# Investor Behaviour, Financial Markets and the International Economy

by

Markus Konrad Brunnermeier

*The London School of Economics and Political Science*

Thesis submitted in partial fulfilment of the requirements of the degree of

**Doctor of Philosophy**

at the University of London

## Abstract

This dissertation focuses on analysing investor behaviour and price processes in asset markets. It consists of four self-contained essays in the areas of market microstructure, risk attitude of boundedly rational investors, and international finance.

Chapter 2 provides a review of the existing literature on the informational aspects of price processes. A common feature of these models is that prices reflect information that is dispersed among many traders. Dynamic models can explain crashes and illustrate a rationale for technical/chart analysis. The second emphasis of this survey is on herding models.

In Chapter 3, I have developed a multi-period trading-game that analyses the impact of information leakage. I find that a trader who receives a signal about a future public announcement can exploit this information twice. First, when he receives his signal, and second, at the time of the public announcement. Furthermore, I show that the investor trades very aggressively on the rumour in order to manipulate the price. This enhances his informational advantage after the correct information is made public. He also trades for speculative reasons, i.e. he buys stocks that he plans to sell after the public announcement.

Chapter 4 provides a theoretical rationale for experimental results such as loss aversion and diminishing sensitivity. A decision maker is considered to be boundedly rational if he can not find his new optimal consumption bundle with certainty when he is faced with a new income level. This makes him more risk averse at his current reference income level. It also makes him less risk averse for a range of incomes below his reference income level.

Chapter 5 considers a two country economy similar to that in Obstfeld and Rogoff (1995). We find that conclusions about whether monetary shocks lead to exchange rate overshooting and spillovers on foreign production and consumption depend crucially on the form of price stickiness. Sticky retail prices not only allow for a profitable 'Beggar Thy Neighbour Policy' but also lead to exchange rate overshooting. This is not the case under sticky wholesale prices and sticky wages.

## Acknowledgements

I enjoyed the help and support of numerous friends and colleagues while working on my dissertation. My greatest debt, however, is to my wife, Smita. She assisted me at each stage of the dissertation process, starting with early discussions of ideas and first modelling attempts through reading every draft of every chapter. Her unfailing support is especially noteworthy given that our graduate studies kept us apart in two continents separated by the Atlantic Ocean.

I was fortunate to have Margaret Bray as my supervisor. In countless meetings, she guided me, gave me useful insights, and suggested helpful modelling alternatives. I am also indebted to David Webb for his friendship and for teaching me to search for the economic implications that lie behind mathematical results. Discussions with Sudipto Bhattacharya allowed me to draw connections between different branches of economics. He suggested that I write the literature survey that resulted in Chapter 2. I appreciate Bob Nobay's comments on this chapter as well as his overall support over the years. I also had very insightful discussions with John Hardman Moore concerning the theoretical aspects of Chapter 3 and with Charles Goodhard about its empirical implications. For Chapter 4, I particularly appreciate Kevin Robert's suggestions. I would also like to thank Martin Hellwig, Antoine Faure-Grimaud, Douglas Gale, Steven Huddart, John Hughes, Bart Lipman, Ben Pollak, Sven Rady, Ailsa Röell, Ariel Rubinstein, Klaus Schmidt, Paolo Vitale and David Wildasin for their helpful comments on various aspects of this dissertation. The comments of numerous other faculty members and visitors at the LSE helped make this dissertation what it is.

I am also indebted to several institutions for their support. I gained a strong foundation in theoretical economics during my first year at the University of Bonn. The LSE in general, and the FMG in particular, provided an exciting and stimulating research environment during my later years. I also appreciate the hospitality extended to me by Duke University during the summers of 1997 and 1998.

I am also appreciative of my fellow students including, Volker Nocke, Thomas de Garidel, Dominik Hotz, Geoffrey Shuetrim, Mimi, Tam, Sönje Reiche, Ulrike Hoffman-Burchardi, Phillip Kalmus and Jörn Rothe, with whom I had many stimulating discussions and who read preliminary drafts of the chapters. In particular, I would like to thank Clemens Grafe for his excellent teamwork during the derivation of the results of Chapter 5 of this thesis. Hopefully our collaboration will lead to many more exciting research projects in the future.

I am especially grateful to my parents, Margareta and Markus, for always supporting my decisions and for the confidence they have put in me. And finally, I would like to thank all the other relatives, friends, and teachers who have positively influenced my life.

# Contents

# List of Figures

# Chapter 1

# Introduction

The economic importance of financial markets has increased steadily in recent decades. Transaction volume has reached new heights and more and more individuals are buying and selling assets.

There are many reasons why investors are active in financial markets. First, professional financial investors often trade in order to exploit their superior information about the value of an asset. Their trading affects the price process and hence the price partly reflects their information. This allows uninformed investors to learn more about the fundamental value of the stock. Second, traders sometimes trade for manipulative reasons in order to confuse uninformed investors who are trying to infer information from the price. While this may be costly for the manipulator in the short run, he hopes to recuperate early losses in later trading rounds. Third, individuals trade assets in order to insure themselves against risk. This trading behaviour depends crucially on their risk attitude. The more risk averse people are, the higher is their willingness to pay for assets which allow them to reduce their exposure to risk. Fourth, exporters and importers are active in the foreign exchange market. They have to convert their foreign revenues from exporting domestic goods or need foreign currency for importing foreign goods. Exchange rate volatility depends crucially on the flexibility of producer and consumer prices.

Each of the four main chapters of this dissertation addresses a certain aspect of these four reasons for trade. Chapter 2 provides an extensive and up to date review of the recent theoretical developments in the literature on price processes. This chapter casts light on models that explain the informational aspects of price changes. A common feature of these models is that the price communicates information which is dispersed among many traders.

After contrasting the Rational Expectation Equilibrium concept with the Bayesian Nash Equilibrium concept, a connection between market completeness and information revelation is drawn. The No-Speculation Theorem is explained. It states that if all traders behave rationally, then differences in information alone will not lead to more trade. On the other hand, differential information combined with possible gains from trade can lead to large trading volumes. No-Trade Theorems describe circumstances where asymmetric information can lead to market breakdown even though there are gains from trade. Two different kinds of No-Trade Theorems are covered in this survey. One is due to the adverse selection effect and the other one is due to too much information revelation through prices which leads to reduced risk sharing. Lastly, situations are described under which bubbles can occur even though all traders are rational and forward looking.

The next section of the survey classifies the standard market microstructure models into five groups. A distinction is drawn between models where traders can trade conditional on the future prices (limit order models) and models in which traders can only submit market orders. They can be further subdivided into models with strategic and competitive traders, respectively. The main focus of this section is on dynamic models covering a whole price process. A new rationale for the technical/chart analysis can be illustrated by means of these models. Dynamic models can also explain stock market crashes. In a setting with widely dispersed information, even relatively unimportant news can lead to large price swings and crashes.

The final section of Chapter 2 deals with various types of herding models. In general, in sequential decision making, herding behaviour can arise even though all agents behave rationally. Informational cascades arise because a predecessor's action only provides a noisy signal of his information. If traders have short horizons, perhaps due to their risk aversion, they have an incentive to gather the same information as other traders do.

In Chapter 3, I develop a multi-period trading game which explains manipulative trading behaviour. The purpose of price manipulation is to confuse the other traders. In this setting, a trader receives an imprecise signal about a future public announcement. The public announcement provides everybody with better information but at a later point in time. The other traders, however, do not know the extent to which this information is already incorporated in the current price. That is, they do not know the extent to which the early informed insider has already moved the price prior to the public announcement. The insider knows and, therefore, he has an informational advantage again. Thus, he can exploit his information twice. First, when he receives his signal, and second, at the time of the public announcement. The analysis also shows that she has an incentive to trade very aggressively prior to the public announcement in order to move the price such that she has a larger informational advantage at the time of the public announcement. In other words, she has an incentive to manipulate the price. The third feature of her trading strategy is that she also trades for speculative reasons. If she buys based on the rumour, she expects to sell at the time of the public announcement. Therefore, the title of the chapter "Buy on Rumours - Sell on News".

The second part of the chapter analyses how information leakage affects the informational efficiency of the price process. It shows that information leakage makes the price process less informationally efficient even after the public announcement. In other words, the price is further away from its fundamental value. This analysis suggests that firms that care about long-run informational efficiency should keep information secret prior to public announcements.

Chapter 4 focuses on investors' risk aversion. Risk aversion affects individuals' optimal portfolio choice and thus their trading behaviour. Experimental evidence suggests that agents who consume at their familiar income level are reluctant to accept income lotteries in which they face a potential loss. Experimental economists attribute this to a high 'loss aversion'. However, after facing an unanticipated drop in income, these agents are willing to take on lotteries they have earlier rejected. Experiments show that they may even become risk loving at a lower income level. In the terminology of Behavioural Economics, agents' behaviour exhibits 'diminishing sensitivity'. This translates into a von Neumann-Morgenstern utility function with the following three features. First, the utility function depends on the current income level. Second, due to loss aversion there is a steep decline in utility for income levels smaller than the current familiar income level. Third, as the income declines the utility function becomes less concave or might even become convex. The latter reflects diminishing sensitivity. Chapter 4 provides a theoretical explanation for such risk attitudes. A decision maker in this model is considered to be boundedly rational if he does not find his new optimal consumption bundle with certainty when he is faced with a new income level. He is likely to err when choosing from thousands of different commodities. This alters his indirect utility function and makes him more risk averse at his current reference income level. It also makes him less risk averse for a range of incomes below his reference income level.

Asset prices including exchange rates are highly volatile. The market microstructure and informational aspects explain sharp price movements in the short run, especially for high frequency data. To understand long run volatility, however, one has to look at the fundamentals like changes in the money supply. Chapter 5 considers a two country economy similar to the model in Obstfeld and Rogoff (1995). The analysis examines whether the form of price stickiness affects conclusions about whether monetary policy shocks lead to exchange rate overshooting and spillovers on foreign production and consumption. We distinguish between sticky retail prices, sticky wholesale prices and sticky wages. Sticky retail prices not only allow for a profitable 'Beggar Thy Neighbour Policy' but also lead to exchange rate overshooting. Although the outcome is similar to the seminal work by Dornbusch (1976), the driving force in this model is quite different. In Dornbusch (1976) exchange rate overshooting is driven by the enforced international interest rate parity. In our model the exchange rate overshoots even though the interest rate parity may not even hold in equilibrium. These results are in sharp contrast to the outcomes under sticky whole sale prices, where prices are fixed in the producers' currency. Contrary to the spirit of the 'Beggar Thy Neighbour Policy', an unexpected monetary expansion in one country benefits inhabitants of the other country as well. The interest parity always holds in equilibrium and there is no exchange rate overshooting. Similar results hold for the case of sticky wages. The analysis also shows that monetary policy shocks can not explain the J-curve effect under either form of stickiness.

# Chapter 2

# Prices, Price Processes and Information: A Survey of the Market Microstructure Literature

## 2.1 Introduction

Every day a vast number of assets changes hands. Whether these assets are stocks, bonds, currencies, derivatives, real estate or just somebody's house around the corner, there are common features driving the market price of these assets. Asset prices fluctuate more sharply than the prices of ordinary consumption goods do. We observe emerging and bursting bubbles, bullish markets as well as stock market crashes. There are many questions which fascinate academics, professionals as well as laymen. When do bubbles develop and crashes occur? Can price history provide us some hint about future price developments? Does technical or chart analysis make sense? Why is the trading volume in terms of assets so much higher than real economic activity? Can people's herding behaviour be simply attributed to irrational panic? Going beyond positive theory, some normative policy issues also arise. What are the early warning signals indicating that a different policy should be conducted? Can a different design of exchanges and other financial institutions reduce the risk of crashes and bubbles?

If financial crises and huge changes in assets prices only affect the nominal side of the economy, there would not be much to be worried about. However, as illustrated by the recent experiences of the Southeast Asian tiger economies, stock market and currency turmoil can easily turn into full-fledged economic crises. The unravelling of financial markets can spill over and affect the real side of economies.

A good understanding of price processes can help us foresee possible dangers. In recent years, the academic literature has achieved a great leap forward in understanding price processes of assets. This paper offers a detailed and up to date review of the recent theoretical literature in this area. It provides a framework for understanding price processes and emphasises the informational aspects of asset price dynamics. The survey focuses exclusively on models that assume that all agents are rational and act in their own self-interest. It does not cover models which attribute empirical findings to the irrational behaviour of agents.

The distinguishing feature of assets is that they entail uncertain payments, most of which occur far in the future. The price of assets is driven by expectations about these future payoffs. New information causes market participants to re-evaluate their expectations. For example, investors react to news about a company's future dividend prospects in the case of stocks or bonds, and to news of a country's economic prospects in the case of currencies. Depending on their information, market participants buy or sell the asset. In short, their information affects their trading activity and, thus, the asset price. Information flow is, however, not just a one-way street. Traders who do not receive a piece of new information are still conscious of the fact that the actions of other traders are driven by their information set. Therefore, uninformed traders can infer part of the other traders' information from the movement of an asset's price. The models covered in this literature survey demonstrate that past price processes can be studied to infer even more information of other traders. Therefore, technical analysis might not be as unreasonable as some earlier theoretical papers had suggested.

This survey is far from exhaustive. It is restricted to price processes and hence it addresses only a specific niche in the market micro structure literature. For a broader overview we refer the reader to O'Hara (1995). Nöldeke (1993) also provides a rigorous introduction to this literature, including its technical aspects. The reference for the empirical analysis of Financial markets is Campbell, Lo, and MacKinlay (1997)

## 2.2 Theoretical Results in a General Setting

### 2.2.1 Rational Expectations Equilibrium and Bayesian Nash Equilibrium Concept

Hayek (1945) was one of the first to look at the price system as a mechanism for communicating information. Prices play an informational role in a world where information is dispersed throughout the economy. Information affects traders' expectations about the uncertain value of an asset. There are different ways to model the formation of agent's expectations. Muth (1960), (1961) proposed a rational expectations framework which requires people's subjective beliefs about probability distributions to actually correspond to objective probability distributions. This rules out systematic forecast errors. The advantage of the rational expectations hypothesis over ad hoc formulations of expectations is that it provides a simple, general and plausible way of handling expectations. It also makes it possible to analyse the efficiency of markets as transmitters of information. In these models, agents draw inferences from all available information, i.e. from exogenous and endogenous data, and particularly from prices. Thus, investors base their actions on the information conveyed by the price as well as on their private information.

Before going into the details of the models, one needs to understand the underlying equilibrium concepts on which the predicted outcome is based. There are two competing equilibrium concepts: the *Rational Expectations Equilibrium* (REE) concept and the game theoretic *Bayesian Nash Equilibrium* (BNE) concept. In a REE, all traders behave competitively, i.e. they are price takers. They take the price correspondence, a mapping from the information sets of all traders into the price space as given. In a BNE, agents take the strategies of all other players, and not the equilibrium price correspondence, as given. The game theoretic BNE concept allows us to analyse strategic interactions in which traders take their price impact into account.

Both equilibrium concepts are probably best explained by illustrating the steps to derive the corresponding equilibrium. I will provide only a descriptive explanation. For a more detailed exposition one should a consult standard game theory book such as Fudenberg and Tirole (1991).

A possible closed form solution of a *Rational Expectations Equilibrium* can be derived in the following five steps. First, after specifying each traders' beliefs, propose a price function (conjecture) $P : \{\mathcal{F}^1, ..., \mathcal{F}^I\} \rightarrow \mathbb{R}_+^J$. This is a mapping from all $I$ traders' information sets $\{\mathcal{F}^1, ..., \mathcal{F}^I\}$ to the prices of $J$ assets. All traders take this mapping as given. One actually proposes a whole set of possible price conjectures $\mathbf{P} = \{P \mid P : \{\mathcal{F}^1, ..., \mathcal{F}^I\} \rightarrow \mathbb{R}_+^J\}$ (e.g. parameterised by undetermined coefficients) since the true equilibrium price function is not known at this stage of the calculations. Second, given the parameterised price conjectures, all traders draw inferences from the prices and one can derive each trader's posterior beliefs about the unknown variables. These beliefs are represented by a joint probability distribution and depend on the proposed price conjecture, e.g. on the undetermined coefficients of the price conjecture. In the third step, each individual investor derives his optimal demand, based on his (parameterised) beliefs and his preferences. In step four, the market clearing conditions are imposed for all markets and the endogenous market clearing price variables are computed. Since individuals' demands depend on traders' beliefs, so do the price variables. This gives the actual price function $P : \{\mathcal{F}^1, ..., \mathcal{F}^I\} \rightarrow \mathbb{R}_+^J$, the actual relationship between the traders' information sets $\{\mathcal{F}^1, ..., \mathcal{F}^I\}$ and the prices for a given price conjecture. Finally, rational expectations are imposed, i.e. the conjecture price function has to coincide with the actual one. Viewed more abstractly, the REE is a fixed point of the function $f : \mathbf{P} \rightarrow \mathbf{P}$. $f(\cdot)$ which maps the conjectured price relationship $\{\mathcal{F}^1, ..., \mathcal{F}^I\} \rightarrow \mathbb{R}_+^J$ onto the actual one. At the fixed point $f(P(\cdot)) = P(\cdot)$, the conjectured price function coincides with the actual one. If one applies the method of undetermined coefficients the fixed point is found by equating the coefficients.

In a strategic *Bayesian Nash Equilibrium* all players take the strategies of all the others as given. A player chooses his own optimal strategy by assuming the strategies of all the other players as given. A BNE is formed by a profile of strategies of all players from which no single player wants to deviate. A strategy determines player $i$'s action at each decision node. It consists of a sequence of action rules. An action rule is a mapping from player $i$'s information set into his action space at a certain point in time.

By illustrating the steps involved in the derivation of a BNE, we can highlight the differences between a BNE and a REE. To derive a BNE, one first has to conjecture a strategy profile. More specifically, one proposes a whole set of profiles described either by a profile of general functions or by undetermined coefficients. These profiles also determine the joint probability distributions between players' prior beliefs, their information and other endogenous variables like other traders' actions, demand and prices. A single player's deviation from a proposed strategy profile alters this joint probability distribution. Therefore, out-of-equilibrium beliefs also need to be specified if one player's deviation can be detected by the other players. In short, the belief system consists of equilibrium beliefs and out-of-equilibrium beliefs. Second, all players update their beliefs using Bayes' Rule and the joint probability distribution, which depends on the proposed set of strategy profiles, e.g. the undetermined coefficients. In the third step, each individual player derives his optimal response given the conjectured strategies of all other players and the market clearing conditions. If the best responses of all players coincide with the conjecture strategy profile, nobody wants to deviate. Hence, the conjectured strategy profile is a BNE. In other words, the BNE is a fixed point in strategy profiles. If one focuses only on equilibria in linear strategies, the proposed set of strategy profiles can be best characterised by undetermined coefficients. Each player's best response then depends on the coefficients in the conjecture strategy profile. The BNE is then given by equating the conjectured coefficients with the ones from the best response. More recent papers make use of variational calculus. This method allows strategies to take any functional form.

One problem with the BNE concept is the multiplicity of equilibria and its dependency on out-of-equilibrium beliefs. Often, certain out-of-equilibrium beliefs can be found that make a deviation unattractive. One branch of the refinement literature tries to reduce the number of equilibria by ruling out certain out-of-equilibrium beliefs.

In summary, the REE concept refers to a competitive environment where traders take the price function as given, whereas the BNE concept allows us to analyse environments where traders take their price impact into account. As the number of traders increases the price impact of a single trader decreases. Therefore, one might be tempted to think that as the number of traders goes to infinity, the BNE of a trading game where all traders submit demand schedules might converge to the competitive REE. Kyle (1989), however, showed that this need not be the case.

The REE provides a specific outcome for each possible realisation of the signals. The question arises whether this mapping from information sets onto outcomes can be *implemented.* In other words, could an uninformed social planner design a mechanism (game form) that would make it individually rational for all market makers to act as in the REE. If there exists a game form with a unique equilibrium which coincides with the REE outcome, then the REE can be *truthfully implemented.* Blume and Easley (1990) show that the REE outcome is only implementable for the case of finitely many traders if private information satisfies a kind of "smallness". More precisely, the private information of a single individual alone must not have any impact on the equilibrium. For the case of a continuum of traders, Dubey, Geanakoplos, and Shubik (1987) show that no continuous mechanism (including the submission of demand functions to a market maker) can truthfully (uniquely) implement the REE correspondence. This occurs because the demand function game does not specify a unique outcome in the case of several market clearing prices. The actual trading outcome depends on the trading mechanism, which makes it clear that the market structure matters. Laffont (1985) considers a class of economies in which the REE mapping can be implemented by an incentive compatible mechanism. This mechanism provides the right benchmark for welfare analysis.

Both equilibrium concepts also differ in their *epistemological assumptions.* Assumptions about the cognitive capacity of agents are an important part of game theory. In a BNE, it is common knowledge that all agents are rational and update their beliefs in accordance with Bayes' Rule. In other words, everybody knows that everybody is rational, and everybody knows that everybody knows that everybody is rational, and so forth, ad infinitum. In contrast, cognitive assumptions are contrary to the spirit of market equilibrium analysis. In a REE each agent is assumed to know the mapping from traders' information into prices, but nothing is assumed about what each agent knows about the other agents' cognitive capabilities and reasoning.

Both equilibrium concepts require that traders conduct complicated calculations. The question, therefore, arises whether it is possible to describe a plausible *learning* process which ultimately yields rational expectations if traders face the same situation repeatedly. It is shown in Bray and Kreps (1987) that rational learning of REE using a correctly specified Bayesian model is actually a more elaborate and informationally demanding form of REE. In such an extended REE, traders learn the "conventional" REE. Alternatively, if agents are boundedly rational in the sense that they are only using ordinary least square regressions to learn about the relationship between the price and the underlying information, the outcome converges under certain conditions to the REE, Bray (1982). The speed of this OLS-learning is analysed in more detail in Vives (1993).

### 2.2.2 Allocative Efficiency and Informational Efficiency

An *allocation* determines not only the current distribution of commodities, production etc. among all agents but it also specifies their redistribution at any point in time conditional on the state of the world. A current allocation, therefore, pre-specifies many future transactions which depend on the realisation of the state. Agents pre-specify future transactions through standardised security contracts and its derivatives, like futures etc. or through individual contractual arrangements. Pre-specified events trigger transaction determined by the allocation. It is important to distinguish these 'intra-allocation' transactions from trades. In a general equilibrium setting, trades refer only to changes from one allocation to another. The applied finance literature does not draw this distinction and calls all transactions trades.

As long as the term *"state of the world"* is not specified, the definition of an allocation will be somewhat vague. A state (path) of the world comprises a complete dynamic description of a possible outcome over the whole present and future time horizon. The description consists of two parts. The first part describes the real world or fundamentals at each point in time. Fundamental, or payoff-relevant, events are the possible endowment processes for each investor, such as, the dividend processes and the price processes for each asset for the entire history from $t = 0$ to $t = T$. The second, epistemological part summarises the agents' knowledge about the fundamentals as well as their knowledge about other agents' knowledge and so on, ad infinitum. Information can be represented by partitions or distributions over the state space. The epistemological component also describes how the information partitions of each individual investor evolves over time. In models with higher order uncertainty, it also covers higher order knowledge, i.e. knowledge about others' knowledge. Note that the state space can always be enlarged such that higher order knowledge of all agents can be represented by (commonly known) partitions/distributions over the extended state space. A model is called complete if its state space and each individuals' partition over it are common knowledge. The price process is part of the fundamentals as well as of the epistemological component since prices processes are endogenous signals and affect the traders' payoffs. An example of a possible state space is given by

$$\{\{endowments\}_{i\in\mathbb{I}}, \{dividend\, of\, asset\, j\}_{j\in\mathbb{J}}, \{price\, of\, asset\, j\}_{j\in\mathbb{J}}, \{\{signals\}_{j\in\mathbb{J}}\}_{i\in\mathbb{I}}\}_{t=0,\dots,T}$$

A current state in $t$ for trader $i$ is an event grouping all states which cannot be ruled out by the information provided until time $t$. In general, the description of a (current) state can be quite cumbersome. There are two ways of simplifying the (current) state description. First, one can group all states (events) which yield the same fundamentals up to the current time and in addition predict the same future fundamentals. In particular, this is possible if the economy exhibits some symmetry. Second, one can exploit the recursive structure of a system. For example, the state description can be simplified when the past matters only for a certain time. In general, a simpler "sufficient (current) state description" can be found. In this case, the new (current) state description is a sufficient statistic for the corresponding group of states (events) in the more cumbersome state space.

Economists distinguish between two different forms of efficiency. An allocation is *(allocative) Pareto efficient* if there is no other allocation which makes at least one agent strictly better off without making somebody else worse off. The price is (strong-form) *informationally efficient* if the price reveals a sufficient statistic for the information dispersed in the economy, Grossman (1978). If even all dispersed information, i.e. the join, is revealed by the price, then the price is fully revealing.[1][2] Otherwise the price is only partially revealing, i.e. informationally inefficient. The terms semi-strong and weak-form informational efficiency were coined by the empirical finance literature. Unfortunately, there is no unique definition of these terms. The current price is always semi-strong (weak-form) efficient if the current price is a sufficient statistic for all public information (past prices). Sometimes a market is also called semi-strong (weak-form) efficient if the knowledge of all public information (past prices) does not provide a profitable trading opportunity.

There are three forms of allocative efficiency: *ex-ante, interim and ex-post allocation efficiency.* Ex-ante efficiency refers to the time before signals are realised, interim efficiency to the time after signal realisation but before prices are observed, and ex-post efficiency refers to the time after the information is revealed through prices.[3] For informationally efficient REE ex-post allocation efficiency is a direct implication of the First Welfare Theorem. Laffont (1985) provides an example of an informationally efficient REE which is interim inefficient and a partially revealing REE which is ex-post inefficient. Hirshleifer (1971) first noted that the expected revelation of information can prevent risk sharing. The incentives to share risk ex-ante disappears if one knows the price which reveals the price. In other words, price revelation can make ex-ante desirable insurance impossible. Because of the Hirshleifer effect, it may be desirable to have a REE which only partially reveals the information of traders. Trade might be possible when prices reveal less information. On the other hand partially-revealing REE lead to a more severe adverse selection problem as uninformed investors can infer less information from prices. The trade-off between the Hirshleifer effect and the adverse selection effect is formally analysed in Marin and Rahi (1996).

Informationally efficient prices lead to some famous paradoxes. If prices are information-

[1] More formally, let $\overline{\mathrm{S}}$ be a sufficient statistic for $\{\mathcal{F}^1, ..., \mathcal{F}^I\}$. $\overline{\mathrm{S}}$ is a sufficient statistic if the knowledge of $\overline{\mathrm{S}}$ leads to the same sequence of action rules (local strategies). $P(\cdot) : \{\mathcal{F}^1, ..., \mathcal{F}^I\} \xrightarrow{g(\cdot)} \overline{\mathrm{S}} \xrightarrow{f(\cdot)} P$ is informationally efficient if $f(\overline{\mathrm{S}})$ is invertible. If in addition $g(\{\mathcal{F}^1, ..., \mathcal{F}^I\})$ is bijective (i.e. invertible), then $P(\cdot)$ is bijective and the prices are fully revealing.

[2] Informationally efficient REE can be derived by considering the corresponding artificial economy, in which all private information is treated as being public. The equilibrium of this artificial economy is a full communication equilibrium. Having solved for this equilibrium, one has to verify that it is a REE of the underlying diverse information economy.

[3] The different notions of allocative efficiency are discussed in more detail in Section 2.2.4.

ally efficient, i.e. they are a sufficient statistic for all private signals, no trader will condition her demand on her private signal. But if traders' demand is independent of the signals, how can prices be informationally efficient? How do traders know whether the observed price is the rational expectations equilibrium price or an off-equilibrium price? Thus, the Grossman-Paradox arises. In a model with endogenous information acquisition, informational efficiency precludes any costly information gathering. There is no incentive to gather costly signals if the sufficient statistic of all signals can be inferred from the prices for free. The problem that an overall equilibrium with costly, endogenous information acquisition does not exist if markets are informationally efficient is known as the *Grossman-Stiglitz Paradox.* These paradoxes do not arise in a Bayesian Nash equilibrium where the traders take the strategies of others, but not prices, as given. The resolution of these paradoxes is shown in Dubey, Geanakoplos, and Shubik (1987) wherein traders can only submit market orders and in Jackson (1991) wherein traders submit demand schedules, i.e. limit and stop orders. In general, a Bayesian Nash equilibrium in mixed strategies also exists in these settings.

In *partially-revealing* equilibria, incompletely informed traders face a signal extraction problem which does not allow them to infer the true reasons for the price change. There are many reasons for price changes, such as information about the dividend/liquidation value of securities/assets, endowments shocks, preference shocks (e.g. cross-sectional changes in risk aversion), and/or private investment opportunities. As long as the price change is due to symmetric information, each trader knows the true reason for it.

If some traders do not know the reasons for the price change they try to infer the asymmetric/differential information leading to it. Agents can generally only infer the price impact of this asymmetric/differential information, but not of the actual information itself. The question uninformed agents face is whether this information is also relevant to their portfolio choice. In other words, is asymmetric/differential information of *common interest* or only of *private interest* for the other traders. More generally: to what extent is information of common interest? To keep the analysis tractable, information which is partially of common interest and partially of private interest is assumed to be decomposable into these two parts. The literature refers to trade due to information of common interest as informational

Price Changes due to:

asymmetric/differential information | symmetric information

*information of common interest* | *information of common and private interest* | *information of private interest*

decompose

informational trading | non-informational noise trading

INFERENCE PROBLEM
for uninformed traders

Figure 2.1: Reasons for price changes

trading, whereas trade due to information of private interest is called uninformed trading or noise/liquidity trading. For example information about the liquidation value of an asset is of common interest. On the other hand, information about trader $i$'s inventory costs might concern only trader $i$'s evaluation of a certain security as long as trader $i$'s behaviour has no impact on the aggregates. An endowment shock for a whole group of investors might affect the portfolio choice of all investors via a change in the equilibrium prices, yet it primarily concerns only those investors who experience the endowment shock. A further example of information of private interest is provided in Wang (1994). In his model, informed investors receive information about a private investment opportunity in which only they can invest. An equilibrium is partially revealing if less informed traders cannot determine whether the unexpected price changes are due to others' information of common interest or information of their private interest. Figure 2.1 provides an illustration of the different reasons for price changes.

Following Grossman and Stiglitz (1980) most models exogenously introduce noise in order to make the equilibrium price only partially revealing. Section 2.3 and 2.4 will cover these models extensively.

Allen (1981) provides a class of exchange economies where the price is "privately revealing". The traders' private signals combined with the price is a sufficient statistic for the pooled information of all traders. The full communication equilibrium of the artificial economy can still be used in such a setting for proving the existence of a REE. In a more general environment where the asymmetry of information persists in equilibrium, a different proof has to be found. In order to apply a Fixed Point Theorem, expected utility functions and, thus, the excess demand functions must be continuous in prices. Ausubel (1990) presents a set of economies where every trader gets two signals. The first signal is a real number and the second signal is binary. The imposition of some differentiability conditions on marginal utility allows Ausubel (1990) to construct a partially revealing REE.

There are also models where investors only observe a noisy price. In Allen (1985) the market clears only approximately since individuals' demands are based on this noisy price. According to the Dominated Convergence Theorem, the noisy component smoothes out discontinuities in the excess demand function. This allows her to apply the Fixed Point Theorem on excess demand functions (instead of on the price mappings) and show the existence of a partially-revealing REE. In Allen's model, traders know the equilibrium relationships between prices and parameters describing the uncertain environment precisely, but the prevailing price vector is not completely accurate. In other words, agents' models (beliefs) coincide with the true model. This rationality assumption is relaxed in Anderson and Sonnenschein (1982) and McAllister (1990), where agents' beliefs are not only based on the state of the world but also on the realised price. Their approach incorporates elements of bounded rationality and goes beyond the scope of this literature survey.

### 2.2.3 Market Completeness

The number of possible allocations, the trading opportunities as well as the informational efficiency of the equilibrium prices depend on the number and nature of tradable assets.

In a *one-period model* with symmetric information, a market is complete if there are enough assets with linearly-independent payoffs such that each possible state of the world is insurable. A state is insurable if the security structure is such that buying or selling a certain combination of assets only alters the payoffs in this single state. In other words, there exists an alternative security structure with Arrow-Debreu securities for each possible state which leads to the same equilibrium outcome. An Arrow-Debreu security for state $\omega$ pays one unit only in state $\omega$. The asset prices of the original securities are given by the weighted sum of state prices, weighted such that the Arrow-Debreu securities replicate the final payoff of the original asset. Normalising the state prices, i.e. the prices of the Arrow-Debreu securities, such that their sum is equal to one, provides a nice interpretation of these prices in terms of probabilities. Imagine a risk neutral investor whose subjective probability distribution over the states of the world happens to coincide with the normalised state prices. His discounted expected value of any asset's payoff is then equal to the equilibrium asset price. Obviously, in the case of incomplete markets, the state price for the uninsurable states are not determined and only a constrained Pareto efficient allocation can be reached in equilibrium.

In a *dynamic multi-period model* the state space is much larger since each possible path/history is a single state. On the other hand, there is also more than one trading round. Markets are *completely equitisable* (or complete in the sense of Debreu (1959)) when there are enough securities with linearly-independent payoffs such that conditional trading on any possible path/history is possible already at time $t = 0$. In other words, any state in the large state space is insurable and, thus, any payoff stream can be generated through a once-and-for-all trade in $t = 0$.

Trading in later trading rounds has the advantage that one can condition the trading activities on the prior history. Dynamic trading strategies already specify in $t = 0$ issues such as when, in which states, which and how many assets are bought or sold. Using dynamic trading strategies the same payoff stream can be generated as in the case of completely (equitisable) markets with much fewer assets. Markets are *dynamically (synthetically) complete* if all states are insurable through dynamic trading strategies.[4] The number of linearly-independent assets only has to be larger than the maximum splitting index. In a model with symmetric information, the splitting index at time $t$ reports the number of branches the path/history can possibly proceed starting from the current event in $t$ to $t + 1$. As long as markets are dynamically complete, any additional asset is redundant and, thus, would

[4] E.g. if the dividend payments of one asset are normally distributed, then the number of states is already infinite and, therefore, any market with finitely many assets is incompletely equitisable. In a continuous time model, the market can still be dynamically complete.

not alter the economy. The price of this redundant asset can be derived from the prices of the other assets.

This result changes dramatically if traders are asymmetrically informed. Now it makes a big difference whether markets are completely equitisable or only dynamically complete. The state space becomes even larger. It also has an epistemological component. The epistemological component alone would not be a major concern since traders' utility depends on the fundamentals, i.e. the payoff relevant outcomes and, thus, only on the real world part of the state space. The main problem for the uninformed traders under asymmetric information is that they cannot distinguish between whether the opponent traded for informational reasons or whether he just followed his pre-specified dynamic trading strategy. Adding another security increases the number of observable prices. Whether markets are completely equitisable or only dynamically complete can make an enormous difference in information aggregation and revelation of prices. Grossman (1995) illustrates this point by means of an example. He compares an economy with actively traded (one-period) bonds with an economy of passively held annuities. Both economies can have the same payoff stream, since bonds allow agents to synthesise the payoff stream of the annuity. In a world of asymmetric/differential information, dynamic trading strategies require traders to make inferences about the future path of bond prices at each point in time. On the other hand, only the passive strategy of holding the market portfolio is required when all payoff streams are equitisable at $t = 0$. Bonds are traded, either for allocative reasons, i.e. in order to synthesise a certain payoff stream or for informational reasons, e.g. to exploit information about the future interest rate. Thus, for uninformed traders, an inference problem arises when there is asymmetric information at $t = 0$ since the extent of allocative trading need not be common knowledge. A second example is provided in Grossman (1988). He points out the important informational difference between a synthesised and a real option. This is analysed within a general equilibrium setting in Grossman and Zhou (1996). If the option is traded, the implied volatility of the underlying asset can be inferred through the option price. This is not possible if the option is synthesised by dynamic trading strategies. In economies where the average risk aversion is decreasing in income, synthesised options lead to higher volatility and mean reversion in returns.

In models with *asymmetric information* each trader's dynamic trading strategy has to satisfy a measurability condition. The condition states that at any time a trader can only apply different trading strategies for different states when he can distinguish between them. As in the case of symmetric information the markets are 'dynamically complete under asymmetric information' if the number of linearly independent assets exceeds a certain splitting index. However, there are now many individual splitting indices, one for each trader. The splitting index at $t$ for trader $i$ indicates the number of subpartitions his information partition can be split into when he receives new additional information at $t + 1$. In general, the partitions have to be defined over the smallest state space such that the partitions of all traders are common knowledge. Taking the maximum splitting index for all traders $i$ at any time $t$ provides the overall splitting index. A market is dynamically complete under asymmetric information if the number of linearly independent assets is larger than the overall splitting index.

In a dynamically complete market setting, any payoff stream can be synthesised using dynamic trading strategies. Therefore, in an incomplete equitisation setting, non-informational trading, possibly over the whole trading horizon, can occur in order to obtain the desired income stream. Since at the same time insiders trade to make use of their information, uninformed traders face an inference problem. They do not know the extent to which the price change is due to insider trading and, therefore, the price is only partially revealing.

If there are only a few assets, informed agents' trading possibilities are also quite restricted. Thus, there are many information constellations which can lead to the same trading behaviour, and more precisely, to the same price vector. It become obvious that not only the number of traded securities matters, but also which securities are traded is important. The actual security design has a tremendous impact on information and has motivated the optimal security design literature. We do not focus on this strand of literature but direct interested readers to Allen and Gale (1994) and Duffie and Rahi (1995).

The *existence* of an informationally efficient REE also depends on the number of assets and the economy's security structure. Existence problems are attacked from two directions, existence theorems and non-existence examples. The crucial question for the existence of fully revealing REE is whether the mapping from signals onto prices is invertible. The first non-existence example was provided by Kreps (1977). Boundaries on the existence of REE are given by four main results. If there is only a finite number of possible signals (e.g. {high, middle, low}) and prices can be any vector in $\mathbb{R}^n_+$, the invertibility of the mapping from signals onto prices fails only in special circumstances. Radner (1979) concluded that a REE exists and is fully revealing, for a generic set of economies. Thus, the example by Kreps is not robust, since a small change in the parameters would destroy non-existence. If the signal structure is more general, in the sense that a signal realisation can take on any value on $\mathbb{R}$, or even $\mathbb{R}^m$, the dimensionality of the signal space plays a crucial role. Allen (1982) showed that if the number of relative prices is larger than the dimensionality of the signal space, then a REE does exist and is fully revealing for a generic set of economies. In the case where the dimension of the signal space is equal to the number of relative prices there exists an open set of economies with no REE, Jordan and Radner (1982). If the dimension of the signal space is higher than the dimension of the relative price space, then there exists a generic set of economies with non fully revealing REE, Jordan (1983). Similar results may apply for the existence of informationally efficient REE. Instead of the dimensionality of the signal space, the dimensionality of the sufficient statistic for the signals matters for informationally efficient REE. In an economy with complete equitisation, the number of assets with linearly independent payoffs is equal to number of fundamental events/histories. Taking the current consumption price as the numeraire, the number of relative prices also equals this number. Thus, there exists generically a competitive REE which reveals all fundamental histories. Since knowing the fundamental histories/events is sufficient for any trading decision, there exists (generically) an informationally efficient REE in economies with complete equitisation.In economies with incomplete equitisation, this need not be the case and, thus, the actual signal realisations may matter. When the signal space is larger than the number of relative prices, a partially-revealing REE might still exist. Existence proofs for partially-revealing REE are only given for special parameterised economies.

### 2.2.4 No-Speculation Theorems and No-Trade Theorems

In the Aumann structure, information is represented by partitions over different possible states of the world. The state space can be extended such that the partitions of all traders are common knowledge. This survey does not cover non-partitional information dealing with issues of bounded rationality. Representing information as partitions allows us to define knowledge operators which report the states of the world in which a certain event is known. There are two equivalent notions of common knowledge. In the usual terminology an event in a certain state is common knowledge if all agents know that the true state lies in this event and all know that all know and so on ad infinitum. The more tractable notion is that an event in a certain state is common knowledge if and only if the finest common coarsening of all traders' partitions, (i.e. the meet) is a subset of this event. This formal notion of common knowledge allowed Aumann to show that rational players cannot "agree to disagree" about the probability of a given event. In other words, if the posterior probability of a rational player about a certain event is common knowledge, then the other player must have the same posterior probability, Aumann (1976). This result requires that all players use the Bayesian updating rule (i.e. they are rational) from a common prior distribution as well as that rationality of all players is common knowledge. The common prior assumption is also known as the Harsanyi doctrine and is in conflict with the Axioms of Savage (1954). However, it acts as a scientific discipline on possible equilibrium outcomes, Aumann (1987). The common prior doctrine states that differences in probability assessments must be due to differences in information. Aumann's (1976) agreement result says intuitively that if some rational trader, A, has a different probability assessment than trader, B, then trader B must conclude that this can only be due to the fact that trader A has information trader B has not considered yet and/or vice versa. It is important to note that the fact that posterior probabilities are equal, does not mean that all traders followed the same reasoning to get this common posterior. They need not have the same information. The formal proof makes use of the sure-thing principle which states that if the expected value of a random variable conditional on event $E$ is the same as the expected value conditional on event $F$, then the expected value conditional on event $E \cup F$, where $E \cap F = \emptyset$, is the same. Geanakoplos (1994) uses this sure-thing principle to generalise Aumann's result and show that common knowledge of actions negates asymmetric information about events. If action rules, which are mappings from players' partitions into the action space, are common knowledge, then there exists an environment with symmetric information that would lead to the same action. From this "Agreement Theorem" it follows that two rational agents never bet against each other. The same carries over to a situation with many traders. If the net trade vectors of the traders are common knowledge no trader will speculate.[5]

But even when the net trade vectors are not common knowledge, i.e. each trader only knows his trading activity and observes the price and maybe the aggregate trading volume, No-Speculation Theorems may still apply. The No-Speculation Theorem in Milgrom and Stokey (1982) states that if it is common knowledge that all traders are rational and the current allocation is ex-ante Pareto efficient, new asymmetric information will not lead to trade, given traders are strictly risk averse and hold concordant beliefs.[6] If the current allocation is known to be ex-ante Pareto efficient, then there is no incentive to trade in order to share risk. In general, an allocation is Pareto efficient if there does not exist an alternative allocation which yields a strictly higher expected utility for at least one agent without reducing the expected utility level of all others. One has to distinguish between ex-ante, interim and ex-post Pareto efficiency. If an allocation satisfies the Pareto criteria with respect to the expected utility of agents before receiving any signal then it is ex-ante efficient. An allocation is interim Pareto efficient if it satisfies the Pareto criteria with respect to the expected utility conditional on knowing the private signal. Similarly, an allocation is ex-post efficient, if there can be no possible Pareto improvement in the expected utilities of agents even after the private signals and public price signals are observed. If an allocation is ex-post inefficient, i.e. an ex-post Pareto improvement for a signal realisation can be made then an ex-ante Pareto improvement is also possible. Therefore ex-ante efficiency implies interim efficiency, which in turn implies ex-post efficiency.

No-Speculation theorems can either be proved using the sure-thing principle as in Milgrom and Stokey (1982) or by using the fact that more knowledge cannot hurt a Bayesian optimiser in a non-strategic environment.[7] The latter approach was used by Kreps (1977)

[5]We refer to speculation as trading where it is common knowledge that agents trade purely for reasons of asymmetric/differential information.

[6]A more detailed discussion can be found in Section 2.4.1, wherein we discuss Grundy and McNichols (1989).

[7]In a strategic dynamic environment, ignorance may allow an agent to commit himself to a (subgame perfect) action which is Pareto improving.

and Tirole (1982) and makes use of the fact that, with common priors, pure speculation, i.e. trade caused by asymmetric information at a Pareto optimal allocation, is a zero-sum game. Anyone who receives a trading offer can infer that her opponent wants to make money by using her superior information. Since the opponent can only gain if somebody else loses, nobody will be willing to trade except at prices that already incorporate her information. In other words, passive investment is a (weakly) dominant strategy. The No-Speculation Theorems can be applied to Bayesian games as well as to REE models. A crucial assumption is that it is common knowledge that the current allocation is ex-ante Pareto optimal and all agents are rational. Note, in contrast to the Bayesian Nash equilibrium, the Rational Expectations Equilibrium concept does not assume common knowledge of rationality.

No-Trade Theorems describe situations which lead to a no trade outcome although the current allocation is not ex-ante Pareto efficient. Asymmetric information does not generate trade in such circumstances. Worse still, it inhibits trade that would otherwise have occured. If the net trade vector is common knowledge the no-trade result extents to REE where the current allocation is not ex-ante Pareto efficient, Geanakoplos (1994). Another kind of No-Trade Theorem arises from the Hirshleifer effect, as mentioned above, Hirshleifer (1971). In this case the anticipated information revelation through prices prevents agents from risk sharing trade. In a world with uncertainty where one group of risk averse traders is better off in one state and the other group in the other state, trading provides a means for ex-ante Pareto improving risk sharing. After the uncertainty is resolved, the group of traders which is better off is not willing to trade anymore, since any allocation is ex-post Pareto efficient. Consider an information structure such that no trader can distinguish between both states, but the combined information, i.e. the join, provides knowledge about the true state. Now, if the price reveals the true state, knowledge of the price prevents trading. Trade will not take place in the first place in anticipation of the information revelation of the price.

Another group of No-Trade theorems is related to Akerlof's market for lemons, Akerlof (1970). They relate to situations where the current allocation is not ex-ante Pareto efficient and agents want to trade for informational and non-informational reasons, e.g. for risk sharing. As explained in Section 2.2.3, non-informational trading demand also arises in incomplete equitisation settings. A price change can be due to high/low informed or uninformed demand, which then leads to a signal extraction problem. Uninformed traders face an adverse selection problem, which allows the informed traders to extract an information rent from the uninformed. If the number of informed traders or the informational advantage of the insiders is too large, then the loss uninformed traders incur through the information rent for the insiders can outweigh their hedging gains. In these cases they are unwilling to trade and one observes a market break down. Bhattacharya and Spiegel (1991) analyse market breakdowns for the case of a single information monopolist who trades with infinitely many competitive uninformed investors. In this model the information monopolist trades strategically, i.e. he takes into account the fact that his order will have an impact on information revelation through prices. The authors conclude their analysis by providing some justifications for insider trading laws.

No-Speculation theorems and No-Trade theorems even arise in a setting with heterogeneous prior beliefs. Morris (1994) show that incentive compatibility considerations can preclude trading.

