) \right). \quad (2.C.3)$$

Similarly, the second term can be written as

$$\frac{1}{\psi^*} \left( 1 - \epsilon \left( 1 - \frac{1}{A} \right) \left( (m^* - r^*) \int_0^{\tau_i^*} e^{-\psi^* t} dt + (m^* + r^*) \int_{\tau_i^* + \Delta^*}^{\infty} e^{-\psi^* t} dt \right) + o(\epsilon) \right). \quad (2.C.4)$$

Combining 2.C.3 and 2.C.4 we get equation 2.4.15 i.e.

$$c_0 = \bar{y} \frac{\psi^*}{\phi^*} + \epsilon C_W + \epsilon C_S + o(\epsilon)$$

with  $C_W$  and  $C_S$  given by 2.4.17 and 2.4.18. Integrating 2.4.17 by parts we get

$$\begin{aligned} & (m^* - r^*) \left( \left[ -\frac{e^{-r^* t}}{r^*} (e^{-(\lambda + \omega^*) t} - e^{-(\lambda + \delta) t}) \right]_0^{\tau_i^*} \right. \\ & \left. + \int_0^{\tau_i^*} \frac{e^{-r^* t}}{r^*} ((\lambda + \delta) e^{-(\lambda + \delta) t} - (\lambda + \omega^*) e^{-(\lambda + \omega^*) t}) dt \right) \\ & + (m^* + r^*) \left( \left[ -\frac{e^{-r^* t}}{r^*} (e^{-(\lambda + \omega^*) t} - e^{-(\lambda + \delta) t}) \right]_{\tau_i^* + \Delta^*}^{\infty} \right. \\ & \left. + \int_{\tau_i^* + \Delta^*}^{\infty} \frac{e^{-r^* t}}{r^*} ((\lambda + \delta) e^{-(\lambda + \delta) t} - (\lambda + \omega^*) e^{-(\lambda + \omega^*) t}) dt \right). \end{aligned}$$

Using equations 2.4.11 (substituting  $w_t$  from 2.3.5) and 2.4.13, we find that the terms in brackets cancel out which yields equation 2.4.16. This concludes the proof of lemma 2.4.3.

## 2.C.2 Proof of Proposition 2.4.2

Consider  $r^*$  and  $m^*$  where

$$m^* = r^* \frac{1 + e^{-r^* \Delta^*}}{1 - e^{-r^* \Delta^*}}.$$

The results of Appendix 2.B imply that for  $\epsilon$  sufficiently small and for  $r$  and  $m$  belonging to a neighborhood of  $r^*$  and  $m^*$ , the consumption plan defined in Proposition 2.4.1 is indeed well-defined, admissible and optimal. These results also imply

that

$$F : (\epsilon, r, m) \rightarrow \left( \int_0^\infty \lambda e^{-\lambda t} a_t dt, \int_0^\infty \lambda e^{-\lambda t} A_t dt \right)$$

is  $C^\infty$  in these variables.

We know that

$$F(0, r^*, m^*) = \left( (1 - k) \frac{D}{r^*}, k \frac{D}{r^*} \right).$$

It is also easy to see that

$$\frac{\partial F_1}{\partial m}(0, r, m) = \frac{\partial F_2}{\partial m}(0, r, m) \neq 0$$

(if  $\epsilon = 0$ ,  $r$  is kept constant and  $m$  changes, it is as if the total wealth is kept constant and the proportion of the wealth in liquid and illiquid assets changes) and that

$$\frac{\partial F_1}{\partial r}(0, r, m) + \frac{\partial F_2}{\partial r}(0, r, m) \neq 0$$

(a change in interest rate has a non-zero effect on the total stock of wealth, at least in the  $\epsilon=0$  case). Therefore the Jacobian matrix

$$\begin{bmatrix} \frac{\partial F_1}{\partial r} & \frac{\partial F_2}{\partial r} \\ \frac{\partial F_1}{\partial m} & \frac{\partial F_2}{\partial m} \end{bmatrix}$$

is invertible and we can apply the implicit function theorem at the point  $(0, r^*, m^*)$ .

It is thus clear that for  $\epsilon$  small there exists an equilibrium. Moreover,  $r$  and  $m$  are  $C^\infty$  in  $\epsilon$ , which establishes proposition 2.4.2.

We now calculate how total wealth changes when  $\epsilon$  changes for fixed  $r^*$  and  $m^*$  and prove lemma 2.4.4.

### 2.C.3 Proof of Lemma 2.4.4

Adding the equations describing the evolution of  $a_t$  and  $A_t$  we get

$$\begin{aligned}\frac{d(a_t + A_t)}{dt} &= \lambda(a_t + A_t) + i_t + I_t \\ &= (\lambda + r)(a_t + A_t) + (R - r)A_t + y_t - c_t - \epsilon I_t.\end{aligned}$$

Multiplying by  $e^{-\lambda t}$ , integrating from 0 to  $\infty$  and using the fact (established in appendix 2.B) that  $(a_t + A_t)e^{-\lambda t}$  goes to zero as  $t$  goes to infinity, we get

$$\int_0^\infty \lambda(a_t + A_t)e^{-\lambda t} dt = \frac{\lambda}{r} \int_0^\infty e^{-\lambda t} (c_t + \epsilon |I_t| - y_t - (R - r)A_t) dt \quad (2.C.5)$$

Equation 2.C.5 will allow us to calculate total wealth as a function of  $\epsilon$ . We will assume that we are at  $(\epsilon, r^*, m^*)$ .

$$\begin{aligned}\int_0^\infty e^{-\lambda t} \epsilon |I_t| dt &= \int_0^{\tau_1^*} e^{-\lambda t} \epsilon \left( \frac{dA_t}{dt} - \lambda A_t \right) - \int_{\tau_1^* + \Delta^*}^\infty e^{-\lambda t} \epsilon \left( \frac{dA_t}{dt} - \lambda A_t \right) = \\ \epsilon A_{\tau_1^*} e^{-\lambda \tau_1^*} + \epsilon A_{\tau_1^* + \Delta^*} e^{-\lambda(\tau_1^* + \Delta^*)} &= 2\epsilon A_{\tau_1^*} e^{-\lambda \tau_1^*} = 2\epsilon \frac{\bar{y}}{\phi^*} (e^{-(\lambda + \omega^*)\tau_1^*} - e^{-(\lambda + \delta)\tau_1^*}) + o(\epsilon).\end{aligned} \quad (2.C.6)$$

$$\begin{aligned}\int_0^\infty e^{-\lambda t} c_t dt &= c_0 \left( \int_0^{\tau_1^*} e^{-(\lambda + \omega_L)t} dt + e^{-(\lambda + \omega_L)\tau_1^*} \int_{\tau_1^*}^{\tau_1^* + \Delta^*} e^{-(\lambda + \omega^*)(t - \tau_1^*)} dt \right. \\ &\quad \left. + e^{-(\lambda + \omega_L)\tau_1^*} e^{-(\lambda + \omega^*)\Delta^*} \int_{\tau_1^* + \Delta^*}^\infty e^{-(\lambda + \omega_B)(t - (\tau_1^* + \Delta^*))} dt \right) \\ &= c_0 \left( \frac{1 - e^{-(\lambda + \omega_L)\tau_1^*}}{\lambda + \omega_L} + e^{-(\lambda + \omega_L)\tau_1^*} \frac{1 - e^{-(\lambda + \omega^*)\Delta^*}}{\lambda + \omega^*} + e^{-(\lambda + \omega_L)\tau_1^*} \frac{e^{-(\lambda + \omega^*)\Delta^*}}{\lambda + \omega_B} \right) \\ &= \frac{c_0}{\lambda + \omega^*} \left( 1 + \epsilon \frac{m^* - r^*}{\lambda + \omega^*} \frac{1}{A} (1 - e^{-(\lambda + \omega^*)\tau_1^*}) + \epsilon \frac{m^* + r^*}{\lambda + \omega^*} \frac{1}{A} e^{-(\lambda + \omega^*)(\tau_1^* + \Delta^*)} + o(\epsilon) \right)\end{aligned}$$

$$= \frac{c_0}{\lambda + \omega^*} \left\{ 1 + \frac{\epsilon}{A} \left( (m^* - r^*) \int_0^{\tau_1^*} e^{-(\lambda + \omega^*)t} dt + (m^* + r^*) \int_{\tau_1^* + \Delta^*}^{\infty} e^{-(\lambda + \omega^*)t} dt \right) \right\} + o(\epsilon), \quad (2.C.7)$$

Replacing for  $c_0$ , we get

$$\int_0^{\infty} e^{-\lambda t} c_t dt = \bar{y} \frac{\psi^*}{\phi^*(\lambda + \omega^*)}$$

$$\left\{ 1 + \epsilon \left( (m^* - r^*) \int_0^{\tau_1^*} (e^{-\psi^* t} - e^{-\phi^* t}) dt + (m^* + r^*) \int_{\tau_1^* + \Delta^*}^{\infty} (e^{-\psi^* t} - e^{-\phi^* t}) dt \right) + \right.$$

$$\left. \frac{\epsilon}{A} \left( (m^* - r^*) \int_0^{\tau_1^*} (e^{-(\lambda + \omega^*)t} - e^{-\psi^* t}) dt + (m^* + r^*) \int_{\tau_1^* + \Delta^*}^{\infty} (e^{-(\lambda + \omega^*)t} - e^{-\psi^* t}) dt \right) \right\} + o(\epsilon). \quad (2.C.8)$$

Using 2.C.5, 2.C.6 and 2.C.8, it is easy to find the sensitivity of total asset demand  $F_1 + F_2$  to the transactions costs  $\epsilon$ . The equations in lemma 2.4.4 follow.

We next derive the expression for  $b - m^*$ . For this purpose we differentiate the equilibrium condition

$$F_1 + F_2 = \int_0^{\infty} \lambda e^{-\lambda t} (a_t + A_t) dt = (1 - k) \frac{D}{r} + k \frac{D}{r + m\epsilon}$$

with respect to  $\epsilon$ . Note that when  $\epsilon$  changes, the *equilibrium values of  $r^*$  and  $m^*$*  will change. Hence we get

$$\frac{\partial(F_1 + F_2)}{\partial \epsilon} \Big|_{\epsilon=0} + \frac{\partial(F_1 + F_2)}{\partial r} \Big|_{\epsilon=0} (b - m^*) + \frac{\partial(F_1 + F_2)}{\partial m} \Big|_{\epsilon=0} \frac{dm}{d\epsilon} \Big|_{\epsilon=0}$$

$$= -(1 - k) \frac{D}{r} \Big|_{\epsilon=0} (b - m^*) - k \frac{D}{(r + m\epsilon)^2} \Big|_{\epsilon=0} ((b - m^*) + \epsilon \frac{dm}{d\epsilon} \Big|_{\epsilon=0} + m^*).$$

Using the fact that  $F_1 + F_2$  is independent of  $m$  for  $\epsilon = 0$ , we have

$$\frac{\partial(F_1 + F_2)}{\partial \epsilon} \Big|_{\epsilon=0} + \frac{\partial(F_1 + F_2)}{\partial r} \Big|_{\epsilon=0} (b - m^*) = -\frac{D}{r^*} (b - m^*) - \frac{m^*}{r^*} k \frac{D}{r^*}.$$

Using the notation of subsection 2.4.2.3 we get

$$(b - m^*) \left( \frac{\partial(F_1 + F_2)}{\partial r} \Big|_{\epsilon=0} + \frac{D}{r^*} \right) = -(W_W + W_S + W_{supply}).$$

We can calculate

$$\frac{\partial(F_1 + F_2)}{\partial r} \Big|_{\epsilon=0}$$

from

$$F_1 + F_2 = \frac{\delta - \omega}{(\lambda + \omega)\phi}$$

and derive  $(b - m^*)$ .

The right-hand side is negative and it is easy to see that the coefficient of  $(b - m^*)$  is positive. Therefore  $(b - m^*)$  is negative.  $\square$

## 2.D Appendix: Proof of Lemma 2.4.5

If  $k$  increases, obviously  $\Delta^*$  decreases. Moreover  $\tau_1^*$  increases and  $\tau_1^* + \Delta^*$  decreases. These can be seen from equations 2.4.11 and 2.4.12 with very simple algebra. Equation 2.4.13 implies then that  $m^*$  increases.

Simple algebra then shows that  $(b - m^*)$  which is proportional to:

$$-(W_W + W_S + W_{supply})$$

decreases, i.e. the effect of  $\epsilon$  on  $r^*$  is stronger. (Note that the function  $g(\cdot) : t \rightarrow \exp(-(\lambda + \omega)t) - \exp(-(\lambda + \delta)t)$  increases in  $[0, \tau^*]$  and that  $\tau_1 < \tau^*$ .)  $\square$



## 2.E Appendix: The Case $k=1$

The case  $k = 1$  is slightly different. We find, as before, that transactions costs have a first order effect on the rate of return on the illiquid asset. The difference with the case ( $0 < k < 1$ ) is that if we introduce a liquid asset in this economy in zero supply, that cannot be sold short, its return will be lower than the return on the illiquid asset by zero-th order term. (i.e. we have a zero-th order liquidity premium.) The reason for this result is that the minimum holding period has a first order length.

Since all of the consumer's wealth is held in the form of the illiquid asset, we have:

$$A_t = w_t$$

The dynamics of  $w_t$  are described by:

$$\begin{aligned} dw_t &= \lambda w_t dt + I_t dt \\ c_t &= y_t + R_t w_t - I_t - \epsilon |I_t| \end{aligned} \tag{2.E.1}$$

Proposition 2.E.1 describes the optimal policy of the consumer for small transactions costs and for a subset of values of  $R$  that are of interest, i.e. such that its equilibrium value belong to this subset.

**Proposition 2.E.1** *For  $\epsilon$  small and for  $R$  belonging to a subset of its possible values, the optimal policy has the following form: The consumer buys the (illiquid) asset until an age  $\tau_1$ . He does nothing (i.e. he consumes his income  $y_t + R w_t$ ) from  $\tau_1$  until an age  $\tau_1 + \Delta$  when he starts selling the asset until he dies.*

The proof of proposition 2.E.1 is analogous to the proof of proposition 2.4.1 and is therefore omitted.

In what follows, we will (briefly) discuss the implications of proposition 2.E.1. We will show how the consumption  $c_t$  as well as the width of the inaction period  $\Delta$  can be derived.

In the case  $k = 1$ , the expression for consumption  $c_t$  must be changed from equation 2.4.3 to

$$\begin{aligned}
 c_t &= c_0 e^{-\omega(t)} & t < \tau_1 \\
 c_t &= R w_t + y_t & \tau_1 \leq t < \tau_1 + \Delta \\
 c_t &= c_{\tau_1 + \Delta} e^{\omega(\tau_1 + \Delta) - \omega(t)} & \tau_1 + \Delta \leq t.
 \end{aligned} \tag{2.E.2}$$

Again the initial consumption  $c_0$  can be obtained from the intertemporal budget equation which in this case can be written as

$$\int_0^{\tau_1} \frac{y_t - c_t}{1 + \epsilon} e^{-\rho(t)} dt + \int_{\tau_1}^{\tau_1 + \Delta} (y_t - c_t) e^{-\rho(t)} dt + \int_{\tau_1 + \Delta}^{\infty} \frac{y_t - c_t}{1 - \epsilon} e^{-\rho(t)} dt = 0 \tag{2.E.3}$$

where

$$\begin{aligned}
 \hat{\rho}(t) &= R_L + \lambda \text{ for } t < \tau_1 \\
 \hat{\rho}(t) &= R + \lambda \text{ for } \tau_1 \leq t < \tau_1 + \Delta \\
 \hat{\rho}(t) &= R_B + \lambda \text{ for } \tau_1 + \Delta \leq t
 \end{aligned}$$

and

$$\rho(t) = \int_0^t \hat{\rho}(s) ds.$$

This intertemporal budget equation is derived from the equation describing the evolution of  $w_t$ , using our results on the investment policy of the consumer. The analysis is very similar to the case  $0 < k < 1$  and is omitted.

The parameter  $\Delta$  condition is derived by the portfolio decision of a consumer of age  $\tau_1$ . This consumer is indifferent between investing in the illiquid asset and not doing so. Given that he starts selling the illiquid asset at  $\tau_1 + \Delta$ , the change in his utility if he buys one unit of the illiquid asset at  $\tau_1$  is also given by equation 2.4.5 and

is equal to zero. From equations 2.E.1<sup>24</sup> and 2.E.2 we get that the following relation between consumption at age  $\tau_1$  and consumption at age  $\tau_1 + \Delta$

$$\frac{c_{\tau_1+\Delta} - y_{\tau_1+\Delta}}{R} = w_{\tau_1+\Delta} = w_{\tau_1} e^{\lambda\Delta} = \frac{c_{\tau_1} - y_{\tau_1}}{R} e^{\lambda\Delta}. \quad (2.E.4)$$

Equations 2.4.5 and 2.E.4 yield the following relation which shows that the minimum period of holding the illiquid asset is of order  $\epsilon$ :

$$\Delta = \frac{\psi^*}{(\lambda + \omega^*)(\delta - \omega^*)} \frac{2\epsilon}{A} + o(\epsilon) \quad (2.E.5)$$

Having characterized the solution to the consumer's problem we turn to the equilibrium determination of  $R$ .

In equilibrium, asset demand

$$\int_0^{\infty} \lambda e^{-\lambda t} w_t dt$$

equals asset market value,  $D/R$ .

As before, we will consider small transactions costs (small values of  $\epsilon$ ), and find their first order effects on  $R$ . We will thus write:

$$R(\epsilon) = r^* + b\epsilon + o(\epsilon) \quad (2.E.6)$$

and calculate  $b$ . Proposition 2.E.2 gives us  $b$ .

**Proposition 2.E.2** *In equilibrium,  $R$  is uniquely determined. It has the form of equation 2.E.6, with  $b$  having an ambiguous sign.*

The proof of proposition 2.E.2 as well as the analytic expression for  $b$  are again omitted.

The discussion on the determination of  $b$  is similar to the discussion offered at the end of subsection 2.4.2.2, (the effects are similar) and is omitted.

---

<sup>24</sup>Note that  $I_t \equiv 0$  between  $\tau_1$  and  $\tau_1 + \Delta$ .

## References

- Aiyagari, S. Rao and Mark Gertler (1991), "*Asset Returns with Transactions Costs and Uninsured Individual Risk: A Stage III Exercise*", *Journal of Monetary Economics*, 27:309-331.
- Allen, Franklin and Douglas Gale (1988), "*Optimal Security Design*", *Review of Financial Studies*, 1,3:229-263.
- Amihud, Yakov and Haim Mendelson (1986), "*Asset Pricing and the Bid-Ask Spread*", *Journal of Financial Economics*, 17,2:223-249.
- Amihud, Yakov and Haim Mendelson (1990), "*The Effects of a Transaction Tax on Securities Returns and Values*", mimeo.
- Amihud, Yakov and Haim Mendelson (1991a), "*Liquidity, Maturity and the Yields on U.S. Government Securities*", *Journal of Finance*, 46,4:1411-1425.
- Amihud, Yakov and Haim Mendelson (1991b), "*Liquidity, Asset Prices and Financial Policy*", *Financial Analysts Journal*, November-December:56-66.
- Amihud, Yacov, Haim Mendelson and Gang Yu (1992), "*Income Uncertainty and Transaction Costs*", New York University, mimeo.
- Blanchard, Olivier (1985), "*Debt, Deficits and Finite Horizons*", *Journal of Political Economy*, 93,2:223-247.
- Blanchard, Olivier and Stanley Fisher, "*Lectures on Macroeconomics*", MIT Press, 1989.
- Boudoukh, Jacob and Robert Whitelaw (1991), "*The Benchmark Effect in the Japanese Government Bond Market*", *Journal of Fixed Income*, September:52-59.
- Boudoukh, Jacob and Robert Whitelaw (1992), "*Liquidity as a Choice Variable: A Lesson from the Japanese Government Bond Market*", forthcoming, *Review of Financial Studies*.

- Brennan, Michael (1975), "*The Optimal Number of Securities in a Risky Asset Portfolio when there are Fixed Costs of Transacting: Theory and Some Empirical Results*", *Journal of Financial and Quantitative Analysis*, September:483-496.
- Constantinides, Georges (1986) "*Capital Market Equilibrium with Transaction Costs.*", *Journal of Political Economy*, 94,4,December:842-862.
- Davis, Mark and Andrew Norman (1990) "*Portfolio Selection with Transaction Costs*", *Mathematics of Operations Research*.
- Duffie, Darrell and T. Sun (1989), "*Transactions Costs and Portfolio Choice in a Discrete-Continuous Time Setting*", mimeo, Stanford University.
- Duffie, Darrell and Matthew Jackson (1989), "*Optimal Innovation of Futures Contracts*", *Review of Financial Studies*, 2:275-296.
- Dumas, Bernard and Elisa Luciano (1991), "*An Exact Solution to the Portfolio Choice Problem Under Transactions Costs*", *Journal of Finance*, 46,2,June:577-595.
- Fleming, Wendell, Sanford Grossman, Jean-Luc Vila and Thaleia Zariphopoulou (1992), "*Optimal Portfolio Rebalancing with Transaction Costs*", mimeo, Massachusetts Institute of Technology.
- Fremault, Anne (1991), "*Stock and Stock Index Futures Trading in an Equilibrium Model with Entry Costs*", mimeo, Boston University.
- Goldsmith, David (1976), "*Transactions Costs and the Theory of Portfolio Selection*", *Journal of Finance*, 31:1127-1139.
- Grossman, Sanford and Jean-Luc Vila (1992), "*Optimal Investment Strategies with Leverage Constraints*", *Journal of Financial and Quantitative Analysis*, 27,2, June:151-168.
- Heaton, John and D. Lucas (1992), "*Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing*", mimeo, Massachusetts Institute of Technology.

- He, Hua and H. Pages (1991) "*Consumption/Investment Decisions with Labor Income: A Duality Approach*", mimeo.
- Levy, Haim (1978), "*Equilibrium in an Imperfect Market: A Constraint on the Number of Securities in the Portfolio*", *American Economic Review*, September:643-658.
- Mayshar, Joram (1979), "*Transactions Costs in a Model of Capital Market Equilibrium*", *Journal of Political Economy*, 87:673-700.
- Mayshar, Joram (1981), "*Transactions Costs and the Pricing of Assets*", *Journal of Finance*, 36:583-597.
- Mehra, R. and Edward Prescott (1985), "*The Equity Premium: A Puzzle*", *Journal of Monetary Economics*, 15:145-162.
- Michaely, Roni and Jean-Luc Vila (1992), "*Tax Heterogeneity, Transactions Costs and Trading Volume: Evidence from the Ex-Dividend Day*", mimeo, Massachusetts Institute of Technology.
- Ohashi, Kazuhiko (1992,) "*Efficient Futures Innovation with Small Transaction Fee*", mimeo, Massachusetts Institute of Technology.
- Pagano, Marco (1989), "*Endogenous Market Thinness and Stock Price Volatility*", *Review of Economic Studies*, 56:269-288.
- Silber, William (1991), "*Discounts on Restricted Stock: The Impact of Illiquidity on Stock Prices*", *Financial Analysts Journal*, July-August:60-64.
- Tuckman, Bruce and Jean-Luc Vila (1992), "*Arbitrage with Holding Costs: A Utility-Based Approach*", *Journal of Finance*, 47,4:1283-1302.
- Vila, Jean-Luc and Thaleia Zariphopoulou (1990), "*Optimal Consumption and Investment with Borrowing Constraints*", mimeo, Massachusetts Institute of Technology.