He and Wang (1995) show that new asymmetric information need not lead to a no-trade outcome if the information is dispersed among many traders. Dispersed information can even lead to a higher trading volume than the volume that would result under symmetric information. Their model is discussed in more detail in Section 2.4.3. The difference is that the initial allocation is not ex-ante Pareto efficient, or at least it is not common knowledge that it is ex-ante Pareto optimal. Also, since prices are only partially revealing, the Hirshleifer effect does not preclude trade.

### 2.2.5 Bubbles

If the stock price exceeds its fundamental value a bubble occurs.[8] The literature dealing with bubbles is huge. Famous historical bubbles are described in Garber (1990). The major quarrel is concerned with the question of whether large changes in prices are due to shifts in the fundamentals or just bubbles. The empirical finding of excess volatility literature starting with LeRoy and Porter (1981) and Shiller (1989) are arguments in favour of the existence of bubbles. The difficulty lies in determining the fundamental value of an asset. It seems plausible to consider the fundamental value of an asset in normal use as opposed

[8] This section is mostly based on Allen, Morris, and Postlewaite (1993).

to some value it may have as a speculative instrument. This description of a fundamental value depends on the context of a particular equilibrium and is therefore not exogenous. A problem arises when there are multiple equilibria. The same repeated economy can have different equilibria in different periods, if e.g. sunspots are used by the agents as a co-ordination device. Sunspots are random variables whose realisation in each period is common knowledge but have no economic relevance except that they emerged as convention to be used as co-ordination devices.

In the case of certainty, the fundamental value is just the present value of the equilibrium market value of the dividend stream. Taking expectations of the possible fundamental values in the case of uncertainty where all agents have the same information is problematic since this is only correct if all investors are risk neutral and there are complete markets. If traders have asymmetric information about the uncertain dividend stream additional problems arise. It is not clear which probability beliefs one should take into account in order to derive the expected value. Determining the fundamental value is therefore quite problematic. In Allen, Morris, and Postlewaite (1993) a clear upper bound on any reasonable fundamental value is used. If the price exceeds this upper bound of fundamental values a "strong bubble" occurs. More formally, a "strong bubble" occurs if there exists a state of the world in which every agent knows with probability one that the price is strictly above this upper bound. An "expected bubble" exists if there is a state of the world in which the price of the asset exceeds every agent's marginal valuation of the asset. This is in line with the notion that the right to resell an asset makes traders willing to pay more for it than they would pay if they were forced to hold the asset forever. This definition was used by Harrison and Kreps (1978) who attributed it to Keynes (1936). In the case of incomplete equitisation this definition does not make much sense since dynamic trading strategies can lead to 'expected bubbles'. The same price need not be an expected bubble if there is complete equitisation. Thus the definition of expected bubbles hinges on the degree of equitisation.

Economists have used different frameworks to explain the existence of bubbles. One way to generate bubbles is to introduce traders who behave irrationally, De Long, Shleifer, Summers, and Waldmann (1990a). If all trades are rational, backward induction rules out bubbles. There are, however, situations where the backward induction argument fails. First, this is the case in models with infinite horizon combined with myopic investors or an overlapping generation framework. With an infinite horizon, backward induction does not have a determined starting point and bubbles can occur because of a "lack of market clearing at infinity." Fiat money is the classical example of a bubble. Tirole (1982) showed that expected bubbles can exist and follow a Martingale if traders are myopic. Second, bubbles may exist in a finite time horizon model, provided there are infinitely many trading opportunities in the remaining trading rounds, Bhattacharya and Lipman (1995). Third, Allen, Morris, and Postlewaite (1993) show that higher order uncertainty can also lead to bubbles. Higher order uncertainty refers to uncertainty about others' information. In almost all standard economic models information is concerned with "pay-off relevant", or fundamental, events. The partitions other agents have are common knowledge. This allows us to model asymmetric information about fundamentals but not about others' knowledge. Higher order uncertainty deals with the case where one does not know the information structure of others, i.e. there is uncertainty about the information partitions others have. Without loss of generality this can be dealt with using an extended state space covering different possible information structures.

In Allen, Morris, and Postlewaite (1993) an example of a strong bubble is provided. In this example informed traders know the liquidation value of the asset, but they do not know whether other traders know it too. Since Tirole (1982) showed that if the ex-ante allocation is Pareto efficient there cannot even be an expected bubble in a fully dynamic REE, it is necessary for the existence of a bubble that there are gains from trade. The example provided in Allen, Morris, and Postlewaite (1993) illustrates four different ways to generate gains from trade. A further necessary condition for the existence of an expected bubble is that short sale constraints are strictly binding with a positive probability, Harrison and Kreps (1978). For strong bubbles to occur two further requirements are necessary. First, each agent has to have private information. This rules out cases in which prices are fully revealing. Second, agents' net trade vectors must not be common knowledge. As discussed above, even at an ex-ante Pareto inefficient allocation this could rule out any trade activity by the theorem that common knowledge of action negates asymmetric information. Therefore, in their example there are at least three traders. Although every trader knows the true value of the asset in their example, i.e. the true value of the asset is mutual knowledge, its price is higher. This is the case, because each trader thinks that he can resell the asset at a price above the true value.

The notion of higher order knowledge captures not only information about fundamentals and other agents' information, but also information about others' information about information, etc. ad infinitum. By extending the state space, the description of a state has many components capturing the payoff and all the higher orders of knowledge. The state space grows fast with the order of knowledge considered. Given a certain state space and a set of events, Morris, Postlewaite, and Shin (1995), define "depth of knowledge" as the number of iterations of knowledge necessary to distinguish all those states which are distinguishable on the basis of fundamentals and iterated knowledge of fundamentals. The state space in the example of Allen, Morris, and Postlewaite (1993) has depth of knowledge one since it is sufficient to distinguish between the states in which others have or do not have information in addition to the fundamental events. It is important to note that the whole state space and partitions of the state space are assumed to be common knowledge. Therefore agents know much more in state spaces with a lower depth of knowledge.

It is shown in Morris, Postlewaite, and Shin (1995) that a strong bubble can only exist in period $t$ if it is not mutual knowledge that at time $(t+1)$ it will be mutual knowledge, that ..., that in $(T-1)$ the true asset value is mutual knowledge. Because information is revealed with the price process, mutual knowledge in period $t$ already incorporates the information that can be inferred from the price in period $t$. Since knowledge can only improve through time, it follows that a bubble can only exist at or after time $t$ if the true asset value is not $(T-t)$th order mutual knowledge at time $t$. Furthermore, if the state space exhibits only a depth of knowledge of order $n$, then $n$ order mutual knowledge is sufficient to rule out bubbles. If this is the case, the true value is common knowledge anyway. In Tirole (1982) the assumed depth of knowledge is zero and therefore any bubble can be ruled out. In the case where the depth of knowledge of the state space is higher than the order of mutual knowledge of the true asset value, some bounds for the size of the bubble can still be provided. For this purpose a subset of the state space is taken, which is common $p$-belief. In other words this subset is believed to be true with at least probability $p$ by each agent and each agent believes that each agent believes that this subset is true with probability $p$ and so on ad infinitum. Using reasoning similar to the knowledge case (which is similar to the $p = 1$ case[9]) bounds for the size of bubbles can be derived. The minimum bound can be found by varying the subset of the state space and its associated $p$.

## 2.3 Classification of CARA-Gaussian Closed-Form Solution Models

In order to go beyond pure existence proofs, one has to specify the economy in more detail. The loss of generality is compensated by the finding of closed-form solutions which allow some comparative statics. In this section we restrict ourselves to a class of economies where all random variables are normally distributed.

These models can be classified in the following manner:

- Limit Order Models
  - Competitive Limit Order Models
  - Strategic Limit Order Models
- Market Order Models
  - Competitive Market Order Models Without a Market Maker
  - Strategic Market Order Models With a Market Maker in Which Traders Move First
  - Sequential Trade Market Order Models in Which the Market Maker Moves First

Two tools facilitate our analysis. The Projection Theorem can be used to address the signal extraction problem, while the Certainty Equivalence argument can be used for simplifying the utility maximising problem.

The final wealth can be written as the inner product $W = \mathbf{\Pi}'\mathbf{x}$, where the vector $\mathbf{\Pi}$ represents the (random) liquidation values of the $J$ securities and the vector $\mathbf{x}$ describes a portfolio. $W$ is also normally distributed, since the sum of normally distributed random variables is also normally distributed.

[9]Dekel and Gul (1997) discuss the distinction between p=1-beliefs and knowledge.

All traders have a constant absolute risk aversion (CARA) utility function of the exponential form

$$U^i(W) = -\exp(-\rho^i W),$$

where the absolute risk aversion measure $\rho^i = -U^{i\prime\prime}(W)/U^{i\prime}(W)$ is independent of agent $i$'s income. Taking the expectation, a function resembling a moment generating function is obtained.[10]

$$E[U^i(W) \mid \mathcal{F}^i] = \int_{-\infty}^{+\infty} -\exp(-\rho^i W) dW_{\mathcal{F}^i} = -\exp[\underbrace{-\rho^i(E[W \mid \mathcal{F}^i] - \frac{\rho^i}{2} Var[W \mid \mathcal{F}^i])}_{Certainty\ Equivalent}].$$

Therefore, maximising the expected utility conditional on the appropriate information set $\mathcal{F}^i$. is equivalent to maximising the certainty equivalence. The imposed assumptions result in linear demand functions.

Before conducting the maximisation procedure, the conditional expected value and variance have to be derived. The Projection Theorem helps us derive the conditional expectations and the conditional covariance and variance terms. Given our assumption that random variables are normally distributed, the first two moments are sufficient to describe the whole conditional distribution. The Projection Theorem provides the *linear* projection of $S$ on the space of quadratic integrable functions. For normally distributed variables, the linear projection is also the optimal one. The Projection Theorem in any Hilbert space is:

$$E[\Pi \mid \vec{S}] = E[\Pi] + (\vec{S} - E[\vec{S}])' \mathbf{Var}^{-1}[\vec{S}] \mathbf{Cov}[\Pi, \vec{S}],$$

$$Var[E[\Pi \mid \vec{S}]] = \mathbf{Cov}[\Pi, \vec{S}]' \mathbf{Var}^{-1}[\vec{S}] \mathbf{Cov}[\Pi, \vec{S}].$$

If $\Pi$ can be decomposed in $\Pi = E[\Pi \mid \vec{S}] + e$, such that the error term $e$ is orthogonal to $E[\Pi \mid \vec{S}]$, i.e. $Cov[E[\Pi \mid \vec{S}], e] = 0$, we have $Var[\Pi] = Var[E[\Pi \mid \vec{S}]] + Var[e]$ and $Cov[\Pi, E[\Pi \mid \vec{S}]] = Var[E[\Pi \mid \vec{S}]]$. Since $Var[\Pi \mid \vec{S}] = Var[e]$,[11]

$$Var[\Pi \mid \vec{S}] = Var[\Pi] - \mathbf{Cov}[\Pi, \vec{S}]' \mathbf{Var}^{-1}[\vec{S}] \mathbf{Cov}[\Pi, \vec{S}]$$

[10] The certainty equivalent for multinormal random variables is

$$E[\exp(\mathbf{w}^T \mathbf{A} \mathbf{w} + \mathbf{b}^T \mathbf{w} + d)] = |\mathbf{I} - 2\boldsymbol{\Sigma}\mathbf{A}|^{-1/2} \exp(\underbrace{\frac{1}{2} \mathbf{b}^T (\mathbf{I} - 2\boldsymbol{\Sigma}\mathbf{A})^{-1} \boldsymbol{\Sigma} \mathbf{b} + d}_{Certainty\ Equivalent}],$$

where $\mathbf{w} \sim \mathcal{N}(0, \boldsymbol{\Sigma})$ and the (co)variance matrix is $\boldsymbol{\Sigma}$ positive definite. $\mathbf{A}$ is a symmetric $m \times m$ matrix, $\mathbf{b}$ is an $m$-vector and $d$ is a scalar. Note the left-hand side is only well-defined if $(\mathbf{I} - 2\boldsymbol{\Sigma}\mathbf{A})$ is positive definite. See also the discussion of Brown and Jennings (1989) in Section 2.4.1 or Anderson (1984, Chapter 2).

[11] Note, if $\Pi$ is a vector $Var[\Pi \mid \vec{S}]$ is the conditional covariance matrix.

For the special case with signals of the form,

$$S_t^i = \Pi + \epsilon_{S,t}^i,$$

where the distribution of the liquidation value $\Pi \sim \mathcal{N}(0, \sigma_\Pi^2)$ and the error terms $\epsilon_{S,t}^i \sim \mathcal{N}(0, \sigma_{S,t}^2)$ are mutually independent, the Projection Theorem simplifies to:

$$E[\Pi \mid \underbrace{S_1^i, ..., S_T^i}_{:=\mathcal{F}_T^i}] = E[\Pi] + \underbrace{\frac{1}{\frac{1}{\sigma_\Pi^2} + \sum_t \frac{1}{\sigma_{S,t}^2}}}_{Var[\Pi|S_1^i,...,S_T^i]} \sum_t [\frac{1}{\sigma_{S,t}^2}(S_t^i - E[\Pi])].$$

The conditional variance $Var[\Pi \mid S_1^i, ..., S_T^i] =: V^c$ is linked to the original variances through the following expression:[12]

$$\frac{1}{V^c} = \frac{1}{\sigma_\Pi^2} + \sum_t \frac{1}{\sigma_{S,t}^2}.$$

The information set of an agent $i$, $\mathcal{F}_T^i$ is given by $\{S_1^i, ..., S_T^i\}$.

Note, the Kalman filter is also derived from the Projection Theorem. The Kalmam filter technique is especially useful for steady state analysis of dynamic models. The problem has to be brought in state-space form

$$z_{t+1} = Az_t + Bx_t + \epsilon_{t,1},$$

$$S_t = Cz_t + \epsilon_{t,2},$$

where, the error terms are i.i.d. normally distributed. The first equation is the transition equation, which determines how the state vector $z_t$ moves depending on the control vector $x_t$. The second equation is the measurement equation, which describes the relationship between the signal $S_t$ and the current state $z_t$.

### 2.3.1 Limit Order Models

Most of the REE models and all the models discussed so far are limit order models. In these models each trader submits a whole demand schedule. Such a demand schedule can be achieved by combining many stop and limit orders. Investors can therefore trade conditional on future prices. Therefore a trader's information set contains not only her signal but also current prices, which influence the optimal individual demand. Within the class of limit order models one can distinguish between competitive and strategic models.

[12] The reciprocal of the variance is called the precision.

### Competitive Limit Order Models

In competitive models traders take the price as given when forming their optimal demand. Investors neglect that their trading activity influences the price, which in turn serves as a signal. In these models traders do not attempt to manipulate prices. To justify such behaviour one could assume that each trader is only a point in a "continuum of clones" with identical private information.[13]

Grossman (1976) presents one of the first models with a closed form REE solution, where information about the return on a single risky asset is dispersed among traders. Every trader receives a noisy signal about the true payoff $\Pi$, $S^i = \Pi + \epsilon^i_S$, where $(\epsilon^i_S)^I_{i=1}$ are mutually independent and identically normally distributed. The riskless asset is traded at an exogenously fixed price with perfectly elastic supply. Using her signal and the price, each trader is solving the signal extraction problem described above to derive her optimal demand. The market clearing condition then provides the market clearing price. In Grossman's example, the equilibrium is informationally efficient, i.e. the equilibrium price is a sufficient statistic for all signals $\{S^i\}^I_{i=1}$. Therefore individual demand schedules do not depend on their private signal which leads to the Grossman-Paradox discussed earlier. A further consequence of the information revelation of prices is that no one has an incentive to acquire costly information. Thus an overall equilibrium with costly endogenous information acquisition does not exist (Grossman-Stiglitz-Paradox). It seems plausible that individual demand does not depend on traders' incomes since all traders have CARA utility functions. At first sight it is surprising that traders' demands do not even depend on the equilibrium price itself, although it serves as a sufficient statistic for all information in the economy. A price change in a REE has not only an income and substitution effect but also an information effect. A price increase signals a higher expected payoff of this asset in an economy with a single risky asset. With CARA utility functions and one risky asset, the income effect should not play any role. In a symmetric equilibrium with fixed aggregate supply, all traders have to demand the same number of assets, independent of the price, in order to satisfy the market clearing condition. Therefore, in Grossman's model the substitution effect and information effect cancel each other out. Finally as discussed in Section 2.2.3, informationally efficient REE can lead to a no-trade outcome.

[13] An excellent overview of limit order models is given in Admati (1989).

Noisy REE models were developed to address these conceptual problems. A random variable whose unobserved realisation affects the equilibrium price is introduced in these models. This noise term makes prices only partially revealing because traders cannot infer whether a price change is due to the noise component or due to informed trading. The informational content of prices can be measured by a signal to noise ratio. In Grossman and Stiglitz (1980) the aggregate supply of the risky asset is random. In their model there are only two groups of traders: the informed (those who bought an identical signal) and the uninformed. Since the supply of risky assets is random, uninformed traders can only partially infer the signal of the informed. Each trader decides whether to acquire information at a certain cost. Grossman and Stiglitz (1980) derive the overall equilibrium with endogenous information acquisition and determine the fraction of informed traders in equilibrium. Their model captures the partial information transmission role of prices, but not the information aggregation role since information is not dispersed among the traders.

This additional aspect is analysed in Hellwig (1980) and Diamond and Verrecchia (1981). Similar to Grossman (1976) the signals are conditionally independent of each other given the true payoff. Whereas in Hellwig (1980) the aggregate supply of the risky asset is assumed to be random, in Diamond and Verrecchia (1981) each investor's endowments are i.i.d. and therefore aggregate supply is random as long as the number of traders does not converge to infinity.[14] For both models a closed-form solution is found where prices are only partially revealing. Hellwig shows that the REE in the "high noise limit" (where the variance of aggregate endowments goes to infinity) corresponds to the equilibrium in which market participants do not try to learn something from the equilibrium price. On the other hand the REE at the "low noise limit" corresponds to an informationally efficient REE as in Grossman (1976). The same is true when investors are almost risk neutral, since investors do not try to insure against the randomness of aggregate supply.

Incorporating noisy aggregate supply or noisy excess demand through random endowments can be thought of as a simplified reduced form for modelling liquidity traders. Liquidity traders trade for reasons exogenous to the model or due to information of private interest. This is in general the case in a setting with incomplete equitisation. In Wang (1994) informed investors trade not only for informational reasons but also in order to in-

[14] Non-i.i.d. endowments are mathematically untractable and are excluded in our discussion.

vest in private investment opportunities. These private investment opportunities are not equitisable, i.e. conditional trade on their dividend streams is not possible. There will always be some components of trade that are not perfectly predictable by others and not perfectly correlated with the future payoff of the traded assets. The latter is conducted by liquidity/noise/"life-cycle" traders.

Admati (1985) extended Hellwig's setting to a model with multiple risky assets and infinitely many agents. In this model, the price of an asset does not necessarily increase with its payoff or decrease with its actual supply. This is the case because a price change in one asset can provide information about other risky assets. Admati's model illustrates that not only the correlation between financial assets' returns (which is the focus in CAPM), but also the correlation between the prediction errors in traders' information is important for determining equilibrium relations.

The main focus in Pfleiderer (1984) is the role of volume and variability of prices. He analyses how a change in the signal's precision alters expected volume. His results are extended in He and Wang (1995) which is discussed in further detail in Section 2.4.3.

Admati and Pfleiderer (1986), (1990) analyse how an information monopolist should sell his information to competitive limit order traders. The more this information is revealed by the price, the lower is the traders' incentive to pay for this information. Admati and Pfleiderer show that it is optimal for a seller to add noise to his information when his information is very precise. This increases the fraction of market participants that would be willing to pay to become better informed. When the number of traders is large, it is better to sell personalised signals, i.e. signals with an idiosyncratic noise term. In this case, the information monopolist sells identically distributed signals to all traders and not only to a fraction of the market participants. The information monopolist can also sell his information indirectly by using it to create an investment fund. Admati and Pfleiderer (1990) show that the fund manager always makes full use of his information. They also illustrate that the degree to which information is revealed by the market price determines whether an indirect sale or direct sale of information leads to higher revenue for the information seller.

Subsection 2.5.4.2 will discuss competitive limit order models with market makers. In these models the group of risk neutral market makers observe only the limit order book, i.e. the noisy aggregate demand schedule. Since they are risk neutral they act as a competitive fringe and thus their information set determines the equilibrium price.

### Strategic Limit Order Models

Traders behave "schizophrenically" in a competitive REE. On the one hand each traders' signal can influence her demand, but her demand has no impact on the price. On the other hand, this price reveals her signal. The competitive environment can be justified when the number of investors becomes very large and each investor is infinitesimal small.

It is, however, a stylised fact that the best-informed traders are large. The traders take into account the effect that their trading has on prices in strategic limit order models. Each trader knows that when she trades larger quantities *prices will move against her.* She therefore incorporates this effect in forming her optimal demand correspondence. Thus, strategic models allow us to analyse market price manipulation by some large traders.

In Kyle (1989) a strategic REE is derived as a symmetric Bayesian Nash equilibrium in demand schedules. Each trader's strategy is a demand schedule which is submitted to an auctioneer. The auctioneer collects all individual demand schedules and derives the market clearing price. Kyle's model is actually a uniform price auction of a divisible asset, whose supply is random. Given CARA-utility functions and normally distributed random variables all excess demand functions are linear. In Kyle's model the informed traders' information set consists of a private individual signal about the true value of the asset and the price. Each demand correspondence of an informed investor is linear in her individual signal as well as in the price. The demand function of the uninformed traders is also linear in price. The fractions of informed investors and uninformed investors is common knowledge in equilibrium. In a Nash equilibrium each player takes the strategies of all others as given.[15] Therefore each trader faces a residual supply curve. Each informed trader $i \in \{1, ..., I\}$ acts like a monopsonist with respect to the residual supply curve

$$p = p_I^i + \lambda_I x^i,$$

where $p_I^i$ is random and $\lambda_I$ is constant. Since informed traders observe $p_I^i$ they choose their demand $x^i$ to maximise their expected utility conditional on $p_I^i$ and on their private individual signal $S^i$. The reciprocal of $\lambda_I$ can be viewed as "market depth", the liquidity

[15] Equilibrium strategies are mutual knowledge in a Nash equilibrium, Brandenburger (1992).

of the market. Whereas in competitive models the aggressiveness of informed traders is only restricted by their risk aversion, in strategic models risk aversion and consequences of strategic behaviour on price cause agents to react less aggressively. They try to avoid trading their informational advantage away. It is also worth noticing that individuals' demand is no longer independent of their initial endowment in strategic models.

In strategic models the prices reveal less information than in a competitive REE which facilitates costly information acquisition. Even in the limit, when noise trading vanishes or traders become almost risk neutral, prices do not become informationally efficient. However, this does not mean that the profit derived from private information (the information rent) is not driven down to zero. Kyle (1989) also shows that as the number of informed speculators increases to infinity the model converges to a monopolistic competition outcome which need not be the same as in a competitive environment. The case with a single information monopolist and many competitive outsiders is analysed in more detail and a tractable closed-form solution is derived. Bhattacharya and Spiegel (1991) derived the No-Trade Theorem discussed in Section 2.2.4 for this special case.

Jackson (1991) shows that the Grossman-Stiglitz Paradox depends crucially on the price taking behaviour of the traders. He develops a strategic limit order model in which a finite number of risk neutral traders submit demand functions. Thereby he models explicitly the price formation process, which illustrates how the signal is incoporated into the price. For specific parameters, in the Bayesian Nash equilibrium of this game costly information acquisition occurs although the price is informationally efficient. In other words, although some agents bear information acquisition costs, they do not have any informational advantage. In this situation, they acquire information because they are driven by the beliefs of the other agents about their information acquisition. Allowing for mixed strategies in a Bayesian Nash equilibrium also resolves the Grossman-Stiglitz Paradox.

Madhavan (1992) compares this setting with a two stage game developed in Glosten (1989) where dealers first quote a price (schedules) and in the second stage investors submit their orders. He wants to illustrates the difference between a quote-driven market such as NASDAQ or SEAQ and an order-driven market, capturing some features of the NYSE.

### 2.3.2 Market Order Models

When an investor submits a single limit order she faces execution risk. She cannot be sure that her order will be executed, since the price can move beyond the set limit. She can avoid this risk by submitting a menu of limit and stop orders such that at each price a certain quantity of orders will be executed. Alternatively, the execution risk can be avoided by using market orders. However, then only the quantity of trade can be fixed and the agent has to bear the risk of changing prices. This additional price risk can complicate the economic analysis and therefore, in most models, risk neutrality of all traders is assumed. As seen before, risk aversion is not needed in strategic models, since not only risk aversion but also strategic behaviour causes investors to trade less aggressively. It can be shown that limit order and market order models exhibit the same degree of informational efficiency in the case of a single informed risk neutral investor, Nöldeke (1993).

If the trader submits her market order before the market maker sets the price, the price risk is completely borne by the investor. In the case where the market maker first sets the price, he commits himself and therefore bears the risk. We will discuss strategic models in which the investors have to move first before turning to models in which the market maker moves first. Some market order models without an actively trading market maker are briefly summarised, before discussing these models.

#### Competitive Market Order Models Without a Market Maker

Hellwig (1982) uses a market order model in order to resolve the Grossman-Stiglitz Paradox. In Hellwig's dynamic model traders can only trade conditional on the past prices and not on the current price. Therefore statistical inference and market clearing do not occur simultaneously. In Hellwig (1982) a null set of the continuum of traders receives information in advance. In discrete time this information is only revealed by the price one period later. This gives the insider the possibility to make use of their information to achieve a positive return. Therefore, traders have an incentive to acquire information. Even as the time span between the trading rounds converges to zero, the insider can make strictly positive returns and an information efficient outcome can be reached arbitrary closely. In Hellwig (1982), traders are myopic and the individual demands are exogenously given rather than derived from utility maximisation. Blume, Easley, and O'Hara (1994) analyse the informational role of volume within such a framework. This model is covered in Section 2.4.4.

### Strategic Market Order Models With Market Maker in Which the Investors Move First

The classical reference for this class of models is Kyle (1985).[16] In his batch clearing model there are three groups of risk neutral players, a single informed investor, many liquidity traders and a market maker who sets the price. The liquidity traders trade for reasons exogenous to the model. Their demand is given by the random variable $\Theta \sim \mathcal{N}(0, \sigma_\Theta^2)$. The single, risk neutral, information monopolist is the only one who knows the true value of the risky asset, $\Pi$. He trades to maximise his profit which is in the static single auction version of the model, the capital gain $(\Pi - P_1)$ times the quantity of stocks, $x$, that he holds. Since he acts strategically, he takes into account that his demand $x$ will influence the price, $P_1$. The informed trader rationally believes that the market maker follows a price setting rule which is linear in the aggregate net order flow $(x + \Theta)$. Formally he maximises his profit $\pi = (\Pi - P_1)x$, where, according to his beliefs, $P_1 = P_0 + \lambda(x + \Theta)$. The single market maker only observes the aggregate net order flow $(x+\Theta)$ and knows that the true value, $\Pi$, is distributed $\mathcal{N}(P_0, \sigma_{\Pi,0}^2)$. Since he cannot observe the net trade vector, $\mathbf{x}$, of the informed investors the No-Trade-Theorem, explained in Section 2.2.3, does not apply. Kyle assumes that the risk neutral market maker acts competitively and thus sets a fair price given his information, i.e. $P_1 = E[\Pi \mid x + \Theta]$. Since $\Pi$ and $\Theta$ are normally distributed, and with insider's demand $x$ is linear in $\Pi$, the Projection Theorem can be applied to solve the signal extraction problem. The Bayesian Nash equilibrium is obtained by equating the coefficients and is given by

$$P_1 = P_0 + \lambda(x + \Theta), \text{ where } \lambda = \frac{1}{2}\left(\frac{\sigma_\Theta^2}{\sigma_{\Pi,0}^2}\right)^{(-1/2)}.$$

$\lambda$, the amount of noise trading, $(\sigma_\Theta^2)$, together with the original variance of $\Pi$, $\sigma_{\Pi,0}^2$, determines to what extent the market maker reacts to a higher/lower aggregate net order flow. The reciprocal of $\lambda$, $(1/\lambda)$ represents the market depth. If on average a lot of uninformed noise trading is going on, the market maker will not adjust the price so quickly if he observes a high order flow. Therefore in this case markets are deep, i.e. many orders can be absorbed without huge price movements. The expected profit for the insider is given by

[16] In comparision to the original article the notation is: $\Pi = v$, $\sigma_{\Pi,0}^2 = \Sigma_0$, $\Theta = u$.

$E[(\Pi - P_1)x] = \frac{1}{2}(\sigma_\Theta^2 \sigma_{\Pi,0}^2)^{(1/2)}$. The market maker breaks even on average. He loses money to the insider but makes the same amount of money from the noise traders on average. After one trading round information is partly revealed and the new variance of the true value of the stock is only half of the original one. Kyle extended this static model to a series of discrete call markets (a sequential auction). In this dynamic setting, the insider faces the trade-off that taking on a larger position in early periods increases early profits but worsens prices in later trading rounds. She tries not to trade her information advantage away. She therefore exploits her information across time by hiding behind noise trading. A dynamic linear recursive equilibrium is derived in Kyle (1985). The author solves the dynamic problem by porposing an ad hoc value function, which he verifies at a later stage. Note the insider takes the equilibrium $\lambda_t$ as given, since the market maker can not determine whether the observed aggregate order flow is due to a deviation of the insider or due to a different signal realisation or noise trader demand. As the time intervals converge to zero in the continuous auction equilibrium, noise trading follows a Brownian motion and the informed trader continuously pushes the price towards her price valuation. The speed of price adjustment is equal to the difference between her price valuation and the current price divided by the remaining trading time. The market depth, $(1/\lambda)$, is constant over time and the market is "infinitely tight", i.e. it is extremely costly to turnover a position in a very short period of time. This is the case because the insider can break up her informational trade into many tiny pieces. Prices follow Brownian motion (which is a martingale process).

Biais and Rochet (1997) show that for the case where the value of the stock is not continuously distributed, out of equilibrium beliefs have to be specified and one has to deal with the problem of multiple equilibria.

Back (1992) extended Kyle's continuous time model by modelling strategy spaces and information directly in continuous time. In Holden and Subrahmanyam (1992) there are many informed traders who compete against each other. This speeds up information revelation through prices. As in the Cournot case, insiders who have the same information are more aggressive and, therefore, trade more of their insider information away. In a dynamic setting the information is revealed immediately as time becomes continuous. In Holden and Subrahmanyam (1994) the insiders are risk-averse. This further speeds up information revelation. Risk-averse agents trade more aggressive in early periods since future prices are more uncertain. Foster and Viswanathan (1996) develop a model, where informed traders observe different signals, which will be discussed in Section 2.4.6. In all these models the focus is on the price process and information revelation. The Bid-Ask Spread is the focus of the next section.

### Sequential Trade Models in Which the Market Maker Moves First

In limit order models all traders submit whole demand schedules. In the market order models discussed above the market maker sets the price after observing the total net order flow. In this section we discuss models in which the market maker has to set the price before he observes orders. He will therefore set the price conditional on the magnitude of the market order. In other words, the market maker sets a whole supply schedule and then the investors choose their optimal market order. If some traders want to buy a large number of shares then the market maker asks for a higher "ask price", since the investor could have superior information. Similarily if someone wants to sell[17] a large number of shares he offers a lower "bid price". In the sequential trading model in Glosten and Milgrom (1985)[18] the order size is fixed to one share at a point in time. Therefore, there is a single ask and a single bid, where the spread is defined as the difference between ask and bid. In Glosten and Milgrom's model there is a continuum of traders. A fraction $\mu$ is informed and a fraction $(1-\mu)$ is uninformed. Informed traders do want to trade when their expected value of the asset is strictly larger than the ask or strictly smaller than the bid. Uninformed traders trade for reasons exogenous to the model. They buy or sell one stock randomly with equal probability independent of the information. It is further assumed that the market maker and all traders are risk neutral.[19] Note, in this setting, submitting an order does not change the price at which the order will be executed. The market maker had set this price in advance and thus the order can only influence the future price development. Furthermore he does not care about the future price development, since the probability that the same trader has a chance to trade again is zero. Thus each trader would like to trade an infinite

[17] A sell order is considered as a negative buy order.

[18] In this survey we will follow the version in Nöldeke (1993).

[19] Risk neutrality of the market maker abstracts from inventory models. In inventory models the market makers can end up with a non optimally diversified portfolio at the end of the trading day. They therefore demand a spread Ho and Stoll (1981).

amount of the stock or not trade at all if the spread is too large. By assumption each trader is restricted to trade only one share. Moreover, in this simplified version, informed traders know the true value of the asset, which is either 0 or 1.[20] If the true value is 0 they sell when the bid price is larger than zero and accordingly when the true value is 1 they buy when the ask price is smaller than one. The specialist observes the buy and sell orders and consequently updates his beliefs about the asset's value using Bayes' Rule. Since the asset's value is either 0 or 1 his conditional expected value is equal to his probability that the true value of the asset is 1. The Bayesian Rule also exhibits that a "no trade event" will not alter his beliefs. The same is true if he observes the same number of buy and sell orders. Thus the market imbalance (the difference between sell and buy orders), is a sufficient statistic for the whole history of market orders.

Glosten and Milgrom (1985) assume that the market maker sets the ask and bid such that his expected profit on any trade is zero. The existence of at least one potential competitor combined with risk neutrality makes this assumption reasonable. Therefore the specialist sets the price equal to his belief that the true value of the asset is 1. Since he can set the price conditional on the next order he takes it into account. The bid price is therefore his belief, given the current market imbalance plus an additional sell order, whereas he sets his ask according to his beliefs given the current imbalance minus one stock, (an additional buy order). The market maker needs this spread to break even since he faces an adverse selection problem. If the fraction of informed traders increases, the adverse selection problem becomes more severe and therefore a wider spread is needed. On the other hand, a higher number of informed traders also increases the speed of information revelation. This analysis implies that the midpoint between ask and bid is not the current expected value for the market maker unless his current expected value is 0.5. Consider the case where the current expected belief is above 0.5. An additional buy order has less informational content than a sell order and, therefore, the midpoint is biased downwards. As the market imbalance increases in absolute terms, i.e. there are more buy orders than sell orders or more sell orders than buy orders, the market maker becomes more certain about the information of the insiders and therefore the size of the spread falls. The transaction price is a Martingale

[20] Strictly speaking, these models are not CARA-Gaussian models. Agents have a constant absolute risk aversion coefficient of zero since they are risk neutral, but the true asset return is not normally distributed. The return distribution can be easily modified to a normal distribution.

but not the quoted prices (ask and bid). The latter are only Markov but not Martingales since any additional trading round leads in expectation to more information for the market maker which tightens the spread over time.[21] Thus the spread size is not a Martingale and consequently it cannot be the case that bid and ask are Martingales. Moreover this model also exhibits serial correlation of order flows.

Easley and O'Hara (1987) extend Glosten and Milgrom's sequential trading model in two ways. In Glosten and Milgrom (1985) the supply schedule which the market maker posts is reduced to one unit of purchase and one unit of sales. Easley and O'Hara allow two different order sizes, small and large orders. Furthermore they introduce the concept of "event uncertainty". Only with probability $\alpha$ will the information structure be as in Glosten and Milgrom (1985). With probability $(1-\alpha)$ an information event does not occur and only uninformed traders trade with each other. Neither the market maker nor the uninformed traders know the true value of the stock. They also ignore whether some traders are informed or not. This model incorporates higher order knowledge since the depth of knowledge of the state space is higher by one degree.[22]

In Easley and O'Hara (1987) nature chooses once at the beginning of the trading day whether an information event happens or not. If information is released the pool of infinitely many traders contains a fraction $\mu$ of informed and a fraction $(1-\mu)$ of uninformed traders. In the other case only uninformed traders are in the pool. Uninformed investors trade for exogenous reasons and take no information aspect into account. They submit large and small orders in an ex-ante specified probabilistic way. Informed traders always prefer to trade large quantities if both quantities are traded at the same price. Informed traders do not act strategically concerning the future price path. They do not take into account that trading a large quantity can influence the future price process. This is justified since there are infinitely many informed traders and thus the probability that an individual trader has the chance to trade again is zero. However, they choose their optimal quantity which is exogenously restricted to either 1 or 2 units. Since at equal prices informed traders prefer to trade larger quantities, the market maker will set a larger spread for the large trades, e.g.

[21] A price process is Markov if a single state, the current price, can represent the whole history. It is a Martingale if the expected future prices are equal to the current price.

[22] Concerning depth of knowledge see Section 2.2.4.

block trades.

Depending on the parameter constellation two types of equilibria can arise. In a separating equilibrium all informed traders prefer to trade two shares, the large quantity, despite the larger spread. Uninformed traders submit market orders for one and two stocks, as exogenously specified. In this separating equilibrium, the spread for market order of size one is zero. In a pooling equilibrium, informed traders submit small and large orders and the market maker requires a spread for both quantities, the larger spread for the block trades. The market maker's uncertainty about whether there was an information release dictates that both the size and the sequence of trades matters. Incorporating this feature can help to explain the partial price recovery that characterises most block trading sequences. The impact of an "event uncertainty" is discussed in more detail in Easley and O'Hara (1992) for the simpler case where the trade size is restricted to one unit as in Glosten and Milgrom (1985). In the case of event uncertainty, the actual trade or no trade is a signal about whether there was some information released to the insiders at the beginning of the trading day. By observing the sequence of market orders the market maker can update his beliefs, not only about the true value of the asset, but also whether the insiders got information about this true value. Absence of trade, therefore, provides a signal and thus the time per se is not exogenous to the price. If the market maker observes a no-trade outcome, he increases his beliefs that nobody has any information and therefore the quotes will be pulled toward 1/2. If the midpoint is at 1/2, observing no trade makes asymmetric information less likely and therefore leads to a lower spread. If the spread is not straddling 1/2 the effect is not so obvious. Further results of their analysis are that the last transaction price is not a sufficient statistic for the past and thus the transaction price process is a Martingale, but is no longer Markov.

Glosten (1989) relaxes the restriction that order sizes are limited to one or two units. Therefore, the market maker quotes a whole price schedule instead of a single bid and ask price. Glosten (1989) compares the dealership market structure consisting of many competitive market makers with the 'specialist system' in which all investors exclusively trade through a monopolistic specialist. He shows that although the monopolistic market maker makes a positive expected profit, under certain circumstances the 'specialist system' provides a higher market liquidity than the competitive system. Glosten (1989) focuses on a one shot interaction between a strategic risk averse trader and the market maker. The trader has a exponential utility function and faces an (normally distributed) endowment shock. In addition, he receives a noisy signal about the liquidation value of the stock, $S = \Pi + \epsilon$, where $\Pi$ and $\epsilon$ are independently normally distributed. He trades, therefore, for liquidity/insurance as well as informational reasons. Given the price schedule set by the market maker(s), the trader submits his utility maximising market order. Before the trader submits his order, the market maker(s) commit(s) himself (themselves) to a price schedule. In the case of competition among market makers they are forced to set a informationally efficient price schedule $P(x) = E[\Pi|x]$. Note, in contrast to Kyle (1985), in Glosten (1989) the market maker sets a price function for a forthcoming *single* transaction which stems from investors who trades for informational as well as for hedging/liquidity reasons. Glosten shows that the more extreme a position the investor wishes to take, the more likely it is that he trades for informational reasons. The market makers, therefore, have to protect themselves by making the price schedule steeper. For extremely large orders, market makers are unable to protect themselves and, therefore, the market closes down. On average, market makers profit from trading with investors with extreme endowment shocks, since they trade for re-balancing reasons and lose to those traders with small endowment shocks who trade for informational reasons. Rothschild and Stiglitz (1976) suggest that an existence problem exists in a setting where there is a continuum of types of investors. Glosten defends his setting because of its tractability and qualitative similarity to a discretised version. Hellwig (1992) shows that the non-existence of fully revealing outcomes in any signalling model arises because of the unbounded type space.

A monopolistic specialist commits himself to a different price schedule, which is determined by

$$\arg \max_{P(\hat{x}(\cdot))} E[P(\hat{x}(\cdot))\hat{x}(\cdot) - \Pi\hat{x}(\cdot)],$$

where $\hat{x}(\cdot)$ is the optimal order size (function) of the trader depending on his endowment shock and his information. In contrast to the competitive market maker case, the monopolist has the ability to cross-subsidise different order sizes. In equilibrium he earns a larger profit from more likely small trades, but makes losses on unlikely large trades. The large trades are unlikely to occur, but likely to result from information based trading. By keeping the price of large trades relatively low, the specialist guarantees that traders with extreme signals do not reduce their trade size in order to pool with trades with less extreme signals. This is the reason why a market structure with a monopolistic specialist stays open for larger trade sizes than a market with multiple market makers. The problem a single market maker faces can also be viewed as a principal-agent-problem. The principal (specialist) sets a menu of contracts $(x, P(x))$ from which the agent chooses the one which maximises his expected utility. This is a mechanism design problem for the market maker, which was analysed as such in Rochet and Vila (1994) with the difference that they consider exogenous noise trading and allow for more general distributions.

The models considered so far avoid any strategic interaction between market makers. They just assume that competition among market makers leads to zero profit in expectations. Dennert (1994) explicitly analyses this interaction. In the first stage, all market makers set a bid and ask price and commit themselves to trade up to an exogenously specified number of shares. In the second stage, the trader chooses his optimal demand. If he trades for liquidity reasons, he will trade with the market maker(s) who offer(s) the best price. An informed trader, on the other hand, trades with many market makers simultaneously as long as it is profitable for him. An increase in the number of registered market makers leads to an increase in informed trading and, thus, increases the transaction costs for the liquidity trader. In the second part of this paper, market makers set whole price schedules in stage one. The analysis exhibits elements of a first-price auction for divisible goods.

Biais, Martimort, and Rochet (1997) model imperfect competition between market makers within the framework of mechanism design theory and make use of the tool of variation calculus.[23] The risk-neutral market maker(s) set the price schedules in the first stage. The risk-averse insider submits market orders possibly to all market makers. The insider's market order depends on his signal and his endowment shock. While Glosten (1989) analyses only the extreme cases of perfect competition and monopoly, Biais, Martimort, and Rochet (1997) analyse the more general case where the number of market makers is finite. The

[23]This paper departs from the normality assumption.

authors derive a unique equilibrium in which the unit price of the shares is increasing. They show that the equilibrium trading volume is below the optimal risk sharing level but higher than in the monopoly case. Competition among market makers leads to a deeper market, which was not necessarily the case in Dennert (1994). Market makers face limited competition due to adverse selection. Market makers are reluctant to undercut each other since this exposes them to a greater extent to disadvantageous informational trade. The intuition is similar to the winner's curse in auction theory. Even as the number of the market makers tends to infinity, a strictly positive bid-ask spread remains and the sum of market makers' profit is still strictly positive. This limiting case is similar to the analysis of Glosten (1994). Uninformed traders, who submit limit orders to a public limit order book before the insider submits his market orders, face the same problem. Glosten (1994) studies the case of perfect competition in limit orders.

Almost all existing models can be grouped in the five categories outlined above. A nice overview about the existence of linear equilibria in static models is provided by Bagnoli, Viswanathan, and Holden (1994). Most dynamic models have the common feature that the price adjusts instantaneously to public information but only gradually to private information. This gradualism is caused by noisy asset supply, strategic behaviour of informed traders or is exogenously given by assuming a sequential trading mechanism where traders are restrained to trade only a certain quantity.

## 2.4 Further Dynamic Models, Crashes and Technical Analysis

This section covers some dynamic models in more detail. The first subsection deals with competitive two period models in which all traders are price takers and all traders receive their information at the same time. These models shows that past prices sill have informational value. Section 2.4.2 introduces the Infinite Regress problem and it demonstrates how learning can lead to serial correlation of variables. Multi-period models are analysed and the informational content of volume data is also illustrated in this Section 2.4.4. The value of technical analysis in models in which different traders receive information at different times is the focus of Section 2.4.5. Finally, Section 2.4.6 covers strategic market order models based on the seminal work of Kyle (1985).

### 2.4.1 Competitive Two-Period Limit Order Models, Technical Analysis and Crashes

In Grundy and McNichols (1989) and Brown and Jennings (1989) two simple competitive limit order models are developed in which not only the current price, but also the past price is useful in predicting the value of the asset. In these competitive limit models technical analysis has positive value. Grossman's (1976) model suggests that capital markets are strong-form informationally efficient, i.e. the revelation of any private information will not change the equilibrium. The noisy REE discussed below are not even weak-form informationally efficient, following Fama's definition, Fama (1970) (1976), i.e. the current price is not a sufficient statistic for all past prices. There are, however, alternative definitions of weak-form efficiency whose conditions are satisfied by these REE.

Brown and Jennings (1989) extend a model similar to Hellwig (1980) to two periods.[24] In their model there are infinitely many a priori identical investors, denoted by $i \in \mathbb{I} = \{1, 2, 3, ...\}$ who are endowed with $B_0$ units of the riskless asset. $B_0$ can be normalised without loss of generality to zero, since all investors have CARA utility functions. All investors start with the same information set, $\mathcal{F}_0$, with beliefs about the liquidation value of $\mathcal{N}(\mu_{\Pi,0}, \sigma^2_{\Pi,0})$. At $t = 1$ and $t = 2$ each investor gets a private signal about the liquidation value, $\Pi$, of the risky asset in $T = 3$, i.e.

$$S_t^i = \Pi + \epsilon^i_{S,t},$$

where $\epsilon^i_{S,t}$ is normally i.i.d. with $\mathcal{N}(0, \sigma^2_{S,t})$. As the signals are unbiased, by the Law of Large Numbers, the average signal $S_t = \lim_{I \to \infty} \frac{1}{I} \sum_{i=1}^{I} S_t^i$ equals $\Pi$ with probability one in each $t$. Trader $i$'s information set is given by $\mathcal{F}_1^i = \{\mathcal{F}_0, S_1^i, P_1\}$ in $t = 1$ and $\mathcal{F}_2^i = \{\mathcal{F}_1^i, S_2^i, P_2\}$ in $t = 2$. The information sets contains the current price $P_t$ since, in a limit order model, traders can trade conditional on the price of the stock $P_t$. Let trader $i$'s stock holding in $t$

[24] As far as possible we will follow the notation in He and Wang (1995) in order to make the papers comparable.

be denoted by $x_t^i$. His final wealth in period $T = 3$ is then given by

$$W_3^i = B_0 + x_1^i(P_2 - P_1) + x_2^i(\Pi - P_2),$$

where $B_0$ is normalised to zero. Traders' expected utility functions are given by

$$E[-\exp(-\rho W_3^i) \mid \mathcal{F}_0],$$

where the constant absolute risk aversion measure $\rho$ is the same for all traders. Each trader maximises his expected utility, given his information set and his price conjecture. Backward induction allows us to break up this optimisation process into two steps. Given a certain $x_1^i$ the maximum utility value at $t = 2$ is given by

$$J_2^i(x_1^i) = \max_{x_2^i} E\{-\exp[\rho(x_1^i(P_2 - P_1) + x_2^i(\Pi - P_2))] \mid \mathcal{F}_2^i\}.$$

At $t = 1$ the problem is

$$J_1^i = \max_{x_1^i} E\{J_2^i(x_1^i) \mid \mathcal{F}_1^i\}.$$

A REE is then given by the equilibrium prices $P_1$ and $P_2$ which have to coincide with the traders price conjectures as well as the optimal stock holdings $(x_1^i, x_2^i)$ for each investor $i$. The market clearing condition guarantees that demand equals supply in both periods. Whereas the average per capita demand for the risky asset is given by $x_t = lim_{I\to\infty} \sum_{i=1}^{I} x_t^i/I$, the per capital supply is assumed to be random in this noisy REE. The random supply is given by $\Theta_1$ in $t = 1$ and $\Theta_2$ in $t = 2$, where $\Theta_2 = \Theta_1 + \Delta\Theta_2$.[25] Brown and Jennings assume that $\Theta_1$ and $\Delta\Theta_2$ are normally distributed

$$(\Theta_1, \Delta\Theta_2) \sim \mathcal{N}[(0,0), \begin{pmatrix} \sigma^2_{\Theta_1} & \varrho\sigma_{\Theta_1}\sigma_{\Delta\Theta_2} \\ \varrho\sigma_{\Theta_1}\sigma_{\Delta\Theta_2} & \sigma^2_{\Delta\Theta_2} \end{pmatrix}],$$

where $\varrho$ is the correlation between the supply increments $\Theta_1$ and $\Delta\Theta_2$. Conditioning trade in $t = 2$ on $P_1$, i.e. technical analysis, has positive value for two reasons. First, $P_1 = \mathbf{L}[\Pi, \Theta_1]$ provides a useful second signal about the true payoff $\Pi$. $\mathbf{L}[\cdot]$ is a linear operator. This effect is most pronounced in the case, where $\Theta_1$ is independent of $\Theta_2$. Second, if $\Theta_1$ and $\Theta_2$ are correlated the price $P_1$ in $t = 1$ helps to get a better prediction of $\Theta_1$. $\Theta_1$, in turn, is useful in predicting $\Theta_2$. A better prediction of $\Theta_2$ reduces the noise in $t = 2$ and thus allows $P_2$

[25] Note, that $x_t$ denotes holdings rather than additional trading demand, whereas $\Theta_1$ and $\Delta\Theta_2$ refer to additonal supply.

to reveal more about the liquidation value $\Pi$. Furthermore, knowing $P_2$ also allows to get a better prediction of $\Theta_1$. Thus a joint estimation using both price conjectures $P_1$ and $P_2$ enhances information revelation. Grundy and McNichols (1989) show that for the case of perfect correlation, i.e. $\Delta\Theta_2 = 0$, $P_1$ and $P_2$ perfectly reveal $\Pi$. Therefore $P_1$ has predictive value even in $t = 2$. Note, $\Theta_1$ and $\Theta_2$ are still correlated, even if $\varrho = 0$, since $\varrho$ is defined as the correlation coefficient between $\Theta_1$ and $\Delta\Theta_2$. In this case $\Theta_t$ follows a random walk and the prediction of $\Theta_1$ using $P_1$ provides the expectation of $\Theta_2$.

The non-myopic REE is derived in the following steps. Since all random variables in this model are normally distributed one can make use of the Projection Theorem. Thus all conditional expected values are linear in their unconditional expected values and the signal surprise component, $S^i - E[S^i]$. Brown and Jennings derive this simple linear relationship for $E_1^i[\Pi]$, $E_1^i[\Theta_1]$, $E_1^i[\Theta_2] = \varrho\frac{\sigma^2_{\Delta\Theta_2}}{\sigma^2_{\Theta_1}}$, $E_1^i[P_2]$, $E_2^i[\Pi]$, where $E_t^i[\cdot]$ is a simplified notation for $E[\cdot \mid \mathcal{F}_t^i]$. They also show that covariance matrices $V_1^i[\Pi, S_2^i, P_2]$ and $V_2^i[\Pi]$ are constants on $\mathcal{F}_0$, where $V_t^i[\cdot]$ denotes $Var[\cdot \mid \mathcal{F}_t^i]$.

The optimal stock holding can be derived by using backward induction. The value function in $t = 2$ given stock holding $x_1^i$ in $t = 1$ is

$$J_2^i(x_1^i) = max_{x_2^i} E_2^i\{-\exp[-\rho[x_1^i(P_2 - P_1) + x_2^i(\Pi - P_2)]]\}.$$

The optimal $x_2^i$ in $t = 2$ is

$$x_2^i = \frac{E_2^i[\Pi] - P_2}{\rho V_2^i[\Pi]}$$

as in Hellwig (1980). Therefore

$$J_2^i(x_1^i) = E_2^i\{-\exp[-\rho[x_1^i(P_2 - P_1) + \frac{E_2^i[\Pi] - P_2}{\rho V_2^i[\Pi]}(\Pi - P_2)]]\}.$$

The only random variable at $t = 2$ is $\Pi$, which is normally distributed. Therefore, the expectation is given by

$$J_2^i(x_1^i) = -\exp[-\rho[x_1^i(P_2 - P_1)] - \frac{(1/2)(E_2^i[\Pi] - P_2)^2}{V_2^i[\Pi]}].$$

The value function for $t = 1$ can then be rewritten as

$$J_1^i = \max_{x_1^i} E_1^i\{-\exp[-\rho[x_1^i(P_2 - P_1)] - \frac{(1/2)(E_2^i[\Pi] - P_2)^2}{V_2^i[\Pi]}]\}.$$

With respect to the information set, $\mathcal{F}_1^i$, $E_2^i[\Pi]$ and $P_2$ are normally distributed random variables. In order to take expectations we rewrite the equation given above in matrix form by completing squares.[26]

$$J_1^i = \max_{x_1^i} E_1^i\{-\exp[\rho x_1^i P_1 + \underbrace{(-\rho x_1^i, 0)}_{:=L^{i\prime}} \underbrace{\begin{pmatrix} P_2 \\ E_2^i[\Pi] \end{pmatrix}}_{:=M^i} +$$

$$+ \underbrace{(P_2, E_2^i[\Pi])}_{=M^{i\prime}} \underbrace{\frac{1}{2}\begin{pmatrix} +\frac{1}{V_2^i[\Pi]} & -\frac{1}{V_2^i[\Pi]} \\ -\frac{1}{V_2^i[\Pi]} & +\frac{1}{V_2^i[\Pi]} \end{pmatrix}}_{:=\mathbf{N}} \underbrace{\begin{pmatrix} P_2 \\ E_2^i[\Pi] \end{pmatrix}}_{=M^i}]\}.$$

Furthermore, let $Q_i$ be the conditional expected value conditional on $\mathcal{F}_1^i$ of the multinomial random variable $M^i$ and its conditional Covariance matrix $W$, i.e. $Q^i := E_1^i[M^{i\prime}]$, $\mathbf{W} := V_1^i[M^{i\prime}]$.

Taking expectations yields

$$J_1^i = \max_{x_1^i}\{- \mid \mathbf{W} \mid^{-(1/2)} \mid 2\mathbf{N} + \mathbf{W}^{-1} \mid^{-(1/2)} \exp[\rho x_1^i P_1 + L^{i\prime} Q^i - Q^{i\prime}\mathbf{N}Q^i +$$

$$+(1/2)(L^{i\prime} - 2Q^{i\prime}\mathbf{N}) \underbrace{(2\mathbf{N} + \mathbf{W}^{-1})^{-1}}_{:=\mathbf{G}}(L^i - 2\mathbf{N}Q^i)]\},$$

where the term in exp[] is the certainty equivalent. The FOC w.r.t. $x_1^i$ is given by

$$x_1^i = \frac{E_1^i[P_2] - P_1}{\rho G_{11}} + \frac{E_1^i[x_2^i](G_{12} - G_{11})}{\rho G_{11}},$$

where $G_{ij}$ are the elements of the matrix $\mathbf{G}$, and

$$x_2^i = \frac{E_2^i[\Pi] - P_2}{\rho V_2^i[\Pi]}.$$

Given the price conjectures of the trader, $x_1^i = \mathbf{L}[\mu_{\Pi,0}, S_1^i, P_1]$ and $x_2^i = \mathbf{L}[\mu_{\Pi,0}, S_1^i, S_2^i, P_1, P_2]$, where $\mathbf{L}[\cdot]$ denotes a linear operator.

This allows us to derive the market clearing price as a linear function

$$P_2 = \mathbf{L}[\mu_{\Pi,0}, \Pi, \Theta_1, \Theta_2],$$

$$P_1 = \mathbf{L}[\mu_{\Pi,0}, \Pi, \Theta_1].$$

[26] See also Anderson (1984, Chapter 2).

Brown and Jennings show that technical analysis has value as long as $P_2$ depends also on $\Theta_1$. This is consistent with the intuition provided earlier.

Since Brown and Jennings only show existence of a non-myopic dynamic REE for the special cases where $P_2$ or $P_1$ together with $P_2$ are informationally efficient they continue their analysis for myopic-investor economies. Myopic dynamic models were first analysed in Singleton (1987).

In a myopic investor economy[27] the first period stock holding simplifies to

$$x_1^i = \frac{E_1^i[P_2] - P_1}{\rho V_1[P_2]}.$$

The second period stock holding is as before

$$x_2^i = \frac{E_2^i[\Pi] - P_2}{\rho V_2[\Pi]}.$$

Brown and Jennings show that under certain parameter restrictions technical analysis has strict positive value. As mentioned above technical analysis has value if $\Theta_1$ helps to predict $\Theta_2$, and $\Theta_2$ has an impact on the information revelation of $P_2$ and/or $P_1 = \mathbf{L}[\Pi, \Theta_1]$ provides a second noisy observation of $\Pi$. The authors provide three equivalent conditions under which technical analysis has no value. Technical analysis has no value, when individual demand in $t = 2$ is independent of $P_1$, or equivalently $P_2$ is independent of $\Theta_1$ or equivalently $Cov[\Pi, P_1 \mid P_2, S_1^i, S_2^i, \mathcal{F}_0^i] = 0$.

Vives (1995) is able to derive a closed-form solution for the case $\varrho = 0$ even if investors act non-myopically by adding a risk-neutral competitive market-maker sector. Focus of Vives' work is to contrast the informativeness of the price process in an economy with myopic investors with an economy where investors have long horizons. In this multi-period model, scaplers, floor brokers, etc. of the risk-neutral market-maker sector observe only the limit order books, i.e. the aggregate demand. They set the price equal to the conditional expectation of the liquidations value $\Pi$ given the information from the limit order books. Introducing this competitive fringe changes the model quite dramatically. Vives (1995) shows that due to the normal distribution the current limit order book (or equivalently the

[27] Interpreting myopic investor economies as OLG models can be misleading, since the agents in $t = 2$ still condition their demand on their signal in $t = 1$.

current price) is a sufficient statistic for all data from the past limit-order books. In other words the prices are (semi-strong) informationally efficient and thus technical analysis has no value. The importance of the competitive market-maker sector can be illustrated for the case where private information is only released at $t = 1$. A buy and hold strategy is optimal for the informed traders in this case. At $t = 1$ informed traders buy assets as in the static Hellwig-model and hold it till $T$. The aggregate demand (limit order book) in $t = 1$ contains the demand of the insiders and the demand of the noise traders, $\Theta_1$. Market-makers set the price equal the conditional expectation of $\Pi$ given the aggregate demand. At $t = 2$ the holding of informed traders does not change. Therefore the limit order books contains only the additional noise trader demand $\Delta\Theta_2$. Since $\varrho = 0$, $\Delta\Theta_2$ contains no additional information and thus market maker set $P_2 = P_1$ absorbing the additional noise demand. In a model without competitive fringe like in Brown and Jennings (1989) informed traders have to take on the position of the additional noise trading in $t = 2$. Since each demand function of the informed traders depends on his signal, more information is revealed by $P_2$. Having a competitive fringe the only motive to trade is to exploit the information advantage and not to insure each other. This allows Vives (1995) to derive a closed form solution even for the case where private information arrives in every period. He shows that the net trading intensity of insiders in period $t$ depends directly on the precision of period $t$ signals.

In Grundy and McNichols (1989) the signals are distorted by a common error term and for most of the paper the random supply in $t = 1$, $\Theta_1$, is kept equal to $\Theta_2$, i.e. the random absolute supply is perfectly correlated. This clarifies the results in Brown and Jennings (1989). Grundy and McNichols' paper not only focuses on technical analysis but also provides a deeper understanding of the No-Speculation Theorem of Milgrom and Stokey (1982).

In their model, exogenous random supply of a single risky asset is given by endowment shocks for each individual trader, which is similar to Diamond and Verrecchia (1981). These shocks are i.i.d. with $\mathcal{N}(\mu_{\Theta_1}, \sigma^2_{\Theta_1} I)$. As $I \to \infty$ the average per capita supply shock, $\Theta_1$, is still random with $\mathcal{N}(\mu_{\Theta_1}, \sigma^2_{\Theta})$, since the variance of individual endowments depends on the number of traders $I$. Furthermore the overall variance of the total supply shocks goes to infinity and thus the Law of Large Numbers cannot be applied. Moreover, as $I$ converges to infinity the individual endowment shock has almost no impact on the aggregate supply and thus provides no information. Therefore, the only private signal trader $i$ receives is

$$S_1^i = \Pi + \omega + \epsilon_{S,1}^i,$$

which has a common error term $\omega$ and an idiosyncratic error term $\epsilon_{S,1}^i$. Both error terms are independently normally distributed with mean zero and variance $\sigma_\omega^2$ and $\sigma_{\epsilon_{S,1}}^2$. The average signal is given by $S_1 := lim_{I\to\infty}(\sum S_1^i/I) = \Pi+\omega$. In contrast to Brown and Jennings (1989) traders receive their private signal only at $t = 1$. As before, traders maximise CARA utlity functions. The constant absolute risk aversion coefficient of trader $i$ is $\rho^i \in [\rho_L, \rho_U] \subset (0, \infty)$. Grundy and McNichols derive, as a first step, a one period reference model. In this model traders conjecture a linear price relation

$$P_1 = \alpha_{0,1} + \alpha_{S,1}S_1 + \alpha_{\Theta,1}X,$$

where $X$ is the aggregate demand in equilibrium. The optimal stock holding of trader $i$ is therefore[28]

$$x_1^i = \frac{E_1^i[\Pi] - P_1}{\rho^i V_1[\Pi]},$$

where $E_1[\Pi]$ is linear in $P_1$ and $S_1^i$ by the Projection Theorem. Notice, that $V_1[\Pi]$ is higher if the variance of the common error term is higher. Averaging over all traders gives the average per capita demand

$$X = \frac{1}{\overline{\rho}V_1[\Pi]}[\beta_{0,1} + \beta_{P,1}P_1 + \beta_{S,1}S_1],$$

where $\overline{\rho}$ is the harmonic mean[29] of all traders' risk aversion coefficients. Rearranging the trader's price conjecture gives

$$X = -\frac{\alpha_{0,1}}{\alpha_{\Theta,1}} + \frac{1}{\alpha_{\Theta,1}}P_1 - \frac{\alpha_{S,1}}{\alpha_{\Theta,1}}S_1 = \Theta_1.$$

By equating the coefficients one gets the REE. The aggregate demand is downward sloping and the supply is vertical. The important coefficient is $(\alpha_{S,1}/\alpha_{\Theta,1})$. Changes in $S_1$ lead to a parallel shift of the demand curve, whereas changes in $\Theta_1$ shift the vertical supply curve. The size of the demand curve shift as $S_1$ changes is measured by $\alpha_{S,1}$, whereas the size of the supply curve shift caused by a different $\Theta_1$ is captured by $\alpha_{\Theta,1}$. Since traders cannot make

[28] I normalise the risk free interest rate $r = 0$, i.e. $R = 1$.

[29] Harmonic mean is defined as $\frac{I}{\sum_i(1/\rho_i)}$.

out whether a price change is due to a demand shift or a supply shift ($S_1$ or $\Theta_1$ change), $\alpha_{S,1}/\alpha_{\Theta,1}$ measures the simultaneous equation problem.

The basic model is then extended to a two period model where, in the second round, no new private information is released and the random supply $\Theta_2$ is the same as in period one. One might expect that no trader will change his stock holding, since no new information arrived. Grundy and McNichols show that his no-trade outcome is indeed an equilibrium. There is, however, a second equilibrium, where the average signal $S_1$ is revealed. If all trader rationally conjecture

$$P_1 = \alpha_{0,1} + \alpha_{S,1}S_1 + \alpha_{\Theta,1}\Theta_1,$$

$$P_2 = \alpha_{0,2} + \alpha_{S,2}S_1 + \alpha_{\Theta,2}\Theta_2,$$

where $\Theta_1 = \Theta_2$ and if both equations are linearly independent, then $S_1$ can be revealed. This can be the case if $\frac{\alpha_{S,1}}{\alpha_{\Theta,1}} \not\models \neq \frac{\alpha_{S,2}}{\alpha_{\Theta,2}}$, since then we have two linearly independent equations with two unknowns.

Grundy and McNichols prove that an informationally efficient REE, which fully reveals $S_1$, exists as long as the variance of $\omega$ is not too large. Their proof proceeds in two steps. First, they show, given a linear pricing relation in round 1, there exists a $S_1$-revealing equilibrium in round 2. Second, as long as the variance of $\omega$, $\sigma_\omega^2$, is not too large, traders rationally foresee the existence of a $S_1$ revealing equilibrium in round 2. When $0 < \sigma_\omega^2 < \overline{\sigma}_\omega^2$ two $S_1$-revealing REE exist. In these equilibria, there are two sources of uncertainty in the first round: $x_2^i$ and $P_2$. These equilibria show that even when no new information arrives, prices and stock holdings can change if the additional price $P_2$ reveals more of the private information. When $\sigma_\omega^2 = 0$, both equilibria, the $S_1$-revealing and the non-$S_1$-revealing, are identical for the first round.

In the $S_1$-revealing REE, trade also occurs in period two, although the only new public information is $P_2$. This seems striking in light of the No-Speculation Theorem developed in Milgrom and Stokey (1982). The No-Speculation Theorem predicts a null trade outcome in period 2, if the allocation after trade in period 1 is Pareto optimal and the prior beliefs about the signals in $t = 2$ are concordant before the signal becomes known. Beliefs are concordant if traders agree on the conditional likelihood of any given realisation of the signal, i.e.

$$Pr[S_2^i = s \mid \Pi = \pi, \mathcal{F}_1^i] = Pr[S_2^i = s \mid \Pi = \pi, \mathcal{F}_1^1] \; \forall\, i, S, \Pi.$$

Intuitively, beliefs are concordant if traders agree about everything expect the prior probability of payoff-relevant states. Since the only new signal in $t = 2$ is $P_2$, which is public, it is sufficient that beliefs about $P_2$ only are "essentially concordant", i.e.

$$\frac{Pr[S_2^i = s \mid \Pi = \pi, \mathcal{F}_1^i]}{Pr[S_2^i = s \mid \Pi = \pi\prime, \mathcal{F}_1^i]} = \frac{Pr[S_2^i = s \mid \Pi = \pi, \mathcal{F}_1^1]}{Pr[S_2^i = s \mid \Pi = \pi\prime, \mathcal{F}_1^1]} \; \forall\, i, S, \Pi.$$

Pareto optimality is given if the marginal rate of substitution for consumption across any two states is the same for all investors. Grundy and McNichols show that if the investors behave *myopically* they reach a Pareto optimal location after the first round. However, when $P_2$ becomes known this allocation is no longer Pareto efficient since traders' beliefs about $P_2$ are not "essentially concordant" at the end of the first round. Therefore, trade will occur. If traders apply dynamic trading strategies, i.e. they do *not* behave *myopically*, trade can also occur in period 2. This is the case when $\sigma_\omega^2 > 0$, i.e. there is a common unknown noise term in the signal. The trading outcome in round 1 is neither Pareto efficient given information $\mathcal{F}_1^i$, nor are the beliefs about the public signal $P_2$ concordant. When $\sigma_\omega^2 = 0$, the true liquidation value $\Pi$ can be inferred from $P_2$ and trade 1 allocation is Pareto efficient and beliefs about $P_2$ are concordant. In this case the No-Speculation Theorem applies and the only trade which occurs is a swapping of riskless assets.

Grundy and McNichols continue their study by introducing an additional publically observable signal in $t = 2$

$$Y_2 = \Pi + \epsilon_{Y,2}.$$

In this case a $S_1$-revealing REE with trade in $t = 2$ exists as before, but also when the new public information $Y_2$ is informative, i.e. $Cov(\Pi, \omega) \neq Var(\omega)$. The authors also provide necessary and sufficient conditions for the existence for non-$S_1$-revealing REE, in which no trade occurs in the second round. Finally, they consider the case where the random supply $(\Theta_1, \Theta_2)$ is not the same in both periods but correlated as in Brown and Jennings (1989). Both types of equilibria exist in this generalised version. In the non-$S_1$-revealing type, no informational trade will occur; the whole trading volume is determined by the additional noisy supply. In the second type, the sequence of prices $\{P_1, P_2\}$ only partially reveals $S_1$, since the supply shocks are not perfectly correlated anymore. However, the sequence of prices reveals more about $S_1$ than $P_1$ alone. Their paper shows that technical analysis has positive value and that trading can be self-generating. It also makes clear how important the traders price conjectures are in determining the economic outcome.

Romer (1993) provides a rationale for large price movements without news. He shows, within a two period asymmetric information model, that a small commonly known supply shift in the second period can lead to large price movements. The aim of his paper is to give a rational explanation for the stock market crash in 1987. In his model asymmetric information is only partially revealed in the first period, but in contrast to Grundy and McNichols (1989) it incorporates uncertainty about the quality of other investors' signals, i.e. higher order uncertainty. Each investor receives one of possibly three signals about the liquidation value of the single risky asset, $\Pi \sim \mathcal{N}(\mu_\Pi, \sigma^2_\Pi)$.[30]

$$S^j = \Pi + \epsilon_{S^j},$$

where $\epsilon_{S^2} = \epsilon_{S^1} + \delta^2$, $\epsilon_{S^3} = \epsilon_{S^2} + \delta^3$ and $\epsilon_{S^1}$, $\delta^2$, $\delta^3$ are independently distributed with mean of zero and variance $\sigma^2_{\epsilon_{S^1}}$, $\sigma^2_{\delta^2}$, $\sigma^2_{\delta^3}$, respectively. Thus, $S^j$ is a sufficient statistic for $S^{j+1}$. There are two equally likely possible states of the world for the signal distribution. *Either* half of the traders receive signal $S^1$ and the other half signal $S^2$ *or* half of the traders receive signal $S^2$ and the other half signal $S^3$. It is obvious that traders who receive signal $S^1$ (or $S^3$) can infer the relevant signal distribution, since each investor knows the precision of his own signal. Only trader who receive signal $S^2$ do not know whether the other half of traders has received the more precise signal $S^1$ or the less precise signal $S^3$. Finally, as usual, the random supply in period 1 is given by the independently distributed random variable $\Theta_1 \sim \mathcal{N}(\mu_{\Theta_1}, \sigma^2_{\Theta_1})$.[31]

In contrast to Grundy and McNichols (1989), the supply of the risky asset changes in period two. This change is common knowledge making the no-trade outcome of Grundy and

[30] The notation in the article is: $\Pi = \alpha, S_j = s_j, \mu_\Pi = \mu, \sigma^2_\Pi = V_\alpha, \Theta_1 = Q, \mu_{\Theta_1} = \overline{Q}, \sigma^2_{\Theta_1} = V_Q$.

[31] Romer claims that he needs this random supply term in order to avoid an informationally efficient REE in $t = 1$. I do not see the necessity for this term since a single price cannot reveal two facts, the signal and the signals quality. In my opinion the structure is rather similar to the partial-revealing REE analysis in Ausubel (1990).

McNichols (1989) very unlikely. The change in price caused by the supply shock allows the $S^2$-investors to infer more precisely the signal distribution. Thus a small supply change can lead to revelation of 'old' information and can have a huge impact on prices. Alternatively, if in addition an option is traded, the quality of information can be revealed by its price already in $t = 1$. This is only the case as long as the quality of the signals can be summarised in one parameter. The informational difference between traded and synthesised options was discussed in more detail in Section 2.2 Grossman (1988).

The stock holdings in equilibrium of $S^1$-traders, $x^1(S^1)$ can be derived directly from the Projection Theorem. They do not make any inference from the price, since they know that their information is sufficient for any other signal. Traders with $S^3$-signals face a more complex problem. They know the signal distribution precisely but they also know that they have the worst information. In addition to their signal $S^3$, they try to infer signal $S^2$ from the price $P_1$. The equilibrium price in $t = 1$, $P_1$, is determined by $x^2(S^2, P_1) + x^3(S^3, P_1) = \Theta_1$ (assuming a unit mass of each type of investors). Since an $S^3$ trader knows $x^2(\cdot)$, $x^3(\cdot)$ and the joint distribution of $S^2$, $S^3$ and $\Theta_1$, he can derive the distribution of $S^2$ conditional on $S^3$ and $P_1$. Since $x^2(S^2, P_1)$ is not linear in $S^2$, $x^3(S^3, P_1)$ is also nonlinear. $S^2$-investors do not know the signal distribution. Therefore, the $Var[P_1 \mid S^2]$ depends on the higher order information, i.e. whether the other half of traders are $S^1$- or $S^3$-investors. $S^2$-traders use $P_1$ to predict more precisely the true signal distribution, i.e. information quality of other traders. If they observe an extreme $P_1$, then it is more likely that other investors got signal $S^3$. Otherwise, if $P_1$ is close to the expected price $\mu_\Pi$ then it seems that others are $S^1$-traders. $S^2$-investors' demand functions $x^2(S^2, P_1)$ are therefore not linear in $P_1$, since $P_1$ changes not only the expectations about $\Pi$, but also the variance. This nonlinearity forces Romer to restrict his analysis to a numerical example. His simulation shows that $S^2$-investors' demand functions are more responsive to price changes. Using these results Romer tries to explain the market meltdown in October 19, 1987. In Section II, Romer (1993) develops an alternative model with trading costs and widespread dispersion of information which explains the stock market crash. This model is not covered by this survey. Another model which incorporates uncertainty about the signal precision of other trades is Blume, Easley, and O'Hara (1994). Blume, Easley, and O'Hara (1994) avoid these nonlinearity problems by considering a market order model instead of a limit order model. In their model volume rather the price in the next period reveals the quality of information.

Gennotte and Leland (1990) provide a similar explanation for stock market crashes in a 'one period' model.[32] As in Romer (1993) there are no major news events. Gennotte and Leland (1990) consider two groups of informed traders. Each (price-)informed trader receives an individual private signal $S^i = \Pi + \epsilon^i$ about the liquidation value $\Pi \sim \mathcal{N}(\mu_\Pi, \sigma_\Pi^2)$. Supply-informed traders receive a signal about the total supply. Aggregate supply results from dynamic hedging trades like program trading, stop and loss strategies etc. as well as from noise liquidity supply represented by the random variable $\Theta$. Random supply $\Theta$ can be divided into the part $\overline{\Theta}$ which is known to everybody, $\Theta_S$ which is only known to the supply-informed traders, and the liquidity supply $\Theta_L$ which is not known to anybody. Superior knowledge of the supply-informed traders about $\Theta_S$ allows them to infer more information from the equilibrium price $p_1$. Gennotte and Leland (1990) show that the equilibrium is given by $p_1 = f(\Pi - \mu_\Pi - k_1\Theta_L - k_2\Theta_S)$, where $k_1$ and $k_2$ are real constants. Note, that since $\pi(p_1)$ need not be linear, $p_1$ need not be normally distributed. However, $f^{-1}(p_1)$ is still normally distributed. A "crash" is possible if $f(\cdot)$ is discontinuous, i.e. a small change in the argument of $f(\cdot)$ leads to a large price shift. In the absence of any program trading (i.e. $\pi(p_1) = 0$) $f(\cdot)$ is continuous. This rules out crashes. Nevertheless, an increase in the supply can lead to a large price shift. The price change is small if the change in supply is common knowledge, i.e. change in $\overline{\Theta}$. If the supply shift is only observable by supply-informed traders, the price change is still moderate. This occurs because price-informed and supply-informed traders take on a big part of this additional supply even if the fraction of informed traders is low. Supply-informed traders know that the additional excess supply does not result from different price signals while price-informed traders can partially infer this from their signal. If, on the other hand, the additional supply is not observable at all, a small increase in the liquidity supply $\Theta_L$ can have a large impact on the price. Crashes only occur when the program trading is large enough to cause a discontinuous price correspondence $f(\cdot)$. The discontinuity stems from non-linearity of program trading $\pi(p_1)$ in $p_1$ which can lead to the possibility of multiple equilibria. Crashes are much more likely and prices are more volatile if some investors underestimate the supply due to program trading. Gennotte

[32] We adjusted the notation to $S^i = p'_i$, $\Pi = p$, $\mu_\Pi = \overline{p}$, $\sigma_\Pi^2 = \Sigma$, $p_1 = p_0$, $\Theta = m$, $\Theta_L = L$, $\Theta_S = S$.

and Leland (1990) illustrate their point by means of an example of a put replicating hedge strategy (synthetic put). Their analysis is in the spirit of Grossman (1988). However, their paper also explains why stock prices do not immediately rebound after a stock crash.

In order to get a better understanding about price processes, one would like to have models which capture a larger time horizon than essentially two periods. Before discussing dynamic models with differential information, we deal with a simpler information structure, namely asymmetric information. Townsend's (1983) article makes clear what kind of problems can arise from a more general information structure.

### 2.4.2 Serial Correlation Induced by Learning and the Infinite Regress Problem

Townsend (1983) laid bare crucial points in his article "Forecasting the Forecast of Others", viewing rational expectation from a macroeconomic angle. Within a rational expectations framework, decision makers solve dynamic decision problems following their objective function and infer information from well specified information sets, taking the aggregate laws of motion as given. These laws are, in turn, those actually generated in the model. The focus of his article is the characteristics of economic time series. He shows that learning can convert serially uncorrelated shocks into serially correlated movements in economic decision variables. Since agents may respond to variables generated by the decisions of others, time series can display certain cross-correlation and may appear more volatile. In the case of disparate but rational expectations, decision makers forecast the forecasts of others. This can lead to relatively rapid oscillations and can make forecasts, as well as forecast errors, serially correlated.

He analysed the behaviour of time series in a dynamic model with a continuum of identical firms in each of two markets. The demand schedule in each market (island) $i$ is given by

$$P_t^i = -\beta_1 Y_t^i + \xi_t^i,$$

where $P_t^i$ is the price in market $i$, $Y_t^i$ is the aggregate output of all individual production functions $y_t^i = f_0 k_t^i$, and $\xi_t^i$ is a demand shock. This shock consists of a "persistent" economy wide component $\Theta_t$ and a "transitory" market specific shock $\epsilon_t^i$, i.e.

$$\xi_t^i = \Theta_t + \epsilon_t^i,$$

where

$$\Theta_t = a_\Theta \Theta_{t-1} + \nu_t \quad -1 < a_\Theta < 1,$$

follows a AR(1) process with $\epsilon_t^i$ and $\nu_t$ jointly normal and independent. Firms can infer $\xi_t^i$s, but they do not know exactly which part steams from a persistent economywide shock and which part is market specific and transitory. After stating the firm's maximisation problem, Townsend derived the first order conditions using the certainty equivalence theorem. He defines the dynamic *linear* rational expectations equilibrium in laws of motion. Following Sargent (1979) one can derive the law of motion of the aggregate (capital stock) in each market without directly calculating the firm-specific laws of motion. The aggregate law of motions have the advantage that they can be computed without being specific about information sets and forecasting. In Townsend's setting the equilibrium can be found by finding the statistically correct forecasts.

Considering the inference problem, the paper is divided into two parts. In the first part, firms in market 1 cannot observe the price in market 2, whereas market 2 firms observe both prices. Townsend calls this an hierarchical information structure. In the second part, firms in both markets can make inferences from both markets' prices.

In part one the information set of market 1 firms consists of $\mathcal{F}_t^1 = \{\underline{K}_t^1, \underline{P}_t^1, \underline{M}_t^1\}$, i.e. the aggregate capital, the price and the common market 1 mean forecast of $\Theta_t$, $M_t$ where $\underline{Z}_t$ denotes a stochastic process up to and including time $t$. Using only observations in $t$ of this information set allows firms to infer exactly the total shock to the economy $\xi^1$. The inference problem for firms in market 2 is similar, except that their information set also contains the price in market 1, i.e. $\mathcal{F}_t^2 = \{\underline{K}_t^2, \underline{P}_t^2, \underline{M}_t^2, \underline{P}_t^1\}$. The price in market 1 provides additional information to what extend the shock is permanent and is, therefore, used in making the forecasts of the shock components. The components of $\xi^i$, however, cannot be inferred precisely although past data help to get a better forecast. The inference problem of the firms can be solved in two ways. Either one uses the Projection Theorem or one applies Kalman filtering, which derives from the Projection Theorem. Applying the Projection Theorem directly has the disadvantage that the state space increases with the history of time. The latter is a steady state approach and exploits a recursive algorithm. Therefore, it is often assumed that the initial date is $t = -\infty$. It is important to notice that Kalman filtering can only be applied if the state vector[33] in the state space form is of finite dimension. Using these methods one can derive $\hat{\underline{\nu}}_t := E(\nu_t \mid \mathcal{F}_t^i)$, the forecasts of $\underline{\nu}_t$ as a linear combination of $\underline{\nu}_t$, and $\underline{\epsilon}_t^i$. It now becomes obvious that the learning mechanism causes some persistence. Although $\nu_t$, $\epsilon_t^i$ are uncorrelated, their forecasts are correlated, since both forecasts $E(\nu_t \mid \mathcal{F}_t^i)$ and $E(\nu_{t-1} \mid \mathcal{F}_{t-1}^i)$ contain $\underline{\nu}_{t-1}$. In other words, all past $\underline{\nu}_{t-1}$ influence the prediction of $\nu_t$. Similarly $E(\Theta_t \mid \mathcal{F}_t^i)$ and $E(\Theta_{t-1} \mid \mathcal{F}_{t-1}^i)$, as well as the forecast errors $[E(\Theta_t \mid \mathcal{F}_t^i) - \Theta]$ are serially correlated. It is important to notice that the forecast error for past $\Theta_s$ (s<t) decreases as time goes on and more and more observations are available.

So far only market 2 firms were forming inference about the components of the demand shock from an endogenous time series, the price in market 1. The price in the first market also depends on the average beliefs in this market, $\underline{M_t^1}$, i.e. the market 1 expectations. These expectations are well defined and can be expressed in terms of a finite number of states. Therefore, the Kalman filter can be applied. In the second part of the paper the information structure is not hierarchical anymore. Firms in market 1 can also draw inferences from $P_t^2$. Since $P_t^2$ depends on the common market 2 forecasts, $M_t^2$, firms in market 1 must have expectations about $M_t^2$, i.e. $E_t^1(M_t^2)$. But firms in market 2 (firms 2) see $P_t^1$ also. So firm 2 must have expectations on $M_t^1$, i.e. $E_t^2(M_t^2)$. Thus firm 1 needs to know expectations $E_t^1(M_t^2)$ and $E_t^1(E_t^2(M_t^1))$. This chain of reasoning can be continued ad infinitum. Therefore, we face an infinite regress problem. One needs, in the space of mean beliefs, infinitely many state variables. This prevents us from applying the standard Kalman filter formulas. Notice that the infinite regress problem arises although the depth of knowledge is only zero. The infinite regress problem is not due to a high depth of knowledge but due to inference of endogenous variables. Townsend then goes on to discuss a related but different infinite regress problem, in which he analyses the case of infinitely many markets.

[33] A state in this setting is an element of the "sufficient (current) state description" as desribed in Section 2.2.2.

New methods developed by Marcet and Sargent (1989a) (1989b) in convergence of least squares learning to rational-expectations equilibria allow us to tackle this infinite regress problem differently. Sargent (1991) shows in that a solution can be found in self-referential models by defining the state variables in a different way. The idea is to model agents as forecasting by fitting vector *arma* models for whatever information they have available. The state vector for the system as a whole is defined to include the variables and the innovation in the vector *arma* models fit by each class of agents in the model. This contrasts to the former formulation in Townsend (1983) where the state covers a system of infinitely many orders of expectations about exogenous hidden state variables. This new approach - to my knowledge - has not been applied so far in the finance literature. Most of the literature avoids the infinite regress problem by assuming a hierarchical information structure as in Wang (1993),(1994). In models with differential information the problem is elegantly by-passed. This is the case in He and Wang (1995) a competitive model and in Foster and Viswanathan (1996) a strategic model.

### 2.4.3 Competitive Multi-Period Limit Order Models

Wang (1993) is - to my knowledge - the first to use Kalman filters in the financial economics literature. He avoids the infinite regress problem by assuming a hierarchical information structure. In his model the information of the informed investors statistically dominates the information of the uninformed. In other words, all variables the uninformed investors can observe are also known by the informed traders. The main focus of this paper is the impact of information asymmetries on the time series of prices, risk premiums, price volatility and the negative autocorrelation in returns, i.e. the mean reverting behaviour of stock prices. For analysing these questions he uses a dynamic asset-pricing model in continuous time. In his economy, investors can invest either in a riskless bond with constant rate of return $(1+r)$, or in equity which generates a flow of dividends at an instantaneous stochastic growth rate $D$. $D$ is determined by the following diffusion process:

$$dD = (\Pi - kD)dt + \mathbf{b}_D d\mathbf{w},$$

where the state variable $\Pi$ follows an Ornstein-Uhlenbeck process,

$$d\Pi = a_\Pi(\overline{\Pi} - \Pi)dt + \mathbf{b}_\Pi d\mathbf{w},$$

and $\mathbf{w}$ is a (3x1) vector of standard Wiener processes, $a_\Pi(> 0)$, $\overline{\Pi}$, $k(\geq 0)$ are constants and $\mathbf{b}_D$, $\mathbf{b}_\Pi$ are (1x3) constant matrices.
The fraction $w$ of informed traders observe in addition to $\underline{D}_t$, $\underline{P}_t$, also $\underline{\Pi}_t$, whereas the uninformed only know $\underline{D}_t$ and $\underline{P}_t$, i.e. $\mathcal{F}^i(t) = \{D_\tau, P_\tau, \Pi_\tau : \tau \Box t\}$ and $\mathcal{F}^u(t) = \{D_\tau, P_\tau : \tau \Box t\}$.[34] It is clear that the informed can infer the expected growth rate $(\Pi - kD)$. When $k = 0$, $\Pi$ is simply the dividend growth rate. When $k > 0$, $\Pi/k$ can be interpreted as the short-run steady-state level of the dividend rate D, which fluctuates around a long-run steady-state level $\overline{\Pi}/k$.

In this setting, the rational expectations equilibrium would fully reveal $\Pi$ to the uninformed. Although the price would adjust, no trading would occur. Under incomplete markets, there can be motivation other than the arrival of new information that cause investors to trade. In the case where the price is not informationally efficient, the irrelevance of heterogeneous information breaks down and investors will trade. In order to have a incomplete markets setting, Wang introduces an additional state variable by assuming that a stochastic quantity of stock supply. The total amount of stocks $(1 + \Theta)$ should be governed by the stochastic differential equation

$$d\Theta = -a_\Theta \Theta dt + \mathbf{b}_\Theta d\tilde{\mathbf{w}},$$

where $\mathbf{b}_\Theta$ is constant (1x3) matrix and $\mathbf{w}$ are the Wiener Processes mentioned above. In this environment the uninformed face following problem. They cannot distinguish whether a change in $(P_t, D_t)$ is due to a change in the dividend growth rate $\Pi_t$ or due to a change in noise supply $\Theta_t$.

Wang analyses first the benchmark case of perfect information, in which all investors are informed. The equilibrium price takes on the form

$$P^* = \Phi + (p_0^* + p_\Theta^* \Theta),$$

where $\Phi$ represents the net present value of expected future cash flows discounted at the risk free rate $r$ and the second term reflects the risk premium. He shows that the expected

[34]The notation $\underline{Z}_t$ represents a (continuous) process up to and including t.

excess return to one share is independent of the variance of the noise supply. In other words, volatility in prices caused by temporary shocks in supply do not change the risk premium in the case of symmetric information.[35] This is in contrast to De Long, Shleifer, Summers, and Waldmann (1990b) where investors have finite horizons and they face additional risk since the remaining trading periods to rewind their positions are becoming fewer. He and Wang (1995) also consider a finite horizon model, in which the variance of $\Theta$ affects the risk premium.

For the case of asymmetric information, we outline all of the major steps, since they are useful for the analysis of later papers. Wang proceeds in the following way to determine a linear rational expectations equilibrium. First, he defines the primary state variables consisting of all known variables for the informed traders. The state space covers also "induced state variables" reflecting the estimates of the uninformed investors. The actual state description should incorporate all signals which the investors receive. Wang simplified the state space by using equivalently the estimates of uninformed investors.

As a second step he proposes a linear rational expectations equilibrium price

$$P = (\phi + p_0) + p_D^* D + \underbrace{p_\Pi \Pi + p_\Theta \Theta}_{:=\xi} + p_\Delta \hat{\Pi} = \Phi + (p_0 + p_\Theta \Theta) + p_\Delta \Delta,$$

depending on $\hat{\Pi}(t) := E[\Pi \mid \mathcal{F}_t^u]$, the estimate of $\Pi(t)$ by uninformed investor. $\hat{\Pi}(t)$ depends on the whole history of dividends and prices. The equilibrium price reveals to the uninformed traders the sum $\xi := p_\Pi \Pi + p_\Theta \Theta$. Therefore, $\mathcal{F}_t^{D,P} = \mathcal{F}_t^{D,\xi}$. The equilibrium price does not depend additionally on $\hat{\Theta} := E[\Theta \mid \mathcal{F}_t^u]$, since $p_\Pi \hat{\Pi} + p_\Theta \hat{\Theta} = p_\Pi \Pi + p_\Theta \Theta =: \xi$. In other words, the uninformed investors can derive $\xi$ but do not know exactly whether the price change is due to a change in $\Pi$ or $\Theta$.

Given the proposed linear REE one can derive in a third step the estimates, $\hat{\Pi}$, and $\hat{\Theta}$. Focusing on a steady state analysis, the uninformed investors apply the Kalman filter on all past data of dividends $D$ and prices $P$ to infer their estimates $\hat{\Pi}$, and $\hat{\Theta}$.[36] Their joint estimation of $\Pi$ and $\Theta$ based on both $D$ and $P$ generates the *induced correlation* between the estimates of $\hat{\Pi}$ and $\hat{\Theta}$.

[35] $\Theta$ can be inferred by everybody and $\Theta$ describes a economy-wide shock. Together with CARA utility functions this result seems plausible.

[36] For a more detailed discussion see Lipster and Shiryayev (1977).

In the fourth step the process for the estimation error $\Delta := \hat{\Pi} - \Pi$ is derived.[37] It is shown that it follows

$$d\Delta = -a_\Delta \Delta dt + \mathbf{b}_\Delta d\mathbf{w}.$$

This estimation error is mean-reverting to zero and thus only temporary. This is the case since the uninformed investors constantly update their estimates, as in Townsend (1983).

Let us in a fifth step derive the instantaneous excess return process $dQ := (D - rP)dt + dP$. It is well known from static models that, for deriving the demand functions the excess returns are relevant.

In the sixth step the uninformed investors' optimisation problem is derived. As in the static case one, can exploit the CARA utility to derive a mathematically tractable form of expected utility for the Bellman equation. The estimators, $\hat{\Pi}_t$ and $\hat{\Theta}_t$, provide a sufficient statistic for $\mathcal{F}^u(t)$. Therefore, by the Separation Principle, $\hat{\Pi}_t$ and $\hat{\Theta}_t$ can be estimated first and then in a second stage the control problem can be dealt with.[38] The optimal control problem is then solved in a similar manner for the informed investors.

Finally the market clearing conditions are imposed and the above proposed price equation can be derived.

Using simulations, Wang (1993) shows the impact of this information structure on stock prices, the risk premium, price volatility and negative serial correlation in returns. Increasing the number of uninformed traders has two effects in this model. First, there is overall less information in the market and prices become less variable. Second, there exists more uncertainty about future dividend payments. Investors will demand a higher risk premium and, therefore, prices become more sensitive to supply shocks. Asymmetry in information among investors can cause price volatility to increase, because the adverse selection problem becomes more severe. Wang demonstrates that the existence of uninformed investors increases the risk premium, since the risk premium only depends on the fundamental risk of the asset perceived by the investors. When the fraction of uninformed investors increases, the price contains less information about future dividend growth. He also shows that the strong mean reversion in $\Theta(t)$ generates negative serial correlation in stock returns even in the case of symmetric information. This correlation can be enhanced as the fraction of

[37] Note that the estimation error for $\Theta$ is given by $p_\Pi / p_\Theta (\Pi - \hat{\Pi})$.

[38] For a more detailed discussion see Fleming and Rishel (1975).

uninformed investors increases. Finally it is shown that the optimal investment strategy of the informed investors not only depends on the value of underlying true state variables but also on the reaction of uninformed investors. In other words, the informed investors make use of the estimations errors of the uninformed. Wang also found that the trading strategies for less informed investors can look like *trading chasing*, i.e. they rationally buy when the price rises and sell when the price drops. He and Wang conclude their paper with further comments and possible generalisations. One was that the whole economy can be reduced to an effective two persons setup[39] even if all investors have different risk aversion coefficients.

In a similar, but discrete time version, Wang (1994) analysed the behaviour of volume. The other major difference to the continuous time model is that, although no exogenous noise is introduced, the price is only partially revealing. This is due to the modeled incompleteness of the markets. If markets are incomplete and investors are heterogeneous, prices are not only affected by aggregate risk but also by individual risk. In such an environment volume plays an important role. This paper tries to show the link between volume and heterogeneity of investors. Investors differ in their information as well as in their private investment opportunities. In order to avoid the infinite regress problem informed investors have a strictly statistically dominant information set in comparison to uninformed traders. Markets are incomplete, since only informed investors have an additional private investment opportunity, besides stocks and bonds whose rate of return is $R = (1 + r)$. The dividend of a stock consists of a persistent component $F_t$ and an idiosyncratic component $\epsilon_{D,t}$. $F_t$, which is only observable by informed investors, follows an AR(1) process:

$$D_t = F_t + \epsilon_{D,t},$$

$$F_t = a_F F_{t-1} + \epsilon_{F,t}, \; 0 \square a_F \square 1.$$

Whereas informed investors can observe $F_t$, all uninformed traders get the same noisy signal $S_t$ about $F_t$.[40]

$$S_t = F_t + \epsilon_{S,t}.$$

Define, for later reference, the excess *share* return as $Q_t := P_t + D_t - RP_{t-1}$. Informed traders can also invest in their private investment opportunity which yields a stochastic

[39] This follows from the aggregation theorem (see Rubinstein (1974)).