Wang, Jiang (1992), "*A Model of Competitive Stock Trading Volume*", mimeo, Massachusetts Institute of Technology.

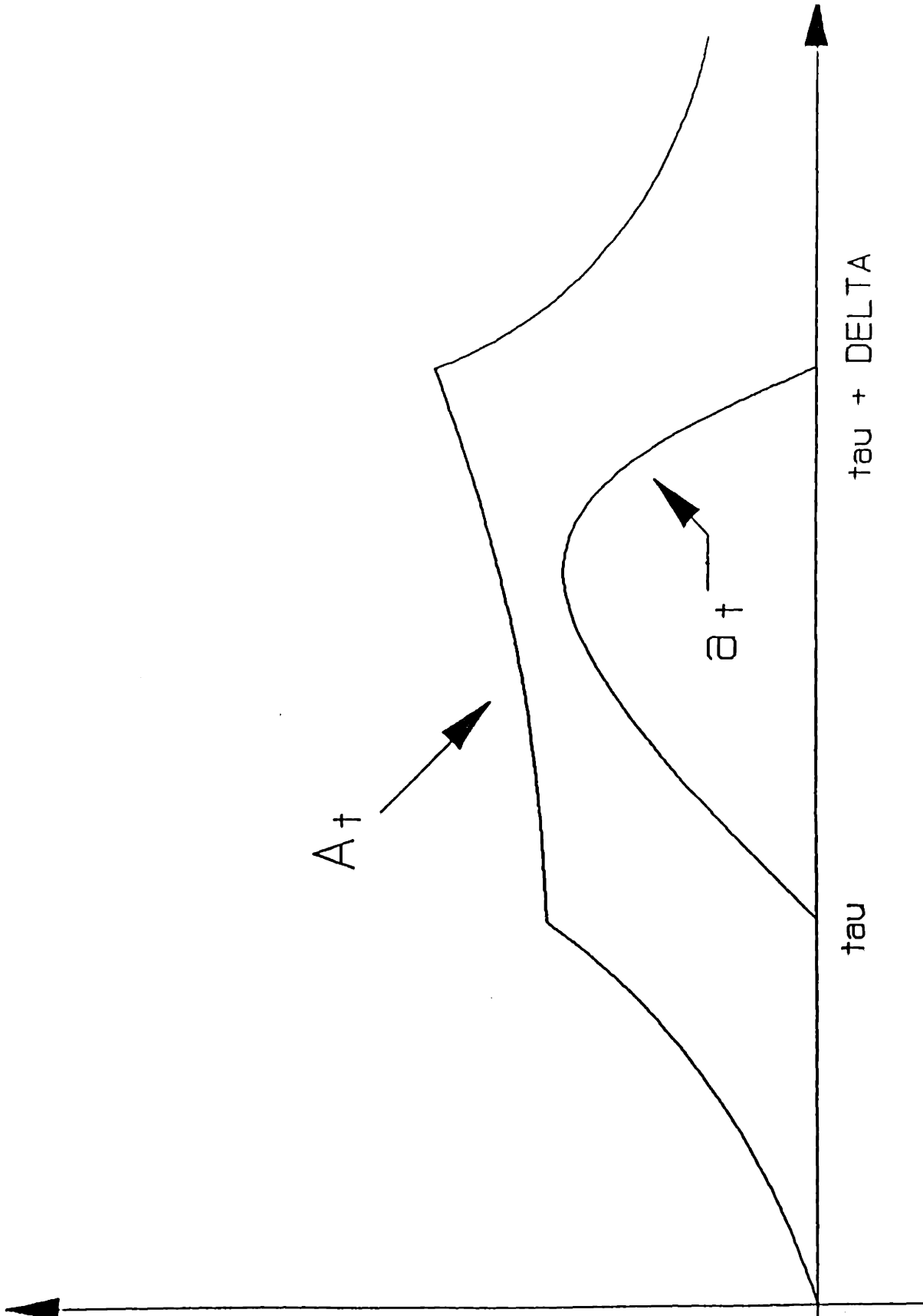


Figure 1: Holdings of the liquid and illiquid assets



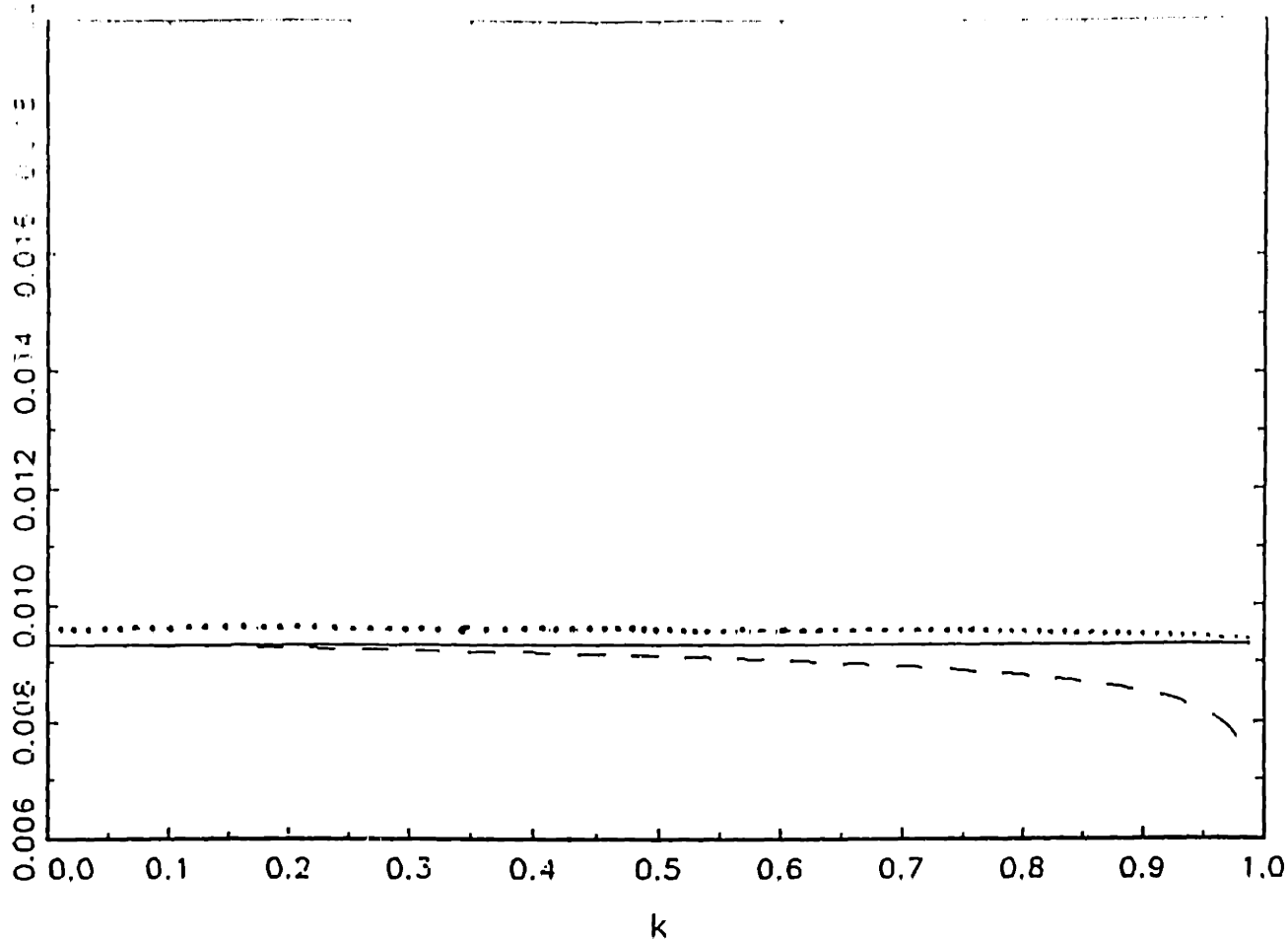
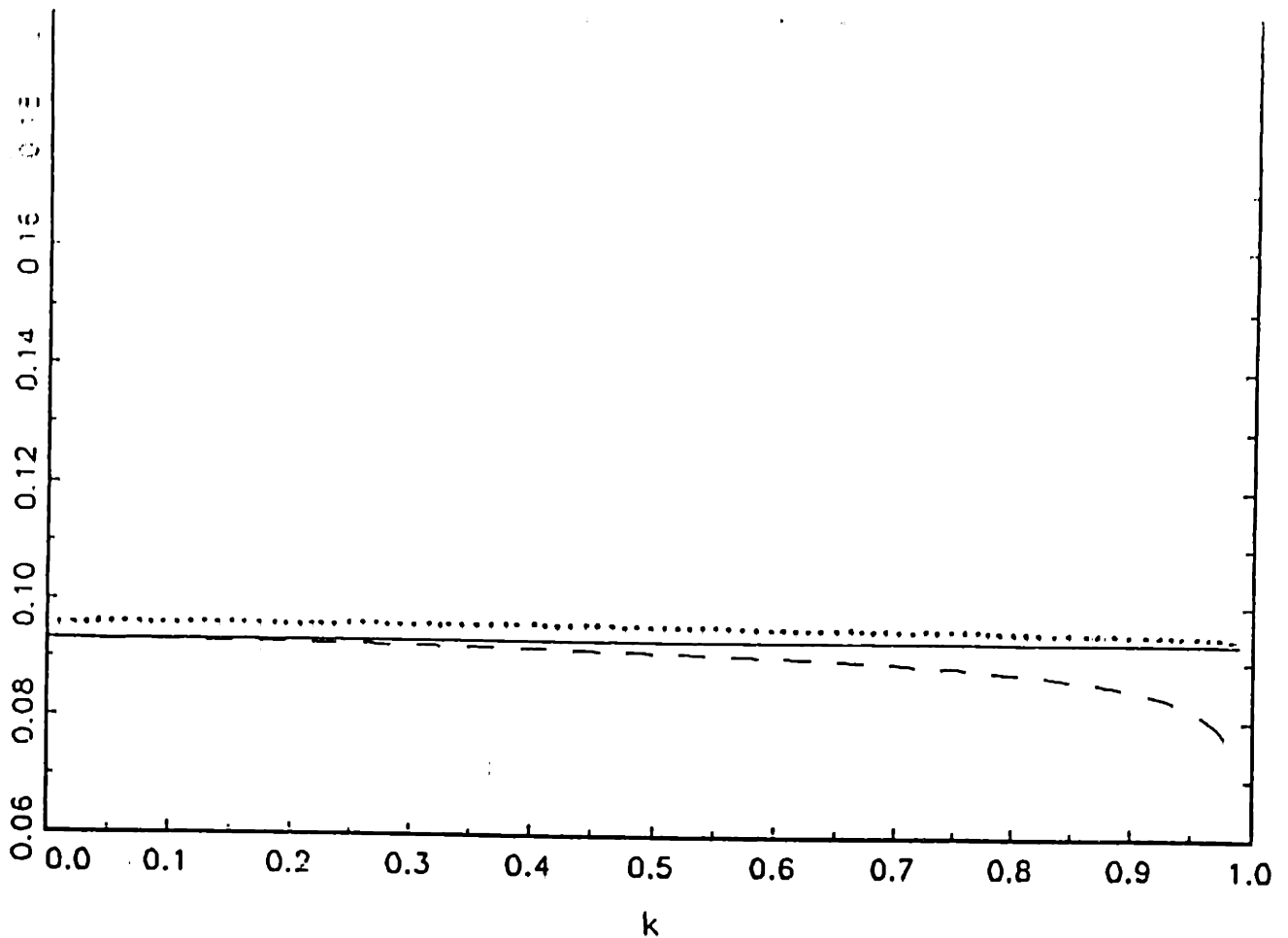


Figure 2: Rates of return as functions of  $k$

This figure plots the rates of return as a function of the fraction,  $k$ , of illiquid assets. The solid line represents the benchmark case where there are no transactions costs. The dotted line represents the rate of return on the illiquid asset while the dashed line represents the rate of return on the liquid asset. For this figure we have used the following parameter values:

$$\lambda = 2\%; \delta = 4\%; \beta = 0.2\%; A = 1; D/Y = 50\%; \epsilon = 3\%.$$



**Figure 3: Rates of return as functions of  $k$**

This figure plots the rates of return as a function of the fraction,  $k$ , of illiquid assets. The solid line represents the benchmark case where there are no transactions costs. The dotted line represents the rate of return on the illiquid asset while the dashed line represents the rate of return on the liquid asset. For this figure we have used the following parameter values:

$$\lambda = 20\%; \delta = 40\%; \beta = 2\%; A = 1; D/Y = 50\%; \epsilon = 3\%.$$

Compared to the previous case, the agent is more impatient and has therefore a shorter horizon. As a result, the interest rate and the liquidity premium are higher than in the previous figure. Qualitative results are however unchanged.