[40] To avoid the infinite regress problem informed traders observe this signal, too.

excess rate of return of

$$q_t = Z_{t-1} + \epsilon_{q,t},$$

where $Z_t$ follows an AR(1) process

$$Z_t = a_Z Z_{t-1} + \epsilon_{Z,t}, \ 0 \square a_Z \square 1.$$

Similar to the stock return the process $\underline{Z_t}$ is only known to the informed traders. Besides making use of their information advantage, hedging the risk reflected by $\epsilon_{q,t}$ is the only incentive for informed traders to trade. All $\epsilon$-terms are normally i.i.d. with the exception that $\epsilon_{D,t}$ and $\epsilon_{q,t}$ can be positively correlated. Wang shows in the case of symmetric information that if $\epsilon_{D,t}$ and $\epsilon_{q,t}$ are uncorrelated a change in expected returns on the private investment will not alter the investors' stock holdings. This changes with a positive correlation between $\epsilon_{D,t}$ and $\epsilon_{q,t}$ since the stock and the private investment are becoming substitutes. The problem the uninformed investors face is that they cannot sort out whether a price increase is due to informed trading, i.e. an increase in $F_t$, or due to uninformed trading, in which case informed traders just want to rebalance their portfolio because of a change in the profatibility of their private investment opportunities. Uninformed traders face, therefore, an adverse selection problem. The analysis of the equilibrium follows the same steps as in Wang (1993). First, the states of the economy $F_t, Z_t, \hat{F}_t = E[F_t \mid \mathcal{F}_t^u]$ are defined. Second, the linear pricing rule

$$P_t = -p_0 + (a - p_F)\hat{F}_t + p_F F_t - p_Z Z_t$$

is proposed. Third, from this equation it is obvious that uninformed traders can infer the sum $\xi_t = p_F F_t - p_Z Z_t$, thus $\xi_t = p_F \hat{F}_t - p_Z \hat{Z}_t$. This explains why $\hat{Z}_t$ is redundant in the state description within the class of linear equilibria. Fourth, using Kalman filtering one derives $\hat{F}_t$, $\hat{Z}_t$ and the estimation errors $\hat{F}_t - F_t =: \Theta_t$.[41] It can be shown that the estimation error $\Theta_t$ follow an AR(1) process, i.e.

$$\Theta_t = a_\Theta \Theta_{t-1} + \epsilon_{\Theta,t}, \ 0 \square a_\Theta < 1.$$

The strict inequality $a_\Theta < 1$ implies that the forecast error is mean reverting. Through time the uninformed traders will learn "old" $F_s$, $Z_s$ better and better but in every period

[41] Hint: $\hat{Z}_t - Z_t$ is determined by $\xi_t = p_F F_t - p_Z Z_t = p_F \hat{F}_t - p_Z \hat{Z}_t$.

new $F_t$, $Z_t$ are appearing. Uninformed investors are *"chasing" a moving target.* The unconditional variance of the estimation error, $Var(\Theta_t) =: \epsilon$ reflects the degree of asymmetry of information.

For determining the optimal stock demand it proves useful to derive the expected excess returns for informed and uninformed traders. The optimal portfolio for each group of investors is a composition of a mean variance efficient portfolio and a hedging portfolio. Investors want to hedge, since expected returns on both the stock and the private investment technology change over time. Since returns on the stock are correlated with changes in expected future returns, it provides a vehicle to hedge against changes in future investment opportunities. Knowing the optimal portfolios, the trading strategies for the informed and uninformed investors can be written as:

$$X_t^i = f_0^i + f_Z^i Z_t + f_\Theta^i \Theta_t,$$

$$X_t^u = f_0^u + f_Z^u \hat{Z}_t.$$

This shows that the optimal stock holding of the uninformed traders only changes when their expectation about the others private investment opportunities changes, i.e. $X_t^u - X_{t-1}^u = f_Z^u(\hat{Z}_t - \hat{Z}_{t-1})$.

$$\hat{Z}_t - \hat{Z}_{t-1} = E_t^u[Z_t] - E_{t-1}^u[Z_{t-1}]$$

can be decomposed into

$$\{E_t^u[Z_{t-1}] - E_{t-1}^u[Z_{t-1}]\} + \{E_t^u[Z_t] - E_t^u[Z_{t-1}]\},$$

where the first component deals with correcting errors of previous periods and the second component induces new positions. Knowing that all trading volume is changes in holding of either the informed or the uninformed investors we can characterise volume by

$$\mathcal{V}_t = (1-w) \mid X_t^u - X_{t-1}^u \mid = (1-w) \mid f_t^u \mid\mid \hat{Z}_t - \hat{Z}_{t-1} \mid,$$

$$E[\mathcal{V}_t] = 2(1-w) \mid f_Z^u \mid \sqrt{2/\pi}.$$

Equipped with these formulae, the effects of asymmetric information on volume by increasing the noise of the signal can be analysed. As the signal of the uninformed becomes less precise the asymmetry of information increases and the adverse selection problem becomes more severe. This reduces the trading volume. This need not be the case if a non-hierarchical information structure is assumed as in Pfleiderer (1984) or He and Wang (1995). Trading volume is always accompanied by price changes, since investors are risk averse. If informed traders face high excess return in private investment they try to rebalance their portfolio by selling stocks. In order to make stocks more attractive to uninformed investors they have to reduce the price. This price reduction needs to be even higher if the adverse selection problem is more severe. This shows that the trading volume is positively correlated with absolute price changes and that this correlation increases with information asymmetry. Volume is also positively correlated with absolute dividend changes. In the case of symmetric information public news announcements about dividends change only the current price, but not the expected return or trading volume. In the case of asymmetric information, different investors update their expectations differently. They respond to public information differently since they interpret it differently. Uninformed investors change their estimates for $\underline{F}_{t-1}$ and $\underline{Z}_{t-1}$ and trade to correct previous errors and establish new positions. Volume in conjunction with current change in dividends or returns can also be used to improve the forecast for expected future excess returns. Under symmetric information, public news (like a dividend change announcement) is immediately reflected in the price. Under asymmetric information, public news can lead to corrections previous trading mistakes. Wang shows that an increase in dividends accompanied by high volume implies high future returns. High volume indicates that the change in dividend was unanticipated. A dividend increase should, therefore, increase prices. The second component of excess returns, the price change, is different because it provides information about noninformational trading as well as the stock's future dividends. Under symmetric information, trade is only done to rebalance portfolios and it is always accompanied by changes in the current price in the opposite direction. In the case of asymmetric information, uninformed investors trade for two reasons: to correct previous errors and to take on new positions if the price adjusts to noninformational trading from the informed investors. The correlation between the current volume and the current returns and expected future returns is ambiguous.

One inconsistency of this analysis is although volume can help to predict future returns uninformed investors in this model do make use of it. A model in which investors also take the information content of volume into account is given in Blume, Easley, and O'Hara

(1994), which we will discuss in Section 2.4.4. First we will refer to a model with generalised information structure, which is provided by He and Wang (1995).

In the dynamic models discussed so far information asymmetry was either strict hierarchical or the information was revealed before the next trading round, as in Admati and Pfleiderer (1988). In He and Wang (1995) this unrealistic assumption is relaxed. They develop a model in which investors have differential information concerning the underlying liquidation value of a stock $\Pi + \delta$. The main focus of their model is the relationship between the pattern of volume and the flow and nature of information. They also analyse the link between volume and price volatility. They find that high volume generated by exogenous private or public information is accompanied by high volatility in prices, whereas high volume generated by endogenous information (like prices) is not accompanied by high volatility.

In their model, there are infinitely many investors, represented by the set $\mathbb{I}$.[42] Each investor can either invest in a bond with a certain gross return rate $R = 1$ or in a stock with a liquidation value $\Pi + \delta$ at the final date $T$. In contrast to Wang (1993) and (1994) this model has a finite horizon and all dividends are paid at the final period $T$. Each investor $i \in \mathbb{I}$ receives a *private* signal $S_t^i$ about the first component of the stock's liquidation value, $\Pi$:

$$S_t^i = \Pi + \epsilon_{S,t}^i,$$

where $\epsilon_{S,t}^i$ is normally i.i.d. with $\mathcal{N}(0, \sigma_{S,t}^2)$ for all investors. They also observe a public signal $Y_t$ about $\Pi$:

$$Y_t = \Pi + \epsilon_{Y,t},$$

where $\epsilon_{Y,t} \sim \mathcal{N}(0, \sigma_{Y,t}^2)$. In addition all traders observe the price $P_t$. The second component of the liquidation value, $\delta$, is never revealed before the terminal date $T$.

Without noisy supply the true value of $\Pi$ would be revealed immediately in $t = 1$. To make the model interesting, the supply of asset is 1 plus a noise term $\Theta_t$. This noise term follows a Gaussian AR(1) process

$$\Theta_t = a_\Theta \Theta_{t-1} + \epsilon_{\Theta,t} \ , \ -1 < a_\Theta < 1.$$

[42] He and Wang make use of charge spaces. For more details about charge spaces see Feldman and Gilles (1985) and Rao and Rao (1983).

This paper provides a way to characterise a linear equilibrium of the above economy in a mathematically tractable way. The vector of state variables of the economy is given by $\Phi_t = (\Pi; \underline{\Theta}_t; \underline{Y}_t; \{\underline{S}_t^i\}_{i \in \mathbb{I}})$ where an underlined capital letter stands for the whole stochastic process up to and including time $t$.

The main goal is to simplify quite dramatically this state space.[43] For characterising the equilibrium price it is useful to derive expected values for $\Pi_t$ and $\Theta_t$ conditional on different information sets. $\hat{\Pi}_t^c$ and $\hat{\Theta}_t^c$ with superscript $c$ is based on publically available information, $\hat{\Pi}_t^{p,i}$, $\hat{\Theta}_t^{p,i}$ with superscript $(p,i)$ are the expected values based *only* on private information, whereas the same terms with superscript $i$ represent the expected values taking all available private and public information of investor $i$ into account. Instead of the hat, $\hat{\cdot}$, the expected value is also written as $E_t^c[\cdot]$, $E_t^{p,i}[\cdot]$, $E_t^i[\cdot]$ whereas the notation for the variance is given by $V_t^c[\cdot]$, $V_t^{p,i}[\cdot]$, $V_t^i[\cdot]$. He and Wang focus on linear REE. Thus $P_t = \mathbf{L}[\Phi_t] = \mathbf{L}[\Pi; \underline{\Theta}_t; \underline{Y}_t; \{\underline{S}_t^i\}_{i \in \mathbb{I}}]$, where $\mathbf{L}[\cdot]$ expresses a linear functional relationship.

Lemma 1 of He and Wang (1995) reduces the necessary state space to $(\Pi; \underline{\Theta}_t; \underline{Y}_t)$, i.e. $P_t = \mathbf{L}[\Pi; \underline{\Theta}_t; \underline{Y}_t]$. This can be shown by using the Law of Large Numbers, since the mean of infinitely many signals converges with probability 1 to $\Pi$.[44] Furthermore, by exploiting the above linear relationship one can replace $\underline{\Theta}_{t-1}$ by $\underline{P}_{t-1}$. We therefore have $P_t = \mathbf{L}[\Pi, \Theta_t, Y_t, \underline{P}_{t-1}, \underline{Y}_{t-1}]$ which can be rewritten as.

$$P_t = a_t \underbrace{(\Pi - \mu_t \Theta_t)}_{=:\xi_t} + \underbrace{b_t Y_t + \mathbf{L}[\underline{P}_{t-1}, \underline{Y}_{t-1}]}_{=\mathbf{L}[\hat{\Pi}_t^c]}.$$

The sum $\xi_t := \Pi - \mu_t \Theta_t$ can be inferred by every investor. Therefore, in a linear REE following information sets are equivalent $\mathcal{F}^c = \{\mathcal{F}_0, \underline{P_t}, \underline{Y_t}\} \Leftrightarrow \{\mathcal{F}_0, \underline{\xi_t}, \underline{Y_t}\}$. He and Wang show in the following steps that the second term $b_t Y_t + \mathbf{L}[\underline{P}_{t-1}, \underline{Y}_{t-1}]$ is equal to $\mathbf{L}[\hat{\Pi}_t^c]$, i.e. it satisfies a specific structure. After this is shown one can conclude that the equilibrium is determined by $\Pi$, $\Theta_t$, and $\hat{\Pi}_t^c$. In order to derive this specific structure we make use of the equivalence between $\mathcal{F}^c = \{\mathcal{F}_0, \underline{P_t}, \underline{Y_t}\}$ and $\{\mathcal{F}_0, \underline{\xi_t}, \underline{Y_t}\}$. The authors apply Kalman filtering to derive the *first order* expectations $\hat{\Pi}^c$, $\hat{\Theta}^c$, i.e. conditional on public information $(\underline{\xi}_t, \underline{Y}_t)$, and $\hat{\Pi}^i$, $\hat{\Theta}^i$, i.e. conditional on all information $(\underline{\xi}_t, \underline{Y}_t, \underline{S}_t^i)$. The stochastic difference

[43] In contrast to the former discussed Wang-papers we have included the signals directly in the state space and not the expected values of $\Pi$.

[44] Remember we have infinitely many investors by using charge spaces.

equations are given by Lemma 2. It is easy to show that $\{\hat{\Pi}_t^i, \hat{\Theta}_t^i, \hat{\Pi}_t^c, \hat{\Theta}_t^c\}$ follows a Gaussian Markov process under filtration $\mathcal{F}_t^i$. Since information is differential in this model, investor $i$'s trading strategy can also depend on *higher-order* expectations, i.e. expectations about the expectations of others etc. He and Wang show in Lemma 3 that higher-order expectations can be reduced to first-order expectations. First, they show that $\hat{\Pi}_t^i$ is a weighted average of $\hat{\Pi}_t^c$ and $\hat{\Pi}_t^{p,i}$. $\hat{\Pi}_t^c$ is given by Lemma using Kalman filtering, where $\hat{\Pi}_t^{p,i}$ follows immediately form the Projection Theorem. The weights $\alpha_t$ and $(1-\alpha_t)$ are independent of $i$ and $\alpha_t$ is given by the ratio $V_t^{p,i}[\Pi_t]/V_t^i[\Pi_t]$. Having represented $\hat{\Pi}_t^i$ as $\alpha_t\hat{\Pi}_t^c + (1-\alpha_t)\hat{\Pi}_t^{p,i}$, one can derive, by integrating over $i$ and by taking conditional expectations, the second-order expectations of $\Pi$ as a weighted average of two first-order expectations. In general, $i$'s higher-order expectations can be expressed as linear function of his first-order expectations. Therefore, it is sufficient if $i$'s optimal trading strategy depends only on his first-order expectations.

For deriving the optimal stock demand it is useful - as usual - to define excess return on one share of stock as $Q_{t+1} := P_{t+1} - P_t$. He and Wang assume for solving the investors' dynamic optimisation problem in Lemma 4 - for the time being - that $Q_t$ and $\Psi_t^i =: E_t^i[\Psi]$, where $\Psi_t^i$ is a simplified state space, follow the Gaussian process

$$Q_{t+1} = A_{Q,t+1}\Psi_t^i + B_{Q,t+1}\epsilon_{t+1}^i,$$

$$\Psi_{t+1}^i = A_{\Psi,t+1}\Psi_t^i + B_{\Psi,t+1}\epsilon_{t+1}^i.$$

They state the Bellman equation, exploit the nice property of CARA utility function in forming expected utility for next period, and derive the optimal stock demand function which is linear in $\Psi_t^i$. Finally they verify that $Q_t$ and $\Psi_t^i$ follow this Gaussian process.

By imposing the market clearing condition the equilibrium price is determined by

$$P_t = [(1-p_{\Pi,t})\hat{\Pi}_t^c + p_{\Pi,t}\Pi] - p_{\Theta,t}\Theta_t = (1 - p_{\Pi,t}\hat{\Pi}_t^c) + p_{\Pi,t}\xi_t.$$

The stock price depends only on $\Pi$, $\Theta_t$, and $\hat{\Pi}_t^c$, so $\mathbf{L}[\hat{\Pi}_t^c]$ summarises the whole history. This is the result that was required to show. The price $P_t$ follows a Gaussian Markov process.

Since $\hat{\Pi}_t^c$ depends on $P_t$, $P_t$ is only determined implicitly. However, finding the explicit solution is trivial, since $\hat{\Pi}_t^c$ is linear in $P_t$. Given the price, one can derive the expected excess return $E_t^i[Q_{t+1}]$ from which it follows that the investor's optimal stock demand is given by

$$X_t^i = d_{\Theta,t}\hat{\Theta}^i + d_{\Delta,t}\underbrace{(\hat{\Pi}_t^i - \hat{\Pi}_t^c)}_{:=\Delta}$$

Given the market clearing condition

$$d_{\Theta,t} = 1,$$

$$d_{\Delta,t} = \frac{\alpha_t}{y_t(1-\alpha_t)}.$$

He and Wang also provide a recursive procedure to calculate the equilibrium. Starting with a guess for the conditional variance of $\Pi$ in $T-1$ they derive coefficients $p_{\Pi,T-1}$, $p_{\Theta,T-1}$, demand, equilibrium price and other parameters. As they proceed backwards they check whether the initial guess of the variance of $\Pi$ was correct. If not they start the procedure with a new initial guess.

Having derived the equilibrium allows us to examine different patterns of trading volume and how private information is gradually impounded into the price. In the benchmark case with *homogeneous information*, i.e. $\sigma_{S,1} = 0$, the true value $\Pi = \hat{\Pi}_t^c = \hat{\Pi}_t^i$ is known immediately and the only remaining risk lies in $\delta$. The equilibrium price in this case is given by $P_t = \Pi - p_{\Theta,t}\Theta_t$ where the second term represents the risk premium. $1/p_{\Theta,t}$ measures the market liquidity in the sense of Kyle (1985). The risk premium increases with the Variance of $\delta$ and over time. The latter increase is due to the fact that the number of trading periods left to unwind speculative positions is decreasing. Furthermore, with only few periods remaining and with $\mid \Theta \mid$ large it becomes less likely that the mean reverting AR(1) process of $\Theta_t$ will reach a value of zero. The volume of trade, $\mathcal{V}^*$ is totally determined by noise trading, which is defined by

$$\mathcal{V}^* = \int_i \mid \Theta_t - \Theta_{t-1} \mid = \mid \Theta_t - \Theta_{t-1} \mid$$

with

$$E[\mathcal{V}^*] = \sqrt{2/\pi Var[\Delta\Theta_t]}.$$

In the case of *differential information* the equilibrium price is given by

$$P_t = [(1 - p_{\Pi,t})\hat{\Pi}_t^c + p_{\Pi,t}\Pi] - p_{\Theta,t}\Theta_t$$

The second component is associated with the risk premium as in the homogeneous information case. The first component reflects investors expectations about the stock's future payoff.

This is not simply proportional to the average of investors' expectations: $\alpha_t\hat{\Pi}_t^c + (1-\alpha_t)\Pi_t$. This is because dynamic trading strategies generate equilibrium prices that differ from those generated by static/myopic strategies since current state variables depend on the history of the economy. The difference between dynamic and myopic strategies also appeared in Brown and Jennings (1989) and in Grundy and McNichols (1989). In particular, the current state depends on past prices. As investors continue to trade, the sequence of prices reveals more information, as shown in their Corollary 1. This tends to decrease $p_{\Theta,t}$, whereas the reduction in remaining trading rounds tends to increase $p_{\Theta,t}$.

The optimal trading strategies consist of two parts. The first represents the supply shock, the second investors' speculative positions. It is important to notice that the trading activity generated by differential information is *not* the simple sum of each investor's speculative investments. This is because, in the case of heterogeneous information, non-informational trade by one investor could be viewed as an informational trade by another. It is also possible that investors on both sides of the trade think that their trades are non-informational, but the trading is purely due to differential information.[45] The paper focuses on the *additional* trading volume generated by differential information.

$$V_t := \mathcal{V}_t - \mathcal{V}_t^*.$$

It's expected value is given by

$$E[V_t] = \frac{1}{\sqrt{2\pi}}(\sqrt{Var[\delta\Theta_t] + Var[\delta x_t^i]} - \sqrt{Var[\delta\Theta_t]}),$$

where $x_t^i := X_t^i - \Theta_t = p_{\Pi,t}/p_{\Theta,t}(\hat{\Pi}_t^{p,i} - \Pi)$ is trader $i$'s total trading activity associated with differential information. In Corollary 2, He and Wang provide a closed form solution for the equilibrium volume in the special case where $\sigma_\delta = 0$. It says that informational trading occurs only as long as investors receive new private information. In this case the individual trader does not know whether the other investors trade because of new information or because of liquidity reasons. It is not common knowledge whether the allocation is Pareto efficient. This is the reason why the dynamic version of the No-Speculation theorem given in Geanakoplos (1994) does not apply in this case.[46] If on the other hand investors receive

[45] For example there is no additional noise in $t$, but half of the traders think $\Theta_t = +0.1$ and the other half thinks $\Theta_t = -0.1$

[46] Similar to Grundy and McNichols (1989) the beliefs about future signals need not be concordant when $\sigma_\delta > 0$.

only private information in $t = 1$ the prices will adjust, but no informational trade will occur. This is the difference to the possible no-trade equilibrium in Grundy and McNichols (1989) which can only occur, if $a_\Theta = 1$ and investors receive information only in the first period.

He and Wang then go on to analyse the behaviour of trading volume after $t = 2$ for the case where only the signal in $t = 1$ is informative. The main findings are that trading persists throughout the whole trading horizon. This is due to the fact that investors establish their speculative position, when they receive their private information in $t = 1$ and then gradually try to unwind their positions. This generates peaks in the volume of trade in the middle of the trading horizon. In the case of public announcements, investors increase their positions right before and close them right after the announcement. Therefore, volume and total amount of information revealed through trading depends on the timing of the announcement. Market liquidity drops right before the announcement and bounces back afterwards. They also find that new information, private or public, generates both high volume and large price changes, while existing private information can generate high volume with little price changes.

### 2.4.4 Inferring Information from Trading Volume in a Competitive Market Order Model

One drawback of the models, discussed so far, is that investors do not extract the predictive power of volume. In Blume, Easley, and O'Hara (1994) a group of traders make explicit use of volume data to improve their prediction of the liquidation value of an asset. In contrast to the models discussed so far (with the exception of Romer (1993)) the quality or precision of each traders' signal is not common knowledge. In the sense of Morris, Postlewaite, and Shin (1995) this model exhibits a higher depth of knowledge by one degree. Information about the fundamentals, i.e. payoff relevant events, is dispersed among the traders, but there is also asymmetry in the information about the quality of the traders' signals. In other words, this model incorporates asymmetric second order information. Every investor receives a private signal about the liquidation value, $\Pi$ of the asset. Each investor knows the quality of his signal, but only a subset of investors, investors in group 1, knows the precision of all signals. The higher order uncertainty about the precision of other investors' signals is the source of noise in their model. This higher order uncertainty provides ground for the predictive power of volume and technical analysis.

Blume, Easley, and O'Hara (1994) start their analysis by showing why the models in Brown and Jennings (1989) and Grundy and McNichols (1989) are not appropriate for analysing the role of volume in predicting the value of an asset.

In the framework of Brown and Jennings (1989) there always exists an informationally efficient REE if trade conditional on price and volume is possible. In this case, trader $i$'s information set is $\mathcal{F}_t^i = (P_t, S_t^i, \mathcal{V}_t, \chi_t)$ where $\mathcal{V}_t$ is the per capita volume

$$\mathcal{V}_t = \frac{1}{2}\frac{1}{I}[\sum_{i=1}^{I} \mid x_t^i \mid + \mid \Theta_t \mid]$$

and $\chi_t^i$ is an indicator function indicating whether the trader $i$ is a buyer or seller. In equilibrium traders demand the same amount of the risky assets $x_t^i = x_t^j =: x_t$. By the market clearing condition $x_t = \Theta_t$. Thus each trader can infer from his demand and $\chi_t^i$ the noisy supply term $\Theta_t$. As the equilibrium price depends only on $\Theta_t$ and the average signal $\overline{S}_t$ can be inferred. Thus, in each period $t$, the tuple $(P_t, \mathcal{V}_t)$ fully reveals $\overline{S}_t$ and $\Theta_t$ and thus technical analysis has no value.

In Grundy and McNichols (1989) each individual is endowed with an i.i.d. random number of risky assets. The variance of this random endowment, $\Theta_t^i$, is given by $Var[\Theta_t^i] = I\sigma_\Theta^2$. In the analysed limit economy with infinitely many traders, i.e. $I \to \infty$, each individual endowment itself has no information content and the variance of the endowments is also infinite. Since the expected trading volume per capital is $\mathcal{V}_t = \frac{1}{2}lim_{I\to\infty}\frac{1}{I}\sum_{i=1}^{I} \mid x_t^i - \Theta_t^i \mid$, it is infinite as well and no inference can be drawn from it.

In contrast to Brown and Jennings (1989) and Grundy and McNichols (1989), Blume, Easley, and O'Hara (1994) develop a market order model with a generalised information structure. More precisely the information set of each trader $i$ contains the whole price and volume process up to but *excluding* the current time $t$. This distinguishes these models, first developed in Hellwig (1982), from limit order models. The normally used limit order models exhibit two drawbacks. First, since volume is not normally distributed, the inference of current volume can be quite cumbersome. Second, revealing REE in a Grundy-McNichols or Brown-Jennings limit order model always exist, since traders can always condition on the information contained in their own net trade: its direction and its magnitude Jordan (1983).

Blume, Easley, and O'Hara (1994) assume the following information structure. The common priors for all traders about the liquidation value are $\Pi \sim \mathcal{N}(\mu_{\Pi,0}, \sigma^2_{\Pi,0})$. Each trader in group 1 receives a signal

$$S^i_t = \Pi + \omega_t + e^i_t,$$

where $\omega_t \sim \mathcal{N}(0, \sigma^2_\omega)$, $e^i_t \sim \mathcal{N}(0, \sigma^2_{e,t})$ and $\omega_t$ and all $e^i_t$ are independent. Note $\sigma^2_{e,t}$) varies over time.
Each trader in group 2 receives a signal

$$S^i_t = \Pi + \omega_t + \epsilon^i_t,$$

where all $\epsilon^i_t$ are i.i.d. $\mathcal{N}(0, \sigma^2_\epsilon)$. It is common knowledge that $I_1 = \nu I$ traders are in group 1 and $I_2 = (1-\nu)I$ traders are in group 2. The asymmetry about the second order information is captured by asymmetric knowledge about the precision of the signals. Traders in group 1 know the precision of group-1 signals, $(1/\sigma^2_{e,t})$ in each $t$ and in addition the precision of the signals, group 2 traders get, $(1/\sigma^2_\epsilon)$. Group 2 trades only know the signal precision of their own group. In contrast to $(1/\sigma^2_\epsilon)$ the precision of group 1 signals, $\sigma^2_{e,t}$, varies randomly over time. This rules out the possibility that traders in group 2 learn over time the group 2 signal precision.
The distribution of the signals is, therefore, given by:
for group 1 signals: $S^i_t \sim \mathcal{N}(\Pi, \sigma^2_{S_{1,t}})$, where $\sigma^2_{S_{1,t}} = \sigma^2_\omega \sigma^2_{e,t}[\frac{1}{\sigma^2_\omega} + \frac{1}{\sigma^2_{e,t}}] =: V[S_{1,t}]$
for group 2 signals: $S^i_t \sim \mathcal{N}(\Pi, \sigma^2_{S_2})$, where $\sigma^2_{S_2} = \sigma^2_\omega \sigma^2_\epsilon[\frac{1}{\sigma^2_\omega} + \frac{1}{\sigma^2_\epsilon}] =: V[S_2]$.

It is obvious from the Strong Law of Large Numbers that the average of the signals in each group, $\overline{S}_{1,t}$ and $\overline{S}_{2,t}$ converges almost surely to $\Pi + \omega_t =: \theta_t$.

Blume, Easley, and O'Hara (1994) restrict their analysis to myopic REE. The individual demand for traders with a constant absolute risk aversion coefficient of unity, i.e. $\rho = 1$, is *approximated* by, for group 1 traders:

$$x_{1,t}^i = \frac{E_{t-1}^{i,1}[\Pi] - P_t}{V_{t-1}[\Pi]} + \frac{S_{1,t}^i - P_t}{V_{t-1}[S_{1,t}]},$$

and for group 2 traders:

$$x_{2,t}^i = \frac{E_{t-1}^{i,2}[\Pi] - P_t}{V_{t-1}[\Pi]} + \frac{S_{2,t}^i - P_t}{V_{t-1}[S_2]}.$$

In contrast to limit order models, there is an additional second term and the expectations are taken with respect to $\mathcal{F}_{t-1}^i$. Adding the demand functions and imposing the market clearing condition gives the equilibrium price. For the limit economy, $P_1$ is

$$P_1 = \frac{\frac{1}{\sigma_{\Pi,0}^2}\mu_{\Pi,0} + [\nu\frac{1}{\sigma_{S_{1,1}}^2} + (1-\nu)\frac{1}{\sigma_{S_2}^2}]\theta_1}{\frac{1}{\sigma_{\Pi,0}^2}\nu\frac{1}{\sigma_{S_{1,1}}^2} + (1-\nu)\frac{1}{\sigma_{S_2}^2}}.$$

Group 1 trader can infer $\theta_1$ from $P_1$, since they know $\sigma_{S_{1,1}}^2$ and $\sigma_{S_2}^2$. $P_1$, however, does not reveal $\theta_1$ for group 2 traders, since they do not know $\sigma_{S_{1,1}}^2$. Note that the conditional distribution of $\theta_1$ given $P_1$ is not normal. Traders in group 2 can infer more information about $\theta_1$ if they include trading volume in their inference calculation. The per capita trading volume in $t = 1$ is:

$$\mathcal{V}_1 = \frac{1}{2}\frac{1}{I}(\sum_{i=1}^{I_1} \mid x_{1,t}^i \mid + \sum_{i=I_1}^{I} \mid x_{2,t}^i \mid)$$

Volume is not normally distributed, because volume is the absolute amount of normally distributed random variables. Blume et al. explicitly characterise the expected per capita volume $\mathcal{V}_1$ in their Proposition 1.

$$\mathcal{V}_1 = \mathcal{V}_1(\theta_1 - P_1, \sigma_{S_{1,1}}^2, \sigma_{S_2}^2, \ldots).$$

Using the equilibrium price relation, one can substitute for $(\theta_1 - P_1)$ a term depending on the signal precisions. The resulting equation shows that volume conveys information about the signal quality of group 1 traders $(1/\sigma_{S_{1,1}}^2)$.

Plotting the derived expression for $\mathcal{V}_1$ with $P_1$ on the abscissa yields a V-shape relationship between price and volume, for any given $(1/\sigma_{S_{1,1}}^2)$. The minimum volume is reached at a price level $P_1 = \mu_{\Pi,0}$. The average traders' posterior means coincide with the prior mean. As $P_1$ deviates from $\mu_{\Pi,0}$, posterior means are changed and the first term of the individual demand functions $x_t^i$ increases the trading volume on average. This results in a strong correlation between volume and price change. The V-shape is very robust. As the signal precision (information quality of group 1 signals) decreases the V-shape becomes more pronounced. The same is true when the quantity of information, i.e. fraction of group 1 traders, decreases.

Keeping the price fixed and differentiating expected capital volume with respect to the precision of trader 1 signals $(1/\sigma^2_{S_{1,1}})$ yields that volume is increasing in the precision of group 1's signals if $(1/\sigma^2_{S_{1,1}}) < (1/\sigma^2_{\omega})$ and decreasing if $(1/\sigma^2_{S_{1,1}}) > (1/\sigma^2_{\omega})$ (provided $(1/\sigma^2_{S_{1,1}}) > (1/\sigma^2_{S_2})$). Intuitively, if trader 1s' signals are very imprecise, their signals are very dispersed and they place little confidence in their signal. They do not trade very aggressively and thus the expected trading volume is low. If on the other hand the signals are very precise, all group 1 traders receive highly correlated signals and thus the trading volume is low again, since trade occurs only between the groups. Therefore, high volume can be a signal for very precise signals but also for very imprecise signals. Volume is first increasing and then decreasing in the signal precision for a given price $P_1$ and, therefore, for a observed price volume pair $(P_t, \mathcal{V}_t)$ two outcomes, high or low precision are feasable. In other words, the functional relationship is not invertible. Therefore, Blume et al. restrict their analysis to the increasing branch of $\mathcal{V}$, i.e. $1/\sigma^2_{S_{1,1}} \in (1/\sigma^2_{S_2}; 1/\sigma^2_{\omega})$. If this is the case the tuple $(P_t, \mathcal{V}_t)$ is revealing $(\theta_1, \sigma^2_{S_{1,1}})$. Since all signals incorporate the common error term $\omega_t$, the liquidation value $\Pi = \theta_1 - \omega_1$ is not known.

In a dynamic setting more realisations of $\theta_t = \Pi + \omega_t$ can be inferred and, therefore, a better estimate about the true liquidation value, $\Pi$ can be made. In each period the precision of the signals for traders in group 1 is drawn randomly and the analysis is similar to the static case. One difference is that priors in period $t$ are not exogenous, but derived from the market statistics up to time $t-1$. Second the volume expression is slightly different, since traders' endowments in $t$ are the equilibrium demands in $t-1$. By the Strong Law of Large Numbers the equilibrium price converges almost surely to $\Pi$, since in each period traders can infer a new $\theta_t$. As time proceeds trade does not vanish, because although traders' beliefs are converging, their precision is diverging at the same rate. Intuitively, in the early trading round traders beliefs are widely dispersed and, therefore, they trade less aggressively. In later trading periods the beliefs are much closer to each other. Since traders are more confident, they take on larger positions. Both effects are offsetting each other and, therefore, volume does not decline with the number of trading rounds.

In the last section Blume, Easley, and O'Hara (1994) compare the utility of a trader who makes use of past market statistics in interpreting the current market statistics with a trader who bases his trading activity only on current market statistics and his priors in $t = 0$. The value of technical analysis is then defined by the amount of money the latter trader, who forgets all past market data, would be willing to pay to know the forgotten past market statistics. Past market data have value because of the common error term $\omega_t$ in the signals. Blume et al. show that the value of technical analysis is decreasing in $\sigma_\omega^2$ and increasing in $\sigma_{\Pi,0}^2$. They conclude that technical analysis has higher value for small, less widely followed stocks.

Note, in Blume, Easley, and O'Hara (1994) all traders trade purely for informational reasons. Nobody faces liquidity shocks and there are no noise traders. There are no gains from trade since agents' endowments and preferences are identical. One might think that the No-Speculation Theorem of Milgrom and Stokey (1982) should apply. The No-Speculation Theorem requires at least higher order mutual knowledge of rationality by all agents. The No-Speculation Theorem need not hold if rationality of all traders is not common knowledge. Blume, Easley, and O'Hara (1994) apply the Rational Expectations Equilibrium (REE) concept. In a REE each agent is only assumed to know the mapping from traders' information into prices. In contrast to a game theoretic equilibrium concept, the REE concept does not specify the cognitive capacity, an agent assumes his opponent players have. In particular, REE does not require common knowledge of rationality. In Blume, Easley, and O'Hara (1994) all traders behave rationally, but they might not be sure whether their opponents are rational. This higher order uncertainty about rationality makes trading possible and the No-Speculation Theorem cannot be applied.

### 2.4.5 Sequential Information Arrival Models and Technical Analysis

In the models with differential information covered so far, all traders receive their information at the same time. Likewise, in the asymmetric information models we discussed informed investors received the same private information at the same time. In this section we will discuss asset trading in situations in which traders receive their information sequentially.

The impact of sequential information arrival on price process and volume was first analysed in Copeland (1976). In his model traders not only take the price of the single risky asset as given but also neglect that the price can convey information. In each period one trader receives a signal. After receiving his signal he alters the intercept of his linear demand function. Hence, in each period of time one trader adjust his demand curve and thus the aggregate demand curve changes. In every period (temporary) "incomplete" equilibria determine the price until a new "complete equilibrium" is reached when all traders have received their signal. Copeland introduces short sales constraints which influence price and volume.

In Section 1 of Copeland (1976) all traders get the same signal (or equivalently they interpret the signal equally) one after the other. The price and volume changes are the same in each period as long as the short sales constraint is not binding for any trader. When the short sales constraint becomes binding, price change and volume decrease over time. After how many trading rounds short sale constraint starts binding, depends on the number of traders, the strength of information, i.e. the intercept shift of the individual demand caused by the new information, and on the supply of the assets. Copeland (1976) generalises the setting to one in which traders can receive different, i.e. either a positive or a negative signal. Stated equivalently they interpret the received signal either optimistically or pessimistically. In this case it is less likely that short sales constraints become binding, since optimists increase their demand, whereas pessimists decrease it. The price change in this case depends crucially on the random order of the signal arrival. There are many different possible ordering of positive and negative signals. Copeland shows that there is still a strong correlation between volume and absolute price change. He also compares the price adjustment process in the sequential information arrival model with one in which information is revealed simultaneously to all traders, a tâtonnement model.

Jennings, Starks, and Fellingham (1981) analyse the relationship between price-change and volume by introducing margin requirements for traders who want to sell short instead of prohibiting short sales altogether. They demonstrate that margins have a significant impact on the relationship between price changes and volume. As in Copeland (1976) the relationship depends on the arrival of signals, on the number of investors and in addition on the implicit cost of the imposed margin requirement.

Treynor and Ferguson (1985) analyse the decision problem of an individual investor who receives information about the value of a stock. The investor roughly knows the price impact of his information, but not whether it is really new information or not, i.e. if other traders have already received it before him. In other words, the order of the sequential information arrival is random and not known. If all other traders, i.e. the markets has already received his information, then it is already incorporated in the price and, therefore, he should not trade. If, on the other hand, he is the first to receive this signal, he should buy or sell the relevant assets.

Treynor and Ferguson (1985) consider a stylised situation in which events are becoming public quite fast. More formally, the event gets known to trader $i$ and to all other traders before a new event occurs. Let $t_E$ be the time of the event, $t_i$ the time when trader $i$ receives the signal and $t_M$ the time when the market knows the information. The authors consider only cases in which $(t_M - t_E)$ and $(t_i - t_E)$ are very short in comparision to the time between events, $(t_{E_j} - t_{E_{j-1}})$.

The investor wants to know the probability of $t_i > t_M$ versus $t_i \square t_M$. For deriving these probabilities he makes use of his

(1) prior distribution about the information dissemination and

(2) the observed price process, combined with his knowledge about the

(a) underlying stochastic process of the price path

(b) price impact of information at $t_M$.

In Treynor and Ferguson (1985) the prior probability that an event occurred in $t_E$ is uniformly distributed within a certain time span with length $\delta$, i.e. the densitiy is $(1/\delta)$. $\gamma$ is the probability that all other traders will receive the information in the next period, provided they have not received it so far. Similarly, $\alpha$ is the probability that investor $i$ will receive the information in the next period. This determines the transition probabilities for the Markov process with the following four possible states: $\omega_1$ nobody, $\omega_2$ only trader $i$, $\omega_3$

only all other traders, and $\omega_4$ all traders, received this signal.

Investor $i$ makes use of his knowledge about the underlying price process governed by $P_t = (1 + r_t)P_{t-1}$, where all $r_t$ are i.i. normally distributed with mean zero and variance $\sigma^2$. The price changes by a multiplicative factor $\exp(V)$, at $t_M$, the time when the event becomes known to all traders. If $t_i$ and $t_M$ are known the distribution of possible price paths is given by

$$\Pr(\underline{r}_{t_i} \mid t_M, t_i) = \prod_{t=-\infty}^{t_i} \{\frac{1}{\sqrt{2\pi}\sigma} \exp[\frac{-r_t^2}{2\sigma^2}]\} \ \ \forall t_M > t_i,$$

$$\Pr(\underline{r}_{t_i} \mid t_M, t_i) = \prod_{t=-\infty}^{t_i} \{\frac{1}{\sqrt{2\pi}\sigma} \exp[\frac{(-r_t - \delta_t \mathrm{t_M V})^2}{2\sigma^2}]\} \ \ \forall t_M \Box t_i.$$

where $\underline{r}_{t_i}$ denotes the whole process of returns $r_t = \frac{P_t - P_{t-1}}{P_{t-1}}$ up to and including $t_i$.

Investor $i$ is interested in the probability distribution $\Pr(t_M \mid \underline{r}_{t_i}, t_i)$, which by Bayes' Rule is

$$\Pr(t_M \mid \underline{r}_{t_i}, t_i) = \frac{\Pr(t_M \mid t_i) \ \Pr(\underline{r}_{t_i})}{\sum_{t_m=-\infty}^{\infty} \Pr(t_M \mid t_i) \ \Pr(\underline{r}_{t_i} \mid t_M, t_i)}.$$

All terms are known except $\Pr(t_M \mid t_i)$. $\Pr(t_M \mid t_i)$ can be rewritten as

$$\Pr(t_M \mid t_i) = \sum_{t_E} \Pr(t_M \mid t_i, t_E) \Pr(t_E \mid t_i),$$

where $\Pr(t_M \mid t_i, t_E) = \Pr(t_M \mid t_E)$, since, given an event occurred, the probability that all other agents get the information only at $t_M$ is independent of when agent $i$ (will or) has received the information. All these probabilities can be directly derived from the given prior information structure.

Treynor and Ferguson provide a numerical example where the trader $i$ infers from the past price process that, with probability of 70 percent, all other traders have not yet received the same information.

In their last section, the authors deliver an optimal portfolio strategy which allows the investor $i$ to capitalise on his information. Their article shows that technical analysis, i.e. inferring information from past prices, helps in the evaluation of new private information.

### 2.4.6 Strategic Multi-Period Market Order Models with a Market Maker

Admati and Pfleiderer (1988) analyse a strategic dynamic market order model.[47] Their model is essentially a dynamic repetition of a generalised version of the static model in Kyle (1985). However, their focus is on intraday price and volume patterns. They attempt to explain the U-shape of the trading volume and price changes, i.e. the abnormal high trading volume and return variability at the beginning and at the end of a trading day. In their model the value of a single risky asset follows the exogenous process

$$\Pi = \overline{\Pi} + \sum_{t=1}^{T} \delta_t$$

where $\delta_t$ is a i.i.d. random variable, whose realisation becomes common knowledge only at $t$. As usual there are two motives for trading: information and liquidity. All $I_t$ informed traders observe the same signal $S_t = \delta_{t+1} + \epsilon_t$ at time $t$, where $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$. In other words, informed traders observe a noisy version of the public information one period in advance. Since $\delta_{t+1}$ is known publically in $t+1$ the informational advantage is only *short-lived.* Informed traders have, therefore, no incentive to restrict their trading in order to have a larger informational advantage in the next period. This simplifies the analysis dramatically. However, analysing only short-lived information neglects interesting aspects. In Admati and Pfleiderer (1988) there are two types of liquidity traders, whose demand depends neither on the price nor on their information. Whereas $J_t$ *discretionary liquidity traders* can choose a period within $[T', T'']$ in which to trade, *nondiscretionary liquidity traders*, must trade a given amount at a specific time. For simplicity it is assumed that the market maker as well as all traders are risk neutral. As in Kyle (1985) the market maker observes the total net order flow $Y_t$, in addition to $\underline{\delta}_{t-1} := (\delta_0, \delta_1, ..., \delta_{t-1})$. The total net order flow in $t$ is given by

$$Y_t = \sum_{i=1}^{I} x_t^i + \sum_{j=1}^{J_t} z_t^j + \Theta_t$$

where the first component represents the aggregated demand from informed traders, the second the aggregated demand from discretionary liquidity traders and the third the aggregated demand from nondiscretionary liquidity traders. The market maker tries to infer the information of the insiders from $Y_t$. The variance of total liquidity trading, $\Psi_t =$

[47] To be consistent with our notation we denote: $\Pi = F$, $\overline{\Pi} = \overline{F}$, $I_t = n_t$, $J_t = m_t$, $Y_t = \omega_t$, $z_t^j = y_t^j$, $\Theta_t = z_t$, and $\sigma_\epsilon^2 = \phi_t$.

$Var(\sum_{j=1}^{J_t} y_t^j + \Theta_t)$, is in contrast to Kyle (1985) endogenously determined, as it depends on the strategic decision of the discretionary traders. As in Kyle (1985) the zero profit condition together with risk neutrality implies that the market maker sets the price equal to his expected value. A linear pricing rule is, therefore, given by

$$P_t = \overline{\Pi} + \sum_{\tau=1}^{t} \delta_\tau + \lambda_t Y_t,$$

where $\frac{1}{\lambda_t}$ again measures the market depth. Given this pricing rule the equilibrium value of $\lambda_t$ is decreasing with the number of informed traders $I_t$. As in Kyle (1985) the market depth $\frac{1}{\lambda_t}$ is increasing with $\Psi$, the variance in liquidity traders demand. Admati and Pfleiderer assume that discretionary liquidity traders take $\lambda_t$ as given, although their trading intensity affects $\lambda_t$. The costs of trading for the liquidity traders, which equals the profit for insiders, is the difference between what the liquidity traders pay and the expected value, i.e. $E[(P_t(\underline{\delta}_t, \underline{Y}_t) - \Pi)(\sum_{j=1}^{J_t} z_t) \mid \underline{\delta}_t, \underline{Y}_{t-1}, \sum_{j=1}^{J_t} z_t]$ which is equal to $\lambda_t(\sum_{j=1}^{J_t} z_t)^2$. Therefore, discretionary liquidity traders would trade when $\lambda_t$ is smallest, i.e. when the market is deepest. This is the case when $\Psi_t$ is high and thus it is optimal for them to "clump" together, which increases $\Psi_t$ even more. High variance in noise trading, $\Psi_t$ allows insiders to hide more of their trade behind noise trade. Their demand in equilibrium is given by $x_t^i = \beta_t^i S_t$, where $S_t$ is the signal about $\delta_{t+1}$ and $\beta_t = \sqrt{\frac{\Psi_t}{Var(S_t)}}$. Thus, at times when liquidity traders clump together, informed traders also trade more aggressively. This increases the overall trading volume in this trading period. The problem is, however, that discretionary traders have to coordinate when to trade and, therefore, many equilibria can arise. It is plausible that the convention arose that they all trade in the beginning and at the end of the trading day. Admati and Pfleiderer (1988) show that equilibria in which discretionary traders clump together exist. They also apply a refinement criterium, which shows that these equilibria are the only ones that are robust to small perturbations in the vector of variances of the discretionary liquidity demands. As in Kyle (1985), the amount of information revelation by prices is independent of the total variance of liquidity trading. More noise trade would suggest less informative prices. On the other hand, more noise allows insiders to be more aggressive in their trade which making the price more informative. The aggressiveness of the insiders is such that both effects will balance out.