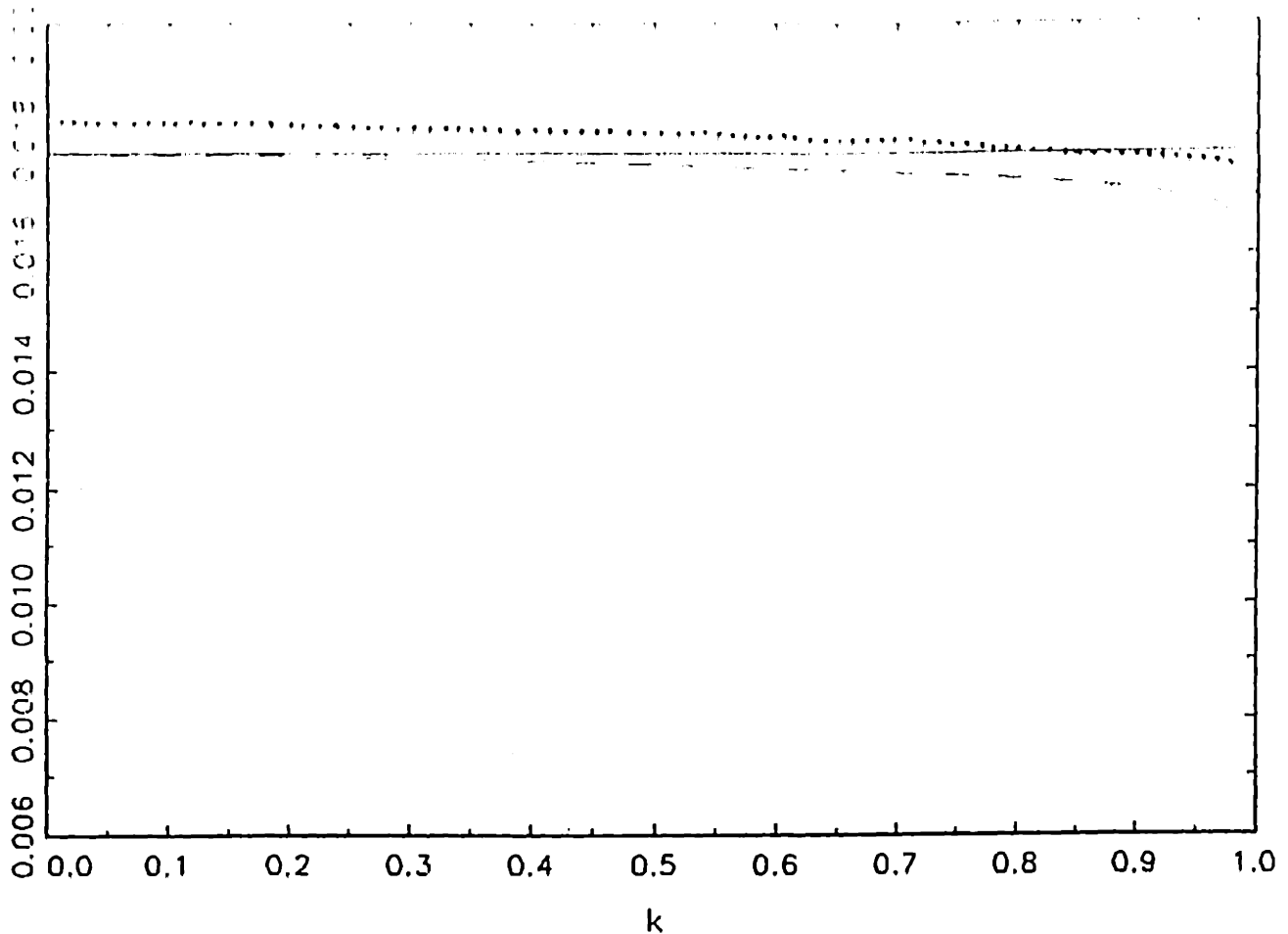


Figure 4: Rates of return as functions of  $k$

This figure plots the rates of return as a function of the fraction,  $k$ , of illiquid assets. The solid line represents the benchmark case where there are no transactions costs. The dotted line represents the rate of return on the illiquid asset while the dashed line represents the rate of return on the liquid asset. For this figure we have used the following parameter values:

$$\lambda = 2\%; \delta = 4\%; \beta = 0.2\%; A = 1; D/Y = 300\%; \epsilon = 3\%.$$

In this case (where financial income is much more important than labor income) the following paradoxical phenomenon occurs: transactions costs lower the rates of return on *both* assets.

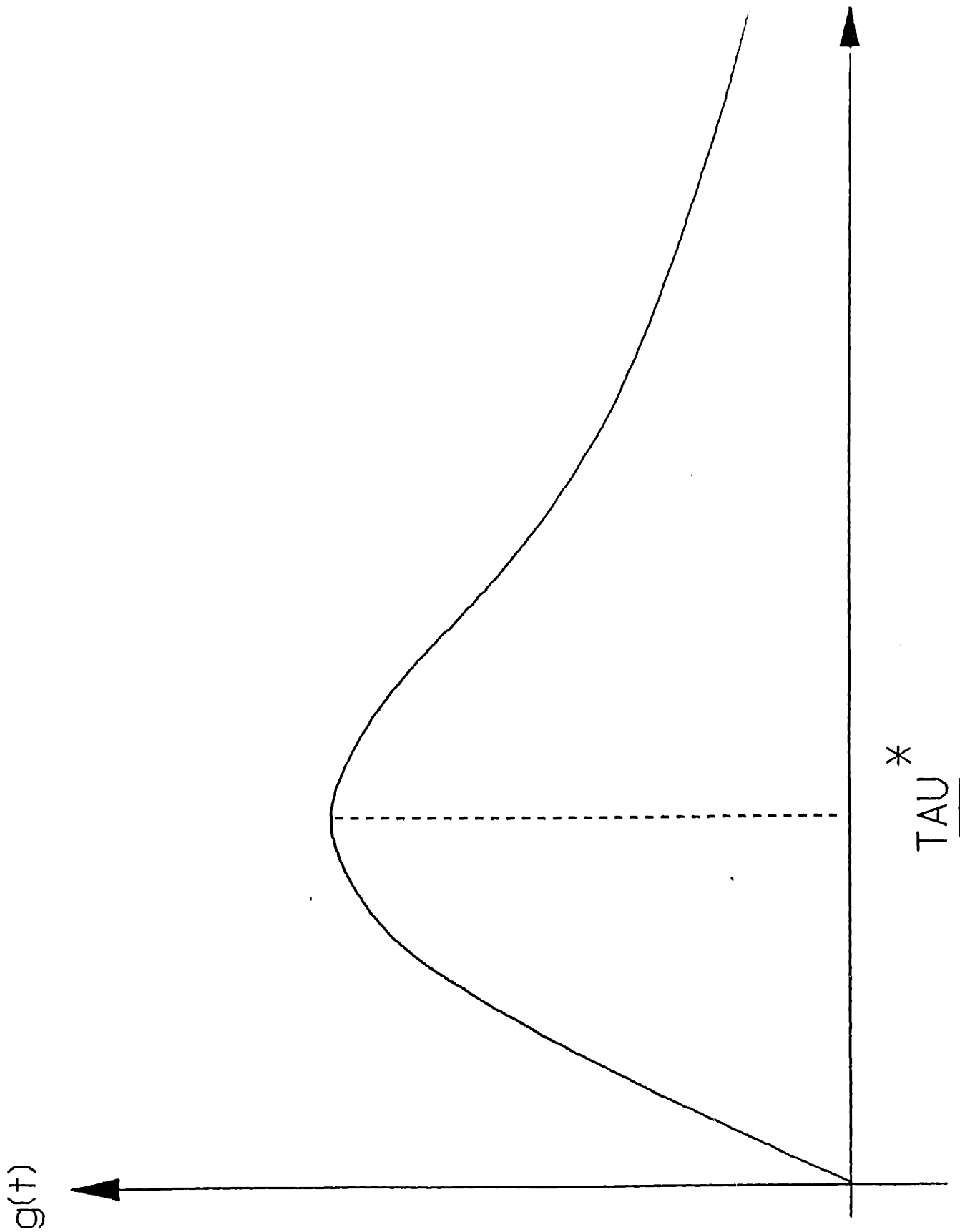


Figure 5: The graph of  $g(t)$

# Chapter 3

## Decentralization and the Management of Competition

### 3.1 Introduction

Almost any large firm is a multiproduct firm. Some of its products are, in addition, horizontally or vertically differentiated i.e. are targeted to different customers but are at the same time close substitutes for some of these customers. The design and pricing of these products require thus considerable coordination among the firm's units.

Lack of coordination in product design was a major concern of top management at GM, before the radical reform of the company in the early 20's.<sup>1</sup> According to Sloan (1963), GM was extremely decentralized and was offering, as a result, too many competing products.

*In the middle [segment of the market], where we were concentrated with duplication, we did not know what we were trying to do, except to sell cars which, in a sense, took volume from each other.<sup>2 3</sup>*

---

<sup>1</sup>See Sloan (1963) chapter 4.

<sup>2</sup>Sloan (1963) p.60

<sup>3</sup>Another example of a firm centralizing decision-making to reduce competition among units is described in Smith (1992) (see section 2 below). A major commercial bank in California centralized

At the same time large firms may seek to promote competition among their units. For instance, in the last decade IBM drastically decentralized its internal structure allowing units to bargain on transfer prices and to trade with outside partners.<sup>4</sup> In addition, under decentralization:

*[Units] are free to compete with each other for business, even if this means cannibalizing the sales of other IBM units.*<sup>5</sup>

As the above examples suggest, the existence of units offering substitute products within a firm raises a number of questions for organizational design. When should top management control the design and pricing of these products, thus centralizing decision-making, and when is it better to allow the units decide which products to make and how to sell them? In the case where it is optimal to let the units decide (and eventually compete) what prevents top management from intervening and restricting competition when the latter threatens to dissipate too many of the firm's profits?

In this chapter we develop a simple model to address the above questions. We consider a firm which is composed by top management and two divisions. Divisions design horizontally differentiated products, characterized by their "location" and their "quality", and sell them to customers. For simplicity, we focus on the allocation of decision rights at the price setting stage. More precisely, top management can restrict the set of customers with whom a division can deal, thus prohibiting competition for customers, or can alternatively allow both divisions to approach all customers.

We argue that when a division is given the right to deal with all customers, it offers a higher quality product. Indeed, by increasing the quality of its product it not only can charge its own customers higher prices, but it also becomes a better competitor of the other division. Using competition as a threat, it can extract concessions from the other division, such as dealing with some of its customers.<sup>6</sup> However, higher

---

its operations, removing branch managers' autonomy to sell loans to customers. Branch managers had, instead, to refer customers to a center that coordinated trade.

<sup>4</sup>See Ferguson (1993) for a detailed description of the evolution of IBM.

<sup>5</sup>Ferguson p.61 and 62.

<sup>6</sup>In the model we assume, for simplicity, that divisions can make side-transfers so that concessions can simply be these transfers.

quality comes at a cost. A division aiming also at attracting the customers of the other division, will opportunistically locate its product too close to the product of the other division.

Top management will allow divisions to compete, i.e. will *decentralize* decision making, when the benefits due to higher product quality outweigh the costs of inefficient locations. We argue that top management's information about customers' characteristics (i.e. their locations) crucially determines the credibility of its commitment to competition among divisions. If it knows customers' locations it can efficiently allocate them to divisions (since in equilibrium it also knows the characteristics of both products) without letting the divisions engage in bargaining and eventually compete.<sup>7</sup>

We show that centralization of decision-making may or may not be optimal. If divisional managers are very risk-averse, centralization dominates decentralization since the benefit of increased incentives is small compared to the cost of inefficient locations. Interestingly, the same conclusion holds when divisional managers' risk-aversion is low. Indeed, high-powered incentive schemes for improving product quality can be provided at a small cost under centralization. Decentralization thus brings small additional benefits and inefficient locations. Decentralization is a better solution if the set of customers who desire both products is small so that divisional incentives to locate products closer are small and locations are close to their first-best values. Costs due to inefficient locations are in the second order, while product quality increases in the first-order.

As we argued above, when a division is given the right to deal with all customers it will offer a higher quality product. Indeed, by becoming a better competitor of the other division, it can extract some concessions from it. Taking surplus away from the other division does not have any social value and if divisional managers receive very high-powered incentive schemes so that they become residual claimants, they will overprovide quality. The idea that competing agents may overinvest in some

---

<sup>7</sup>In our model bargaining is costless, for simplicity. However, top management prefers to intervene even for a very small cost of bargaining (which can easily be introduced in the model).

activities is familiar in industrial organization; firms may offer too many products because they are only “stealing” each other’s business and may spend too much on R&D (Tirole (1988) Chs. 7 and 10). In our model competition is valuable because agents’ incentives are initially *too low*.

The channel through which competition affects incentives in our work is different than the one studied in the principal-agent literature, namely that competition improves the information available for contracting. Papers in that literature have shown that the performance of an agent can better be assessed if performance measures of competing agents are available and that, as a result, this agent can be given a higher-powered incentive scheme.<sup>8</sup> Even if the principal does not have access to performance measures of competing agents, competition may still improve the information available for contracting by making the agent’s profits less variable (Hart (1983), Scharfstein (1987)).

Rey and Tirole (1986) also address the question of whether a principal (in their model a manufacturer) prefers to restrict competition among agents (retailers). A monopoly retailer will charge a price which is too high from the manufacturer’s viewpoint, as long as he does not buy from the manufacturer at marginal cost. Competition allows the manufacturer to resolve this moral-hazard problem (since retailers will charge the price at which they are buying from the manufacturer) and at the same time provides insurance to retailers making their profits less variable, as in Hart (1983). When retailers compete they will thus charge lower prices (keeping the price at which they are buying from the manufacturer constant) and will on average supply the same quality. By contrast in our work, when divisions are allowed to compete, they offer higher quality products and charge higher prices as a result. A second difference with Rey and Tirole (1986) is that the cost of competition in their model is the inability of retailers to price-discriminate, while in our model is the choice of inefficient locations by the divisions.

Our work is also related to a recent paper by Holmstrom and Tirole (1991) study-

---

<sup>8</sup>For summaries of the literature on relative performance evaluation see, for instance, Holmstrom and Tirole (1989) and Tirole (1988), Ch. “The Theory of the Firm”.



ing transfer pricing within firms. They compare a regime of “exchange autonomy” where top management allows divisions to trade with outside partners to a regime where divisions can only trade internally (but can refuse to trade).<sup>9</sup> They show that if divisions are allowed to trade with outside partners, they will offer higher quality products but may choose a product design which is general purpose and not specific to the needs of the other division. Holmstrom and Tirole’s (1991) work and ours, although considering different issues, share the feature that if divisions have more decision rights, they will have more incentives but may also take decisions which are not optimal from the firm’s viewpoint.

The remainder of the chapter is organized as follows: In section 3.2 we describe the model and in section 3.3 we derive the first-best. In sections 3.4 and 3.5 we study the outcomes under decentralization and centralization. Section 3.6 compares the two organizational forms. All proofs are in the appendix.

## 3.2 The Model

### 3.2.1 Supply and Demand

We consider a firm which is composed by top management and two divisions. Each division makes one product. The potential buyers of these products (the “customers”) are uniformly distributed in the interval  $[0, 1]$ . The product of division  $i$  ( $i = 1, 2$ ) (referred to as product  $i$ ) is characterized by its location,  $l_i$ , and its “quality”,  $v_i$ . We assume that “transportation” costs are linear, so that the valuation of a customer located at  $x$  for product  $i$  is

$$v_i(x) = v_i - t|x - l_i|. \quad (3.2.1)$$

We also assume that the location of each customer is known to the divisions and that the divisions can charge customer-specific prices.<sup>10</sup>

---

<sup>9</sup>They also study the case where one division can order the other division to trade.

<sup>10</sup>We plan to study the case where customers’ locations are not known and divisions cannot price-discriminate in future work.

Product 1 can be located anywhere in  $[0, 1/2]$  and product 2 can be located anywhere in  $[1/2, 1]$ . The manager of division  $i$  can improve the quality of product  $i$  by exerting effort  $e_i$  at a non-monetary cost  $C(e_i)$ . We will assume that

$$v_i = v_0 + e_i, \quad v_0 > 0, \quad (3.2.2)$$

and

$$C(e_i) = \frac{1}{2} C e_i^2. \quad (3.2.3)$$

The marginal cost of each of the two products is zero.

For simplicity, we will restrict the parameter space assuming that only one product cannot cover the whole market even if the manager of the corresponding division exerts first best effort. Since the optimal effort level when the whole market is covered by one product is  $\frac{1}{C}$  and the optimal location in order to cover the market is  $\frac{1}{2}$ , our assumption (“assumption A”) is equivalent to

$$\frac{v_0}{t} + \frac{1}{Ct} \leq \frac{1}{2}.$$

In words, product quality when effort is zero,  $v_0$ , has to be small compared to transportation costs,  $t$ , and the cost of exerting effort and improving quality,  $C$ , has to be large.

### 3.2.2 Information Structure and Contracts

We assume that a divisional manager’s effort as well as the actions that he undertakes to determine the location of his product are observed by the other divisional manager. However, neither of these actions is verifiable, i.e. divisional managers cannot write short-term contracts with each other that bind them to exert a given effort level or to choose a particular location for their product. This will naturally be the case if, as we will assume, top management does not perfectly understand product design (quality and location) and cannot enforce such contracts. The non-verifiability of actions

will imply in particular that divisions may not be able to coordinate on an efficient choice of product locations. Finally, divisions are assumed to know the location of all customers.

Top management is assumed to be less informed about both divisions' environments than each divisional manager. In particular by not perfectly understanding product design, top management does not observe effort levels (product quality) and product location. We however assume that it knows the location of customers. Since in equilibrium it also knows product quality and product location, it can infer which product is better for a given customer. We briefly consider the case where top management does not know customers' locations in section 3.6.

Divisional profits are partly generated by the revenues from dealing with customers. We assume that they contain an additional component which is random, normally distributed and independent across divisions,<sup>11</sup> and are thus a noisy signal of the actions taken by divisional managers. Profits are the only variables in our model which are verifiable by outside parties (such as courts).

We assume that divisional managers are given incentive schemes which depend only the profits of their own division (and not on the profits of the other division) (This assumption is more innocuous than it may seem. See the analysis of the centralized firm in Section 3.5.) are linear and have the same slope which is between 0 and 1. Denoting by  $\pi_i$  the profits of division  $i$ , the manager of division  $i$  thus receives

$$A\pi_i + B_i. \tag{3.2.4}$$

We assume that both managers are risk-averse and have a negative exponential VNM utility function with coefficient of absolute risk-aversion equal to  $\alpha$ . Their reservation levels are normalized to zero.

We consider two organizational forms, the decentralized firm and the centralized firm. Top management allows divisions to compete for a given customer in the decentralized firm while it allocates customers to divisions in the centralized firm. We

---

<sup>11</sup>It can be due, for example to other divisional operations.

assume that under decentralization divisions bargain efficiently so that competition only defines the status quo in the bargaining game, and that bargaining power is equal. Since the incentive schemes given to divisional managers have the same slope, customers are allocated to divisions efficiently from the firm's viewpoint.

The timing is thus

**Decentralized Firm**

Incentive contracts signed.	Divisional managers exert effort, $e$ , choose location, $l$ .	Bargaining under threat of competition.	Product sales. Profits realized.
-----------------------------	--	---	----------------------------------

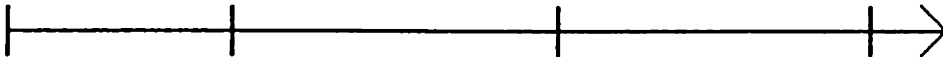


Figure 1: Timing under Decentralization

**Centralized Firm**

Incentive contracts signed.	Divisional managers exert effort, $e$ , choose location, $l$ .	Top management assigns customers to divisions.	Product sales. Profits realized.
-----------------------------	--	--	----------------------------------

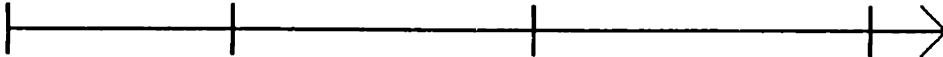


Figure 2: Timing under Centralization

To illustrate our model, we consider the example of commercial banks in a large city. There is demand for bank loans but different individuals are interested in

loans with different characteristics (business are interested in short-term lending, entrepreneurs in funding a start-up venture, consumers need cash for acquiring durable goods).<sup>12</sup>

Managers sell loans to customers and can undertake two kinds of investments in order to become more efficient:

- **“Supply” Effort** They can acquire a skill that is not specific to any kind of loan, but rather improves the chances of striking deals in good terms (like learning state of the art software to gather information in a timely fashion, or information technology that improves readiness of on-line help to customers, whatever the kind of loan they want to purchase). (The supply effort corresponds to the investment  $e$  above.)
- **“Demand” Effort** They can become specialized in being particularly efficient in customizing loans to a particular segment of the market (by learning detailed information on a particular sector of the population or following more closely the trends of an industry. (This corresponds to the location choice,  $l$ , above.)

The bank has two divisions in the city, that deal directly with customers. Each division is run by one manager. Each manager has expertise in the business and is able to assess the investment levels of the other manager. Top management is however unable to measure those investment levels.<sup>13</sup>

Top management can centralize or decentralize the firm’s operations. gathering more or less precise information about customers.

Under decentralization, divisional managers are not required to obtain approval from top management to strike deals with customers. Under centralization divisional managers transmit information on a potential deal with a customer to top management, who either approves the deal or allocates the customer to the other division.<sup>14</sup>

---

<sup>12</sup>Heterogeneity is *large* and it is therefore difficult to segment the market in non-arbitrary intervals of potential customers.

<sup>13</sup>A manager can infer the ability (investment) of the other manager because customers transmit information from one office to another.

<sup>14</sup>The center is unable to separate the market geographically if, for instance, business have multiple locations.

Smith (1992) provides a case study of the reorganization of several divisions in a major investment bank in California. She examines in particular<sup>15</sup> the transition from relative autonomy to a regime of centralization in a division of the bank. The manager of the division (referred to as L.) is being deprived of authority over some actions that she previously controlled. The control is being recovered by the next superior level in the hierarchy, the area management group (referred to as AMG.)