Admati and Pfleiderer (1988) then incorporate endogenous information acquisition. Traders can buy the signal $S_t$ at a fixed costs $c$. This makes the number of informed traders, $I_t$, endogenous. The authors apply two different equilibrium concepts. In the second concept each insider's strategy depends on $I_t$. When $I_t$ is high, insiders compete with each other and therefore their profits will be lower, or equivalently, the trading costs for liquidity traders will be lower. At times when discretionary traders clump together, $\Psi_t$ is high and, therefore, many insiders will enter the market. This reduces the trading costs for liquidity traders even more since the insiders are competing against each other. Thus, endogenous information acquisition intensifies the effects explained above and one would expect large trading volume at certain times.[48]

Admati and Pfleiderer's model is a very simplified picture of reality. First, information is only asymmetric for one period, i.e. it is short-lived. Long-lived asymmetric information was considered in Foster and Viswanathan (1990) and in Holden and Subrahmanyam (1994) for risk averse traders. The second simplification is that information is only asymmetric but not differential, i.e. all insiders observe the same signal. This assumption is relaxed in the next paper discussed.

Foster and Viswanathan (1996) extend the model in Kyle (1985) to the case with $I$ risk neutral informed investors.[49] Each investor gets a long-lived individual private signal at $t = 0$. In contrast to Admati and Pfleiderer (1988), there are no discretionary liquidity traders. (Nondiscretionary) liquidity traders demand $\Theta_t \sim \mathcal{N}(0, \sigma_\Theta^2)$ shares in period $t \in [1, ..., T]$. As in Kyle (1985) the market maker only observes the total net order flow $Y_t = \sum_{i=1}^{I} x_t + \Theta_t$ and sets the price at time $t$ according to

$$P_t = E[\Pi \mid \underline{Y}_t],$$

[48]Pagano (1989a) provides a model which illustrates the negative correlation between trading volume and market thinness as well as volatility. In this model risk-averse investors' value the stock for hedging reasons differently and have to pay a fixed transaction cost to enter the market. Each additional trader who enters the market reduces the market thinness and thus volatility. This generates for the other risk-averse traders a positive externality. Pagano (1989a) shows that there are multiple "bootstrap" equilibria, some with low trading volume and high price volatility, and others with high trading volume and low volatility. The latter are Pareto superior. Pagano (1989b) shows that, in presence of different transaction costs, traders may be unable to co-ordinate on a single market.

[49]For consistency we adjust the notation to $I = M$, $t = n$, $Y_t = y_n$, $\Pi = v$, $S_0^i = s_{i,0}$, $\overline{S} = \hat{v}$, $P_t = p_n$, $\hat{S}_{0,t}^i = t_{in}$, $S_t^i = s_{in}$ and $\because^i = \cdot^i/$.

where his prior distribution of $\Pi$ is given by $\mathcal{N}(P_0, \sigma_\Pi^2)$ and $\underline{Y}_t$ denotes the whole process $(Y_1, ..., Y_t)$. Informed traders $i \in \mathbb{I} = \{1, 2, ..., I\}$ have to submit their market orders $x_t^i$ before $Y_t$ becomes known. Since each trader $i$ knows his individual demand, $\underline{x}_t^i$ and the whole history of $\underline{Y}_{t-1}$ he can infer the net order flow of *all other* traders $\underline{z}_{t-1} = \underline{Y}_{t-1} - \underline{x}_{t-1}^i$. Each informed trader receives an individual signal $S_0^i$ at the starting point of trading. The joint distribution of all individual signals with the asset's true value is given by

$$(\Pi, (S_0^1, ..., S_0^I)) \sim \mathcal{N}((P_0, \vec{0})), \begin{pmatrix} \sigma_\Pi^2 & \Delta_0 \\ \Delta_0 & \Psi_0 \end{pmatrix}),$$

where $\Delta_0$ is a vector with $I$ identical elements, i.e. $\Delta_0' = (c_0, c_0, ..., c_0)$ and $\Psi_0$ is the variance-covariance matrix of the signals given by

$$\Psi_0 = \begin{pmatrix} \Lambda_0 & \Omega_0 & ... & \Omega_0 \\ \Omega_0 & \Lambda_0 & ... & \Omega_0 \\ ... & ... & ... & ... \\ \Omega_0 & \Omega_0 & ... & \Lambda_0 \end{pmatrix}.$$

This signal structure imposes a strong symmetry assumption, since (a) all signals have the same covariance $c_0$ with the true asset value, (b) all signals have the same variance $\Lambda_0$ and (c) the cross variance between signals is $\Omega_0$ for all signals. It covers also the special cases $\Omega_0 = \Lambda_0$ where all insiders get the same signal and $\Omega_0 = 0$ where all signals are independent.

By applying the Projection Theorem one gets

$$E[\Pi - P_0 \mid S_0^1, ..., S_0^I] = \Delta_0'[\Psi_0]^{-1} \begin{pmatrix} S_0^i \\ ... \\ S_0^I \end{pmatrix}.$$

By the imposed symmetry assumptions, all elements of the vector $\Delta_0'[\Psi_0]^{-1}$ are identical, say to $\square$ Therefore, the inner product can be rewritten as

$$E[\Pi - P_0 \mid S_0^1, ..., S_0^I] = \underbrace{\square I}_{:=\theta} \underbrace{\frac{1}{I} \sum_{i=1}^{I} S_0^i}_{:=\overline{S}} = \theta \overline{S}.$$

$\overline{S}$, the average of all signals $S_0^i$ is a sufficient statistic for all signals. It follows that the market maker and the informed traders need not infer each individual signal $S_0^i$ but only the average signal $\overline{S}$. This allows us to simplify the sufficient state description dramatically.

The market maker's estimate of $S_0^i$ at $t$ is given by

$$\hat{S}_{0,t}^i := E[S_0^i \mid Y_1, ..., Y_t] = E[S_0^i \mid \underline{Y_t}].$$

The market maker sets a competitive price $P_t = E[\Pi \mid \underline{Y_t}]$. Since $(S_0^1, ..., S_0^I)$ is a sufficient statistic for $Y_t$ and for[50] $P_0 = 0$

$$P_t = E[E[\Pi \mid S_0^1, ..., S_0^I] \mid Y_t] = \theta E[\overline{S} \mid \underline{Y_t}] = \theta \frac{1}{I} \sum_{i=1}^{I} \hat{S}_{0,t}^i.$$

The informational advantage of informed trader $i$ in period $t$ is the difference

$$S_t^i := S_0^i - \hat{S}_{0,t}^i.$$

Foster and Viswnathan further define the following conditional variances and covariances

$$\Sigma_t := Var(\theta\overline{S} \mid \underline{Y_t}) = Var(\theta\overline{S} - P_t) = Var(E[\Pi \mid S_0^1, ..., S_0^I] \mid \underline{Y_t}),$$

$$\Lambda_t := Var(S_0^i \mid \underline{Y_t}) = Var(S_t^i),$$

$$\Omega_t := Cov(S_0^i, S_0^j \mid \underline{Y_t}) = Cov(S_t^i, S_t^j),$$

and derive the following relationships (using the Projection Theorem):

$$\Sigma_t = \frac{\theta^2}{I}[\Lambda_t + (I-1)\Omega_t],$$

which implies

$$\Lambda_{t-1} - \Lambda_t = \Omega_{t-1} - \Omega_t,$$

$$\Sigma_{t-1} - \Sigma_t = \theta^2[\Lambda_{t-1} - \Lambda_t],$$

and therefore

$$\Lambda_t - \Omega_t = \chi \;\; \forall t.$$

Since the market maker will learn the average signal $\overline{S}$ much faster than any individual signal, the correlation between the informational advantage of insiders, $\Omega_t$, must become negative after a sufficient number of trading rounds. This negative correlation between $S_t^i$ will cause the waiting game explained below. Foster and Viswanathan use a Bayesian Nash equilibrium concept given the price setting behaviour of the market maker and restrict their analysis to linear Markov equilibria. As in Kyle (1985) the authors apply Bayesian Nash

[50] In equations (2) and (5) in the paper the prior mean of $P_0$ is missing. Thus they apply only for $P_0 = 0$.

equilibrium concept. The equilibrium is represented by a tuple $(\mathbf{X}^1, ..., \mathbf{X}^I, \mathbf{P})$ where $\mathbf{X}^i$ is a vector of demand correspondences for trader $i$ for each date, $t$, i.e.

$$\mathbf{X}^i = (x_1^i, ..., x_T^i), \ where \ x_t^i = x_t^i(S_0^i, \underline{Y}_{t-1}^i, \underline{x}_{t-1}^i),$$

and $\mathbf{P}$ is a vector of price setting functions for each $t$, i.e.

$$P_t = P_t(\underline{Y}_t) = E[\Pi \mid \underline{Y}_t].$$

$x_t^i(\cdot)$ is the stock holding of trader $i$ at time $t$ which maximises his profits from time $t$ until $T$. $\mathbf{X}^i(\cdot)$ is optimal by backward induction. Foster and Viswanathan impose a Markov Perfect refinement criterion on the possible set of equilibria. How restrictive this criterion is depends on which state space the (trade) strategies can be based on. There are, therefore, two different state spaces: The first state space is given by the choice of nature, whereas the second covers events of the original state space, on which traders can base their trading strategies. The smaller the latter state space is, the more restrictive is the Markov Perfect refinement criterion. The state space given by the choice of nature is $(\Pi, \{S_0^i\}_{i \in \mathbf{I}}, \underline{\Theta}_T)$. Incorporating the choice of each trader, one can consider the following extended state space$(\Pi, \{S_0^i\}_{i \in \mathbf{I}}, \{\underline{x}_T, \underline{\Theta}_T\}_{i \in \mathbf{I}})$. Knowing $\Theta_t = \sum_{i=1}^I x_t^i - Y_t$, $(\Pi, \{S_0^i\}_{i \in \mathbf{I}}, \{\underline{x}_T, \underline{Y}_T\}_{i \in \mathbf{I}})$. This state space can be rewritten as $(\Pi, \{\underline{S}_T^i\}_{i \in \mathbf{I}}, \{\underline{\hat{S}}_{0,T}^i\}_{i \in \mathbf{I}}, \{\underline{x}_T, \underline{Y}_T\}_{i \in \mathbf{I}})$, as $S_t^i = S_0^i - \hat{S}_{0,t}^i$. All strategies have to satisfy the measurability condition, i.e. traders can condition their strategies only on states they can distinguish, i.e. on partitions. The author focus on linear recursive Markov perfect equilibria which satisfy following conditions:

$$x_t^i = \beta_t S_{t-1}^i,$$

$$\hat{S}_{0,t}^i = \hat{S}_{0,t-1}^i + \zeta_t Y_t,$$

$$P_t = P_{t-1} + \lambda_t Y_t,$$

where $Y_t = \sum_i^I x_t^i + \Theta_t$ and it is shown that $\lambda_t = \theta \zeta_t$ and $\hat{S}_{0,t}^i = \hat{S}_{0,t-1}^i + \zeta_t Y_t$ is necessary to guarantee that the forecasts of the others' forecasts is linear.

Foster and Viswanathan show that the dimensionality of the state space can be reduced, since a sufficient statistic for the past can be found for this equilibrium concept. Trader $i$ bases his strategy on his information set $(S_0^i, \underline{Y}_{t-1}, \underline{x}_{t-1}^i)$. Since in equilibrium his optimal demand is given by $x^i_\tau = x^i_\tau(S^i_0, \underline{Y}_{\tau-1})$ $\forall\tau$ his information set can be simplified to $(S^i_0, \underline{Y}_{t-1})$. This makes it clear that trader $i$ can only manipulate trader $j$'s beliefs about the true value, $\Pi$, via $\underline{Y}_t$. Foster and Viswanathan show that $S^i_{t-1}$, the information advantage at $t-1$, is a sufficient statistic for trader $i$ to predict $E[\Pi - P_{t-1} \mid \mathcal{F}^i_{t-1}] = \eta_t S^i_{t-1}$, since $P_{t-1}$ is common knowledge and all random variables are normal. $\eta$ and $\phi$ are constant regression coefficients. As this is true for all traders, it is also sufficient for trader $i$ to forecast $S^j_{t-1}$ in order to forecast the forecasts of others, i.e. $E[S^j_{t-1} \mid \mathcal{F}^i_{t-1}] = \phi_t S^i_{t-1}$. By induction the $t^{th}$order forecast, the forecast of trader $i$ about the forecast of trader $j$ about the forecast of trader $i$, etc. is also a linear function of $S^i_{t-1}$. Thus the hierarchy of forecasts is not history dependent. In other words the infinite regress problem, which we will discuss in detail later, is avoided. Their analysis shows that, in equilibrium, the dimensionality issue can be resolved.

In order to check whether this is really a Nash equilibrium one has to show that no trader has an incentive to deviate. For analysing deviation a larger state space is needed and the dimensionality issue arises again. Suppose only trader $i$ deviates from the equilibrium strategy and submits arbitrary market orders $(x^i_1, ..., x^i_t)$ in the first $t$ periods. Let $Y^{i}_t{}', P_t{}', \hat{S}^{i,i}_{0,t}{}', S^{j,i}_t{}'$ with the additional superscript ${}^{i}{}'$, be the corresponding variables when traders play the equilibrium strategies. By construction $S^{i,i}_t{}'$, the informational advantage, is orthogonal to $(\underline{Y}^{i}{}'_{t-1})$. Note that $(\underline{Y}^{i}{}'_{t-1})$ is in $i$'s information set because $i$ also knows the strategy he would have followed in equilibrium and thus he can also derive the change in other traders' expectations caused by his strategy change. Therefore trader $i$'s information set also captures $S^{i,i}_{t-1}{}'$, $P^i_{t-1}{}'$, $S^{j,i}_{0,t-1}{}'$. A sufficient statistic for his information set is given by $S^{i,i}_{t-1}{}'$ together with the deviation from the equilibrium price $(P^i_{t-1}{}' - P_{t-1})$. Therefore $E[\Pi - P_{t-1} \mid \mathcal{F}^i_{t-1}] = E[\Pi - P^i_{t-1}{}' \mid S^{i,i}_{t-1}{}'] + (P^i_{t-1}{}' - P_{t-1})$. Foster and Viswanathan conjecture the value function for trader $i$

$$V^i[S^{i,i'}_{t-1}, P^{i'}_{t-1} - P_{t-1}] = \alpha_{t-1}(S^{i,i'}_{t-1})^2 + \psi_{t-1} S^{i,i'}_{t-1}(P^{i'}_{t-1} - P_{t-1}) - \mu_{t-1}(P^{i'}_{t-1} - P_{t-1})^2 + \delta_{t-1}$$

and derive the optimal market order size for a certain time period. The resulting conditions for the Markov Perfect linear recursive equilibrium allow to verify that the proposed value function was indeed correct. Finally, they relate their results to less general models, like Kyle (1985), Holden and Subrahmanyam (1992) and others.

For calculating numerical examples, Foster and Viswanathan apply a backward induction algorithm for the case of three traders and four trading rounds. They compare four different correlations between the initial signals $S_0^i$; very high, low positive, zero and low negative correlation. The major findings are that (1) the lower the signal correlation, the less informative is the price process, (2) profit for insiders is lowest with identical information and highest with positive but not perfect correlation, (3) with positive signal correlation $\lambda_t$, the market maker's sensitivity falls over time, whereas with negative correlation $\lambda_t$ rises with the trading rounds, and finally (4) the conditional correlation of the remaining information advantage $S_t^i$ is decreasing over time and becomes negative provided there are enough trading rounds.

These results are the outcome of two effects. First, with heterogeneous signals the competitive pressure is reduced since each trader has some monopoly power. Second, when the $S_t^i$s become negatively correlated traders play a waiting game. This is driven by the fact that the market maker learns more about the average signal than about the individual signals. With negatively correlated $S_t^i$, traders are more cautious and more reluctant to take on large positions early. Foster and Viswanathan then go to analyse the effects of increasing the number of trading rounds keeping the total liquidity variance, $T\sigma_\Pi^2$, constant. With more trading rounds the speed of information revelation is higher and a U-shape pattern of $\lambda_t$ arises and becomes more pronounced. This U-shape of the market maker's sensitivity results from the waiting game. Their analysis suggests that dynamic competition with heterogeneously informed traders can be quite distinct. Whereas insiders with identical information trade very aggressively, (i.e. they are in a "rat race") insiders with heterogeneous information trade less aggressively since they play a waiting game.

Back, Cao, and Willard (1997) conduct the same analysis in continuous time. They prove that there is a unique linear equilibrium when signals are imperfectly correlated and derive a "closed form" expression for the equilibrium. There does not exist a linear equilirium when signals are perfectly correlated.

Vayanos (1996) studies a strategic dynamic limit order model à la Kyle (1989). He shows that the forgone gains from trade lost due to strategic behaviour, increase as the time between trades shrinks.

## 2.5 Herding Models

In the preceeding sections we dealt with situations in which information became known gradually. In many situations some agents could act earlier than others. Sequential decision taking may cause a phenomena like herding even if all agents behave rationally. Herding in sequential decision making can occur for at least three different reasons:

- Payoff Externalities
- Reputational Effects in Principal-Agent Models
- Informational Externalities

The latter two effects can not only lead to herding but also to delays in decision making in a world where the order of decision making is endogenous. No decision may be made at all in the extreme case. Herding models refer to an environment in which each agent makes one irreversible decision. This distinguishes the herding literature from the experimentation literature.

### 2.5.1 Herding Due to Payoff Externalities

In an environment in which payoffs are higher if all agents choose the same action, it is obvious that herding occurs in at least one possible equilibrium. Standard coordination failure games provide one possible environment.[51] The model in Admati and Pfleiderer (1988) is an example of herding caused by payoff externatility. All discretionary liquidity traders try to trade at the same time, i.e. clump together. Bank runs are further examples of herding in finance. Froot, Scharfstein, and Stein (1992) and Hirshleifer, Subrahmanyam, and Titman (1994) show that myopic investors try to acquire information about the same event as others. These papers will be discussed in greater detail below. In this class of models herding is defined as "doing the same." However, agents do not neglect their own information. Whether herding is socially optimal depends on the whole payoff structure of the game.

[51] However, mixed strategy equilibria are also possible.

### 2.5.2 Reputational Cascades in Principal-Agent-Models

In this section we will refer to models in which herding among agents is induced by reputational effects towards the principal. Herding in this case is only individually optimal and the First Best cannot be achieved. In Scharfstein and Stein (1990) two risk neutral agents (managers) invest sequentially in two identical investment objects. Each manager is either smart or dumb. Neither the principal nor the agents themselves know the types. Each agent receives a binary signal $\{S_H^i, S_L^i\}$ about the true liquidation value $\Pi \in \{\Pi_H, \Pi_L\}$ of the projects. The signal structure satisfies following conditions:

(1) $\Pr(S_H \mid \Pi_H, smart) > \Pr(S_H \mid \Pi_L, smart)$,

(2) $\Pr(S_H \mid \Pi_H, dumb) = \Pr(S_H \mid \Pi_L, dumb)$,

(3) $\Pr(S_H \mid smart) = \Pr(S_H \mid dumb)$,

(4) smart agents' signals are (perfectly) correlated.

(1) states that a smart agent gets with higher probability the right signal, whereas (2) says that dumb managers get with equal probability the high signal, independent of whether the project is good or bad. Condition (3) guarantees that the signal is purely about the investment project and cannot be used by a single agent to improve his knowledge about his type. Condition (4) states that smart agents have the same (correlated) forecast error. The payoffs $\{\Pi_H, \Pi_L\}$ are such that the first agent invests if he gets the high signal and does not otherwise. Knowing this, the second agent can infer the signal of the first agent and can base his decision on both signals. First Best, however, is not obtained in this environment, since agent two cares about his reputation with respect to the principal, i.e. he wants to appear as a smart type. Agent two's reputation increases when he makes the right decision. If he makes the wrong choice, it is better for his reputation if the first agent has chosen the wrong alternative, too. Combining this with the assumption that it is more likely that two smart agents get the same wrong signal causes him to follow the others' decisions. Agent two's information set contains both signals. If his signal coincides with the one of agent one, he will make the same investment decision as agent one. If his signal is different from the first one he will still follow the first agent's decision. Intuitively, the reasoning is that if he is wrong so is agent one. This could also be due to the possibility that two smart agents

accidentally received a wrong signal. If he would follow his own signal then the principal thinks that it is more likely that at least one agent is dumb. If it turns out that his signal is wrong, he would be considered a dumb manager. The authors show that a separating equilibrium does not exist and agent two always employs a herding strategy in equilbrium given plausible beliefs, Cho and Kreps (1987). This result hinges on the assumption that the prediction error of smart agents is correlated, because only then the principal's updating rule becomes a function of agent one's investment decision. The herding effect disappears for the case where the smart signals are independent.[52] Trueman (1994) generalised this analysis and applied it to analyst forecasting. Another model with reputational cascades is Zwiebel (1995). In this model managers may prefer choosing a common action rather than a superior innovation, since the common action provides a more accurate benchmark for relative performance evaluation.

### 2.5.3 Statistical Cascades Due to Information Externalities

In models in this section, an agent takes neither reputational effects nor the fact that his true payoff for a given action depends on the actions of others into account. Precluding pre-play communication, herding is only generated by the positive informational externalities a predecessor generates for his successors. Inefficient information cascades can occur if the successor can only partially infer the predecessors' information from his action but he still ignores his own signal.

#### Exogenous Sequencing

This strand of papers began with Banerjee (1992) and Bikhchandani, Hirshleifer, and Welch (1992). In Banerjee (1992) $I$ risk neutral agents choose an asset $j \in [0,1]$ on an interval of the real line. All assets' payoffs are zero, with the exception of asset $j^*$, whose certain payoff is $\Pi$. All agents have uniform priors. With probability $\alpha < 1$ an agent gets a signal, which is true with probability $\beta$ and false with probability $(1-\beta)$. If the signal is wrong, then it is uniformly distributed on the interval. Agents make their decision sequentially. Successors can observe the predecessors' decisions, but not their signals.

[52] In my opinion herding may also occur if agents are very risk averse, in particular if the principal employs relative performance evaluation.

Banerjee derives the Bayesian Nash equilibrium after he has assumed three tie-breaking rules - which are in favour of a non-herding outcome. In equilibrium, the first agent follows his signal or chooses $j = 0$ as assumed decision rule if he has not observed a signal. Given the tie-breaking rule the second decision maker only follows the first agent if he has no signal. Otherwise he follows his signal, which can be identical to agent one's decision. The third decision maker always follow his predecessors if they have chosen the same action $j' \neq 0$, regardless of his signal. This is optimal for him, since both predecessors choose only the same asset $j' \neq 0$ in two cases. Either agent one got signal $j'$ and agent two got no signal and followed agent one or agent one and agent two got the same signal $j'$. In the former case - which occurs with probability $(1 - \alpha)$ - decision maker three is indifferent between following the predecessors' decisions and his own signal. In the latter case, whose likelihood is $\alpha$, $j'$ is $j^*$ with probability one, since the event that agent one and agent two get the same wrong signal $j'$ is a zero probability event. Therefore, decision maker three will follow his predecessors and ignore his signal. Agent four knows that agent three's decision carries no information about his signal. Thus, he faces exactly the same situation as decision maker three and herding will occur. In the case where agent one gets a wrong signal and agent two no signal, i.e. he just follows agent one, the whole crowd runs in the wrong direction. This happens, although asset $j^*$ could be found with probability one, if a large enough number of agents could communicate. This inefficiency only occurs (in sequential decision making) if the predecessors' actions are not a sufficient statistic for their information, i.e. the successors can only partially infer the information of the predecessors. In Banerjee (1992) the one dimensional action space on $[0, 1]$ cannot reflect the signal which is two-dimensional, a variable on $[0, 1]$ and a binary variable $\Xi = \{0, 1\}$, which indicates whether the predecessor received a signal or not.

In Bikhchandani, Hirshleifer, and Welch (1992) this is achieved by a discrete action space: adopt or reject. Although the signal in their basic example is also only binary $\{H, L\}$, the action space cannot capture the whole information of a later decision maker.[53] This information consists of his own signal and of information derived from predecessors' actions.[54] The first decision maker follows his signal in their model. If the second agent

[53] In their generalised version the signals are not binary.

[54] A continuous action space could reveal the posterior of an immediate predecessor which is a sufficient

gets an opposite signal he is indifferent between adopting and rejecting. Bikhchandani, Hirshleifer, and Welch (1992) assume a random tie-breaking rule. In this case the third agent cannot infer the second agent's signal. If the second agent randomly chooses the same action as the first one, the third agent cannot infer the second agent's signal. From the viewpoint of the third agent, it can be that either the second agent got the same signal as the first one or the opposite signal and randomly chose the same action. If the third decision-maker observes that both predecessors have chosen the same action, he will always follow them regardless of his own signal. All following decision makers know that the third decision-maker ignored his own signal and therefore they do not try to infer any information from his action. Actually, they face the same problem as the third decision maker and, therefore, join the crowd. The cascade evolving in this manner prevents the aggregation of information and, therefore, convergence to the correct action need not occur.

The authors show for a special case in their section "Fashion Leaders" that a higher signal precision for the first decision maker can make informationally inefficient cascades sooner and more likely. The reason is that higher signal precision for the first agent makes it more likely that the second agent follows the first one. In a setting with an endogenous decision sequence it seems plausible that agents with the highest precision are willing to decide first. Their analysis about the fragility of cascades to public information releases is also interesting. Public information prior to the first agent's decision can make inefficient cascades even more likely. On the other hand, public information after a cascade has already begun, is always socially beneficial. Already a small amount of public information can shatter a long-lasting cascade. As explained above, a cascade is created by the decision of the first two agents and, thus, the public information need only lift out their information.

Gale (1996) provides a similar example, where in contrast to Bikhchandani, Hirshleifer, and Welch (1992) the *signal space* is continuous, i.e. $S^i \in [-1;+1]$ and the action space is still binary $\{H, L\}$. In this case a cascade can be shattered if an extreme signal arises. For simplicity, lets assume that the payoff of at least $I$ identical investment opportunities is given by the average of all signals, i.e. $\Pi = \frac{1}{I}\sum_{i=1}^{I} S^i$. Given that the signals are uniformly distributed over $[-1,+1]$ the First Best solution is achieved if all agents invest if and only if

statistic for all past signals. In this case no herding occurs.

$\Pi = \frac{1}{I}\sum_{i=1}^{I} S^i > 0$. In sequential decision making, agent one invests if $S^1 > 0$ and agent 2 if $S^2 + E[S^1 \mid action^1] > 0$, etc. If agent 2 observes that agent 1 has invested, he will invest if $S^2 + E[S^1 \mid S^1 > 0] = S^2 + \frac{1}{2} > 0$. If agent 2 also invests, agent 3 will invest if $S^3 + \frac{3}{4} > 0$, and so forth. In other words if agent 1 and 2 have invested then agent 3 needs to receive a really bad signal $S^3 < -\frac{3}{4}$ in order to not invest, i.e. not to follow his predecessors. This means that a cascade becomes more and more stable over time, i.e. the signal necessary to break up a cascade has to be more and more extreme. Although herding behaviour will occur, an informational cascade can never occur in Gale (1996). Smith and Sørensen (1994) define 'herding' as convergence in actions and 'cascades' as convergence in agents' beliefs. In cascades agents ignore their own signal and base the current decision only on historic actions.

Smith and Sørensen (1994) not only consider a continuous signal space but also allow agents' preferences to differ. Incorporating diversity in taste can lead to situations of 'confounded learning'. In such situations the observed history does not provide additional information for decision making and the decision of each type of agent might forever split between two actions.

Smith and Sørensen (1997) relate herding models to the literature of experimentation. This literature stems from Rothschild's (1974) two-armed bandit analysis. Herding models correspond to the experimentation problem faced by single myopic experimenter who forgets his formal signal but remembers his past actions. The incorrect herding outcomes correspond to the familiar failure of complete learning in an optimal experimentation problem.

Lee (1993) shows, how crucial the discretness of the *action space* is. The discretness plays a dual role. (1) It prevents somebody's actions from fully revealing his posteriors and (2) it prevents each agent from fully using of his information. In Lee's model the likelihood of an inefficient cascade decreases as the action space grows. He also claims that Banerjee's model is an exceptional case, since the (degenerated) payoff structure in Banerjee (1992) does not distinguish between small and large errors.[55]

[55] Lee's claim cannot follow from his model, since in his model every agent gets a signal with certainty unlike in Banerjee (1992). It can be shown that the payoff function in Banerjee (1992) can be generalised to a certain degree.

**Endogenous Sequencing and Strategic Delay**

If each decision maker could decide when to decide, everybody would want to be the last one, in order to profit from the positive information externalities generated by his predecessors' decisions. Strategic delays caused by information externalities were first discussed in Chamley and Gale (1994) and Gul and Lundholm (1995).

In Chamley and Gale (1994) time is discrete $t = 1, 2, ..., \infty$ and each of (randomly) $I$ agents has an investment opportunity, i.e. a real option to invest or not to invest. Each investor knows whether he himself has an investment opportunity, but he does not how many investors have this opportunity as well. Therefore, he does not know $I$. The true payoff of the identical investment opportunities is increasing in $I$, in the number of possible investment opportunities, and not in the number of investments actually undertaken. Agents who invest early reveal that they have an investment opportunity. This positive information externality allows the successors to update their beliefs about the true $I$. In order to avoid all agents waiting forever, each agents waiting costs are given by a common discount factor $0 < \delta < 1$. Chamley and Gale (1994) focus on *symmetric* Perfect Bayesian Equilibria in which agents apply behavioural strategies.[56] They show that there are three exclusive possible equilibrium continuation paths given a certain history of past investments. If beliefs about the number of people who got an investment opportunity are sufficiently optimistic, all players immediately invest and the game ends. On the other hand, if these beliefs are sufficiently pessimistic no one will invest and hence no information is revealed. In this case the game ends as well, since one period later the situation has not changed. For intermediate beliefs, given a certain investment history, the remaining players with investment opportunity are indifferent between investing and waiting. Hence, they randomise in this period, i.e. employ a behavioural strategy. The remainig investors who have not yet invested try to update their beliefs about the total number of investment opportunities from the random number of investments in this period. It is obvious that information aggregation is inefficient in such an setting. The authors also show that as the period length increases, the possibility

[56] Action rules determine an action at a certain partition/decision knode. Randomising over different action rules at any partition is a behavioural strategy. A strategy is a sequence of action rules. Randomising over *pure* strategies is a mixed strategy.

of herding disappears. This raises an interesting question concerning continuous trading on stock markets: does continuous trading lead to strategic delays and herding, i.e. to a worse information aggregation than in a single batch auction.[57]

In Gale (1996) the type of an agent is given by his signal $S^i \in [-1, 1]$ about the payoff of the investment opportunities. The payoff of each investment project is $\Pi = \frac{1}{I}\sum_i^I S^i$. Gale considers only the case for two agents, i.e. $I = 2$. With a common discount factor $\delta$, a decision-maker with a higher signal is more impatient to invest than somebody with a lower signal. The aim is to derive the threshold level $\overline{S}$ for the signal value required to motivate an investor into investing in period 1. Whether somebody exercises his real option early depends on the probability that he will regret in the next period that he has invested early. An investor who invests early regrets it if the other investor has not invested and his posterior beliefs about the payoff are $S^i + E[S^{-i} \mid S^{-i} < \overline{S}] < 0$ are negative. The event that the other agent does not invest occurs with probability $\Pr(S^i < \overline{S})$. In equilibrium, an agent with signal $\overline{S}$ is indifferent between waiting and investing in the first period, i.e. the waiting costs are equal to the option value of delay.

$$(1-\delta)\overline{S} = -\delta \Pr(S^i < \overline{S})\{\overline{S} + E[S^i \mid S^i < \overline{S}]\}.$$

There exists a unique equilibrium $\overline{S}$ in which information is not fully revealed and the outcome need not be efficient. E.g. if both signals are $0 < S^i < \overline{S}$ nobody will invest even though it would be socially optimal. Another feature of the equilibrium is that the game ends after two periods. If nobody invested in the first two periods, investment stops forever, i.e. an investment collapse can occur. Similar results carry over to a setting with $I$ agents.

In contrast to Chamley and Gale (1994) in Gul and Lundholm (1995) time is continuous. In Bikhchandani, Hirshleifer, and Welch (1992) and Banerjee (1992) agents (partially) ignore their own information which consequently leads to inefficient information aggregation, i.e. to information cascades. In Gul and Lundholm (1995) the timing when to act as well as when *not* to act improves the information aggregation. In their model, endogenous sequencing leads to information efficient *clustering* as opposed to informationally inefficient

[57] As seen in Section 2.4.1, Grundy and McNichols (1989) show that with an increasing number of trading rounds per day, and thus observable prices, the price sequence can reveal the noise term or even higher dimensional signals.

information cascades. In Gul and Lundholm's model agents maximise a utility function, which captures a tradeoff between the accuracy of a prediction and how early the prediction is made (waiting costs). Each agent observes a signal $S^i \in [0,1]$, which helps him to forecast $\Pi = \sum_{i=1}^{I} S^i$. The authors show that the strategy of each player can be fully described by a function $t^i(S^i)$. $t^i(S^i)$ reports the latest possible time at which agent $i$ with signal $S^i$ will make his forecast given the other players have not done so already. Since $t^i(S^i)$ is continuous and strictly decreasing, i.e. $t^i(S^i)$ is invertible, the time when the first agent acts fully reveals his signal to the succeeding decision maker. In a two agent setting, the second agent will make his prediction immediately afterwards. Whereas in the former models only the succeeding decision makers profit from positive information externalities, in Gul and Lundholm (1995) the first agent learns from the others' inaction. The first decision maker can *partially* infer the signals of his successors by noticing that they have not acted before him. This biases his decision towards the successor's forthcoming decisions. Consequently agents tend to cluster, i.e. their forecasts are closer together in a setting with endogenous sequencing than in a setting with exogenous ordered forecast. Gul and Lundholm call this effect *'anticipation'*.[58] There is a second source of clustering, labelled *'ordering'*. This occurs because (1) agents with the most extreme signal realisations have higher waiting costs and thus act first and (2) the signals of predecessors are revealed fully, whereas inaction of the successors only partially reveals their signals. More pronounced signals have a larger impact on the true value $\Pi = \sum_{i=1}^{I} S^i$. Since more pronounced signals are fully revealed first, whereas the signals with lower impact are fully revealed later, forecasts are 'on average' closer together than in the case where the less pronounced signals would be fully revealed first.

The distinctive feature of Zhang's (1997) model is that the precision (quality) of the private signal, and not just its content is private information. His model incorporates higher order uncertainy. The signal is binary, and reports with probability $p^i$ which of the two investment projects is the good one. The quality (precision) of the signal is measured by $p^i$, where each $p^i$ is drawn from a continuous probability distribution over $[1/2, \overline{p}]$, with $\overline{p} < 1$. The realisation of the signal as well as its quality $p^i$ is only known to agent $i$. The

[58] Note the similarity to (descending) Dutch auctions.

agents' action space at each point in time is either to wait (which discounts the payoffs by $\delta$) or to invest either in investment project 1 or 2. As in Gul and Lundholm (1995) time is continuous.

Zhang derives the unique equilibrium in pure strategies in closed form. The equilibrium exhibits an initial delay of action till the agent with the highest precision (highest $p^i$) invests. Given the binary investment choice and binary signal space the second decision maker will always ignore his signal, since it is of worse quality. He will immediately mimic the first movers investment decision. Consequently the second agent's investment choice carries no additional information and therefore all other agents will follow immediately the first mover, as well. In summary, after a certain initial delay one can observe a sudden onset of investment cascades. In contrast to Gul and Lundholm (1995) the outcome is not informationally efficient, since everybody's investment decision depends only on the signal with the highest precision. Moreover, the initial delay incurs waiting costs, which is another source of inefficiency. As the number of agents increases the per capita efficiency loss is bounded away from zero. Furthermore each player tends to wait longer, since it is more likely that someone has a more precise signal and will invest before him.

Gale (1996) discusses the problems which arise in herding models in continuous time. For a more detailed discussion of the 'closure problem' see Harris, Stinchcombe, and Zame (1997).

Neeman and Orosel (1998) analyse a sequential common value auction, where the seller can determine the order in which he will approach potential buyers. Potential buyers submit a bid, i.e. their action space is continuous. The bid of a rival bidder creates a information externality as well as a payoff externality. In contrast to the English common value auction, potential buyers can only bid when they are approached by the seller. This makes the winner's curse in this auction less severe, thus leading to better information aggregation as well as to a (weakly) higher revenue for the seller. Neeman and Orosel's analysis can also be viewed as a search problem for the seller where buyers' bids are correlated for three reasons. First, their signals are correlated. Second, they bid for an object of common value, and third, all buyers condition their bid on the publicly observed history of bids.

Allowing agents to act more than once and/or revising their decision would be a natural extension to the herding literature. This leads us to models of experimentation. See Bolton and Harris (1993), Bergemann and Välimäki (1996) and Leach and Madhavan (1993) for useful expositions.

### 2.5.4 Herding in Financial Economics

Prices in financial markets are almost continuous if one neglects the small tick size. The same is true for the possible order size. Given a one dimensional signal space, herding is thus very unlikely to occur, Lee (1993). Traders in financial markets, therefore, do not blindly follow investment decisions of others. On the other hand, herding results can be obtained by introducing higher order uncertainty and fixing the order size for each trader, as in Glosten and Milgrom (1985). Herding also occurs in information acquisition when all investors search for the same piece of information. This is the case when they have short horizons.

#### Herding, Bubbles and Higher Order Uncertainty

In finance, trading causes not only informational externalities, but also changes in prices. This changes the payoff structure for all successors. Avery and Zemsky (1995) use a information structure similar to Bikhchandani, Hirshleifer, and Welch (1992). They show in a Glosten-Milgrom setting (see Section 2.3.2.3) that the price adjusts exactly in such a way that it offsets the incentives to herd. This is the case, because the market maker and the insiders learn the same from past trading rounds. Avery and Zemsky (1995) distinguish between herding behaviour and informational cascades. Herding occurs if traders imitate the decision of their predecessors even though their own private signal advises them to take a different action. In informational cascades no additional information is revealed to the market, since the distribution over the observable actions is independent of the state of the world. An informational cascade never occurs in an extended Glosten-Milgrom setting in which insiders get different noisy signals. In addition, traders never engage in herding behaviour provided signals are monotonic. The price converges to the true asset value and the price process exhibits no "excess volatility" given its Martingale property. Similar to Easley and O'Hara (1992), Avery and Zemsky introduce "event uncertainty." Insiders receive either a perfect signal that no new information has arrived or a noisy signal which reports with probability $p$ the true asset value, $\Pi \in \{0,1\}$. In other words, all insiders receive signals with the same precision, $p' = 1/2$ (no information event) or $p' = p \in (1/2, 1]$. $p$ is known to the insiders, but not to the market maker, i.e. the market maker does not know whether an information event occurred or not. This asymmetry in higher order information between insiders and the market maker allows insiders to learn more about the price process (trading sequence) than the market maker. Since the market maker sets the price, the price adjustment is slower. This can lead to herding behaviour. This is consistent with the results in (Bikhchandani, Hirshleifer, and Welch 1992) where prices are esentially 'fixed.' Note, that event uncertainty can lead to herding behaviour, but not to informational cascades, because the market maker can gather information about the occurrance of an information event. Surprisingly, herding increases the market maker's awareness of information events. However, herding does not distort the asset price and thus it does not explain bubbles. In order to explain bubbles and excess volatility a more complex information structure is needed. Avery and Zemsky consider a setting with two types of informed traders. One group of traders receives its signals with low precision $p_L$, whereas the other receives them with high precision $p_H = 1$, i.e. they receive a perfect signal. This information structure incorporates higher order uncertainty, since it is not known whether the proportion of insiders with the precise signal is high or low. This makes it difficult for the market maker to differentiate between a market composed of well informed traders following their perfect signal and one with poorly informed traders who herd. In the latter case, bubbles[59] can arise.

Gervais (1995) shows that uncertain information precision can lead to a cascade state. In this case, insider's information precision gets never revealed and thus the bid-ask spread does not reflect the true precision. In Gervais (1995) all agents receive a signal with the same precision, $p_H$, $p_L > 1/2$, or $p_{no} = 1/2$. If $p_{no} = 1/2$, no information event occurs. In contrast to Avery and Zemsky (1995), the signals do not refer to the liquidation value of the asset, $\Pi$, directly, but only to a certain aspect $\pi_t$ of $\Pi$. More formally, every trader receives

[59] Bubbles are possible, because trade is restricted to one stock at a time in a Glosten-Milgrom setting. Therefore, the results of Section 2.2.4 for limit order models do not apply.

a noisy signal about aspect $\pi_t$, which takes on a value $\frac{1}{T}$ or $-\frac{1}{T}$ with equal probability of 1/2. The final liquidation value of the asset is then given by $\Pi = \sum_{t=1}^{T} \pi_t$. Note for each $\pi_t$, there is only one signal. As in Glosten and Milgrom (1985) the risk-neutral market maker sets competitive quotes. If the bid-ask spread is high, insiders trade only if their signal precision is high. The trade/no-trade sequence allows the market maker to updates his beliefs about the quality of the insider's signals. Furthermore, he updates his beliefs about the true asset value $\Pi$. Therefore, the competitive spread has to decrease over time. Note the trading/quote history is more informative for insiders because they already know the precision of the signal. When the competitive bid-ask spread decreases below a certain level, insiders will engage in trading independent of the precision of their signal. This prevents the competitive market maker to learn more about the signals' precision, i.e. the economy ends up in a cascade state with respect to higher order uncertainty.

### Herding in Information Acquisition and Short Horizons

Brennan (1990) noted the strong interdependence of individual information acquisition decisions. In a market with many investors the value of information about a certain (latent) asset may be very small if this asset pays a low dividend and no other investor acquires the same information. If on the other hand many investors collect this information the share price adjusts and rewards those traders who gathered this information first. Coordinating information collection activities can therefore be mutually beneficial.[60] Brennan (1990) formalises his argument using an overlapping generation model where agents live only for three periods.

In Froot, Scharfstein, and Stein (1992) herding in informational acquisition is due to investors' short horizons, i.e. their myopia. This behaviour effects the asset price. In general, backward induction rules out any alteration of the price process caused by the short horizons of traders. This is true in models with exogenous information acquisition and a finite number of time periods. However, as shown in Tirole (1982) myopic behaviour can lead to bubbles in infinite horizon models. In Froot, Scharfstein, and Stein (1992) the asset price is influenced by the endogenous information decision. Each individual trader has to decide whether to

[60] Note, for stock price manipulation coordination is also required if there are many investors in the market.

receive a signal about event A or event B. All traders worry only about the short-run price development, since they are short-sighted. They can only profit from their information if it is subsequently reflected in the price. Since this is only the case if enough traders observe the same information, each trader's optimal information acquisition depends on the others' information acquisition. The resulting positive information spillovers explain why traders care more about the information of others than about the fundamentals. In Keynes (1936) words, "skilled investment today is to "beat the gun"...". Observing this behaviour of traders led Keynes to compare the stock market with a beauty contest. A judge in a beauty contest who wants to support the winning candidate, has to be more concerned about the opinion of the other judges than about the relative beauty of the contestants. This phenomenon does not arise in the stock market, if all traders take the whole future into account, i.e. if they only care about the final liquidation value. If this is the case, information spillovers are negative. Thus, it is better to have information others do not have. Therefore in this case every investor will try to collect information for different events.

In Froot, Scharfstein, and Stein (1992) traders are assumed to be short sighted and then the picture changes. In the articles Hirshleifer, Subrahmanyam, and Titman (1994) and Holden and Subrahmanyam (1996) short horizons of traders are not exogenously assumed, but result endogenously form the model specification. We will discuss these two articles after providing the intuition of Froot, Scharfstein, and Stein (1992). Froot, Scharstein and Stein base their model on Kyle (1985). The asset's liquidation value is given by two components, $\nu$ and $\delta$, i.e.

$$\Pi = \nu + \delta,$$

where $\nu \sim \mathcal{N}(0, \sigma_\nu^2)$ refers to event A and $\delta \sim \mathcal{N}(0, \sigma_\delta^2)$ to the independent event B. Each trader can decide whether to observe either $\nu$ or $\delta$, but not both. After observing $\nu$ or $\delta$ he submits a market order to the market maker at $t = 1$. The authors assume that half of the submitted market orders are executed at $t = 1$ and the second half at $t = 2$. The period in which an order is processed is random. Furthermore, liquidity traders submit market orders of aggregate random size $\Theta_t$ in each period. As in Kyle (1985) the risk neutral market maker sets a competitive price in each period based on the observed total net order flow. Thus, the price partially reveals the information collected by the informed traders. At $t = 3$ all traders, i.e. insiders and liquidity traders, unwind their position by assumption. In other words, the risk neutral market maker takes on all risky positions. The trading price at $t = 3$, $P_3$ is $\Pi = \nu + \delta$ with probability $\alpha$ for the case where $\nu$ and $\delta$ are publicly announced at $t = 3$. $\nu$ and $\delta$ are announced only at $t = 4$ with probability $(1 - \alpha)$. Therefore the trading price does not change, i.e. $P_3 = P_2$. Thus, $\alpha$ provides a measure of the degree of short-sightedness of the traders. The traders' horizons are very short if $\alpha$ is close to zero. In that case, traders only care about $P_2$ and not about $\Pi$. On the other extreme, if $\alpha$ is close to one, traders' horizons are long.

The expected profit per share for an insider is $\frac{P_2 - P_1}{2}$ if $\nu$ and $\delta$ are only announced at $t = 4$, i.e. $P_3 = P_2$. This case occurs with probability $\alpha$. With probability $(1/2)$ the trader is lucky and his order is processed early, i.e. he gets his shares for $P_1$. He can then sell it at $t = 3$ for $P_3 = P_2$. With probability $\alpha$, however, $\nu$ and $\delta$ are announced already at $t = 3$, i.e. $P_3 = \Pi$. In this case a trader who submitted an order at $t = 1$ also buys a share for $P_1$ or $P_2$ with equal probability, but sells it at $t = 3$ for $P_3 = \Pi$. His expected profit in this case is given by $\Pi - \frac{1}{2}[P_1 + P_2]$. Thus, the overall expected profit per share for an informed trader is

$$E\{\alpha[\Pi - \frac{P_1 + P_2}{2}] + (1 - \alpha)[\frac{P_2 - P_1}{2}]\}.$$

In both cases the profit is determined by $P_3$, the price at which the informed trader unwinds his position. $P_3 = \Pi$ with probability $\alpha$. Thus, $\nu$ and $\delta$ are equally important, with probability $\alpha$. With probability $(1 - \alpha)$, $P_3 = P_2$. Since $P_2$ depends on the information set of all informed traders, each insider cares about which information the other traders are collecting. Let's consider for illustrative reasons the extreme case $\alpha = 0$, i.e. $\nu$ and $\delta$ are only publicly announced in $t = 4$. If in this case all other investors collect information $\nu$, then information $\delta$ is worthless, since $\delta$ will only enter into the price in $t = 4$, i.e. after the investors have already unwinded their positions. In this case all investors will herd to gather information $\nu$ and nobody will collect information $\delta$.[61] Thus, short horizons of traders creates positive informational spillovers which lead to herding in information acquisition.