*Branch top management centralized lending into specialized divisions, reflecting the bank's commitment to targeting and profiting from stratified market segments. The differentiation of the lending function into specialized units removed a significant source of L.'s authority. She no longer had authority to make loans, nor did she manage loan personnel. [...] L. forwarded [loan application] to the appropriate consumer, real state or commercial loan center.[...] Her role was [now] strictly one of referral.<sup>16</sup>*

*Even when L. did tabulate statistical information, the role she once played in evaluating and acting on that information had changed. Whereas formerly L. would have tracked branch information and used it to develop an integrated picture of branch performance, the job of tracking various kinds of information had been transferred into specialized sections in the AMG.<sup>17</sup>*

*The new PPCE and the extensive documentation it contained provided the area manager with the information required for indirect management: in other words, it gave area managers more control over individual branches.*

---

<sup>15</sup>See Smith (1992), pp.95-101.

<sup>16</sup>In addition, it becomes clear that the transfer of authority required a redefinition of the information flowing from lower level (L.) to upper level (the AMG). The information flow is defined in a document of the corporation, the Performance Evaluation Plan (PPCE in the text).

<sup>17</sup>Smith (1992) p.97

### 3.3 The First Best

In this section we derive the effort levels and product locations that maximize total welfare. Total welfare (which is equal to top management's payoff since the reservation levels of the divisional managers are normalized to zero) is

$$\begin{aligned} & \int_0^1 \max(v_1(x), v_2(x), 0) dx - \frac{1}{2}C(e_1^2 + e_2^2) \\ &= \int_0^1 \max(v_1 - t|x - l_1|, v_2 - t|x - l_2|, 0) dx - \frac{1}{2}C(e_1^2 + e_2^2). \end{aligned} \quad (3.3.1)$$

Proposition 3.3.1, proven in Appendix 3.A, gives us the first best effort levels,  $e_1^*$  and  $e_2^*$ , and locations  $l_1^*$  and  $l_2^*$ .

**Proposition 3.3.1** *The maximum of expression 3.3.1 is achieved by choosing*

$$l_1^* = \frac{1}{4} \qquad l_2^* = \frac{3}{4}.$$

*If  $2v_0/t + 1/Ct < 1/2$ ,  $e_1^*$  and  $e_2^*$  are given by*

$$e_1^* = e_2^* = v_0 / \left( \frac{Ct}{2} - 1 \right)$$

*and the firm does not sell to all customers (i.e. the market is not "covered").*

*If  $2v_0/t + 1/Ct \geq 1/2$ ,  $e_1^*$  and  $e_2^*$  are given by*

$$e_1^* = e_2^* = \frac{1}{2C}.$$

*and the market is covered.*

Assumption A makes the problem sufficiently concave so that it is not optimal to require different effort levels (and product locations) from the two managers.

If the market is not covered (i.e. if product quality is low compared to trans-

portation costs and the cost of improving it is large), total welfare is maximum for any locations which imply non-overlapping sets of customers willing to buy a given product and which are not too close to the extremes, 0 and 1. (So that each product is located in the middle of its market segment - the set of customers who are buying the product.)  $1/4$  and  $3/4$  always belong to the set of optimal locations. If the market is covered,  $1/4$  and  $3/4$  are the unique optimal locations and each product is located in the middle of its market segment. ( $[0, 1/2]$  for product 1 and  $[1/2, 1]$  for product 2.) In addition, since more customers buy the products if the market is covered, the optimal effort levels are higher than if the market is not covered.

### 3.4 The Decentralized Firm

We now come to the analysis of the decentralized firm. In section 3.2 we assumed that top management allows divisions to compete for a given customer. We then assumed that divisions bargain efficiently so that competition only defines the status quo of the bargaining game, and that bargaining power is equal.

Considering a customer located at  $x$  and assuming, for example, that  $v_1(x) > v_2(x) \geq 0$ , it is easy to see that division 1 deals with the customer and receives  $v_1(x) - v_2(x)/2$  while division 2 receives  $v_2(x)/2$ .<sup>18</sup>

Aggregating over customers, division 1's profits are given by

$$\begin{aligned} & \int_{S_1} \left( \max(v_1(x), 0) - \frac{1}{2} \max(v_2(x), 0) \right) dx + \int_{[0,1] \setminus S_1} \frac{1}{2} \max(v_1(x), 0) dx = \\ & \int_{S_1} \left( \max(v_1 - t|x - l_1|, 0) - \frac{1}{2} \max(v_2 - t|x - l_2|, 0) \right) dx \\ & \quad + \int_{[0,1] \setminus S_1} \frac{1}{2} \max(v_1 - t|x - l_1|, 0) dx \end{aligned} \quad (3.4.1)$$

where the set  $S_1$  is defined by

---

<sup>18</sup>If there is competition, division 1 serves the customer charging the price  $v_1(x) - v_2(x)$ . Its payoff is thus  $v_1(x) - v_2(x)$  while the payoff of division 2 is 0. Since bargaining power is equal, i.e. each divisional manager makes a take-it-or-leave-it offer to the other divisional manager with probability  $1/2$ , it is clear that the payoffs of the two divisions are given by the previous expressions.



$$S_1 = \{x \in [0, 1] : v_1(x) > v_2(x)\} = \{x \in [0, 1] : v_1 - t|x - l_1| > v_2 - t|x - l_2|\} \quad (3.4.2)$$

and division 2's profits are given by the symmetric expression.

Using equation 3.2.4 which gives managers' incentive schemes, we get for the payoff of the manager of division 1

$$A \left( \int_{S_1} \left( \max(v_1 - t|x - l_1|, 0) - \frac{1}{2} \max(v_2 - t|x - l_2|, 0) \right) dx + \int_{[0,1] \setminus S_1} \frac{1}{2} \max(v_1 - t|x - l_1|, 0) dx \right) + B_1 - \frac{1}{2} C e_1^2. \quad (3.4.3)$$

To determine the outcome under decentralization we proceed in two steps. In Proposition 3.4.1, proven in appendix 3.B, we characterize pure-strategy equilibria of the effort and location game taking incentive schemes as given. We then determine the optimal value of  $A$ , the slope of the incentive schemes.

**Proposition 3.4.1** *There exists a pure-strategy equilibrium in which locations and effort levels are given by*

**case I**  $2v_0/t + A/Ct < 1/2$

$$l_1 = \frac{1}{4} \qquad l_2 = \frac{3}{4},$$

$$e_1 = e_2 = v_0 / \left( \frac{Ct}{2A} - 1 \right)$$

*and the market is not covered. In any other pure-strategy equilibrium effort levels and payoffs are the same.*

**case II**  $2v_0/t + A/Ct \geq 1/2$

$$l_1 = 1 - l_2 = \left( \frac{1}{4} + \frac{v_0}{2t} \right) / \left( \frac{3}{2} - \frac{A}{Ct} \right) \geq \frac{1}{4},$$

$$e_1 = e_2 = (2A/C)l_1 = \left(\frac{t}{2} + v_0\right) / \left(\frac{3Ct}{2A} - 1\right)$$

and the market is covered. This is the unique pure-strategy equilibrium.

In case I product quality,  $v_0$ , is low compared to transportation costs,  $t$ , the cost of improving it,  $C$ , is high and the benefit,  $A$ , that divisional managers get is small. The market is not fully covered and the divisions do not have incentives to offer products that appeal to overlapping sets of customers.

By contrast, in case II divisions offer products which appeal to overlapping sets of customers and whose distance is too small compared to the first best ( $l_1 \geq 1/4$ ,  $l_2 \leq 3/4$ ). The private gain of each division to locate its product closer to the product of the other division is larger than the social gain. Indeed, by locating its product closer, the division better “penetrates” the market of the other division and becomes a more serious potential competitor.<sup>19</sup> The incentive to locate closer and becoming a more serious potential competitor is clearly larger if there is a significant overlap already ( $v_0$  high,  $C$  small,  $A$  large).

Effort levels are higher in case II than in case I. Indeed, the manager of one division by exerting more effort supplies a product which is valued more by the customers of his division (who now have a larger mass) as well as by the customers of the other division (so that this product can better compete with the product of the other division).

It is straightforward to show that the top management’s payoff is given by

$$\int_0^1 \max(v_1 - t|x - l_1|, v_2 - t|x - l_2|, 0)dx - \frac{1}{2}C(e_1^2 + e_2^2) - A^2\alpha\sigma_u^2 \quad (3.4.4)$$

for the values of  $e_1$ ,  $e_2$ ,  $l_1$ ,  $l_2$  determined in proposition 3.4.1. Expression 3.4.4 is derived from expression 3.3.1 by subtracting the costs of having the divisional man-

---

<sup>19</sup>Of course, (price) competition does not take place in equilibrium.

agers bearing risk. The slope of the incentive scheme,  $A \in [0, 1]$ , is determined by the maximization of expression 3.4.4. Expression 3.4.4 is not generally concave in  $A$  since the responsiveness of effort to  $A$  increases with  $A$ .<sup>20</sup> Although the characterization of the optimal  $A$  is involved, it is easy to show (see appendix 3.B) that  $A$  decreases with managers' risk aversion,  $\alpha$ , and the variance of the noise,  $\sigma_u^2$ . Comparative statics w.r.t.  $v_0$ ,  $t$  and  $C$  are ambiguous.

### 3.5 The Centralized Firm

In this section we study the centralized firm in which the general office allocates customers to divisions. Each division charges its customers their valuations.

Since the sets

$$S_1 = \{x \in [0, 1] : v_1 - t|x - l_1| > v_2 - t|x - l_2|\}$$

$$S_2 = \{x \in [0, 1] : v_1 - t|x - l_1| = v_2 - t|x - l_2|\}$$

and

$$S_3 = \{x \in [0, 1] : v_1 - t|x - l_1| < v_2 - t|x - l_2|\}$$

are ordered, in the sense that  $\forall (x_1, x_2, x_3) \in S_1 \times S_2 \times S_3$   $x_1 < x_2 < x_3$ <sup>21</sup>, top management can achieve its maximum payoff by choosing a dividing point  $\bar{x} \in [0, 1]$  such that customers with locations  $< \bar{x}$  deal with division 1 and customers with locations  $\geq \bar{x}$  deal with division 2. For simplicity we will consider only these strategies for top management.

We will rule out a class of equilibria that exhibit an extreme lack of coordination. In these equilibria top management allocates all customers to one division, say division 1, "shutting" division 2. The manager of division 1 is then very motivated while the manager of division 2 does not exert any effort and may locate his product close

---

<sup>20</sup>If  $A$  is higher, increasing  $A$  will induce the managers to exert more extra effort since they will receive the benefit of selling a better product to a larger market.

<sup>21</sup>This is easy to check, using  $l_1 \leq l_2$ .

enough to product 1 so that all customers prefer product 1 to product 2.

In proposition 3.5.1, proven in appendix 3.C, we characterize the remaining pure-strategy equilibria.

**Proposition 3.5.1** *There exists a pure-strategy equilibrium in which locations and effort levels are given by*

**case I**  $2v_0/t + A/Ct < 1/2$

$$l_1 = \frac{1}{4} \qquad l_2 = \frac{3}{4},$$

$$e_1 = e_2 = v_0 / \left( \frac{Ct}{2A} - 1 \right)$$

*and the market is not covered. In any other pure-strategy equilibrium, effort levels and payoffs are the same.*

**case II**  $2v_0/t + A/Ct \geq 1/2$

$$l_1 = \frac{1}{4} \qquad l_2 = \frac{3}{4},$$

$$e_1 = e_2 = \frac{A}{2C}$$

*and the market is covered. This is the unique pure-strategy equilibrium.*

The outcome in case I is exactly the same as in proposition 3.4.1. In case II where divisional products appeal to overlapping sets of customers, divisions choose first best locations for their products. By being unable to deal with the customers of the other division their private gain to locate their product closer to the product to the other division coincides with the social gain. At the same time effort levels in case II are lower than in proposition 3.4.1. When deciding to exert more effort or not, the manager of a division considers only the increase in valuation of the customers of

his own division and not of the customers of the other division.