However, even if all investors herd on some noise term $\zeta$, which is totally unrelated to the fundamental value $\Pi = \nu + \delta$, a rational investor is better off if he also collects information $\zeta$ rather than only information about fundamentals. If $\alpha = 0$ and all other investors are

[61] In a (Nash) equilibrium it is mutual knowledge which information the other traders are collecting, Brandenburger (1992).

searching for $\zeta$, the fundamentals $\nu$ and $\delta$ are only reflected in $P_4$. The price at which the traders have to close their position, $P_3(=P_2)$ depends on $\zeta$, given their strategies. In the case where it is not sure whether $\nu$ and $\delta$ will only be announced at $t=4$, i.e. $\alpha>0$, herding in information acquisition still occurs if $\alpha$ is sufficiently small, i.e. traders are sufficiently short sighted. Note for the case $\alpha=1$ demands are "strategic substitutes", while for $\alpha=0$ they are "strategic complements".

One might argue that the above reasoning may break down in an overlapping generations (OLG) framework in which a new generation of short-sighted traders enters the market in each period. Inefficient herding still occurs in the following OLG setting. Generation $t$ speculators can study one of $k$ pieces of information. At the end of period $t$, one of these pieces will be randomly drawn and publicly announced. In the following period $t+1$ a new additional piece of information can be studied. Thus, each trader in each generation can study one of $k$ pieces of information. For each generation it pays off to have accidentally studied the information, which gets publicly announced at the end of the period. Since one is only lucky with a probability $(1/k)$ the price movement will be determined more by what the other traders have studied. Hence herding in information acquisition may occur.

In the preceeding papers, the herding behaviour was due to the exogenously assumed short horizons of the traders. In the following two papers short horizons of traders are endogenously derived. In contrast to Froot, Scharfstein, and Stein (1992) in Hirshleifer, Subrahmanyam, and Titman (1994) and Holden and Subrahmanyam (1996) competitive limit order models are employed to derive endogenously myopic behaviour from agents' risk aversion.

In Hirshleifer, Subrahmanyam, and Titman (1994) a continuum of competitive risk averse investors search for the *same* information $\delta$ about the liquidation value $\Pi$ of a single risky asset.

$$\Pi=\overline{\Pi}+\delta+\epsilon,$$

where $\overline{\Pi}$ is known and $\delta\sim\mathcal{N}(0,\sigma_\delta^2)$, $\epsilon\sim\mathcal{N}(0,\sigma_\epsilon^2)$. Some investors, whose mass is $M$, receive information $\delta$ accidentally early, i.e. already in $t=1$, whereas the others, whose mass is $(N-M)$ are informed later. Both groups of traders receive the same information $\delta$, but at different times. All traders maximise CARA utility functions of the final wealth $W_3$, i.e. $U = -\exp(-\rho W_3)$. The demand for the risky asset by the early-informed is denoted by $x_t^e(\delta, \cdot)$, whereas that by the late-informed is $x_t^l(\cdot, \cdot)$. The aggregate demand of liquidity traders is modelled by the random variables $\Theta_1 \sim \mathcal{N}(0, \sigma_{\Theta_1}^2)$ in $t = 1$ and $\Delta\Theta_2 \sim \mathcal{N}(0, \sigma_{\Delta\Theta_2}^2)$ in $t = 2$.[62] Finally, there is also a group of risk neutral competitive market makers (scalpers, floor brokers, etc.) who observe the limit order book, i.e. the noisy aggregate demand schedules, but not the information $\delta$. The noisy aggregate demand function is $D_1(\cdot) = Mx_1^e(\delta, \cdot) + (N-M)x_1^l(\cdot) + \Theta_1$ in $t = 1$ and $D_2(\cdot) = Mx_2^e(\delta, \cdot) + (N-M)x_2^l(\delta, \cdot) + \Theta_1 + \Delta\Theta_2$ in $t = 2$. Given risk neutrality and competitiveness of the market makers, the market makers set a semi-strong efficient price with respect to their information sets, i.e. $P_1 = E[\Pi \mid D_1(\cdot)]$ and $P_2 = E[\Pi \mid D_1(\cdot), D_2(\cdot)]$.

In equilibrium investors conjecture the following linear price relations:

$$P_2 = \overline{\Pi} + a\delta + b\Theta_1 + c\Delta\Theta_2,$$

$$P_1 = \overline{\Pi} + e\delta + f\Theta_1.$$

The equilibrium is derived by backward induction. At $t = 2$ both groups of investors, early and late informed, know $\delta$ and their stock holding is therefore as usual

$$x_2^e(\delta, P_2) = x_2^l(\delta, P_2) = \frac{\overline{\Pi} + \delta - P_2}{\rho\sigma_\epsilon^2}.$$

At $t = 1$ only the group of early-informed investors knows $\delta$. Their stock holding is

$$x_1^e(\delta, P_1) = \frac{E[P_2 \mid \mathcal{F}_1^e] - P_1}{\rho}\left[\frac{1}{Var[P_2 \mid \mathcal{F}_1^e]} + \frac{1}{\sigma_\epsilon^2}\right] + \frac{\overline{\Pi} + \delta - E[P_2 \mid \mathcal{F}_1^e]}{\rho\sigma_\epsilon^2}.$$

The demand of early-informed trades consists of two components. The first term captures the speculative demand due to an expected price change. The second term is the expected final stock holding, which the early-informed traders try to achieve at the "on average" better price $P_1$. Investors who receive their signal only at $t = 2$, demand nothing at $t = 1$, i.e. $x_1^l = 0$. This is due to the fact that they do not have superior information to the market makers. Since the market makers are risk neutral (1) no risk premium is offered and (2) expected $P_2$ is unbiased. In other words, risk averse late-informed traders cannot hedge their date 2 demands already at $t = 1$.

[62] All demand functions are expressed in stock holdings, therefore, the additional demand in $t = 2$ is given by $\Theta_2 - \Theta_1 = \Delta\Theta_2$.

There are five equilibrium configurations for the coefficients of the price relations in this economy. In the fully revealing equilibrium no investor holds any stocks. In addition there are two equilibria where prices do not move, i.e. $P_1 = P_2$. Hirshleifer, Subrahmanyam, and Titman (1994) focus on the remaining two equilibria in which trading occurs and the price is not the same in both periods. In these equilibria both price moves, $P_1 - P_0$ and $P_2 - P_1$ are positively correlated with $\delta$, i.e. $P_2$ reveals "on average" more about $\delta$ than $P_1$. This is due to the fact that the market makers' information set, which determine the price, is improving by observing two noisy aggregate demand curves. Furthermore, both aggregate demand curves depend on information $\delta$, given that both groups of traders observe the same information. This is supported by the correlation between $\Theta_1$ and $\Theta_2$, since $\Delta\Theta_2$ is independent of $\Theta_1$. The price changes $P_1 - P_0$ and $P_2 - P_1$ themselves, however, are uncorrelated and thus prices follow a Martingale process, given the market makers' filtration.

The trading behaviour of the early-informed investors exhibits speculative features. They take on large positions in $t = 1$ and "on average" unwind partially in $t = 2$ at a more favourable price $P_2$. More precisely, their trading in $t = 1$, $x_2^e$ is positively correlated with the price change $P_2 - P_1$ in $t = 2$, whereas their trading in $t = 2$ is negatively correlated with this price change. Therefore, they partially unwind their position and realise capital gains "on average." The intuition for this result is as follows. Since no risk premium is paid due to the market makers' risk neutrality, risk averse traders would be unwilling to take on any risky stock in the absence of any informational advantage. Early-informed investors are willing to take on risk since they receive a signal $\delta$ in $t = 1$. Their informational advantage, together with the existence of noise traders compensates them for taking on the risk represented by the random variable $\epsilon$. However, the informational advantage of early-informed traders with respect to the late-informed traders vanishes in $t = 2$, since both now receive the same signal $\delta$. Thus, early-informed traders share the risk with late-informed traders in $t = 2$, i.e. $Cov(x_2^l, x_2^e) > 0$. In addition, the informational advantage of the early-informed with respect to the market makers shrinks as well, since at $t = 2$ market makers can observe a second different limit order book. This limit order book carries information for the market makers, especially since the stock holding of the noise traders is correlated in both periods. This allows the market makers to get a better idea about $\delta$ and, thus, $P_2$ should be "on average" closer to $\overline{\Pi} + \delta$ than $P_1$. Therefore, in period two, both these effects cause early-informed traders to partially unwind the position they built up in the previous preiod. The unwinding behaviour of early-informed traders in this sequential information arrival models also stimulates trading volume.

The fact that early-informed traders on average unwind their position in $t = 2$ is in sharp contrast to models based on Kyle (1985). In these models the risk neutral insider tries to buy the stocks in small pieces in order to hide behind noise trading, i.e. his stock holding over time is positively correlated.

Having analysed the second stage, Hirshleifer, Subrahmanyam, and Titman (1994) show that herding can occur in the information acquisition stage. At the time when traders decide which information to collect they do not know whether they will find the information early or late. Hirshleifer, Subrahmanyam, and Titman (1994) derive expressions for utilities of the early-informed and late-informed individuals. The authors provide a numerical example, in which the ex-ante utility before knowing when one receives the information is increasing in the total mass of informed traders. If this is the case, it is worthwhile for traders to concentrate on the same informational aspects, i.e. gather information about the same stocks. In other words they will herd in information acquisition. Whether a higher mass of informed traders really increases their ex-ante utility depends on the parameters, especially on $\sigma_\epsilon^2$. More informed traders lead to more late-informed traders, which makes it easier for early-informed traders to unwind larger positions in $t = 2$. Thus, there are more traders in $t = 2$ willing to share the risk resulting from $\tilde{\epsilon}$. This is disadvantageous for the late-informed, since there is tougher competition among them. This is the case since the extent of noise trading does not change. Increasing the mass of informed traders also increases the number of early-informed traders. This decreases the utility of both, early-informed and late-informed traders. In order to obtain herding the first effect has to outweigh the latter three. This requires that $\sigma_\epsilon^2$ is sufficiently high. The authors try to extend their analysis by introducing some boundedly rational elements which lies outside the scope of this literature survey.

In Hirshleifer, Subrahmanyam, and Titman (1994) all traders search for the same piece of information, which they randomly receive earlier or later. In Holden and Subrahmanyam (1996) traders can decide whether to search for short-term information or for long-term information. They choose between two signals which are reflected in value at different points of time. Holden and Subrahmanyam show that under certain conditions all risk averse traders focus exclusively on the short-term signal.

In their model the liquidation payoff of a single risky asset is

$$\Pi = \overline{\Pi} + \delta + \eta + \nu + \epsilon,$$

where $\delta$, $\eta$, $\nu$, and $\epsilon$ are mutually independent normally distributed and $\overline{\Pi}$ is normalised to zero without loss of generality. Traders who acquire short-term information observe $\delta$ at $t = 1$. At $t = 2$, $\delta$ becomes publicly known, as well as $\eta$, which no trader has known before. Traders who search for long-term information observe $\nu$ already in $t = 1$, which gets known to the public only in date 3. At $t = 3$ $\epsilon$ is also realised and known by all traders, i.e. $\Pi$ is common knowledge at $t = 3$. Note that since the components of $\Pi$ cannot be traded directly the markets are incomplete. This is essential for this analysis.

A competitive limit order model is employed as in Hirshleifer, Subrahmanyam, and Titman (1994). A mass $M$ of long-term informed traders and a mass of $N = 1 - M$ of short-term traders submit limit orders to the limit order book. The aggregate order size of the liquidity traders is random and given by $\Theta_1$ in $t = 1$ and $\Delta\Theta_2$ in $t = 2$. A group of risk neutral market makers observes only the publicly available information and the noisy aggregate demand schedule, i.e. the limit order book. Since the market makers act competetively and they are risk neutral, their information sets determines the price.[63]

Analysing the overall equilibrium backwards, the mass of short-term traders, $N$, and of long-term traders $M$, is kept fixed at the second stage and is endogenised at the first stage. Backward induction is also applied within the second stage for deriving the optimal stock holdings of informed risk averse traders. At $t = 2$, the stock holding demand are standard for the long-term informed traders,

$$x_2^l = \frac{\nu + \delta + \eta - P_2}{\rho\sigma_\epsilon^2}$$

and for the short-term informed,

$$x_2^s = \frac{E[\nu \mid \mathcal{F}_2^s] + \delta + \eta - P_2}{\rho[\sigma_\epsilon^2 + Var[\nu \mid \mathcal{F}_2^s]]} = 0.$$

[63]For a similar model with differential information see Vives (1995).

$x_2^s = 0$, since the market makers have the same information set as the short-term-informed and therefore the numerator in the above equation is zero. In economic terms, it would not make a lot of sense for risk averse short-term investors to hold risky stocks if the risk neutral market makers have the same information. Since $x_2^s$ is zero, $x_1^s$ is standard, i.e.

$$x_1^s = \frac{E[P_2 \mid \mathcal{F}_1^s] - P_1}{\rho Var[P_2 \mid \mathcal{F}_1^s]}.$$

Short-term informed traders try to exploit the expected price change $(P_2 - P_1)$ and at $t = 2$ they close their position. Long-term traders stock holding at $t = 1$ is

$$x_1^l = \frac{E[P_2 \mid \mathcal{F}_1^l] - P_1}{\rho \mathcal{S}_1} + \varrho E[x_2^l \mid \mathcal{F}_1^l],$$

where $\mathcal{S}_1$ and $\varrho$ are nonstochastic quantities.

Holden and Subrahmanyam derive the REE only for a special case and continue their analysis with numerical simulations. In equilibrium long-term traders reduce their date 1 demand if the variance of $\eta$, whose realisation will be announced at $t = 2$, is very high. They do not want to expose themselves to the announcement risk generated by $\eta$ (and reflected in $P_2$). They engage in heavier trading after a large part of uncertainty about the asset's value is resolved.

Holden and Subrahmanyam go to endogenise $M$ and, thus, $N = 1 - M$. The equilibrium mass $M$ can be derived by comparing the ex-ante utilities of short-term informed traders with the utility of long-term informed traders. They show that for certain cases the ex-ante utility from collecting short-term information is higher for $M \in [0, 1]$ than the utility from gathering the long-term signal. Thus, all traders search for the short-term signal in equilibrium. This is the case if the traders are sufficiently risk averse and $\sigma_\epsilon^2$ is substantially high.[64] Intuitively, short-term informed investors can only make use of their information from the price change $(P_2 - P_1)$, provided there are noise trader in $t = 1$, distorting $P_1$. Since $\eta$ makes $P_2$ risky, high variance in $\eta$ reduces their aggressiveness. Long-term informed traders can exploit their information from both price changes, $(P_2 - P_1)$ and $(P_3 - P_2)$. As described above, high variance of $\eta$ makes long-term informed agents delay their purchase. Therefore, they are more active at $t = 2$ and they exploit $(P_3 - P_2)$ to a greater degree. If the variance of $\epsilon$ is very high, i.e. speculating at $t = 2$ is very risky, long-term informed

[64] The assumption that $\sigma_\epsilon^2$ has to be very high is hidden in the legend of figure 3.

traders are very cautious at $t = 2$. Thus, they cannot make as much money out of their information as short-term informed traders can.

Holden and Subrahmanyam further show that as the degree of liquidity trading increases, both types of information are more valuable. Short-term investors profit more from higher variance, at least for the case where the variance of noise trading is the same in both periods.[65]

Another question they address is whether long-term information can be made more valuable by making it short-term. In other words, is it profitable for long-term informed investors to disclose their information already in $t = 2$? The impact of early credible disclosures is discussed in their last section.

Shleifer and Vishny (1990) provide further reasons why investors might be short-sighted. Incomplete markets which prevent complete risk sharing, credit constraints and other market imperfections make arbitrage cheaper for short-term assets than long-term assets resulting in less mispricing in short-term assets. In other words, it leads to systematically less accurate pricing of long-term assets. This, in turn, affects investment decisions of managers in the firms. Managers who are averse to mispricing of their equity because of potential takeovers etc. therefore, tend to conduct more short-term investments whose returns can be verified quickly. Alternatively, short-term behaviour of managers can also arise in agency models. Since this is true not only for the managers of listed firms but also for the managers of investment firms, pension funds etc. even institutional traders behave mypocially. Brandenbruger and Polak (1996) have a model where managers ignore their superior information and follow the opinion of the market. This is strictly less informationally efficient than herding behaviour among profit maximising firms.

## 2.6 Conclusion and Summary of Literature Survey

This survey covers a large section of rational models, in which difference in information drives prices. We begin with the concept of rational expectation equilibria and go on to study different partially-revealing REE. These models are summarised and the the limitations of

[65] In my opinion this is only true if $\sigma_\epsilon^2$ is sufficiently high.

the REE concept is discussed. We provide the intuition for two proofs of the No-Speculation Theorem. This theorem states that at a Pareto-efficient allocation and given concordant beliefs, information will not lead to further trade. Furthermore, two different kinds of No-Trade Theorems are illustrated. They refer to no-trade outcomes even when the current allocation is not Pareto optimal. The possible occurrence of bubbles even in situations where all traders are rational is explained.

The second part of the survey is restricted to the class of CARA-Gaussian models. The main focus is on dynamic REE models and on providing some rationale for technical analysis. The final section covers models of sequential information arrival. We catalog herding models and illustrate the distinct features of these models. The survey closes with herding models in information acquisition caused by short horizons of traders.

Most of this literature is quite recent and further major developments can be expected. The papers covered in this survey are by no means conclusive, nor do I claim that I have chosen the most important ones. Moreover, I have neglected a broad range of areas (like price experimentation, the analysis of disclosure of private information, effects driven by bounded rationality, extensive discussion about endogenous information acquisition etc.). There are many factors affecting the price process of an asset. This is surely one reason why this merits interest and examination.

# Chapter 3

# Buy on Rumours - Sell on News: A Manipulative Trading Strategy

## 3.1 Introduction

Investors base their expectations about the future payoffs of an asset on their information. This information affects their trading activity and, thus, the asset price. Information flow is, however, not just a one-way street. Traders who have not received new information are conscious of the fact that the actions of other traders are driven by their information. Thus, uninformed traders can infer part of the other traders' information from the movement of an asset's price.

But even when a trader does receive information he faces a problem if he wants to exploit it in the stock market. To determine his optimal trading moves, he has to figure out how much of this information is new. The current price already reflects this information if prices are (partially) revealing and other traders traded on this or similar information in previous trading rounds. Therefore, the question boils down to whether other market participants have already received related information. The same situation arises in the case of a public announcement. All market participants - with the exception of the trader who acquired this information early - do not know the extent to which this information is already incorporated in the current price.

To address this issue I develop a model where a trader receives an imprecise signal about a forthcoming public announcement. Even if the signal is imprecise, the trader trades on it and moves the price. Only this trader will know the price impact of his trading activity in the trading round prior to the public announcement. Thus he has an additional informational advantage at the time of the public announcement. I show that an early-informed trader can exploit his private information twice. First, when he receives his signal, and second, at the time of the public announcement. This result only holds if other traders draw inferences from the past price even after the public announcement, i.e. if they conduct technical analysis. Other traders face an additional error term in interpreting the past price. Since they do not know the price impact of the early informed traders action, they can not isolate the informational content of the past price from the part resulting from the imprecision of his signal. Paradoxically, it is the imprecision of the early-informed trader's signal which gives him the informational advantage at the time of the public announcement.

In addition to showing that the early-informed trader can exploit his information twice, I show that he trades for speculative reasons. I define 'speculative trading' as trading that is undertaken with the intent to unwind the acquired position after the public announcement. In this context, the early-informed trader can exploit his knowledge of the others' error and 'on average' reverse the position that he built up in the previous trading round. In short, after receiving a positive (negative) imprecise signal the trader buys (sells) stocks that he expects to sell (buy) at the time of the public announcement. In other words he follows the well known trading strategy: "Buy on Rumours - Sell on News".

Trading with the intention of moving the price such that the informational advantage is enhanced at the time of the public announcement is referred to as 'manipulative trading'. This paper also demonstrates that the trader who receives the information leakage trades in order to manipulate the price in his favour. If the early-informed agent trades very aggressively in the first trading round, the imprecision of his signal has a larger impact on the current price. This imprecision causes the other market participants to make an error while inferring information from the price. He 'throws sand in the eyes of the others'. The early-informed trader's future capital gains result from correcting the others' misinterpretation.

Hence, by trading more aggressively in the first trading round he increases his expected future capital gains in later trading rounds. Even though manipulative trading reduces capital gains prior to the public announcement, the trader more than makes up for this with additional profits afterwards. Manipulative trading behavior contrasts sharply with Kyle (1985) where the insider trades less aggressively today in order to save his informational advantage for future trading rounds. In my setting, he trades more aggressively now in order to enhance his future informational advantage. Therefore, the trading strategy should more appropriately be called "Trade 'Aggressively' on Rumours - Sell on News".

This paper also presents an alternative explanation for aggressive behavior. Recent experimental findings suggest that traders overestimate the importance of their private information (De Bondt and Thaler 1995). This literature attributes this behavior to irrational overconfidence on the part of the traders. My analysis, however, provides a rational explanation for their 'overactivism'.

This paper also highlights the importance of other traders' information in the interpretation of prices and runs counter to the notion of informational efficiency of markets. It illustrates that in some situations, knowledge about what other market participants know can be more important than knowledge about the fundamental value of a stock. This is in the spirit of Keynes' well known beauty contest argument (Keynes 1936). If it is important to know other traders' information in order to interpret the price, then the price cannot be a sufficient statistic for all individual signals. This sheds new light on the strong-form informational efficiency of markets. For the Grossman-Stiglitz Paradox[1] to arise, it is, therefore, not only necessary that all traders are price takers, as illustrated in Jackson (1991), but also that each market participant knows how his information is related to the information of other agents. Rumours are especially detrimental for achieving informationally efficient markets. Even after the truth is announced, rumours still distort the price and should there-

[1]The Grossman-Stiglitz Paradox refers to the non-existence of an overall equilibrium with endogenous information acquisition when prices are informationally efficient. If a price is informationally efficient, it reflects all private information, i.e. one can infer a sufficient statistic for all private signals by observing the price. Consequently, no trader has an incentive to gather costly information. However, if nobody collects information, the price cannot be informative and it would be worthwhile to buy a signal.

fore be avoided. This finding lends a rationale for crisis management wherein early public announcements are always recommended.

The remainder of the paper is organised as follows. The related literature is briefly summarised in Section 2. Section 3 outlines the model. It shows that early-informed traders still have an informational advantage at the time of the public announcement and that they trade for speculative as well as manipulative reasons. The impact of information leakage on informational efficiency is illustrated in Section 4. Section 5 extends the analysis to mixed strategies and examines a setting where many informed traders hear a rumor. Conclusion and topics for future research are presented in Section 6.

## 3.2 Related Literature

The prior literature focuses primarily on determining the conditions under which traders can manipulate asset prices. Price manipulation can have detrimental implications if it causes the price of an asset to depart even further from its fundamental value, i.e. if it causes the price to become less informationally efficient. The literature distinguishes between trade-based, information-based and action-based stock price manipulation (Allen and Gale 1992).

Classical examples of trade-based manipulation are market 'corners' and short 'squeezes'. Illiquid markets allow some market participants to temporarily exercise some monopoly power and move the price in their favour. However, not too much can be gained because the unwinding of the established position causes the market to move in the opposite direction again. Manipulative trading strategies are profitable if the spot market is less liquid than the futures market. Liquidity allows a trader to go long into futures without affecting price of futures significantly. The trader can then buy the underlying stocks after having established his futures position. If the spot market is illiquid, the price rises and he can short squeeze other traders. Other traders who are short in futures have to buy the underlying stock in order to deliver. Kumar and Seppi (1992) illustrate price manipulation if futures are settled by cash rather than by physical delivery. The intuition is that 'cash settlement' acts as an infinitely liquid market in which pre-existing futures positions are closed out relative to the less liquid spot market. In Allen and Gorton (1992) trade-based manipulation is possible since buy orders are more likely to be from informed traders than sell orders. Therefore, the market is less liquid for upswings than for downturns. Allen and Gale (1992) present a model about trade-based manipulation with higher order uncertainty where all traders are price takers except for one large trader, who is either an informed trader or an uninformed manipulator. Similarly, in Chakraborty (1997) there is also a potentially well informed insider. In addition, there are less informed followers and uninformed liquidity traders. Chakraborty's model illustrates manipulation by the potentially informed insider within a Glosten and Milgrom (1985) setting. Fishman and Hagerty (1995) show that mandatory disclosure of individual trading activities can lead to manipulation.

Information-based manipulation involves the release of false information or the spreading of rumours in order to achieve a favourable stock price. Vila (1989) presents a simple model of information-based manipulation in which the trader, after going short, releases false information in order to buy the stock back at a cheaper price. Benabou and Laroque (1992) focus on the credibility of insiders and gurus who profitably manipulate prices by sometimes publicising incorrect statements. They have an incentive to make false announcements in order to move the price in their favour. Since it is known that their information is noisy and thus manipulation cannot be detected with certainty, their reputation is not destroyed completely. Nevertheless their credibility is hurt, thereby rendering future information-based manipulations less effective.

Action-based manipulation results when corporate insiders entangle corporate decisions with their private stock market activities. They can take stock value enhancing or reducing actions within a firm with the objective of making private gains from speculation in the stock market.

Treynor and Ferguson (1985) address the problem faced by a trader who does not know whether his information is already known to all the other market participants or not. They demonstrate that the past price process can help the trader answer this question. That is, they illustrate the usefulness of technical analysis for this problem.

In my paper only the early-informed trader knows the extent to which the information revealed to the public is already reflected in the price. All other market participants try to infer this from the past price changes. In addition, this analysis focuses on the strategic behaviour of the early-informed trader who trades for manipulative and speculative reasons. The complete analysis highlights the necessary conditions that generate the inference problem for the public. Furthermore, in contrast to most of the other models in the literature manipulative trading is derived without the imposition of any restrictions on the traders' order size.

## 3.3 Analysis

### 3.3.1 Model Setup

There are two assets in the economy: a risky stock and a risk-free bond. For simplicity we normalise the interest rate of the bond to zero. Market participants include risk-neutral informed traders, liquidity traders and a market maker. Informed traders' sole motive for trading is to exploit their superior information about the fundamental value of the stock. Liquidity traders buy or sell shares for reasons exogenous to the model. Their demand typically stems from information which is not of common interest such as from their need to hedge against endowment shocks or private investment opportunities in an incomplete market setting.[2] A single competitive risk-neutral market maker observes the aggregate order flow and sets the price.

Traders submit their market orders to the market maker in two consecutive trading rounds taking into account the price impact of their orders. The market maker sets the price in each round after observing the aggregate order flow and trades the market clearing quantities. As in Kyle (1985) the market maker is assumed to set informationally efficient prices; thus his expected profit is zero. The underlying Bertrand competition with potential rival market makers is not explicitly modelled in this analysis.[3] Informed traders receive

[2] See Brunnermeier (1997) for a detailed discussion of the different reasons why liquidity traders trade, and for a discussion on the distinction between information of common versus private interest.

[3] Alternatively, one could also employ a setting where the competitive risk-neutral market maker sector of scalpers, floor brokers etc. submit limit orders, i.e. demand schedules. The price is then determined by

their signal before trading begins in $t = 1$. The public announcement occurs prior to trading in $t = 2$. The timeline in Figure 3.1 illustrates the sequence of moves.

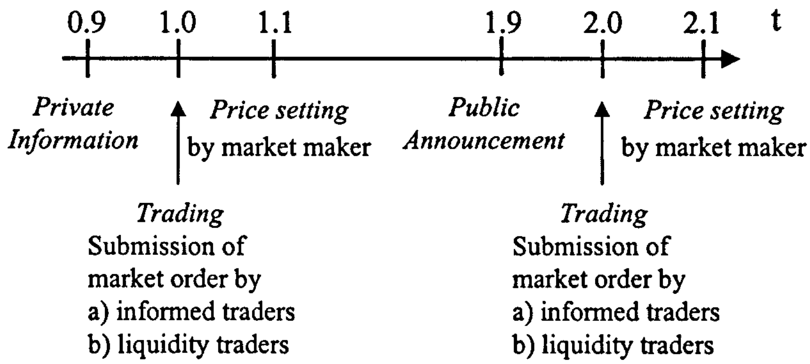



Figure 3.1: Timeline

Traders face a price risk submitting market orders since they do not know the price at which their trade will be executed. In contrast, limit orders allow the trader to specify a price at which the order will be executed. By combining many limit and stop orders traders can create demand schedules which allow them to trade conditionally on the current price. Unfortunately, limit order models make the analysis less tractable without adding more insight. Therefore, I have opted for a market order setting similar to Kyle (1985), Admati and Pfleiderer (1988) and Foster and Viswanathan (1996).

Many different events can provide information about the equity value of a company. Events like earnings announcements, a major contract with a new client, legal allegations, a new CEO, macroeconomic news etc. can have a significant impact on the market value of a stock. Let us restrict our attention to only two events, $A$ and $B$. Their impact on the value of the stock is modelled by the two random variables $\delta^A$ and $\delta^B$, which are independently normally distributed with mean zero. The liquidation value of the stock $v = \delta^A + \delta^B$ is paid out in $t = 3$. Event $A$ is publicly announced before the second trading round and the price impact $\delta^A$ of event $A$ becomes common knowledge to all market participants. Prior to the announcement some information about the event $A$ leaks already to trader $A$. This informa-

market clearing.

tion, $\delta^A+\varepsilon$, leaks possibly in the form of a rumour. The error term $\varepsilon$ reflects the imprecision of the rumour. Trader $B$ knows event $B$'s impact on the stock value already in period one. The price in $t=1$, $p_1$, still carries information even after the public announcement due to trader $B$'s information. In period three the true value of the stock $v=\delta^A+\delta^B$ is known to everybody. That is, trader $B$'s early information $\delta^B$ is made public in $t=3$. Liquidity traders do not receive any information and their aggregate trading activity is summarised by the random variables $u_1$ in period one and $u_2$ in period two. The information structure is common knowledge, i.e. I assume that all market participants know that trader $A$ has received some noisy information about a forthcoming public announcement but they do not know its content.[4]

The information structure is summarised in the following table:

| player $i$ | in period $t=1$ | in period $t=2$ | in period $t=3$ |
|---|---|---|---|
| market maker | $X_1$ | $\delta^A, p_1, X_2$ | $\delta^B, p_2$ |
| trader $A$ | $\delta^A+\varepsilon$ | $\delta^A, p_1$ | $\delta^B, p_2$ |
| trader $B$ | $\delta^B$ | $\delta^A, p_1$ | $p_2$ |

where $X_1=x_1^A+x_1^B+u_1$ is the aggregated orderflow in $t=1$ and $X_2=x_2^A+x_2^B+u_2$ is the orderflow in $t=2$. For notational simplicity I will denote the signal of trader $i\in\{A,B\}$ at time $t$ by $S_t^i$. The random variables $\delta^A$, $\delta^B$, $\varepsilon$, $u_1$ and $u_2$ are independently normally distributed with mean zero. For symmetry reasons let $Var[\delta^A]=Var[\delta^B]$.

An analysis of price manipulation is ruled out in a Rational Expectations Equilibrium setting because all traders are assumed to be price-takers. In a Perfect Bayesian Nash equilibrium setting, however, all traders take the strategies of all other players as given. That is,

[4]This problem can also be captured in a model with higher order uncertainty, i.e. information leakage occurs only with a certain probability. Trader 1 receives then two pieces of information. In addition to the actual signal he knows whether some information leaked or not. Trader 1's informational advantage at the time of the public announcement stems from his knowledge of whether he received an early signal or not. Such models were not pursued in this paper because they are either very simplistic or intractable and do not provide much additional insight.

they are aware that their trade affects the price. All informed traders submit their market orders, $x_t^i$, based on their information in each trading round to the market maker. Note that in period two, each trader $i$ knows not only his signal, the price $p_1$ and the public information $\delta^A$ but also his demand in $t = 1$, $x_1^i$. After observing the aggregate net order flow the risk-neutral market maker sets the execution price $p_t$. The price is semi-strong informationally efficient, i.e. the price is the best estimate given the market maker's information. A different price would lead to an expected loss or an expected profit for the market maker. The latter is ruled out because the market maker faces Bertrand competition from potential rival market makers. For ease of exposition the strategy for the market maker, is exogenously specified. He has to set informationally efficient prices in equilibrium, i.e. $p_1 = E[v|X_1]$ and $p_2 = E[v|X_1, \delta^A, X_2]$ due to potential Bertrand competition.

A Perfect Bayesian Nash equilibrium of this trading game is given by strategy profile $\{\{x_1^{i,*}(\cdot), x_2^{i,*}(\cdot)\}_{i=\{A,B\}}, p_1^*(\cdot), p_2^*(\cdot)\}$ such that

(1) $x_2^{i,*} \in \arg\max_{x_2^i} E[x_2^i(v - p_2)|S_1^i, x_1^i, p_1, \delta^A]\ \forall i \in \{A, B\}$,
(2) $x_1^{i,*} \in \arg\max_{x_1^i} E[x_1^i(v - p_1) + x_2^{i,*}(v - p_2)|S_1^i]\ \forall i \in \{A, B\}$, and
prices $p_1^* = E[v|X_1^*]$ and $p_2^* = E[v|X_1^*, \delta^A, X_2^*]$,

where the conditional expectations are derived using Bayes' Rule to ensure that the beliefs are consistent with the equilibrium strategy.

### 3.3.2 Characterization of Linear Equilibrium

Proposition 1 characterises a Perfect Bayesian equilibrium in linear pure strategies. It has the nice feature that each trader's demand is the product of his trading intensity (or aggressiveness) and the difference in the trader's and market maker's expectations about the value of the stock. Linear strategies have the advantage that all random variables remain normally distributed. In addition, the pricing rules are linear as a consequence of the Projection Theorem.[5] In period one the market maker's pricing rule is $p_1 = \lambda_1 X_1$ and in

[5]Since all variables are normally distributed the orthogonal projection of $v$ on the space of linear-affine functions of $S$ is equal to the projection of $v$ (in the sense of $\mathcal{L}^2$) on the space $\mathcal{L}^2(S)$ of quadratic integrable

period two it is $p_2 = \delta^A + E[\delta^B|X_1, \delta^A] + \lambda_2 X_2$ in equilibrium. As in Kyle (1985) $\lambda_t$ reflects the price impact of an increase in market order by one unit. This price impact restricts the trader's optimal order size. Kyle interpreted the reciprocal of $\lambda_t$ as market depth. If the market is very liquid, i.e. $\lambda_t$ is very low, then an increase in the trader's demand has only a small impact on the stock price. For expositional clarity I denote the regression coefficient of $y$ on $x$ by $\phi_x^y := \frac{Cov[x,y]}{Var[x]}$.

**Proposition 1** *A Perfect Bayesian Nash equilibrium in which all pure trading strategies are of the* ***linear*** *form*

$x_1^i = \beta_1^i(S_1^i)$,

$x_2^i = \beta_2^i(E[v|S_1^i, p_1, \delta^A] - E[v|p_1, \delta^A])$,

*and the market maker's pricing rule*

$p_1 = E[v|X_1] = \lambda_1 X_1$,

$p_2 = E[v|X_1, \delta^A, X_2] = \delta^A + \phi^{\delta^B}_{S_2^{p_1}} S_2^{p_1} + \lambda_2 X_2$, *with* $S_2^{p_1} = \frac{X_1 - \beta_1^A \delta^A}{\beta_1^B}$,

*is given by the fixed points of the following system of equations*

$$\begin{aligned} \beta_1^{A,*} &= \frac{1}{2\left(\lambda_1 - \lambda_2 \left(\gamma_2^A\right)^2 \phi^{\varepsilon}_{S_1^A}\right)} \phi^{\delta^A}_{S_1^A} \\ \beta_1^{B,*} &= \frac{1}{2\lambda_1}\left(1 - 2\lambda_2 \beta_2^B \gamma_2^B \left(1 - \phi^{\delta^B}_{S_2^{p_1}}\right)\right) \end{aligned}$$

*where*

$$\lambda_1 = \frac{\beta_1^{A,*} Var[\delta^A] + \beta_1^{B,*} Var[\delta^B]}{Var[\beta_1^{A,*}\left(\delta^A + \varepsilon\right) + \beta_1^{B,*}\left(\delta^B\right) + u_1]} \qquad \lambda_2 = \frac{Cov[\delta^B, X_2 | S_2^{p_1}]}{Var[X_2 | S_2^{p_1}]}$$

*with*

$$\beta_2^A = \frac{1}{2\lambda_2}\frac{1}{2}\frac{1}{1 - \frac{1}{4}\phi_w^{S_2^{A,w}}\phi^w_{S_2^{A,w}}}\phi^w_{S_2^{A,w}} \qquad \gamma_2^A := \frac{\beta_2^A}{\beta_1^B}\frac{\phi^{\delta^B}_{S_2^{p_1}}}{\phi^w_{S_2^{A,w}}}$$

$$\beta_2^B = \frac{1}{2\lambda_2}\frac{1 - \frac{1}{2}\phi_w^{S_2^{A,w}}\phi^w_{S_2^{A,w}}}{1 - \frac{1}{4}\phi_w^{S_2^{A,w}}\phi^w_{S_2^{A,w}}} \qquad \gamma_2^B := \frac{1}{2\lambda_2}\frac{1}{\beta_1^B}\phi^{\delta^B}_{S_2^{p_1}} - \frac{1}{2}\frac{\beta_2^A}{\beta_1^B}\frac{\phi^{\delta^B}_{S_2^{p_1}}}{\phi^w_{S_2^{A,w}}}$$

*if the second order conditions* $\lambda_2 > \lambda_1 \max\left\{\left[\frac{b_2^A}{b_1^B}\frac{\phi^{\delta^B}_{S_2^{p_1}}}{\phi^w_{S_2^{A,w}}}\right]^2, \left[\frac{1}{b_1^B}\phi^{\delta^B}_{S_2^{p_1}} - \frac{1}{2}\frac{b_2^A}{b_1^B}\frac{\phi^{\delta^B}_{S_2^{p_1}}}{\phi^w_{S_2^{A,w}}}\right]^2\right\}$, $\lambda_2 > 0$ *(with* $b_t^i := 2\lambda_t \beta_t^i$*) are satisfied.*

---

functions of $S$. Consequently, $E[v|\mathbf{S}] = E[v] + (\mathbf{S} - E[\mathbf{S}])^\tau \mathbf{Var}^{-1}[\mathbf{S}] Cov[v, \mathbf{S}]$, which allows us to calculate the conditional expectations.

The interested reader is referred to the Appendix for a complete proof of the proposition. The proof makes use of backward induction. In order to solve the continuation game in $t = 2$, the information structure prior to trading in $t = 2$ has to be derived. For this purpose propose an arbitrary action rule profile, $\{\{\beta_1^i\}_{i\in\{A,B\}}, p_1(X_1)\}$ for $t = 1$, which is mutual knowledge and is considered to be an equilibrium profile by all agents. In $t = 2$ all market participants can derive the aggregate order flow $X_1 = \beta_1^A(\delta^A + \varepsilon) + \beta_1^B \delta^B + u_1$ from price $p_1$. After knowing $\delta^A$ the price signal is $S_2^{p_1} = \delta^B + \frac{\beta_1^A}{\beta_1^B}\varepsilon + \frac{1}{\beta_1^B}u_1$. Since all market participants know $S_2^{p_1}$, it is useful to state each traders' information relative to the publicly known symmetric information, i.e. orthogonalize the signals with respect to $S_2^{p_1}$. The stock is split into an expected part $E[v|S_2^{p_1}, \delta^A] = \delta^A + E[\delta^B|S_2^{p_1}]$ and an unexpected part $w := \delta^B - E[\delta^B|S_2^{p_1}]$. This 'virtual' split of the stock $v$ into a risk-less bond and a risky asset $w$ is possible, without loss of generality, as long as all traders are risk-neutral or have exponential utility functions. Note that the stock split depends on the proposed action rule profile $\{\{\beta_1^i\}_{i\in\{A,B\}}, p_1(X_1)\}$ for $t = 1$ and not on the one actually chosen. In other words, the stock split is not affected if some trader deviates. Trader 1's information is $S_2^{A,w} = w + \frac{1}{\beta_1^B}\vartheta_1$, where $\vartheta_1 = u_1 - E[u_1|S_2^{p_1}]$. Trader $B$ knows the fundamental value of the risky component $w$ but has to forecast trader $A$'s forecast $S_2^{A,w}$ in order to predict trader $A$'s order size in $t = 2$. Knowing trader $A$'s demand in $t = 2$ would help trader $B$ predict the price $p_2$ at which his orders will be executed.[6] In $t = 2$ traders face a generalised static Kyle-trading-game with the usual trade-off. On the one hand, a risk-neutral trader wants to trade very aggressively in order to exploit the gap between his estimate of the fundamental value of the stock and the price of the stock. On the other hand, very aggressive trading moves the price at which his order will be executed towards his estimate of the asset's value since it allows the market maker to infer more of the trader's information from the aggregate order flow. This latter price impact reduces the value-price gap from which the trader can profit and restrains the traders from trading very aggressively.

[6] Note that the infinite regress problem discussed in Townsend (1983) can be avoided because of the linearity of the trading strategies $x_t^i = \beta_t^i S_t^i$, where $S_t^i$ is the difference between trader $i$'s and the market maker's expectation of $v$. Linearity also preserves the normality distribution of the (conditional) random variables.

Using backward induction one has to check whether a single player wants to deviate in $t = 1$ from the proposed action rule profile, $\{\{\beta_1^i\}_{i\in\{A,B\}}, p_1(X_1)\}$. Trading in $t = 1$ affects not only the capital gains in $t = 1$ but also the future prospects for trading in $t = 2$. If trader $i$ deviates in $t = 1$ from the proposed action rule $x_1^i(\cdot)$ to $x_1^{i,di}(\cdot)$ by trading more or less aggressively his expected future capital gains at $t = 2$ are affected in two ways. His optimal trading intensity at $t = 2$ is affected, as is the price in $t = 2$, $p_2$. The change in $p_2$ is due to the misperception by all other market participants. All other players, thinking that trader $i$ did not deviate, still play their equilibrium strategy. Thus, they infer the wrong signal from the aggregate order flow in $t = 1$, $X_1^{di}$ or $p_1^{di} = \lambda_1 X_1^{di}$, where the superscript $di$ indicates that trader $i$ deviated. This alters the other traders' order size in $t = 2$ and the market maker's price setting in $t = 2$. The impact of both effects on the price schedule is given by $2\lambda_2\gamma_2^i(x_1^i - x_1^{i,di})$. Consequently, the deviant adjusts his market order in $t = 2$ by $\gamma_2^i(x_1^i - x_1^{i,di})$ as stated in the proposition.

The value function

$$V_2^i(x_1^{i,di}) = \kappa^i \left(S_2^{i,w}\right)^2 - \tau^i S_2^{A,w}\left(x_1^{i,di} - x_1^i\right) + \psi^i \left(x_1^{i,di} - x_1^i\right)^2 \forall i \in \{A, B\}$$

captures the impact of trading in $t = 1$ on the (maximal) expected gains in $t = 2$. Capital gains are the product of the optimal order size and the estimated value-price gap in $t = 2$. The value function is quadratic since deviation in $t = 1$ affects the optimal order size as well as the estimated value-price gap in $t = 2$ linearly. The value function can be decomposed into four parts. The first component is the expected capital gains in the proposed equilibrium captured by $\kappa^i \left(S_2^{i,w}\right)^2$. The second component reflects the impact of the deviation in $t = 1$ on the optimal order size in $t = 2$ while ignoring the impact on the execution price. The third component holds the order size fixed while taking into account the fact that a deviation in $t = 1$ changes the price in $t = 2$. This occurs because of the misperception of other market participants and the deviant's demand adjustment in $t = 2$. Finally, the fourth component isolates the effects which are solely due to the induced change in the optimal order size and in the price schedule not covered by the other effects. The second and third component are summarised by the coefficient $\tau^i$ and the fourth by coefficient $\psi^i$. An equilibrium is reached if no trader wants to deviate from the proposed action rule profile in $t = 1$. In other words, the Perfect Bayesian Nash Equilibrium is given by the fixed point described in Proposition 1.

Proposition 1 also presents two inequality conditions. They result from the second order conditions in the traders' maximisation problems. They guarantee that the quadratic objective functions for each period have a maximum rather than a minimum. In economic terms, they require that the market is sufficiently liquid/deep in trading round one relative to trading round two. These inequality restrictions rule out the case where it is optimal to trade an unbounded amount in $t = 1$, move the price, and make an infinitely large capital gain in $t = 2$.

### 3.3.3 Exploiting Information Twice

Information about the fundamental value of the stock as well as information about other traders' demand affects the traders' optimal order size. In period two, traders can infer some information from the past price, $p_1$. Brown and Jennings (1989) call this inference 'technical analysis'. If a trader's prediction of the stock's liquidation value is more precise than the market maker's prediction, then the trader has an informational advantage. Proposition 2 shows that trader $A$ still has an informational advantage in period two because both trader $B$ and the market maker employ technical analysis. Trader $A$ can, therefore, *exploit his private information twice.* First, when he receives his signal, and second, at the time of the public announcement. This is surprising since one might think that the public announcement is a sufficient statistic for trader $A$'s private information.

**Proposition 2** *Trader $A$ (almost surely) retains an informational advantage in period two in spite of the public announcement in period two. Technical analysis is more informative about the value of the stock for trader $A$ than for trader $B$. Trader $A$'s informational advantage in period two is increasing in his trading intensity and decreasing in the trading intensity of trader $B$ in $t = 1$.*

In period two, only trader $A$ knows the exact extent to which the price, $p_1$ already reflects the new public information, $\delta^A$. It is interesting to note that the informational advantage of trader $A$ in period two is a consequence of the technical analysis conducted by trader $B$ and by the market maker. Both trader $B$ and the market maker try to infer information in $t = 2$ from the past price $p_1$. In general, technical analysis serves two purposes.