Top management's payoff is again given by expression 3.4.4 for the values of  $e_1$ ,  $e_2$ ,  $l_1$ ,  $l_2$  determined in proposition 3.5.1. In appendix 3.C we determine  $A$  and show, as before, that it decreases with  $\alpha$  and  $\sigma_u^2$ .

Before leaving this section we discuss our assumption that divisional managers' incentive schemes depend only on the profits of their own division and not on profits of the other division. Consider the extreme case where each manager's incentive scheme depended only on the *sum* of the profits of the two divisions. In this case, a manager would not attribute any value to his division becoming a more serious competitor of the other division and to profits being transferred from that division to his division. Locations and effort levels in both the decentralized and the centralized firm would then be given by proposition 3.5.1. Since managers would have to be compensated for bearing more risk, giving incentive schemes that depend on joint profits is dominated by centralizing and giving incentive schemes that depend on individual profits.

### 3.6 Comparison of Organizational Forms

In this section we compare the decentralized and the centralized firm. As the analysis of the previous sections showed, decentralization brings more incentives and higher quality products together with inefficient locations for these products. Due to the non-concavity of expression 3.4.4 w.r.t  $A$  in both the decentralized and the centralized case, the analysis of top management's welfare for general values of the parameters  $v_0$ ,  $t$ ,  $C$ ,  $\alpha$  and  $\sigma_u^2$  is somewhat involved. However, the following intuitive results hold

- If divisional managers are very risk-averse (or the variance of the noise is large) so that the slope of the incentive schemes is small, centralization improves welfare. Indeed, the benefit of increased incentives is small compared to the cost of inefficient locations (provided that even if divisional managers exert zero effort, products, when optimally located, appeal to overlapping sets of

customers),<sup>22</sup>

- If divisional managers' risk-aversion is small (or the variance of the noise is small) centralization improves welfare. Indeed, high-powered incentives for improving product quality can be provided at a small cost in the centralized firm. Decentralization thus brings small additional benefits (there may even be over-investment compared to the first best if incentives remain high-powered) at the cost of inefficient locations.
- If for the optimal incentive scheme in the centralized case, the set of customers who desire both products is small, decentralization improves welfare. Indeed, decentralization will induce (slightly) more effort while locations will deviate little from their first best values so that the costs induced by inefficient locations are very small (in the second order).

These intuitive results are made precise and proven in appendix 3.D.

We finally discuss the feasibility of the two organizational forms. Until now we have assumed that top management knows customers' locations. Since in equilibrium it also knows product quality and product location it can infer which product is better for a given customer. If bargaining between the divisions is costly,<sup>23</sup> top management would strictly prefer to allocate customers to divisions without letting them bargain. We thus see that decentralization may not be feasible.

Decentralization may become feasible if top management faces a cost to intervene, for instance if it does not know customers' locations. Bargaining among divisions may then be a more attractive option than mandating a suboptimal allocation of customers to divisions. We plan to study the feasibility of decentralization when top management is less informed, in future work.

---

<sup>22</sup>This result holds because divisional managers can choose any location for their products at the same cost. If, instead, some locations are costlier to choose than others and incentive schemes have a small slope, equilibrium locations may not differ much in the two organizational forms.

<sup>23</sup>We can easily introduce a small cost of bargaining in the model.

### 3.A Appendix: Proof of Proposition 3.3.1

Writing that the left-derivative of expression 3.3.1 w.r.t  $e_1$  is  $\geq 0$  at  $e_1^*$  we get

$$measure(S'_1) - Ce_1^* \geq 0 \Rightarrow e_1^* \leq \frac{1}{C} \min(1, \frac{v_0 + e_1^*}{t}) \leq \frac{1}{C}$$

where

$$S'_1 = S_1 \cap \{x \in [0, 1] : v_1^* - t|x - l_1^*| \geq 0\}$$

and  $S_1$  is the set defined in section 3.4 (for  $e_1^*, e_2^*, l_1^*, l_2^*$ ).

Therefore

$$v_1^* = v_0 + e_1^* \leq v_0 + \frac{1}{C} < \frac{t}{2}$$

(and similarly for  $v_2^*$ ) by assumption A, and one product even if located at  $1/2$  cannot cover the whole market. It is thus not optimal to offer two products such that all customers (weakly) prefer one to the other.

We will now fix  $v_1$  and  $v_2$  (smaller than  $t/2$ ) and determine the optimal locations,  $l_1$  and  $l_2$ .<sup>24</sup> We will then determine the optimal effort levels. Our previous analysis implies that  $l_1$  and  $l_2$  are such that each product is strictly better than the other for some set of customers

Suppose that products do not overlap i.e.

$$l_1 + \frac{v_1}{t} \leq l_2 - \frac{v_2}{t}. \tag{3.A.1}$$

Suppose then that  $l_1 - v_1/t < 0$ . It is then easy to show that the right-derivative of expression 3.3.1 w.r.t  $l_1$  ( $l_1 < 1/2$  since  $l_1 - v_1/t < 0$ ) is  $-tl_1 + v_1 > 0$ .<sup>25</sup> Therefore

$$l_1 - \frac{v_1}{t} \geq 0 \tag{3.A.2}$$

---

<sup>24</sup>These optimal locations exist since we maximize a continuous function over the compact set  $[0, 1/2] \times [1/2, 1]$ .

<sup>25</sup>We have to distinguish two cases  $l_1 + v_1/t < l_2 - v_2/t$  and  $l_1 + v_1/t = l_2 - v_2/t$ .

and similarly

$$l_2 + \frac{v_2}{t} \leq 1. \quad (3.A.3)$$

Inequalities 3.A.1, 3.A.2 and 3.A.3 imply that  $(v_1 + v_2)/t \leq 1/2$ . Expression 3.3.1 is then equal to

$$\frac{v_1^2}{t} + \frac{v_2^2}{t} - \frac{1}{2}C(e_1^2 + e_2^2). \quad (3.A.4)$$

Defining  $l_1 = 1/4 + (v_1 - v_2)/2t$  and  $l_2 = 1/2 + l_1$ <sup>26</sup> and using  $(v_1 + v_2)/t \leq 1/2$ , inequalities 3.A.1, 3.A.2 and 3.A.3 are satisfied and payoff is the same as when locating products at  $l_1$  and  $l_2$ .

Suppose now that products overlap i.e.

$$l_1 + \frac{v_1}{t} > l_2 - \frac{v_2}{t}. \quad (3.A.5)$$

We denote by  $x_s$  the location of the customer who is indifferent between the two products i.e.  $x_s \equiv (l_1 + l_2)/2 + (v_1 - v_2)/2t$ . ( $l_1 < x_s < l_2$ )

Suppose then that  $l_1 - v_1/t \geq 0$ . It is then easy to show that the left-derivative of expression 3.3.1 w.r.t  $l_1$  ( $l_1 > 0$  since  $l_1 - v_1/t \geq 0$ ) is  $-v_1 + (x_s - l_1) > 0$ .<sup>27</sup> Therefore

$$l_1 - \frac{v_1}{t} < 0 \quad (3.A.6)$$

and similarly

$$l_2 + \frac{v_2}{t} > 1. \quad (3.A.7)$$

Inequalities 3.A.5, 3.A.6 and 3.A.7 imply that  $(v_1 + v_2)/t > 1/2$ .

The derivative of expression 3.3.1 w.r.t  $l_1$  is  $-tl_1 + t(x_s - l_1)$  (therefore  $l_1 > 0$  and  $l_1 < 1/2$  and this derivation is meaningful). Similarly the derivative w.r.t  $l_2$  is  $-t(l_2 - x_s) + t(1 - l_2)$ . Setting these derivatives equal to zero we find  $l_1 = 1/4 + (v_1 - v_2)/2t$

---

<sup>26</sup> $l_1 \in [0, 1/2]$  since  $|v_1 - v_2|/t < 1/2$ .

<sup>27</sup>We again have to distinguish two cases  $l_1 - v_1/t > 0$  and  $l_1 - v_1/t = 0$ .



$(l_1 \in [0, 1/2])$  and  $l_2 = 1/2 + l_1$ .

Expression 3.3.1 is then equal to

$$\frac{v_1 + v_2}{2} - \frac{t}{8} + \frac{(v_1 - v_2)^2}{2t} - \frac{1}{2}C(e_1^2 + e_2^2). \quad (3.A.8)$$

To summarize, our above analysis implies that  $l_1 = 1/4 + (v_1 - v_2)/2t$  and  $l_2 = 1/2 + l_1$  are always optimal locations. In addition if  $(v_1 + v_2)/t \leq 1/2$  payoff is given by expression 3.A.4 while if  $(v_1 + v_2)/t > 1/2$  payoff is given by expression 3.A.8.

Assumption A implies  $1/Ct < 1/2$  which in turn implies the concavity of expressions 3.A.4 and 3.A.8 w.r.t  $e_1$  and  $e_2$ . Since these expressions are symmetric in these variables, the optimal  $e_1$  and  $e_2$  are equal and straightforward maximization gives us proposition 3.3.1.

### 3.B Appendix: Proof of Proposition 3.4.1

We will first characterize pure-strategy equilibria and then show that they exist.

#### Characterization

Taking the left-derivative of expression 3.4.3 w.r.t  $e_1$  and writing that the derivative is  $\geq 0$  for the equilibrium value of  $e_1$  we get

$$\begin{aligned} A(\text{measure}(S'_1) + \frac{1}{2}\text{measure}(S'_2 \cup S'_3)) - Ce_1 &\geq 0 \\ \Rightarrow e_1 &\leq \frac{A}{C}\min(1, \frac{v_0 + e_1}{t}) \leq \frac{A}{C} \end{aligned} \quad (3.B.1)$$

where

$$S'_k = S_k \cap \{x \in [0, 1] : v_1 - t|x - l_1| \geq 0\}$$

and the  $S_k$ 's are the sets defined in sections 3.4 and 3.5 (for the equilibrium values of  $e_1, e_2, l_1, l_2$ ).

Therefore one product even if located at  $1/2$  cannot cover the whole market and it is easy to check that no equilibrium in which all customers (weakly) prefer one

product to the other, exists. (Otherwise the manager of the inferior product would change its location so that customers strictly prefer his product.)

Suppose that in equilibrium products do not overlap i.e. inequality 3.A.1 holds. A similar argument to that given in the proof of proposition 3.3.1. implies that inequalities 3.A.2 and 3.A.3 hold.

The derivative w.r.t  $e_1$  is  $A(v_0 + e_1)/t - Ce_1 = 0$ .<sup>28</sup> Therefore  $e_1$  is given by its expression in proposition 3.4.1, case I (as well as  $e_2$ ). Moreover, combining inequalities 3.A.1, 3.A.2 and 3.A.3 and using the expressions for  $e_1$  and  $e_2$ , we find that an equilibrium with non-overlapping products exists only if

$$\frac{2v_0}{t} + \frac{A}{Ct} \leq 1/2. \quad (3.B.2)$$

Suppose now that products overlap i.e. inequality 3.A.5 holds. Defining  $x_s$  as before we can show that inequalities 3.A.6 and 3.A.7 hold.

The derivative of expression 3.4.3 w.r.t  $l_1$  is  $-tl_1 + t(x_s - l_1) + 1/2(tl_1 + v_1 - tx_s)$  (therefore  $l_1 > 0$  and  $l_1 < 1/2$  and this derivation is meaningful). Similarly the derivative w.r.t  $l_2$  is  $-1/2(tx_s - (tl_2 - v_2)) - t(l_2 - x_s) + t(1 - l_2)$ . The derivatives w.r.t  $e_1$  and  $e_2$  are  $x_s + 1/2(l_1 + v_1/t - x_s)$  and  $1/2(x_s - (l_2 - v_2/t)) + (1 - x_s)$ . Setting these derivatives equal to zero we get the expressions in proposition 3.4.1 (case II).

Inequalities 3.A.5, 3.A.6 and 3.A.7 together with the expressions for  $e_1$  and  $e_2$  imply that an equilibrium with overlapping products exists only if

$$\frac{2v_0}{t} + \frac{A}{Ct} > 1/2. \quad (3.B.3)$$

### Existence

Suppose that inequality 3.B.2 holds so that the only possible equilibria involve non-overlapping products located in the middle of their market segments. We first note that if we choose  $l_1 = 1/4$  and  $l_2 = 3/4$ , inequalities 3.A.1, 3.A.2 and 3.A.3

---

<sup>28</sup>We again have to distinguish two cases according to whether  $l_1 + v_1/t < l_2 - v_2/t$  or  $l_1 + v_1/t = l_2 - v_2/t$ .

are satisfied. We now show that a manager, say manager 1, does not get anything by deviating. Consider a deviation  $(\bar{e}_1, \bar{l}_1)$ . Equation 3.B.1 implies that  $\bar{e}_1$  is always smaller than its equilibrium value. Since a manager always prefers to offer a product which does not overlap with the other product (provided that the product is always located in the middle of its market segment and that product quality is kept constant) the choice  $(\bar{e}_1, \bar{l}_1)$  is dominated by  $(\bar{e}_1, 1/4)$  which is obviously dominated by  $(e_1, 1/4)$ .

Suppose now that inequality 3.B.3 holds so that the only possible equilibrium is the one described in proposition 3.4.1 (case II). Let us consider a deviation  $(\bar{e}_1, \bar{l}_1)$  by manager 1. If products do not overlap, using inequalities 3.A.1, 3.A.2 and 3.B.3 we find that  $\bar{v}_1/t \leq 1/4$ . This last inequality is not compatible with the first-order condition for  $e_1$  and inequality 3.B.3. Moreover, it is clearly not optimal for manager 1 to offer a product which all customers prefer to product 2 (since it would then have to cover the whole market), nor to offer a product which all customers find inferior to product 2. Finally, it is easy to check that manager 1's payoff is concave in  $(e_1, l_1)$  in the domain of overlap, so that the first-order conditions are sufficient for optimality.

### Incentive Schemes

If inequality 3.B.2 holds, the derivative of expression 3.3.1 (where  $e_1, e_2, l_1, l_2$  are given by proposition 3.4.1, case I), w.r.t  $A$  is

$$4(1 - A) \frac{\frac{v_0^2}{t} \left(\frac{Ct}{2}\right)^2}{\left(\frac{Ct}{2} - A\right)^3} - 2\alpha\sigma_u^2 A. \quad (3.B.4)$$

By contrast if inequality 3.B.3 holds the derivative is

$$\begin{aligned} & 2 \frac{\left(\frac{t}{2} + v_0\right) \left(\frac{Ct}{2}\right)^2}{\left(\frac{3Ct}{2} - A\right)^3} \left[ 3 \left( \frac{3}{2}(1 - A) - A \left( \frac{2v_0}{t} + \frac{1}{Ct} - \frac{1}{2} \right) \right) - \left( \frac{2v_0}{t} + \frac{A}{Ct} - \frac{1}{2} \right) \right] \\ & - 2\alpha\sigma_u^2 A. \end{aligned} \quad (3.B.5)$$

The derivative is not always decreasing in  $A$ . (It is easy to see that a sufficient condition for the derivative to be decreasing in  $A$  is that  $1/Ct < 1/6$ .) However, the derivative is decreasing in  $\alpha\sigma_u^2$  and Topkis's (1978) monotonicity theorem implies

that the optimal  $A$  is decreasing in  $\alpha\sigma_u^2$ .

### 3.C Appendix: Proof of Proposition 3.5.1

#### Equilibrium

Suppose that in equilibrium all customers (weakly) prefer the product of one division, say division 1, to the product of the other division. Suppose first that the two products are not the same. Then  $\bar{x} \geq l_2$ . If  $\bar{x} < 1$ , and since product 1 does not cover the whole market (by assumption A), the manager of division 2 would have an incentive to increase  $l_2$ . Therefore  $\bar{x} = 1$ . If both products are the same, we can assume that  $\bar{x} \geq 1/2$  and the same conclusion holds. Ruling out these equilibria where top management allocates all customers to one division, consider an equilibrium where products do not overlap. We then have

$$l_1 + \frac{v_1}{t} \leq \bar{x} \leq l_2 - \frac{v_2}{t}. \quad (3.C.1)$$

Following similar steps as in the proof of proposition 3.4.1, we can show that equations 3.A.2 and 3.A.3 hold and that  $e_1$  and  $e_2$  are given by their expressions in proposition 3.4.1 (case I). Combining inequalities 3.C.1, 3.A.2 and 3.A.3 and using the expressions for  $e_1$  and  $e_2$ , we find that an equilibrium with non-overlapping products exists only if inequality 3.B.2 holds. Clearly, any values  $l_1$ ,  $l_2$  and  $\bar{x}$  that satisfy inequalities 3.C.1, 3.A.2 and 3.A.3 are equilibrium values, and it is easy to check that setting  $l_1 = 1/4$ ,  $l_2 = 3/4$  and  $\bar{x} = 1/2$ , inequalities 3.C.1, 3.A.2 and 3.A.3 hold.

Suppose now that products overlap i.e. inequality 3.A.5 holds. Defining  $x_s$  as before, we have  $x_s = \bar{x}$  and we can show that inequalities 3.A.6 and 3.A.7 hold. The derivatives of agents' payoffs w.r.t  $l_1$  and  $l_2$  are  $-tl_1 + t(x_s - l_1)$  and  $-t(l_2 - x_s) + t(1 - l_2)$  and w.r.t  $e_1$  and  $e_2$  are  $Ax_s/t - Ce_1$  and  $A(1 - x_s)/t - Ce_2$ . Setting these derivatives equal to zero we get the expressions in proposition 3.5.1 (case II). As before, inequalities 3.A.5, 3.A.6 and 3.A.7 together with the expressions for  $e_1$  and  $e_2$  imply that an equilibrium with overlapping products exists only if inequality 3.B.3

holds. It is clear that managers cannot do better by deviating from equilibrium.

### Incentive schemes

If inequality 3.B.2 holds, the derivative of expression 3.3.1 (where  $e_1$ ,  $e_2$ ,  $l_1$ ,  $l_2$  are given by proposition 3.5.1, case I), w.r.t  $A$  is the same as in expression 3.B.4.

By contrast if inequality 3.B.3 holds the derivative is

$$\frac{1}{2C}(1 - A) - 2\alpha\sigma_u^2 A. \quad (3.C.2)$$

As before, the derivative is decreasing in  $A$  if  $1/Ct < 1/6$ , and the optimal  $A$  is always decreasing in  $\alpha\sigma_u^2$ .

## 3.D Appendix: Comparison of Organizational Forms

We will prove the following proposition

### Proposition 3.6.1

**Part 1** *Suppose that  $v_0/t > 1/4$ . Then if  $\alpha\sigma_u^2$  is sufficiently large, welfare is higher under centralization.*

**Part 2** *Suppose that  $2v_0/t + 1/Ct > 1/2$ . Then if  $\alpha\sigma_u^2$  is sufficiently small, welfare is higher under centralization.*

**Part 3** *Suppose that the optimal  $A$  under centralization verifies  $2v_0/t + A/Ct = 1/2$ . Then, welfare is higher under decentralization and under some cases is strictly higher.*

### Proof:

#### Part 1

Since  $v_0/t > 1/4$ , inequality 3.B.3 holds for any value of  $A$ . Expression 3.B.5 shows then that the optimal  $A$  can be made arbitrarily small if  $\alpha\sigma_u^2$  is sufficiently large. Proposition 3.4.1 implies then that  $e_1$  and  $e_2$  can be made arbitrarily close to zero and  $l_1$  and  $1 - l_2$  arbitrarily close to  $1/6 + v_0/3t$  (which is strictly larger than

1/4 since  $v_0/t > 1/4$ ). Since in the centralized firm effort levels are positive, welfare is higher under centralization for  $\alpha\sigma_u^2$  sufficiently large.

### Part 2

Welfare under centralization is larger than its value for  $A = 1$ , which in turn can be made arbitrarily close to welfare under the first best for  $\alpha\sigma_u^2$  sufficiently small. By contrast, welfare under decentralization cannot be made arbitrarily close to welfare under the first best for  $\alpha\sigma_u^2$  sufficiently small. Indeed, since  $2v_0/t + 1/Ct > 1/2$ , effort levels are strictly smaller than first best effort levels for  $A$  such that  $2v_0/t + A/Ct \leq 1/2$ , and by continuity for  $A$  slightly larger. For  $A$  even larger, effort levels may approach first best effort levels but the welfare loss due to inefficient locations is bounded away from zero.

### Part 3

If the optimal  $A$  under centralization, say  $A_c$ , verifies  $2v_0/t + A/Ct = 1/2$ , decentralization can only improve welfare since welfare under decentralization for  $A_c$  is equal to the maximum welfare under centralization. To provide parameter values such that welfare is strictly higher in the decentralized firm assume that  $1/Ct < 1/6$  so that welfare in the centralized firm is concave in  $A$  and assume that the right-derivative of welfare in the centralized firm w.r.t  $A$ , i.e. expression 3.C.2, is zero for  $A_c$ .<sup>29</sup> Simple algebra shows that expression 3.B.5 is always strictly larger than expression 3.C.2 for values of  $A$  such that  $2v_0/t + A/Ct = 1/2$  (and for  $A_c$  in particular). Therefore for  $A$  slightly higher than  $A_c$  welfare under decentralization is strictly higher than the maximum welfare under centralization. (Obviously this result is not “knife-edge” and holds even if the optimal  $A$  is close to an  $A$  that verifies  $2v_0/t + A/Ct = 1/2$ .)

---

<sup>29</sup>Welfare in both the centralized and the decentralized firm is not differentiable for values of  $A$  such that  $2v_0/t + A/Ct = 1/2$  so if such a value is optimal, the right-derivative is not necessarily zero.

## References

- Ferguson, Charles H. and C. R. Morris (1993), *"Computer Wars"* Times Books.
- Hart, Oliver D. (1983),, *"The Market Mechanism as an Incentive Scheme"*, Bell Journal of Economics, Autumn, 14,2:366-382.
- Holmstrom, Bengt R. and Jean Tirole (1989), *"The Theory of the Firm"*, (in R. Schmalensee and Willig, R., eds.), *Handbook of Industrial Organization*, Ch. 2.
- Holmstrom, Bengt R. and Jean Tirole (1991), *"Transfer Pricing and Organizational Form"*, Journal of Law, Economics and Organization, 7,2:201-228.
- Rey, Patrick and Jean Tirole (1986), *"The Logic of Vertical Restraints"*, American Economic Review, 76:921-939.
- Scharfstein, David (1987), *"Product Market Competition and Managerial Slack"*, Rand Journal of Economics, 19,1:147-155.
- Shleifer, Andrei (1985), *"A Theory of Yardstick Competition"*, Rand Journal of Economics, 16:319-327.
- Sloan, Alfred (1963), *"My Years at General Motors"*, Doubleday.
- Smith, Vicky (1992), *"Managing in the Corporate Interest"*, University of California Press.
- Tirole, Jean (1988), *"The theory of Industrial Organization"*, MIT press.
- Topkis, Donald M. (1978), *"Minimizing a Submodular Function on a Lattice"*, Operations Research, 26:305-21.