First, traders try to infer from the past price more about the fundamental value of the stock, and second they use the past price to forecast the forecasts of the others. Knowing others' estimates is useful for predicting their market orders in $t = 2$. This in turn allows traders to estimate the execution price $p_2$ more precisely. Trader $B$ trades conditional on $p_1$ in $t = 2$ in order to improve his forecasts of trader $A$'s market order in $t = 2$. The price, $p_1 = \lambda_1 X_1$ depends on the individual demand of trader $A$, $x_1^A$, and thus on the signal $\delta^A + \varepsilon$.

When conducting technical analysis trader $B$ is aware that price $p_1$ is affected by the error term $\varepsilon$. Trader $A$'s informational advantage in $t = 2$ is his knowledge of the error $\varepsilon$. He can infer $\varepsilon$ from the difference between his signal in $t = 1$ and the public announcement in $t = 2$. If both trader $B$ and the market maker would ignore the price signal, $p_1$, the public announcement would be a sufficient statistic for trader $A$'s private information, $\delta^A + \varepsilon$. For the case in which they employ technical analysis, trader $B$ would like to know the extent to which the trading activities of trader $A$ changed price, $p_1$. Knowledge not only of $\delta^A$ but also of $\varepsilon$ would allow them to infer even more information from the price, $p_1$. The public announcement in $t = 2$ is not a sufficient statistic of $\delta^A + \varepsilon$ for interpreting the past price, $p_1$.

Trader $A$ applies technical analysis in order to infer more information about the fundamental value of the stock. This information is also valuable for predicting trader $B$'s demand in $t = 2$. The additional information about the value of the stock provided by technical analysis is higher for trader $A$ than for the market maker. For trader $B$ technical analysis only provides information about trader $A$'s forecast since trader $B$ already knows the liquidation value $v$ in $t = 2$. Since trader $A$ knows his own demand, he can infer $\frac{1}{\beta_1^B}(\frac{p_1}{\lambda_1} - x_1^A) = \delta^B + \frac{1}{\beta_1^B} u_1$, which is trader $B$'s signal perturbed by the demand of the noise traders. The market maker can infer $(\delta^B + \frac{1}{\beta_1^B} u_1) + \frac{\beta_1^A}{\beta_1^B}\varepsilon$, which is trader $A$'s price signal perturbed by the error term, $\varepsilon$. Trader $A$'s informational advantage is, therefore, given by $\frac{\beta_1^A}{\beta_1^B}\varepsilon$, which increases with his trading intensity, $\beta_1^A$, and decreases with the trading intensity of trader $B$, $\beta_1^B$. Intuitively, if trader $A$ trades more aggressively in $t = 1$ his signal's imprecision has a higher impact on the price, $p_1$.

### 3.3.4 Speculative and Manipulative Trading

In general, trading occurs for risk sharing purposes or for informational reasons. Since all traders are risk-neutral in this setting, their only motive to trade is to exploit their informational advantage. As illustrated in Proposition 2 current trading affects future informational advantages. In Kyle (1985) the single insider reduces his trading intensity in order to save information for future trading rounds. The single insider faces a trade-off. Taking on a larger position in period one can result in higher profits today but also leads to worse prices for current and future trading rounds. Thus in a Kyle (1985) setting the insider restrains his trading activity with the objective of not trading his informational advantage away.

In contrast to the literature based on Kyle (1985), trader $A$ in my model trades more aggressively in period one. He incurs myopically non-optimal excessive trades in period one and then recuperates the losses and makes additional profit in period two. Trading more aggressively in period one changes the price in such a way that his informational advantage in the next trading round is enhanced. Trading with the sole intention of increasing one's informational advantage in the next period is defined as *manipulative trading*. *Speculative trading* is defined as trading with the expectation to unwind one's position in the next period. The following definitions restate the two trading objectives:

**Definition 1** *Speculative trading is carried out with the expectation of unwinding the acquired speculative position in the next period.*

Speculative trading can also be manipulative.

**Definition 2** *Manipulative trading is intended to move the price in order to enhance the informational advantage in the next period.*

Manipulative trading is excessive in the sense that it is the component of trading intensity which exceeds the optimal myopic trading intensity, holding the other market participants' strategies fixed. The myopic trading intensity does not take into account the fact that by trading more aggressively trader $A$ could enhance his informational advantage in period two.

Proposition 3 shows that trader $A$ trades for speculative reasons since he expects to unwind part of his accomplished position in period two. Furthermore, he trades excessively with the objective of manipulating the price.

**Proposition 3** *In period one, trader $A$ trades conditional on his current information in order to build up a long-term position, and also for speculative and manipulative reasons.*

*Speculative trading is given by $\gamma_2^A \phi_{S_1^A}^{\varepsilon} \beta_1^A S_1^A$.*

*Manipulative trading is given by $\lambda_2 \left(\gamma_2^A\right)^2 \phi_{S_1^A}^{\varepsilon} \beta_1^A S_1^A$,*
*where the coefficients in front of $S_1^A$ are strictly positive.*

The proof in the appendix shows that if trader $A$ receives a positive signal, all trading objectives induce the trader to take a long position in the stock. Similarly, if trader $A$ receives a negative signal he sells the stock.

All traders trade conditional on their signal in period one. Therefore, the price $p_1$ reflects not only the signal about $\delta^B$ but also the signal about $\delta^A + \varepsilon$. The main motive for technical analysis is to infer more information about $\delta^B$ and about the others' forecasts. All market participants can separate the impact of $\delta^A$ on $p_1$, but only trader $A$ can deduce the impact of the $\varepsilon$ error term on $p_1$. The market maker's inference about $\delta^B$ from the price $p_1$ is perturbed by $\frac{\beta_1^A}{\beta_1^B}\varepsilon$. He overestimates (underestimates) $\delta^B$ if $\varepsilon$ is positive (negative). Since trader $A$ can infer $\varepsilon$ in period two, he can make money by correcting the market maker's error. If $\varepsilon$ is positive (negative), trader $A$ sells (buys) stock in period two. In period one, not even trader $A$ knows $\varepsilon$. His prediction of $\varepsilon$, given his signal $S_1^A = \delta^A + \varepsilon$, $E[\varepsilon|S_1^A] = Var[\varepsilon](Var[\delta^A] + Var[\varepsilon])^{-1}S_1^A$ is always of the same sign as his trade in period one $x_1^A = \beta_1^A S_1^A$. Therefore, trader $A$ expects to trade in the opposite direction in period two. 'On average', he partially unwinds his position in period two. This is solely due to informational reasons since trader $A$ expects the price to overshoot in $t = 2$. Given, however, the information of the market maker or of any other outsider who only observes the past prices and the public announcement, the price follows a Martingale process, i.e. it neither overshoots nor undershoots.

The purpose of manipulative trading is to extend the informational gap in the second trading round. It relates to the literature on signal-jamming. By trading excessively in $t = 1$, trader $A$ worsens the other market participants' price signal $S_2^{p_1}$ in $t = 2$ about the fundamental value $\delta^B$. The reason is that by trading more aggressively the imprecision of trader $A$'s signal $\varepsilon$ has a larger impact on $p_1$. Consequently, the price signal $S_2^{p_1} = \delta^B + \frac{\beta_1^A}{\beta_1^B}\varepsilon + \frac{1}{\beta_1^B}u_1$ reveals more information about $\varepsilon$ and less about the fundamental value $\delta^B$. This increases trader $A$'s informational advantage in $t = 2$ with respect to the market maker. It also makes trader $B$'s forecast about trader $A$'s $\delta^B$ forecast worse. Trader $B$ already knows the fundamental value $\delta^B$ and his only motive for conducting technical analysis is to achieve a better prediction of trader $A$'s market order and thus the execution price in $t = 2$, $p_2$. Trader $A$'s market order is based on his information, $\delta^B + \frac{1}{\beta_1^B}u_1$ and $S_2^{p_1}$. The only term trader $B$ does not know is $u_1$, trader $A$'s error in predicting the fundamental value $\delta^B$. The price signal $S_2^{p_1}$ allows him to derive $\beta^A\varepsilon + u_1$. This helps him to forecast trader $A$'s order size. However, he can not perfectly forecast it since his signal is perturbed by $\beta_1^A\varepsilon$, the imprecision of the rumor times trader $A$'s trading intensity in period one. In short, if trader $A$ trades more aggressively in period one, he builds up a larger informational advantage with respect to the market maker and also reveals less of his informational advantage to his competitor trader $B$. Overall, more aggressively trading in period one increases trader $A$'s expected future capital gains. The proof in the appendix shows that in equilibrium the trading intensity of trader $A$ is higher if he takes the impact on future expected capital gains into account, given the strategies of all other players. It is the expected knowledge of the $\varepsilon$-term in $t = 2$ which induces manipulative trading.

Speculative trading is also caused by the imprecision of trader $A$' signal, $\varepsilon$. Consequently, an increase in trading intensity in period one due to manipulative behaviour also leads to more speculation. The trader expects to unwind a larger position in $t = 2$.[7]

Hirshleifer, Subrahmanyam, and Titman (1994) appeal to traders' risk-aversion and thus provide a very distinct explanation for speculative behaviour. In their model, early-informed traders receive the same piece of information one period prior to the late-informed traders, while the competitive risk-neutral market makers observe only the limit order book. Traders submit limit orders, in the form of whole demand schedules, to the risk-neutral market

[7]Note if $\delta_A$ and $\delta_B$ could be traded separately neither speculative nor manipulative trading would arise.

makers who set the price. All traders have to be risk-averse in their analysis. Furthermore, since they are also competitive price-takers, manipulative trading is ruled out. The intuition for speculative trading in their model is as follows. Since no risk premium is paid due to the market makers' risk-neutrality, risk-averse traders would be unwilling to take on any risky stock in the absence of any informational advantage. Early-informed investors are willing to take on risk since they receive a signal in period one. Their informational advantage, together with the existence of noise traders, compensates them for taking on the risky asset. However, the informational advantage of early-informed traders with respect to the late-informed traders vanishes in period two since both now receive the same signal. Thus, early-informed traders share the risk with late-informed traders in period two. In addition, the informational advantage of the early-informed traders with respect to the market makers shrinks as well since market makers can observe a second limit order book. Therefore, in period two, both these effects cause early-informed traders to partially unwind the position they built up in the previous period.

## 3.4 Impact of Information Leakage on Informational Efficiency

In the information structure analysed above, trader $A$ already received in period one some information about the forthcoming public announcement in period two. In other words, some news about the public announcement leaked to trader $A$ before it was made public. This section compares this to the benchmark case where the announcer manages to keep the content of his public announcement secret. That is, no information leaks and thus trader $A$ has no informational advantage in period one. It addresses the question of whether information leakage makes the price more or less informationally efficient.

A market is (strong-form) informationally efficient if the price is a sufficient statistic for all the information dispersed among all market participants. In this case the market mechanism perfectly aggregates all information which is available in the economy, and the price reveals it to everybody. In general, if traders trade for informational as well as non-informational reasons, the price is not informationally efficient. This is also the case in my setting where some traders try to exploit their superior information and others trade for liquidity reasons. Nevertheless, one can distinguish between more and less informationally efficient markets. A measure of informational efficiency should reflect the degree to which information dispersed among many traders can be inferred from the price (process) together with other public information. Consider the forecast of the fundamental value of the stock $v$, given the pool of all available information in the economy at a certain point in time. If the price (process) is informationally efficient then the price(s) and other public information up to this time yields the same forecast. Consequently, the variance of this forecast conditional on prices and other public information is zero. This conditional variance increases as the market becomes less informationally efficient. Therefore, I choose the reciprocal of this conditional variance, i.e. the precision, as a measure of the degree of informational efficiency. Note that the degree of informational efficiency depends crucially on the pool of information in the economy. To illustrate this, consider a world without asymmetric information. In this setting any price process is informationally efficient even though it is uninformative. The conditional variance of the stock value itself captures how informative the price (process) and the other public information are.[8] This variance is zero if all public information, including the price process, allows one to perfectly predict the liquidation value of the stock. In this case everybody knows the true stock value. This variance term, therefore, also measures the risk a liquidity trader faces when trading this stock.

The following definitions define both measures more formally.

**Definition 3** *The reciprocal of the variance* $Var[E[v|\{p_t, S_t^{public}, \{S_t^i\}_{i \in \mathbf{I}}\}_{t \square \tau}]|\{p_t, S_t^{public}\}_{t \square \tau}]$ *conditional on the public information,* $S_\tau^{public}$*, and the pool of private information up to time* $\tau$ *measures the degree of informational efficiency at time* $\tau$*.*
*The reciprocal of the conditional variance* $Var[v|\{p_t, S_t^{public}\}_{t \square \tau}]$ *measures how informative the price (process) and the public information are.*

Equipped with these measures, one can analyse how the information leakage of $\delta^A + \varepsilon$ to trader $A$ affects informational efficiency and informativeness of the price (process). In

[8] Note that all public information at the beginning of the trading game is incorporated in the common priors.

addition it allows us to address the role of the imprecision of the rumour.

Since these definitions are time dependent, let us analyse informational efficiency and informativeness at the time after the first trading round, after the public announcement of $\delta^A$, and after the second trading round. Let us assume for the following proposition that there is a sufficient amount of liquidity trading in $t = 1$. More precisely, $Var[u_1] > \frac{6}{25}\sqrt{\frac{2}{5}Var[u_2]Var[\delta^B]}$.

**Proposition 4** *In $t = 1$ information leakage makes the price $p_1$ more informative but less informationally efficient if the information leakage is sufficiently precise. However, after the public announcement in $t = 2$ both informativeness as well as informational efficiency are reduced.*

Leakage of information makes the price $p_1$ in $t = 1$ more informative, if $Var[\varepsilon]$ is not too high. Trader $A$ trades on his information $\delta^A + \varepsilon$ and thus price $p_1$ reveals information about not only $\delta^A$ but also about $\delta^B$. Trader $A$'s market activity increases informed trading relative to liquidity trading. This allows the market maker as well as the public to infer more information from the aggregate order flow $X_1$. Note that for very high $Var[\varepsilon]$ this might not be the case since aggressive manipulative trading activity could increase the non-informative component of the aggregate order flow.

On the other hand, information leakage makes the market less informationally efficient in $t = 1$. If there is no leakage, $p_1$ reveals more about $\delta^B$ than it reveals about $E[v|\delta^B, \delta^A + \varepsilon] = \delta^B + \phi^{\delta A}_{S_1^A}(\delta^A + \varepsilon)$ in the case of a leakage. The reason is that sufficiently precise information leakage leads to a higher $\lambda_1$ which reduces the trading intensity of trader $B$, $\beta_1^B$. Therefore less information can be inferred about $\delta^B$. In addition, $\delta^A + \varepsilon$ can only be partly inferred from the price $p_1$. Both effects together result in a lower informational efficiency for $p_1$ in the case of a precise leakage.

After the public announcement in $t = 2$ $\delta^A$ as well as $\delta^B$ are known by some traders in the economy, (i.e. the best forecast of $v$ given the pooled information is $v$). Consequently, the measures of informational efficiency and informativeness coincide from that moment onwards. Since $\delta^A$ is common knowledge, the conditional variance stems solely from the uncertainty about $\delta^B$. The proof in the appendix shows that sufficiently precise information leakage leads to a less liquid market, i.e. to a higher $\lambda_1$. As illustrated above, the leakage of information increases $\lambda_1$. This reduces $\beta_1^B$ and thus makes the price signal about $\delta^B$ less precise. In addition, the price signal $S_2^{p_1} = \delta^B + \frac{\beta_1^A}{\beta_1^B}\varepsilon + \frac{1}{\beta_1^B}u_1$ is perturbed by the $\varepsilon$-error term. Therefore, information leakage makes the price $p_1$ after the public announcement less informative and less informationally efficient. The same is true after the second trading round for the price process $\{p_1, p_2\}$.

In summary, information leakage reduces informational efficiency at each point in time. It makes the price (process) more informative prior to the public announcement and less informative afterwards.

## 3.5 Extensions

### 3.5.1 Analysing Mixed Strategy Equilibria

The propositions in Section 3 showed that trader $A$'s informational advantage at the time of the public announcement as well as speculative and manipulative trading result from the imprecision of the rumour. This raises the question of whether the trader could generate some (additional) imprecision himself by trading above or below his optimal level in period one. Intuitively, the market maker will decrease the price, if the trader sells more today. Consequently, all other traders lower their evaluation of the stock tomorrow. This allows cheaper purchases tomorrow. Such manipulation was conducted by the Rothschild brothers during the Napoleonic wars. At the beginning of the 19th century, stock and bond prices in London depended crucially on news of the war. Despite knowing about Napoleon's fate at Waterloo, the Rothschild brothers sold English shares with the intent to drive prices down and repurchase cheaper shares later. Such a pure strategy cannot arise in equilibrium. In any Nash equilibrium, strategies of all players are mutual knowledge, i.e. all other traders would know the Rothschilds' true motivation for selling stocks.[9] It can, however, be the random realization of a mixed strategy.

[9] In a Kyle (1985) setting such behavior is not optimal.

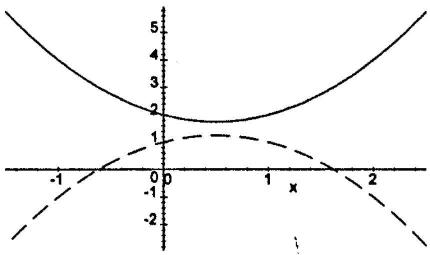
Let us focus on mixed (or behavioral) strategies for trader $A$ of the form $x_1^A = \beta_1^A(\delta^A + \varepsilon) + \gamma_1^A\zeta$, i.e. trader $A$ adds some noisy component $\gamma_1^A\zeta$ to his optimal demand.[10] In order to preserve normality of all random variables, assume $\zeta \sim \mathcal{N}(0,1)$. Adding random demand $\gamma_1^A\zeta$ in trading round one makes the market more liquid in $t=1$, but less liquid in $t=2$. This occurs because trader $A$ trades in $t=2$ on information generated by $\gamma_1^A\zeta$. The changes in the liquidity measure, $\lambda_t$ also alter the trading intensities, $\beta_t^i$. All this affects the new price signal $S_2^{p_1} = \delta^B + \frac{\beta_1^A}{\beta_1^B}\varepsilon + \frac{\gamma_1^A}{\beta_1^B}\zeta + \frac{1}{\beta_1^B}u_1$, which has the additional error term $\frac{\gamma_1^A}{\beta_1^B}\zeta$. This additional term is known to trader $A$, but not to the other market participants. Therefore, trader $A$'s informational advantage in $t=2$ consists of his knowledge of $\frac{\beta_1^A}{\beta_1^B}\varepsilon$ as well as of $\frac{\gamma_1^A}{\beta_1^B}\zeta$. The two error terms differ in two respects. First, whereas trader $A$ knows $\zeta$ already in $t=1$, he learns the precise value of $\varepsilon$ only at the time of the public announcement. Second, if trader $A$ wants to increase the importance of the error term $\frac{\beta_1^A}{\beta_1^B}\varepsilon$ by varying $\beta_1^A$, he must also trade more aggressively on his information in $t=1$. In contrast, trader $A$ can control the impact of the error term $\frac{\gamma_1^A}{\beta_1^B}\zeta$ on the price signal $S_2^{p_1}$ separately by adjusting $\gamma_1^A$. The trade-off is that while on the one hand he acts like a noise trader in $t=1$ incurring trading costs, on the other hand he also increases his informational advantage in $t=2$.

The analysis of the continuation game in $t=2$ is analogous to the one in Proposition 1. The only difference stems from the less informative price signal $S_2^{p_1}$. This alters the stock split and trader $B$'s forecast of traders 1's forecast. Formally, $\phi_{S_2^{p_1}}^{\delta^B} = Var[\delta^B](Var[\delta^B] + (\frac{\beta_1^A}{\beta_1^B})^2 Var[\varepsilon] + (\frac{\gamma_1^A}{\beta_1^B}) + (\frac{1}{\beta_1^B})Var[u_1])^{-1}$ and $\phi_w^{S_2^{A,w}} = \frac{(\beta_1^A)^2 Var[\varepsilon] + (\gamma_1^A)^2}{(\beta_1^A)^2 Var[\varepsilon] + (\gamma_1^A)^2 + Var[u_1]}$ change due to the additional $\gamma_1^A$-terms. This affects $\beta_2^i$, $\gamma_2^i$ and $\lambda_2$. In $t=1$ trader $A$ expects a larger informational advantage for the second trading round due to randomization, $E[S_2^{A,w}|S_1^A] = -(\phi_{S_2^{p_1}}^{\delta^B} + \frac{1}{\beta_1^B}\phi_{S_2^{p_1}}^{u_1})\frac{1}{\beta_1^B}[\beta_1^A\phi_{S_1^A}^{\varepsilon}S_1^A + \gamma_1^A\zeta]$. Trader $A$'s trading rule only exhibits the proposed form $x_1^A = \beta_1^A S_1^A + \gamma_1^A\zeta$, if $\lambda_1 = \psi^A = \lambda_2(\gamma_2^A)^2$. This implies $\beta_1^A = \frac{1}{2\lambda_1}$.

For a mixed strategy to sustain in equilibrium, trader $A$ has to be indifferent between

[10]Pagano and Röell (1993) conjecture a mixed strategy equilibrium in a model which analyzes front-running by brokers. Investors submit their orders to the broker who forwards it to the market maker. Prior to trading the broker observes the aggregate order flows for the next two trading rounds. Hence he has more information than the market maker in the first trading period. In the first trading round he front-runs by adding his own (possible random) orders.

any realised pure strategy, i.e. between any realisation of $\zeta$. Since the random variable $\zeta$ can lead to any demand with positive probability, he has to be indifferent between any $x_1^A$ in equilibrium. This requires that the marginal trading costs in $t = 1$ exactly offset the expected marginal gains in $t = 2$. Trader $A$'s objective function consists of two parts: the expected capital gains in $t = 1$ and the expected value function for $t = 2$. They are illustrated in Figure 3.2.




Components of Trader $A$'s Objective Function

Trader $A$ is only indifferent between all realizations of $\zeta$ if his $(x_1^{A,dA})^2[-\lambda_1 + \psi^A] + x_1^{A,dA}[\phi^{\delta A}_{S_1^A} S_1^A - \tau^A E[S_2^{A,w}|S_1^A] - 2\psi^A x_1^A] + C_1$ reduces to a constant, $C_1$. Trader $A$ faces no additional trading costs in $t = 1$ for the pure strategy given by the realization $\zeta = 0$ but he still has an informational advantage in $t = 2$. Therefore, the expected overall profits have to be strictly positive since he is indifferent between any realization of $\zeta = 0$. Note that even if trader $A$ receives no signal in $t = 1$ his informational advantage in $t = 2$ in the case of $\zeta = 0$ is $S_2^{A,w} = E[\delta^B|\delta^B + \frac{1}{\beta_1^B}u_1] - E[\delta^B|S_2^{p_1}] = (\phi^{\delta^B}_{\delta^B+\frac{1}{\beta_1^B}u_1} - \phi^{\delta^B}_{S_2^{p_1}})(\delta^B + \frac{1}{\beta_1^B}u_1)$. In summary, the necessary conditions for a mixed strategy equilibrium are that $\lambda_1 = \psi^A$and $\phi^{\delta A}_{S_1^A} S_1^A - \tau^A E[S_2^{A,w}|S_1^A] - 2\psi^A x_1^A = 0$. The second necessary condition simplifies to $1 - 2\lambda_2(\gamma_2^A)\beta_1^A = 0$ and is equivalent to the first one.

Proposition 5 exploits the facts that in any mixed strategy equilibrium the second order condition of trader $A$ is binding ($\lambda_1 = \psi^A$) and that the second order condition for trader $B$ ($\lambda_1 > \psi^B$ ) also has to be satisfied. The proposition shows that no mixed strategy equilibrium exists except if the market is very liquid in $t = 2$.

**Proposition 5** *There does not exist a mixed strategy equilibrium for sufficiently small* $Var[u_2]$.

See Appendix A.5 for a proof of this proposition. Note that the second order conditions also require that the trading round one is sufficiently liquid relative to trading round two, as stated in Proposition 1. Huddart, Hughes, and Levine (1998) analyse a mixed strategy equilibria in a setting where the insider trades are disclosed after each period.

### 3.5.2 Increasing the Number of Traders

In reality there are many informed traders active in the market. One question which might arise is whether the results of Section 3 also hold in a setting with many informed traders. Before increasing the number of traders let us investigate what distinguishes trader $A$ who received a signal about $\delta^A$ from trader $B$ who received a signal about $\delta^B$. In the setting described above, trader $A$'s prior knowledge about event A causes him to trade for speculative as well as manipulative reasons. However, trader $B$ does not act speculatively or manipulatively in $t = 2$ despite his prior knowledge of the forthcoming public announcement about event B in $t = 3$. Neither the timing per se nor the fact that trader $B$ got a precise signal about the public announcement in $t = 3$ can explain the difference. Trader $B$ still would not speculate or try to manipulate the price even if his signal is imprecise, i.e. $\delta^B + \varepsilon^B$. This is in spite of the fact that the imprecision of trader $A$'s signal is necessary for trader $A$'s behaviour. The distinctive feature is that when $\delta^A$ is publicly announced in $t = 2$, $p_1$ still carries some information for market participants, which induces them to conduct technical analysis. This, in turn, makes it worthwhile for trader $A$ to manipulate $p_1$. On the other hand, when $\delta^B$ is announced, neither $p_1$ nor $p_2$ carry any additional information. Since everybody knows the true value of the stock, $v = \delta^A + \delta^B$, nobody trades conditional on $p_2$. Thus trader $B$ has no incentive to manipulate the price in $t = 2$.

Having understood this crucial distinction let us first analyse the impact of increasing the number of traders who receive some information about $\delta^B$ in $t = 1$, and then increase the number of traders who can potentially act manipulatively in equilibrium. If there are many informed $B$-traders who receive different signals $\delta^B + \varepsilon^{Bi}$, $i \in \{1, ..., I\}$ they have an additional incentive to conduct technical analysis. In $t = 2$ they not only draw inference from price $p_1$ in order to improve their forecast of trader $A$'s forecast, they also try to learn more about the fundamental value $\delta^B$. They try to infer each others' signal from $p_1$ although they know that the past price $p_1$ is perturbed by the $\varepsilon$-error term. This makes manipulation of $p_1$ even more effective and consequently trader $A$ speculates and trades to manipulate the price.[11]

In the context of rumours, it may be hard to envision an information structure where only a single trader receives some vague information about a forthcoming public announcement. Instead, there could be many traders who receive some signal. In a setting in which all early-informed traders receive the same signal with a common noise component, $\delta^A + \xi$, manipulative trading and speculation still occur, but to a smaller extent. The reason is that all $A$-traders try to free ride on the manipulative activity of the other manipulators. Manipulation is costly but benefits all other $A$-traders in the second trading round. Such an information structure is, however, not very plausible since every recipient of a rumour can interpret it slightly differently. Even if all agree on the informational content of the rumour, they can still disagree on how it impacts the fundamental value of the stock. Therefore, the information structure that best fits the description of a rumour is one where many traders receive a signal $\delta^A + \xi + \varepsilon^{Ai}$ with a common and a private noise term. The private noise term $\varepsilon^i$ alleviates the free rider problem. On the other hand, as the number of traders who hear about the rumor increases, the importance of the $\varepsilon^i$-terms diminishes. In addition, $\varepsilon^i$ distorts trader $i$'s estimate of $\xi$. This discussion suggests that rumours lead to more manipulative and speculative trading as long as they are not widely spread among many traders.

## 3.6 Conclusion

The objective of this paper is to model how traders respond to a public announcement. Traders have to figure out how much of this public information is really new relative to the information already reflected in the price. A trader who receives an imprecise signal prior to the public announcement trades on it and moves the price. He has an additional informational advantage in the second trading round if the other market participants draw some inference from the past price. The early-informed trader can better interpret the past price. His trading also has a speculative feature. If he buys (sells) stocks when he receives

[11] Proof of these propositions in a two $\delta_B$-trader setting is available from the author on request.

his rumour, then he expects to sell (buy) it at the time when the news is made public. His strategy follows the well-known trading rule: "Buy on Rumours - Sell on News". By trading more aggressively prior to the public announcement, he can also affect the price and worsen the price signal for the other traders. This manipulative trading enhances his informational advantage at the time of the public announcement. Thus, his trading strategy is really to "Buy Aggressively on Rumours - Sell on News".

This paper adds to the literature by explaining the behaviour of an early-informed trader who trades for both manipulative and speculative reasons. It also provides an alternative, rational explanation for the overactivism of traders. By demonstrating how rumours can reduce the informational efficiency of markets, the paper also provides support for the use of early public announcements as a tool for crisis management.

Some further extensions come to mind. A higher order uncertainty model could be used to address the same questions. However, these models tend to be either very simplistic or very intractable. The same analysis could also be conducted in a different framework where the market maker sets bid and ask prices before the order of a trader arrives, e.g. a setting á la Glosten (1989). Preliminary analysis suggests that such a setting would yield similar outcomes. Additional insights could be obtained by endogenising the information acquisition process. For example, traders might like to commit themselves to purchase less precise signals. It would also be interesting to determine when it is more profitable to buy imprecise information about a forthcoming announcement and when it is more lucrative to acquire some long-lived information. The paper illustrates that information leakage reduces informational efficiency, but it does not make any normative welfare statements. In order to conduct a welfare analysis, one has to endogenise the trading activities of the liquidity traders. For example, one could consider risk-averse uninformed investors who are engaged in a private investment project. If the returns of these private investment projects are correlated with the value of stock, they trade for hedging reasons even though they face trading costs. A thorough welfare analysis would allow us to evaluate insider trading laws. But these are all tasks for the future.

## 3.7 Appendix of Chapter 3

### 3.7.1 Proof of Proposition 1

**Propose an arbitrary action rule profile for** $t = 1$, $\{\{x_1^i(S_1^i)\}_{i\in\{A,B\}}, p_1(X_1)\}$. This profile can be written as $\{\{\beta_1^i\}_{i\in\{A,B\}}, p_1(X_1)\}$ since I focus on linear pure strategy Perfect Bayesian Nash Equilibria. Suppose that this profile is mutual knowledge among the agents and they all think it is an equilibrium profile.

**Equilibrium in continuation game in** $t = 2$.

**Information structure in** $t = 2$.

After $\delta^A$ is publicly announced, $\delta^B$ is the only uncertain component of the stock's value.

The *market maker* knows the aggregate order flow in $t = 1$, $X_1 = \beta_1^A(\delta^A + \varepsilon) + \beta_1^B(\delta^B) + u_2$ in addition to $\delta^A$. His price signal $S_2^{p_1}$ (aggregate order flow signal, $X_1$) can be written as $S_2^{p_1} = \frac{X_1 - \beta_1^A \delta^A}{\beta_1^B} = \delta^B + \frac{\beta_1^A}{\beta_1^B}\varepsilon + \frac{1}{\beta_1^B}u_1$.

Since all market participants can invert the pricing function $p_1 = \lambda_1 X_1$ in $t = 2$, they all know $S_2^{p_1}$. For expositional clarity let us 'virtually' *split the stock* $v$ into a risk-free bond with payoff $\delta^A + E[\delta^B|S_2^{p_1}]$ and a risky asset $w$. In equilibrium $E[w|S_2^{p_1}] = 0$ and $Var[w|S_2^{p_1}] = \left(1 - \phi_{S_2^{p_1}}^{\delta^B}\right)Var[\delta^B]$. The 'virtual' split of the stock $v$ into a risk-free bond and a risky asset $w$ is possible, without loss of generality, as long as all traders are risk neutral or have CARA utility functions.

*Trader A* can infer $\varepsilon$ in $t = 2$ and thus his price signal is $\delta^B + \frac{1}{\beta_1^B}u_1$. After orthogonalizing it to $S_2^{p_1}$, his signal can be written as $S_2^{A,w} := w + \frac{1}{\beta_1^B}\vartheta_1$, where $w = \delta^B - E[\delta^B|S_2^{p_1}]$ and $\vartheta_1 = u_1 - E[u_1|S_2^{p_1}]$. Trader $A$'s forecasts of the fundamental value of $w$ is $E[w|S_2^{A,w}] = \phi_{S_2^{A,w}}^w S_2^{A,w}$, $\phi_{S_2^{A,w}}^w = \frac{Var[\delta^B]}{Var[\delta^B] + \frac{1}{(\beta_1^B)^2}Var[u_1]} = \phi_{\delta^B + \frac{1}{\beta_1^B}u_1}^{\delta^B}$. Trader $A$'s forecast of trader $B$'s forecast is also $E[w|S_2^{A,w}]$.

*Trader B* knows the fundamental value $w$. His forecast of trader $A$'s forecast is $E[w + \frac{1}{\beta_1^B}\vartheta_1|w, S_2^{p_1}] = E[w + \frac{1}{\beta_1^B}\vartheta_1|w] = \phi_w^{S_2^{A,w}} w$, where $\phi_w^{S_2^{A,w}} = \frac{(\beta_1^A)^2 Var[\varepsilon]}{(\beta_1^A)^2 Var[\varepsilon] + Var[u_1]}$.

**Action (trading) rules in** $t = 2$.

Due to potential Bertrand competition the risk-neutral *market maker* sets the price $p_2 = E[v|X_1, X_2] = \delta^A + E[\delta^B|S_2^{p_1}] + \lambda_2 X_2$. The first two terms reflect the value of the bond from the stock split and the last term $\lambda_2 X_2 =: p_2^w$ is the price for $w$. Note that $\lambda_2 = \frac{Cov[w, X_2|S_2^{p1}]}{Var[X_2|S_2^{p1}]}$.

*Trader A*'s optimisation problem in $t = 2$ is $\max_{x_2^A} x_2^A E[w - p_2^w|S_2^{A,w}]$. The first order condition of $\max_{x_2^A} x_2^A E[w - \lambda_2 \left(x_2^A + \beta_2^B S_2^{B,w} + u_2\right) p_2^w|S_2^{A,w}]$ leads to $x_2^{A,*} = \beta_2^A S_2^{A,w}$, where $\beta_2^A = \frac{1}{2\lambda_2}\left(1 - \lambda_2 \beta_2^B\right)\phi^w_{S_2^{A,w}}$.

*Trader B*'s optimisation problem is $\max_{x_2^A} x_2^A E[w - p_2^w|S_2^{B,w}]$. The first order condition translates into $x_2^{B,*} = \beta_2^B S_2^{B,w}$, where $\beta_2^B = \frac{1}{2\lambda_2}\left(1 - \lambda_2 \beta_2^A \phi_w^{S_2^{A,w}}\right)$. The second order condition for both traders' maximisation problem is $\lambda_2 > 0$.

The *equilibrium* for a given action (trading) rule profile in $t = 1$ is given by

$$\beta_2^A = \frac{1}{2\lambda_2}\frac{1}{2}\frac{1}{1 - \frac{1}{4}\phi_w^{S_2^{A,w}}\phi^w_{S_2^{A,w}}}\phi^w_{S_2^{A,w}},\ \beta_2^B = \frac{1}{2\lambda_2}\frac{1 - \frac{1}{2}\phi_w^{S_2^{A,w}}\phi^w_{S_2^{A,w}}}{1 - \frac{1}{4}\phi_w^{S_2^{A,w}}\phi^w_{S_2^{A,w}}},$$

$$\lambda_2 = \Big\{\frac{1}{2}\frac{Var[\delta^B|S_2^{p1}]}{Var[u_2]}\left(\left(b_2^A + b_2^B\right) - \frac{1}{2}\left(b_2^A + b_2^B\right)^2\right) + \frac{1}{2}\frac{Cov[\delta^B, u_1|S_2^{p1}]}{Var[u_2]}\left(\frac{b_2^A}{\beta_1^B}\left(1 - \left(b_2^A + b_2^B\right)\right)\right)$$
$$- \frac{1}{4}\frac{Var[u_1|S_2^{p1}]}{Var[u_2]}\left(\frac{b_2^A}{\beta_1^B}\right)\Big\}^{\frac{1}{2}},$$

where $b_2^i := 2\lambda_2 \beta_2^i$. $b_2^i$ depends only on the regression coefficients $\phi_w^{S_2^{A,w}}$ and $\phi^w_{S_2^{A,w}}$ which are determined by the proposed action rule profile in $t = 1$.

### Equilibrium in $t = 1$.

The proposed arbitrary action rule profile is an equilibrium if no player wants to deviate given the strategies of the others.

The *market maker*'s pricing rule in $t = 1$ is always given by $p_1 = E[v|X_1] = \lambda_1 X_1$ with $\lambda_1 = \frac{Cov[v, X_1]}{Var[X_1]}$. He has to set an informationally efficient price due to (potential) Bertrand competition.

#### Trader A's best response.

Deviation of trader $A$ from $x_1^A(S_1^A) = \beta_1^A S_1^A$ to $x_1^{A,dA}(S_1^A)$ will not alter the subsequent trading intensities of the other market participants, i.e. $\lambda_1, \beta_2^B, \lambda_2$. They still believe that trader $A$ plays his equilibrium strategy since they cannot detect his deviation. Nor does his deviation change his own price signals since he knows the distortion his deviation causes. The definition of $w$ is also not affected by this deviation.

*Other market participants' misperception in $t = 2$.*

Trader $A$'s deviation, however, distorts the other players price signal, $S_2^{p1}$ to $S_2^{p1,dA}$. This occurs because the other market participants attribute the difference in the aggregate order flow in $t = 1$ not to trader $A$'s deviation, but to a different signal realisation or different noise trading. Deviation to $x_1^{A,dA}(\cdot)$ distorts the price signal by $S_2^{p_1,dA} - S_2^{p_1} = \frac{1}{\beta_1^B}\left(x_1^{A,dA} - x_1^A\right)$. Trader $B$'s signal prior to trading in $t = 2$ is not $w$ but $w - \phi_{S^{p_1}}^{\delta^B}\left(S_2^{p_1,dA} - S_2^{p_1}\right)$. His market order in $t = 2$ is, therefore, $\beta_2^B w - \beta_2^B \phi_{S^{p_1}}^{\delta^B}\frac{1}{\beta_1^B}\left(x_1^{A,dA} - x_1^A\right)$. Price $p_2$ is also distorted. The market maker's best estimate of $w$ prior to trading in $t = 2$ is $\phi_{S^{p_1}}^{\delta^B}\left(S_2^{p_1,dA} - S_2^{p_1}\right)$ and after observing $X_2^{dA}$, $p_2^{w,dA} = \phi_{S^{p_1}}^{\delta^B}\frac{1}{\beta_1^B}\left(x_1^{A,dA} - x_1^A\right) + \lambda_2\left(x_2^{A,dA} + \beta_2^B w - \beta_2^B\phi_{S^{p_1}}^{\delta^B}\frac{1}{\beta_1^B}\left(x_1^{A,dA} - x_1^A\right) + u_2\right)$. Since $\beta_2^A = \frac{1}{2\lambda_2}\left(1 - \lambda_2\beta_2^B\right)\phi_{S_2^{A,w}}^w$, $p_2^{w,dA} = \lambda_2\left(x_2^{A,dA} + \beta_2^B w + u_2\right) + 2\lambda_2\frac{\beta_2^A}{\beta_1^B}\frac{\phi^{\delta^B}_{S_2^{p_1}}}{\phi^w_{S_2^{A,w}}}\left(x_1^{A,dA} - x_1^A\right)$.

*Trader A's optimal trading rule in $t = 2$ after deviation in $t = 1$* results from the adjusted maximisation problem $\max_{x_2^{A,dA}} E[x_2^{A,dA}\left(w - p_2^{w,dA}\right)|S_2^{A,w}]$. It is given by $x_2^{A,dA,*} = \beta_2^A S_2^{A,w} - \gamma_2^A\left(x_1^{A,dA} - x_1^A\right)$, where $\gamma_2^A := \frac{\beta_2^A}{\beta_1^B}\frac{\phi^{\delta^B}_{S_2^{p_1}}}{\phi^w_{S_2^{A,w}}}$, if the second order condition $\lambda_2 > 0$ is satisfied.

*Trader A's value function* $V_2^A(x_1^{A,dA}) = x_2^{A,dA,*}E[w - p_2^w|S_2^{A,w}]$. After replacing $x_2^{A,dA,*}$ with $\beta_2^A S_2^{A,w} - \gamma_2^A\left(x_1^{A,dA} - x_1^A\right)$ and noting that $(1 - \lambda_2\beta_2^B) = 2\lambda_2\beta_2^A$ it simplifies to $V_2^A(x_1^{A,dA}) = \psi^A\left(x_1^{A,dA} - x_1^A\right)^2 - \tau^A S_2^{A,w}\left(x_1^{A,dA} - x_1^A\right) + \kappa^A\left(S_2^{A,w}\right)^2$, with $\psi^A = \lambda_2\left(\gamma_2^A\right)^2$, $\tau^A = 2\lambda_2\beta_2^A\gamma_2^A$, $\kappa^A = \lambda_2\left(\beta_2^A\right)^2$. In $t = 1$, trader $A$ forms expectations $E[V_2^A(x_2^{A,dA})|\delta^B]$ of the value function in $t = 2$. $S_2^{A,w}$ is random in $t = 1$.
$E[S_2^{A,w}|S_1^A] = -\left(\phi^{\delta^B}_{S_2^{p_1}} + \frac{1}{\beta_1^B}\phi^{u_1}_{S^{p_1}}\right)\frac{\beta_1^A}{\beta_1^B}E[\varepsilon|S_1^A] = -\frac{\beta_1^A}{\beta_2^A}\gamma_2^A E[\varepsilon|S_1^A] = -\frac{1}{\beta_2^A}\gamma_2^A\left(1 - \phi^{\delta^A}_{S_1^A}\right)x_1^A$ since $\phi^{\delta^B}_{S^{p_1}} + \frac{1}{\beta_1^B}\phi^{u_1}_{S^{p_1}} = \frac{\phi^{\delta^B}_{S_2^{p_1}}}{\phi^w_{S_2^{A,w}}}$.

*Trader A's optimization problem in $t = 1$* is thus
$\max_{x_1^{A,dA}} E[x_1^{A,dA}\left(v - p_1^{dA}\right) + V_2^A(x_1^{A,dA})|S_1^A]$, where $p_1^{dA} = \lambda_1 X_1^{dA} = \lambda_1\left(x_1^{A,dA} + \beta_1^B S_1^B + u_1\right)$. Since $S_1^A$ is orthogonal to $S_1^B$ the first order condition is $E[\delta^A|S_1^A] - 2\lambda_1 x_1^{A,dA} + 2\psi^A\left(x_1^{A,dA} - x_1^A\right) - \tau^A E[S_2^{A,w}|S_1^A] = 0$. Therefore, $x_1^{A,dA,*} = \frac{1}{2\left(\lambda_1 - \lambda_2\left(\gamma_2^A\right)^2\right)}(\phi^{\delta^A}_{S_1^A} + 2\lambda_2\left(\gamma_2^A\right)^2\beta_1^A)S_1^A$. The second order condition is $\lambda_1 > \lambda_2\left(\gamma_2^A\right)^2$. In Equilibrium $\beta_1^A = \frac{1}{2\left(\lambda_1 - \lambda_2\left(\gamma_2^A\right)^2\phi^{\varepsilon}_{S_1^A}\right)}\phi^{\delta^A}_{S_1^A}$.

**Trader $B$'s best response.**

*Other market participants' misperception in $t = 2$.*

Deviation from $x_1^B(S_1^B) = \beta_1^B S_1^B$ to $x_1^{B,dB}(S_1^B)$ distorts the price signal by $S_2^{p_1,dB} - S_2^{p_1} = \frac{1}{\beta_1^B}\left(x_1^{B,dB} - x_1^B\right)$. Trader $A$'s signal prior to trading in $t = 2$ is not $w + \frac{1}{\beta_1^B}\vartheta_1$ but

$w - \phi^{\delta^B}_{S^{p_1}}\left(S^{p_1,dB}_2 - S^{p_1}\right) + \frac{1}{\beta^B_1}\left(\vartheta_1 - \phi^{u_1}_{S^{p_1}}\left(S^{p_1,dB}_2 - S^{p_1}\right)\right)$. His market order is, therefore, $x^{A,dB}_2 = \beta^A_2\left(w + \frac{1}{\beta^B_1}\vartheta_1\right) - \beta^A_2\left(\phi^{\delta^B}_{S^{p_1}} + \frac{1}{\beta^B_1}\phi^{u_1}_{S^{p_1}}\right)\frac{1}{\beta^B_1}\left(x^{B,dB}_1 - x^B_1\right)$. Price $p_2$ is also distorted. The market maker's best estimate of $w$ prior to trading in $t = 2$ is $\phi^{\delta^B}_{S^{p_1}}\left(S^{p_1,dB}_2 - S^{p_1}_2\right)$ and after observing $X^{dB}_2$, $p^{w,dB}_2 = \phi^{\delta^B}_{S^{p_1}}\left(S^{p_1,dB}_2 - S^{p_1}_2\right) + \lambda_2\left(x^{A,dB}_2 + x^{B,dB}_2 + u_2\right)$. Let $\gamma^B_2 := \frac{1}{2\lambda_2}[\phi^{\delta^B}_{S^{p_1}} - \lambda_2\beta^A_2\left(\phi^{\delta^B}_{S^{p_1}_2} + \frac{1}{\beta^B_1}\phi^{u_1}_{S^{p_1}_2}\right)]\frac{1}{\beta^B_1}$ then $p^{w,dB}_2 = \lambda_2\left(\beta^A_2\left(w + \frac{1}{\beta^B_1}\vartheta_1\right) + x^{B,dB}_2 + u_2\right) + 2\lambda_2\gamma^B_2\left(x^{B,dB}_1 - x^B_1\right)$.

*Trader B's optimal trading rule in $t = 2$ after deviation in $t = 1$* is the result of $\max_{x^{B,dB}_2} E[x^{B,dB}_2\left(w - p^{w,dB}_2\right)|S^{B,w}_2]$. The optimal order size in $t = 2$ is $x^{B,dB,*}_2 = \beta^B_2 S^{B,w}_2 - \gamma^B_2\left(x^{B,dB}_1 - x^B_1\right)$, if the second order condition $\lambda_2 > 0$ is satisfied. Note that if we replace $\beta^A_2$ with $\frac{1}{2\lambda_2}\left(1 - \lambda_2\beta^B_2\right)\phi^w_{S^{A,w}_2}$, $\gamma^B_2$ simplifies to $\frac{1}{2}\frac{1}{\beta^B_1}\phi^{\delta^B}_{S^{p_1}_2}\frac{3 - \phi_w^{S^{A,w}_2}\phi^w_{S^{A,w}_2}}{4 - \phi_w^{S^{A,w}_2}\phi^w_{S^{A,w}_2}}$.

*Trader B's value function* $V^B_2(x^{B,dB}_1) = x^{B,dB,*}_2 E[w - p^w_2|S^{B,w}_2]$. After replacing $x^{B,dB,*}_2$ with $\beta^B_2 S^{B,w}_2 - \gamma^B_2\left(x^{B,dB}_1 - x^B_1\right)$ and noting that $(1 - \lambda_2\beta^A_2\phi_w^{S^{A,w}_2}) = 2\lambda_2\beta^B_2$ it simplifies to $V^B_2(x^{B,dB}_1) = \psi^B\left(x^{B,dB}_1 - x^B_1\right)^2 - \tau^B S^{B,w}_2\left(x^{B,dB}_1 - x^B_1\right) + \kappa^B\left(S^{B,w}_2\right)^2$, with $\psi^B = \lambda_2\left(\gamma^B_2\right)^2$, $\tau^B = 2\lambda_2\beta^B_2\gamma^B_2$, $\kappa^B = \lambda_2\left(\beta^B_2\right)^2$. In $t = 1$, trader $B$ forms expectations $E[V^B_2(x^{B,dB}_2)|\delta^B]$ of the value function in $t = 2$. $S^{B,w}_2 = w$ is random in $t = 1$. The expectation of $S^{B,w}_2$ is given by $E[S^{B,w}_2|S^B_1] = E[w|\delta^B] = \left(1 - \phi^{\delta^B}_{S^{p_1}_2}\right)\delta^B$.

*Trader B's optimisation problem in $t = 1$* is thus $\max_{x^{B,dB}_1} E[x^{B,dB}_1\left(v - p^{dB}_1\right) + V^B_2(x^{B,dB}_1)|S^B_1]$, where $p^{dB}_1 = \lambda_1 X^{dB}_1 = \lambda_1\left(\beta^A_1 S^A_1 + x^{B,dB}_1 + u_1\right)$. Since $S^B_1$ is orthogonal to $S^A_1$ the first order condition reduces to $x^{B,dB,*}_1 = \frac{1}{2\left(\lambda_1 - \lambda_2\left(\gamma^B_2\right)^2\right)}\left(\left(1 - 2\lambda_2\beta^B_2\gamma^B_2\left(1 - \phi^{\delta^B}_{S^{p_1}_2}\right)\right) - 2\lambda_2\left(\gamma^B_2\right)^2\beta^B_1\right)S^B_1$. The second order condition is $\lambda_1 > \lambda_2\left(\gamma^B_2\right)^2$.

**Perfect Bayesian Nash Equilibrium** is given by a fixed point in $(\beta^{A,*}_1, \beta^{B,*}_1)$.

$$\beta^{A,*}_1 = \frac{1}{2\left(\lambda_1 - \lambda_2\left(\gamma^A_2\right)^2\phi^{\epsilon}_{S^A_1}\right)}\phi^{\delta^A}_{S^A_1}$$

$$\beta^{B,*}_1 = \frac{1}{2\lambda_1}\left(1 - 2\lambda_2\beta^B_2\gamma^B_2\left(1 - \phi^{\delta^B}_{S^{p_1}_2}\right)\right),$$

where

$$\lambda_1 = \frac{\beta^{A,*}_1 Var[\delta^A] + \beta^{B,*}_1 Var[\delta^B]}{Var[\beta^{A,*}_1\left(\delta^A + \epsilon\right) + \beta^{B,*}_1\left(\delta^B\right) + u_1]}$$

with

$$\beta_2^A = \frac{1}{2\lambda_2}\frac{1}{2}\frac{1}{1-\frac{1}{4}\phi_w^{S_2^{A,w}2}\phi^w_{S_2^{A,w}}}\phi^w_{S_2^{A,w}} \quad \gamma_2^A := \frac{\beta_2^A}{\beta_1^B}\frac{\phi^{\delta^B}_{S_2^{p_1}}}{\phi^w_{S_2^{A,w}}}$$

$$\beta_2^B = \frac{1}{2\lambda_2}\frac{1-\frac{1}{2}\phi_w^{S_2^{A,w}2}\phi^w_{S_2^{A,w}}}{1-\frac{1}{4}\phi_w^{S_2^{A,w}2}\phi^w_{S_2^{A,w}}} \qquad \gamma_2^B := \frac{1}{2\lambda_2}\frac{1}{\beta_1^B}\phi^{\delta^B}_{S_2^{p_1}} - \frac{1}{2}\frac{\beta_2^A}{\beta_1^B}\frac{\phi^{\delta^B}_{S_2^{p_1}}}{\phi^w_{S_2^{A,w}}}$$

$$\lambda_2 = \{\frac{1}{2}\frac{Var[\delta^B|S_2^{p_1}]}{Var[u_2]}\left(\left(b_2^A + b_2^B\right) - \frac{1}{2}\left(b_2^A + b_2^B\right)^2\right) + \frac{1}{2}\frac{Cov[\delta^B, u_1|S_2^{p_1}]}{Var[u_2]}\left(\frac{b_2^A}{\beta_1^B}\left(1 - \left(b_2^A + b_2^B\right)\right)\right)$$
$$-\frac{1}{4}\frac{Var[u_1|S_2^{p_1}]}{Var[u_2]}\left(\frac{b_2^A}{\beta_1^B}\right)\}^{\frac{1}{2}},$$

where $b_2^i := 2\lambda_2\beta_2^i \; i \in \{A, B\}$ if the second order conditions

$\lambda_2 > \lambda_1 \max\{\left[\frac{b_2^A}{b_1^B}\frac{\phi^{\delta^B}_{S_2^{p_1}}}{\phi^w_{S_2^{A,w}}}\right]^2, \left[\frac{1}{b_1^B}\phi^{\delta^B}_{S_2^{p_1}} - \frac{1}{2}\frac{b_2^A}{b_1^B}\frac{\phi^{\delta^B}_{S_2^{p_1}}}{\phi^w_{S_2^{A,w}}}\right]^2\}$, $\lambda_2 > 0$ are satisfied. ■

### 3.7.2 Proof of Proposition 2

For the market maker as well as for trader $B$ the $p_1$-price signal is $S_2^{p_1} = \delta^B + \frac{\beta_1^A}{\beta_1^B}\varepsilon + \frac{1}{\beta_1^B}u_1$. Trader $A$ can infer $\varepsilon$ and thus his price signal is more precise. Trader $A$'s informational advantage $\frac{\beta_1^A}{\beta_1^B}\varepsilon$ increases in $\beta_1^A$ and decreases in $\beta_1^B$. ■

### 3.7.3 Proof of Proposition 3

#### Speculative Trading

Trader $A$ expects to trade $\beta_2^A E[S_2^{A,w}|S_1^A]$ in $t = 2$.
Since $E[S_2^{A,w}|S_1^A] = -\left(\phi^{\delta^B}_{S_2^{p_1}} + \frac{1}{\beta_1^B}\phi^{u_1}_{S_2^{p_1}}\right)\frac{\beta_1^A}{\beta_1^B}\phi^{\varepsilon}_{S_1^A}S_1^A = -\frac{\beta_1^A}{\beta_2^A}\gamma_2^A\phi^{\varepsilon}_{S_1^A}S_1^A$ and $\beta_1^A, \beta_2^A > 0$, trader $A$ expects to sell (buy) stocks in $t = 2$ if he buys (sells) stocks in $t = 1$.

#### Manipulative Trading

Trader $A$ trades excessively for manipulative reasons if $\beta_1^A > \beta_1^{A,\text{myopic}}$ (given the strategies of the other market participants).
$\beta_1^A = \frac{1}{2\left(\lambda_1 - \lambda_2\left(\gamma_2^A\right)^2\phi^{\varepsilon}_{S_1^A}\right)}\phi^{\delta^A}_{S_1^A}$ whereas $\beta_1^{A,\text{myopic}} = \frac{1}{2\lambda_1}\phi^{\delta^A}_{S_1^A}$. Thus manipulative trading is given by $\frac{\lambda_2\left(\gamma_2^A\right)^2\phi^{\varepsilon}_{S_1^A}}{2\lambda_1\left(\lambda_1 - \lambda_2\left(\gamma_2^A\right)^2\phi^{\varepsilon}_{S_1^A}\right)}\phi^{\delta^A}_{S_1^A}S_1^A$. The second order condition requires that $\lambda_1 > \lambda_2\left(\gamma_2^A\right)^2\phi^{\varepsilon}_{S_1^A}$.

Note that for $Var[\varepsilon] = 0$, $\phi^{\varepsilon}_{S_1^A} = 0$ neither speculative nor manipulative trading will occur. ■

### 3.7.4 Proof of Proposition 4

This proposition compares two different equilibria: one with information leakage and one without. Let us denote all variables of the former equilibrium with upper bars and all variables of the latter equilibrium with tilde.

If Proposition 4 holds for $Var[\varepsilon] = 0$ with strict inequalities it also holds for positive $Var[\varepsilon]$ in the environment around $Var[\varepsilon] = 0$.

As $Var[\varepsilon] \to 0$ the equilibrium strategies converge continuously to $\overline{\beta}_2^A \to 0$, $\overline{\beta}_2^B \to \frac{1}{2\lambda_2}$, $\overline{\phi}_w^{S_2^{A,w}} \to 0$, $\overline{\phi}^{\varepsilon}_{S_1^A w} \to 0$, $\overline{\gamma}_2^1 \to \frac{1}{2\lambda_2}\frac{1}{2\overline{\beta}_1^B}\overline{\phi}^{\delta^B}_{S_2^{p_1}}$, and thus $\overline{\beta}_1^A \to \frac{1}{2\lambda_1}$ ($\overline{b}_1^A \to 1$). For any given $\lambda_1$, $\lambda_2$ simplifies to $\frac{1}{2}\sqrt{\frac{(1-\phi^{\delta^B}_{S_2^{p_1}})Var[\delta^B]}{Var[u_2]}}$ and trader $B$'s trading intensity is $\beta_1^B = \frac{1}{2\lambda_1} - \frac{1}{\beta_1^B}\phi^{\delta^B}_{S_2^{p_1}}\frac{3}{4}\sqrt{(1-\phi^{\delta^B}_{S_2^{p_1}})\frac{Var[u_2]}{Var[\delta^B]}}$. That is, the corresponding $\lambda_1$ for a given $\beta_1^B$ is given by $\lambda_1 = \frac{1/2}{\beta_1^B + \frac{3}{4}\frac{1}{\beta_1^B}\phi^{\delta^B}_{S_2^{p_1}}\sqrt{(1-\phi^{\delta^B}_{S_2^{p_1}})\frac{Var[u_2]}{Var[\delta^B]}}}$.

**Prior to public announcement**

*Informativeness*

$\overline{p}_1$ is more informative than $\widetilde{p}_1$, i.e.

$Var[\delta^A + \delta^B | \overline{\beta}_1^A(\delta^A+\varepsilon) + \overline{\beta}_1^B\delta^B + u_1] < Var[\delta^B|\widetilde{\beta}_1^B\delta^B + u_1] + Var[\delta^A]$

$Var[\delta^A + \delta^B] - \overline{\lambda}_1\overline{\beta}_1^A Var[\delta^A] - \lambda_1\overline{\beta}_1^B Var[\delta^B] < Var[\delta^A+\delta^B] - \widetilde{\lambda}_1\widetilde{\beta}_1^B Var[\delta^B]$

$-\frac{1}{2}\overline{b}_1^A Var[\delta^A] - \frac{1}{2}\overline{b}_1^B Var[\delta^B] < -\frac{1}{2}\widetilde{b}_1^B Var[\delta^B]$

For $Var[\delta^A] = Var[\delta^B]$ and $Var[\varepsilon] \to 0$, $\overline{b}_1^A \to 1$ the inequality above simplifies to $1 > \widetilde{b}_1^B - \overline{b}_1^B$. Since $\overline{b}_1^B, \widetilde{b}_1^B \in ]0,1[$ this is always satisfied.

*Informational Efficiency*

$\overline{p}_1$ is less informationally efficient than $\widetilde{p}_1$, i.e.

$Var[\phi^{\delta^A}_{S_1^A}(\delta^A+\varepsilon) + \delta^B | \overline{\beta}_1^A(\delta^A+\varepsilon) + \overline{\beta}_1^B\delta^B + u_1] > Var[\delta^B|\widetilde{\beta}_1^B\delta^B + u_1]$

Since $Cov[\phi^{\delta^A}_{S_1^A}(\delta^A+\varepsilon) + \delta^B, X_1] = Cov[v, X_1] = \lambda_1 Var[X_1]$

$\phi^{\delta^A}_{S_1^A}(Var[\delta^A+\varepsilon]) - \overline{\lambda}_1\overline{\beta}_1^A V[\delta^A+\varepsilon] - \overline{\lambda}_1\overline{\beta}_1^B Var[\delta^B] > -\widetilde{\lambda}_1\widetilde{\beta}_2^B Var[\delta^B]$

$Var[\delta^A] - \frac{1}{2}b_1^A Var[\delta^A+\varepsilon] > \frac{1}{2}(\overline{b}_1^B - \widetilde{b}_1^B)Var[\delta^B]$. For $Var[\delta^A] = Var[\delta^B]$ and $Var[\varepsilon] \to 0$ the inequality to $1 > \overline{b}_1^B - \widetilde{b}_1^B$. This is always true.

**Prior to trading in $t = 2$**

If $\widetilde{\beta}_1^B > \overline{\beta}_1^B$ then $\widetilde{S}^{p_1} = \delta^B + \frac{1}{\widetilde{\beta}_1^B}u_1$ is more informative than $\overline{S}^{p_1} = \delta^B + \frac{1}{\overline{\beta}_1^B}u_1$, even if $Var[\varepsilon] = 0$.

In the $(\beta_1^B, \lambda_1)$-space the equilibrium is determined by the intersection of $\lambda_1 = \frac{1/2}{\beta_1^B + \frac{3}{4}\frac{1}{\beta_1^B}\phi_{S_2^{p_1}}^{\delta B}\sqrt{(1-\phi_{S_2^{p_1}}^{\delta B})\frac{Var[u_2]}{Var[\delta^B]}}}$ (1) with $\widetilde{\lambda}_1 = \frac{\widetilde{\beta}_1^B Var[\delta^B]}{\left(\widetilde{\beta}_1^B\right)^2 Var[\delta^B]+Var[u_1]}$ (2) in the case where no information leaks and with $\overline{\lambda}_1 = \frac{\frac{1}{2\lambda_1}Var[\delta^A]+\overline{\beta}_1^B Var[\delta^B]}{\left(\frac{1}{2\lambda_1}\right)^2+\left(\overline{\beta}_1^B\right)^2 Var[\delta^B]+Var[u_1]}$ (3) in the case of information leakage. (3) can be simplified to $\overline{\lambda}_1 = \frac{\overline{\beta}_1^B Var[\delta^B]+\sqrt{\left(\overline{\beta}_1^B\right)^2 Var[\delta^B]^2+\left(\overline{\beta}_1^B\right)^2 Var[\delta^B]Var[\delta^A]+Var[u_1]Var[\delta^A]}}{2\{\left(\overline{\beta}_1^B\right)^2 Var[\delta^B]+Var[u_1]\}}$. Note that we can restrict our attention to the positive root only because of the second order condition.

*Claim 1:* $\overline{\lambda}_1(\overline{\beta}_1^B) > \widetilde{\lambda}_1(\widetilde{\beta}_1^B)$ for all $\overline{\beta}_1^B = \widetilde{\beta}_1^B$ follows immediately

*Claim 2:* $\lambda_1(\beta_1^B) = \frac{1/2}{\beta_1^B + \frac{3}{4}\frac{1}{\beta_1^B}\phi_{S_2^{p_1}}^{\delta B}\sqrt{(1-\phi_{S_2^{p_1}}^{\delta B})\frac{Var[u_2]}{Var[\delta^B]}}}$ (1) is strictly decreasing in $\beta_1^B$ as long as $Var[u_1] > \frac{6}{25}\sqrt{\frac{2}{5}Var[\delta^B]Var[u_2]}$.

Its derivative is negative if the denominators' derivative is positive. The denominator can be rewritten as

$\beta_1^B + \frac{3}{4}\left(\beta_1^B\right)^{-2}\left(Var[\delta^B] + \left(\beta_1^B\right)^{-2}Var[u_1]\right)^{-1.5}\left(Var[\delta^B]Var[u_1]Var[u_2]\right)^{0.5}$. Its derivative w.r.t. $\beta_1^B$ is $1+\frac{3}{4}\frac{\sqrt{Var[\delta^B]Var[u_2]}}{Var[u_1]}\left(1-\phi_{S_2^{p_1}}^{\delta B}\right)^{1.5}\left(2-3\left(1-\phi_{S_2^{p_1}}^{\delta B}\right)\right)$. The global minimum for $\left(1-\phi_{S_2^{p_1}}^{\delta B}\right)^{1.5}\left(2-3\left(1-\phi_{S_2^{p_1}}^{\delta B}\right)\right)$ at $\phi_{S_2^{p_1}}^{\delta B} = \frac{2}{5}$ is $-\frac{2}{5}\sqrt{\frac{2}{5}}\frac{4}{5}$. From this it follows immediately that for $Var[u_1] > \frac{6}{25}\sqrt{\frac{2}{5}Var[\delta^B]Var[u_2]}$, $\frac{\partial\lambda_1}{\partial\beta_2^A} < 0$.

*Claim 3:* $\widetilde{\lambda}_1(\widetilde{\beta}_1^B)$ is weakly increasing in $\widetilde{\beta}_1^B$, i.e. $\frac{\partial\widetilde{\lambda}_1}{\partial\widetilde{\beta}_1^B} \geq 0$.

$\frac{\partial\widetilde{\lambda}_1}{\partial\widetilde{\beta}_1^B} = \frac{Var[\delta^B]\left(Var[u_1]-\left(\widetilde{\beta}_1^B\right)^2 Var[\delta^B]\right)}{\left(\left(\widetilde{\beta}_1^B\right)^2 Var[\delta^B]+Var[u_1]\right)^2}$. $\frac{\partial\widetilde{\lambda}_1}{\partial\widetilde{\beta}_1^B} \geq 0$ if $Var[u_1] \geq \left(\widetilde{\beta}_1^B\right)^2 Var[\delta^B]$. Replacing $\widetilde{\beta}_1^B$ with $\frac{1}{2}\sqrt{\frac{Var[u_1]}{Var[\delta^B]\frac{1}{2}\widetilde{b}_1^B\left(1-\frac{1}{2}\widetilde{b}_1^B\right)}}\widetilde{b}_1^B$ the condition simplifies to $\widetilde{b}_2^A$ □ 1. This is always the case in equilibrium.

From Claim 1 to 3 it follows that $\overline{\lambda}_1 > \widetilde{\lambda}_1$ and $\overline{\beta}_1^B < \widetilde{\beta}_1^B$ in the corresponding equilibria.

**After trading in $t = 2$**

The continuation game in $t = 2$ corresponds to a static Kyle (1985) model with a risky asset $w$. Since for lower $\beta_1^B$, the variance $Var[w] = Var[\delta^B|S_2^{p_1}]$ is higher, the price process $\{p_1, p_2\}$ reveals less information. ■

### 3.7.5 Proof of Proposition 5

In any mixed strategy equilibrium player 1 has to be indifferent between any $x_1^A$, i.e. $\lambda_1 = \lambda_2 \left(\gamma_2^A\right)^2$. In addition the second order condition of trader $B$, $\lambda_1 \geq \lambda_2 \left(\gamma_2^B\right)^2$ must hold. Thus a necessary condition for a mixed strategy equilibrium is

$$\gamma_2^B \leq \gamma_2^A$$

$$\frac{1}{2}\frac{1}{\beta_1^B}\phi_{S_2^{p1}}^{\delta^B}\frac{3-\phi^w_{S_2^{A,w}}\phi_w^{S_2^{A,w}}}{4-\phi^w_{S_2^{A,w}}\phi_w^{S_2^{A,w}}} \leq \frac{1}{\beta_1^B}\phi_{S_2^{p1}}^{\delta^B}\frac{1}{2\lambda_2}\frac{\frac{1}{2}}{1-\frac{1}{4}\phi^w_{S_2^{A,w}}\phi_w^{S_2^{A,w}}}$$

$$\lambda_2 \leq \frac{2}{3-\phi^w_{S_2^{A,w}}\phi_w^{S_2^{A,w}}} < 1,$$

$$\text{where } \phi^w_{S_2^{A,w}}\phi_w^{S_2^{A,w}} = \frac{\left(\beta_1^A\right)^2 Var[\varepsilon]+\left(\gamma_1^A\right)^2 Var[\zeta]}{\left(\beta_1^A\right)^2 Var[\varepsilon]+Var[u_1]+\left(\gamma_1^A\right)^2 Var[\zeta]}\frac{Var[\delta^B]}{Var[\delta^B]+\left(\frac{1}{\beta_1^B}\right)^2 Var[u_1]} < 1.$$

Since $\lambda_2$ is strictly decreasing in $Var[u_2]$ with $Var[u_2] \to 0 \Rightarrow \lambda_2 \to \infty$, for $Var[u_2]$ strictly smaller than the constant $C^*_{Var[u_2]}$ there exists no mixed strategy equilibrium. ■

# Chapter 4

# On Bounded Rationality and Risk Aversion

## 4.1 Introduction

Choosing the optimal consumption bundle out of thousands of commodities for a given income is a difficult task. This model describes a boundedly rational decision maker who does not always find his optimal consumption bundle with certainty. By facing the problem repeatedly, he can figure out more precisely which is his optimal bundle. It is therefore more likely that he will choose the optimal one at an income level he is used to, called the reference income level. This alters the decision maker's attitude towards income lotteries. Extending the standard model in this way helps to explain experimental findings like loss aversion, status quo bias, diminishing sensitivity etc. It also provides a theoretical reasoning as to why people become less risk averse after they faced an unexpected loss. We do not claim that the decision makers know the exact reasons for their risk preferences over lotteries. We believe instead that decision makers base their risk preferences on rules of thumb generated by a developmental-cognitive process.

Aim of this study is to provide a theoretical explanation for experimental findings summarized by Tversky and Kahneman (1986). For these experiments the certainty equivalence method was applied. They show that there are some common reaction patterns in choices involving risk. The following value function (developed by Tversky and Kahneman) captures these patterns.

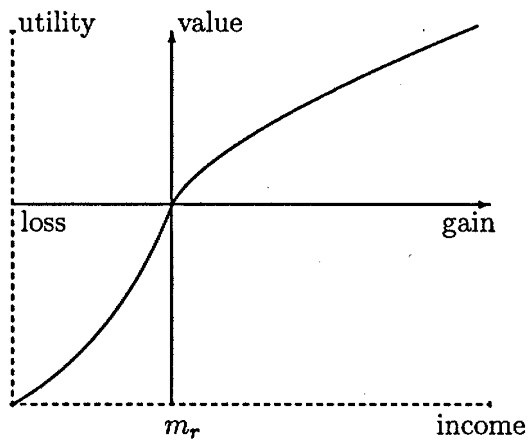



Figure 4.1: Value function of Tversky and Kahneman

One of the reaction patterns is that preferences are quite insensitive to small changes of wealth, but highly sensitive to corresponding changes in the reference income level, $m_r$. The reference income plays an important role in determining whether a lottery will be accepted or not. It is therefore useful to consider losses and gains with respect to this reference income level. A significant property is that individuals exhibit loss aversion, i.e. they are much more responsive to losses than to gains. In terms of a utility function, this means that a certain income decrease results in a much higher loss of utility than the utility gain associated with the same income increase. In other words the value function is steeper in the loss region than in the gain region. This leads to high risk aversion at the reference income level. Another important stylised fact is 'diminishing sensitivity' to losses. An increase in a small monetary loss leads to a far higher decrease in utility than an increase in a monetary loss that was already large. A consequence of diminishing sensitivity is that decision makers are often risk-loving over losses, i.e. the corresponding utility function becomes convex in the loss region.

In order to explain these findings we will base our theoretical analysis on the classical model framework, where the decision maker has a complete, reflexive and transitive preference ordering over the space of commodity bundles. We show that the Tversky-Kahneman value function can be explained without relaxing the von Neumann-Morgenstern utility axioms. These axioms allow us to represent both the preferences over lotteries and commodity bundles by an affine transformable utility function over the commodity space. Keeping the analysis within these axioms highlights the impact of our approach on the risk attitude of decision makers. Although our analysis can explain the value function developed by Tversky and Kahneman, it cannot explain all experimental results. To achieve this a departure of the von Neumann-Morgenstern axioms and/or a different approach (e.g. Prospect Theory Kahneman and Tversky (1979)) are necessary. Nice summaries of this fast growing literature are given in Camerer (1995) and Hey (1997). In our analysis we will relax the rationality assumptions only slightly in order to derive the relationship between risk aversion and bounded rationality. Boundedly rational decision makers are not only restricted by the availability of information but also in their ability to learn, i.e. processing the available information. In order to save information processing costs or time, boundedly rational decision makers apply simplified thinking and calculation procedures. The application of a heuristic, e.g. a cut off learning rule, in choosing the consumption bundle affects individual's risk aversion behaviour. A fully dynamic model should capture the whole learning process. As long as the optimal consumption bundle cannot be derived without cost/effort, our results hold for any possible learning process. Rather than modelling a certain learning process, which does not provide any additional insight, we restrict our analysis to the static consumption problem after the learning process is completed and the reference point is determined. This highlights the ingredients necessary to explain the experimental findings. In the random choice approach, as described in Section 2, the decision maker has to learn "how to choose the most preferred consumption bundle". Since he has chosen his consumption bundle at his reference income level many times before, he is very familiar with this choice. Furthermore, it was worthwhile for him to put a lot of thought into choosing the right consumption bundle at this income level, since he had expected to consume at this reference income level many times over. It is therefore reasonable to assume that he will not make errors (or at least he will not make "larger" errors) at the reference income level. At a different income level, there is scope for the decision maker to err. In Section 3 we employ a different approach. The decision maker has not processed all the information perfectly to figure out his 'true' preference ordering for the entire commodity space. Using reasoning similar to that in Section 2, he knows his optimal consumption bundle at his reference income level. Therefore he can exclude all utility functions which do not lead to this optimal consumption bundle at this reference income level. Both approaches result in reference-dependent preference models, which were introduced by Tversky and Kahneman (1991). We show that in both approaches bounded rationality increases risk aversion at the reference income level and there exists a range of income levels below the reference income where bounded rationality reduces risk aversion. In summary, all three aforementioned properties of the Tversky-Kahneman value function can be explained by our analysis. In Section 4 we draw some conclusions and give some general applications and implications of these results in other economic fields.

## 4.2 Random Choice Approach

In the random choice approach we relax the rationality assumptions slightly by assuming that the decision maker has to learn how to choose the optimal commodity bundle. Since he is boundedly rational, he will apply a heuristic. This heuristic allows him to choose a consumption bundle faster and with less effort, but at a cost of possible deviations from the optimal bundle. At his reference income level, he is willing to put much more effort into thinking and is willing to invest more time in finding the optimal consumption bundle, since the decision maker expects to frequently face the same maximisation problem. Therefore, the deviations from the optimal consumption bundle at the reference income level are much smaller. For simplicity we assume that he chooses the optimal bundle, $x^*(m_r, p_r)$ at the reference income level, $m_r$, given reference prices, $p_r$.[1] A decision maker is aware of the fact that if he accepts a lottery over income he has to choose a new consumption bundle at a possibly different income level. At this new income level he will choose the optimal consumption bundle with less than probability one. Therefore if a boundedly rational decision maker accepts a lottery he actually faces a compound lottery. At the first stage a lottery outcome over income is drawn and at the second stage he faces another lottery caused by the error he makes. The latter one has always negative expected value, because making errors can only worsen his situation.

The effect of applying heuristics on risk aversion is formalized in the following way.

[1] As we keep prices constant in our analysis, we will drop $p$ as argument in the functions.

**Assumptions**

**A 1** $u(x) : \mathbb{R}^k_+ \mapsto \mathbb{R}$, the utility function represents a complete, reflexive and transitive preference ordering over the commodity space $\mathbb{R}^k_+$

**A 2** $u(x)$ is a von Neumann Morgenstern utility function,
i.e. $\tilde{x} \succeq \tilde{x}' \iff E\,u(\tilde{x}) \geq E\,u(\tilde{x}')$,
where $\tilde{x}$ and $\tilde{x}'$ are random commodity bundles

**A 3** $u(x)$ is weakly increasing, and strictly increasing in at least one of its arguments

**A 4** $u(x) \in C^2(x)$ such that all resulting indirect utility functions are also well defined and twice continuously differentiable

**A 5** The actual consumption bundle chosen at reference income $m_r$ is
$x^E(\hat{x}, m_r) := \hat{x} + \underline{\tilde{e}}(\hat{x}, m_r)$

where

$\hat{x}$ is the target bundle the decision maker tries to achieve

$\underline{\tilde{e}}$ is the error captured by a k-dimensional random variable (function) in the state space $\mathbb{S} = \{1, ..., S\}$ with a subjective probability distribution $\Pi$

An optimal target bundle to aim for is given by
$\hat{x}^*(m, m_r) \in \arg\max\{E_s u(\hat{x} + \underline{\tilde{e}}(\hat{x}, m_r))$ s.t. $p\hat{x} \leq m\}$
whereas an optimal bundle is given by
$x^*(m) \in X^*(m) := \arg\max\{u(x)$ s.t. $px \leq m\}$
To simplify notation let $\tilde{e}_x(m, m_r) := \underline{\tilde{e}}(\hat{x}^*(m, m_r), m_r)$,
$\tilde{e}_x(\cdot)$ is such that

(i) $x^E \subset \mathbb{R}^k_+$ *(no negative consumption)*

(ii) $px^E \leq m \; \forall x^E$ *(affordability)*

(iii) there exists $\hat{x}^*(m_r, m_r) \in X^*(m_r)$
s.t. $\tilde{e}(m_r, m_r) = 0$ *(no error at $m_r$)*

(iv) if possible under (i) - (iii)

for given $m_r$, there exists for each $\hat{x}$
at least one $s'$ with $\pi_{s'} > 0$ s.t.
$x^E_{s'}(\hat{x}, m_r) \notin X^*(m)$ *(error possibility)*

(v) $\exists\, \hat{x}^*(m, m_r)$
s.t. $e_{x,s}(m, m_r) \in C^2(m)\ \forall s$ *(smoothness)*

(vi) decision maker can choose for all $\underline{m} < m$
$x^E(\underline{m}, m_r) + x_i \frac{m-\underline{m}}{p_i}$, where for $x_i$, $\frac{\partial u}{\partial x_j} > 0$.
This implies that he strictly *prefers higher income.*

A1 - A4 are standard utility assumptions. We suppose that the decision maker is still fairly rational since he has a transitive preference ordering which satisfies the von Neumann Morgenstern axioms. The assumption that the indirect utility function is twice continuously differentiable together with A5(v) allows us to apply the Arrow-Pratt risk aversion measure. The difference from standard microeconomic models lies in A5, which states that the decision maker makes errors if he is not consuming at his reference income level. He tries to consume target bundle $\hat{x}$ but ends up consuming $x^E$.[2] Note that the decision maker need not necessarily aim for an optimal bundle, since he knows the distribution of the error term. This can especially be the case when the error term is biased. If e.g. the decision maker will always end up buying accidently too much chocolate, it is probably useful for him to aim at a consumption bundle with less chocolate than in the optimal consumption bundle. A5(i) rules out negative consumption for any commodity. A5(ii) guarantees that the errors are such that the decision maker spends not more than his income. A5(iii) assumes that the decision maker knows how to choose the optimal consumption bundle at the reference income level. In other words he makes no errors at the reference income level, $m_r$. This can be easily relaxed to the case where the decision maker makes only errors at the reference income level, which have no utility impact. A5(iv) states that the boundedly rational decision maker can make significant errors if he is not at his reference income level. In the case of

[2]There are several possible explanations why the decision maker might not consume his target consumption bundle $\hat{x}$, but $x^E$. Consider for example a setting where the decision maker does not buy all commodities at once but sequentially. Due to some miscalculation he might buy to many of those goods he purchases in the beginning. He then has not enough money left for the remaining commodities. In the random utility approach (Section 3) the error $x^E - \hat{x}$ results from the ignorance of the optimal consumption bundle.

strictly quasiconcave utility functions any possible error reduces the expected utility. A5(vi) rules out that an increase in income makes the decision maker worse off. In other words, the increase of the error due to higher income has a lower impact on the expected utility than the enlargement of the budget set. This assumption is plausible since the decision maker does not need to spend all of his income. Consequently higher income does no harm. As he has by A3 the opportunity to spend the remaining income for the commodity which leads to a strict increase in his utility, he will always strictly prefer higher income and spend all his money.

To clarify the analysis we define some indirect utility functions and the standard Arrow-Pratt risk aversion measure.

**Definitions**

**D 1** Indirect utility function when decision maker makes no errors

$$v(m) := u(x^*(m))$$

where $x^*(m) \in \arg\max\{u(x) \text{ s.t. } px \leq m\}$

**D 2** 'Best' indirect utility function when decision maker is boundedly rational, i.e. he knows that he will make some errors

$$Ev^E(m, m_r) := E_s u(\underbrace{\hat{x}^*(m, m_r) + \tilde{e}_x(m, m_r)}_{x^E(\hat{x}^*, m_r)})$$

where $\hat{x}^*(\cdot)$ and $\tilde{e}_x(\cdot)$ are defined in A5

Since it is not sure whether a boundedly rational decision maker can derive his optimal target consumption bundle, $\hat{x}^*$, $Ev^E(\cdot)$ is only an upper bound for his indirect utilility function. Our analysis can be applied to any possible indirect utility function of the boundedly rational decision maker as long as it is twice continuously differentiable and strictly increasing. Note, by A5(vi) $Ev^E(m)$ is strictly increasing.

To simplify notation we drop $m_r$ as arguments in all of the indirect utility functions, since $m_r$ is constant in our analysis.

**D 3** The functional $f(v)$ relates the indirect utility function $v(m)$ of an identical *rational* decision maker to the indirect utility function $Ev^E(m)$ of a *boundedly rational* decision maker.
By A3 $v(m)$ is strictly increasing in $m$. Therefore there exists the inverse of $v(m)$, $h(v(m)) = m$.
Let

$$f(v) \quad := \quad Ev^E(h(v))$$

Since $v(m)$ is twice continuously differentiable in $m$, so is $h(v)$; and since $Ev^E(m) \in C^2(m)$, $f(v) \in C^2(v)$.

**D 4** Arrow-Pratt measure of (absolute) risk aversion for the indirect utility function of the *rational* and of the *boundedly rational* decision maker

(i) for $v$: $RA^v(m) := -\frac{\partial^2 v/\partial m^2}{\partial v/\partial m}$

(ii) for $Ev^E$: $RA^{Ev^E}(m) := -\frac{\partial^2 Ev^E/\partial m^2}{\partial Ev^E/\partial m}$

**D 5** Risk aversion contribution of bounded rationality
$RAC(m) := RA^{Ev^E}(m) - RA^v(m)$

This definition allows us to separate risk aversion into two parts, one being the actual risk aversion given by the concavity of the utility function and the other being the risk aversion contribution of bounded rationality induced by the optimal consumption bundle not being chosen. By Lemma 1 it is clear that the 'additive' definition of the risk aversion contribution in D5 is reasonable. Lemma 1 has the same flavour as part of Pratt's theorem (Pratt 1964). All proofs are presented in the appendix.

**Lemma 1** $RAC(m) = -\frac{\partial^2 f/\partial v^2}{\partial f/\partial v}\frac{\partial v}{\partial m}$.

Lemma 1 relates the risk aversion contribution term to the functional $f(v)$, which facilitates proving the following propositions.

Lemma 2 shows that an error reduces the decision maker's expected utility, since a non-optimal commodity bundle is consumed with positive probability. This is not the case at a zero income level, since no consumption takes place, and at the reference income level. At the reference income level, $m_r$ the decision maker has learnt how to choose the optimal consumption bundle.

**Lemma 2** *A 'focal point' is an isolated income level where bounded rationality has no impact on the utility level. At all other income levels bounded rationality strictly reduces the indirect utility function.*
*'Focal points' are* $c_1 = 0$ *and* $c_2 = m_r$ *i.e. e.g. for* $Ev^E(m)$
*(i)* $Ev^E(m) = v(m)$ *for* $m \in \{0, m_r\}$,
*(ii)* $Ev^E(m) < v(m)$ *for* $m \in \mathbb{R}_+ \setminus \{0, m_r\}$.

Lemma 2 illustrates that the indirect utility function of an identical rational decision maker (who makes no errors) is an upper envelope for the indirect utility function of the boundedly rational agent. Figure 4.2 illustrates the two focal points 0 and $m_r$. The focal point $c_1 = 0$ depends on the assumption A5(i) $x^E \subset \mathbb{R}^k_+$, which states that consumption of any commodity cannot be negative. This binds the space for the error term. With decreasing income this space decreases and at zero income the possible consumption set is the single point 0, i.e. the error term vanishes. In other words, since at a zero income level only zero consumption is possible, there is no possibility to err. If one allows negative consumption the space for the possible errors need not shrink with decreasing income. Therefore Lemma 2 does not hold in this case. Assumption A5(i) turns out to be important for Proposition 2.

The shape of $Ev^E$ can also be made plausible using a truly dynamic learning model. Before the decision maker has learnt how to choose his optimal consumption bundle at the reference income level, $x^*(m_r)$, his expected utility is strictly below $v(m)$ for all positive income levels. After he knows $x^*(m_r)$ his utility level for incomes around $m_r$ increases,

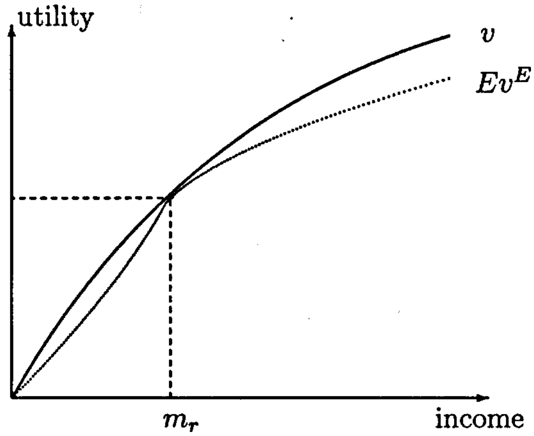



Figure 4.2: Indirect utility functions

leading to an indirect utility function $Ev^E(m)$ illustrated in Figure 4.2.

Using Lemmas 1 and 2 we can show that bounded rationality (defined as making small errors in choosing the optimal commodity bundle) increases risk aversion at the reference income level. This is consistent with experimental results.

**Proposition 1** *(i) Bounded rationality increases absolute risk aversion at the reference income level, i.e.* $RAC(m_r) > 0$.
*(ii) A boundedly rational decision maker strictly prefers the reference income level,* $m_r$*, to any lottery whose certainty equivalence for an identical rational decision maker is* $m_r$.

Lotteries are much less attractive for a boundedly rational decision maker because for each outcome of the lottery, leading to an income different from 0 or $m_r$, he cannot be sure to consume an optimal consumption bundle. The indirect utility function $Ev^E(m)$ can be thought of as resulting from maximisation behaviour subject to an additional constraint, as given by A5. Given that the indirect utility function of an identically rational individual is the envelope of $Ev^E(m)$ with tangent point at the reference income level, Proposition 1 seems obvious. Proposition 1 is in the same vein as the Le Châtelier principle, especially if one assumes that the decision maker learns to choose the new optimal consumption bundle over time. The Le Châtelier Principle states that the response of optimised variables to a small structural change to the system is reduced, the more constraints are added to how the variables can be changed. In our case $x^*(m)$ maximises the utility function for a given income, $m$. The additional constraints how $x^*$ can change as $m$ departs from $m_r$ are given by Assumption A5. The Le Châtelier principle fails globally since the constraint summarised by Lemma 2(ii) is not binding at two distinct income levels (Roberts 1996).[3] This Proposition also shows that a decision maker is more risk averse if he has spent a huge amount of his money on durable goods. A sudden income change constrains him from adjusting to the new optimal consumption bundle. He has still to consume the durable commodities, which he bought in previous periods. It is interesting that the expected riskiness of the income stream together with his risk aversion determines the amount he is willing to spend on durable commodities, which, in turn, influences his risk aversion.

Proposition 1 emphasises the importance of considering the reference point for the analysis of risk behaviour and shows why responses in utility to losses are more extreme than responses to gains. The distinction between loss aversion, "status quo bias" and "endowment effect" is nicely explained in Rabin (1998). In our model loss aversion is due to a change in income which results in costs incurred (effort exerted), since the decision maker has to think about choosing a new optimal consumption bundle. A decision maker must be compensated for the additional costs of thinking arising from a lottery over income, since this induces him to find a new commodity bundle. This is the idea behind Proposition 1.

One might argue that lotteries over income can be diversified by borrowing and lending. Consequently, Proposition 1 will then only apply to uncertainty over someone's wealth level. In a world with many independent uncertainties, they may average out. Nonetheless, Proposition 1 refers only to a (instantaneous) per period utility function. A concave per period utility function causes then diversification over time. In addition, a perfect capital market does not exist in reality, and the decision maker also has to find out (think about) the wealth impact of a lottery outcome.

Since the risk aversion contribution term (a result of bounded rationality) is a real num-

[3] I am indebted to Kevin Roberts for pointing out this similarity.

ber, we can consider it as a measure of bounded rationality.

Whereas bounded rationality increases risk aversion at the reference income level, Proposition 2 claims that bounded rationality decreases risk aversion or even leads to risk loving behaviour at a lower income level. Our theoretical proposition can explain the experimental findings of Tversky and Kahneman (1986). As explained above, the risk-seeking attitude in parts of the loss region can be attributed to 'diminishing sensitivity'.

**Proposition 2** *There exists a range of incomes* $(\underline{m}, \overline{m})$ *between two 'focal points' where bounded rationality reduces risk aversion or leads to risk loving behaviour, i.e.* $RAC(m) < 0$. *For the 'focal points'* $c_1 = 0$ *and* $c_2 = m_r$ *the income range* $(\underline{m}, \overline{m})$ *is within* $(0, m_r)$.

As mentioned above the result is driven by the fact that at a lower income level only smaller errors are possible. Lower income reduces the budget set within the actually chosen consumption bundle, $x^E = \hat{x} + e$ has to lie.

The range of income levels where risk aversion decreases because of bounded rationality is determined by both the error term and the utility function. There are three factors determining the size and locale of that range of income levels. First, since indifference curves at lower income levels are generally more curved, the same error causes a higher disutility at a lower income level. The second factor is that with lower income the error possibility space shrinks. At the extreme, for zero income there is no 'space' left for any error. One can show that the degree to which the error possibility space shrinks depends on the number of available commodities. Third, with lowering the income the distance to the reference income level increases. It is plausible that the variance of the error term increases with this distance. A larger variance in turn leads to a lower utility level for strictly quasiconcave utility functions. All three factors influence the size and locale of this income range. Whereas the first effect suggest that the relevant range should be farther away from the reference income, the second suggests the opposite. The third effect pushes $(\underline{m}, \overline{m})$ closer to $m_r$ if the variance of the error term increases concave as the distance from the reference income level increases.

By Lemma 2 it is obvious that both Propositions are not only true for $Ev^E(m)$ but for any strictly increasing indirect utility function which is twice continuously differentiable.

These results can be generalised to the case where the indirect utility functions are not twice differentiable. Since the traditional Arrow-Pratt risk aversion measure is not defined anymore in such a setting, local risk aversion can be measured by using the preference ordering over $\epsilon$-income lotteries, by comparing their certainty equivalence. It is easy to see that Proposition 1 still holds, and so does a slightly modified Proposition 2.

This approach assumes that the decision maker knows his 'true' preference ordering but is not able to pick his optimal commodity bundle. However, in reality he has to learn his 'true' preference ordering. The implication of not knowing the exact true preference ordering will be the focus of the following section.

## 4.3 Random Utility Approach

Since there are time and effort costs to finding out the 'true' preference ordering for the boundedly rational decision maker, he will focus his learning primarily upon his relevant income level. He will also try to find the true preference ordering over the whole commodity space, but he is not willing to spend too much effort and time learning his 'true' preference which is only relevant for far distant income levels. In our analysis we assume that the decision maker exerts enough effort to find his most preferred commodity bundle at his reference income level, whereas at some different income he has a certain distribution over possible preference orderings. Each possible preference ordering is represented by a utility function $u_s(x)$. The difference from standard information economics lies in the fact that the boundedly rational decision maker does not know his true utility function, not because he is lacking information, but because he is not able to process all his information in time. In a situation where all the information has been perfectly processed and a rational decision maker still cannot figure out his 'true' utility function, we can consider each $u_s(x)$ as being the true expected utility function over the utility functions of the finest partition of the rational decision maker, $I_s \in \Im$. If one accepts the axioms of von Neumann Morgenstern, the decision maker still has a complete preference ordering over the whole commodity space despite whether or not any information has been processed. These preference ordering is represented by his expected utility function.

The above described model is formalised in the following way:

**Assumptions**

**A 1'** The decision maker knows a (strictly positive) probability distribution $\Pi = \{\pi_1, \ldots, \pi_S\} > 0$ over the set of possible utility functions $\mathbb{U} = \cup_s u_s(x)$

**A 2'** All $u_s(x)$ satisfy the assumptions **A1 - A4**

**A 3'** All $u_s(x)$ are such that

(i) $x_s^*(m_r) = x_{m_r}^* \; \forall s,$

where $x_s^*(m_r) = \arg\max\{u_s(x) \text{ s.t. } px \leq m_r\}$

(ii) $\cap_s x_s^*(m) = \emptyset \; \forall m \in \mathbb{R}_+ \setminus \{0, m_r\},$

where $x_s^*(m) = \arg\max\{u_s(x) \text{ s.t. } px \leq m\}$

Assumption 3'(i) states that the decision maker knows his optimal consumption bundle at his reference income level, $m_r$. In other words he can rule out utility functions which do not lead to $x_{m_r}^*$ at $m_r$. We do not assume that he knows the utility level of this consumption bundle, let alone his utility function at his reference income level. The second part of A 3' rules out the case where at a certain income level all possible utility functions lead to the same optimal consumption bundle.

Proposition 3 illustrates the analogy between the effects on risk aversion in the random choice and the random utility approach. The fact that the 'true' optimal consumption bundle is not chosen with certainty drives these effects in both approaches. The proof in the appendix shows that the random utility approach can be reinterpreted in such a way that the assumptions of the random choice approach are satisfied.

**Proposition 3** *The random utility approach leads to the same effects on risk aversion as the random choice approach.*

While these effects on risk aversion due to not choosing the optimal consumption bundle with probability one are the same for both approaches, in the random utility approach the actual 'true' risk aversion also depends on which of the possible utility functions $u_s(x)$ is the 'true' one. The boundedly rational decision maker does not know his true utility function and hence his true risk aversion. This is not the case if all possible utility functions $u_s(x)$ exhibit the same indirect utility function $v_s(m)$ $\forall s$. However, it is important to notice that a slight change in the relative prices will immediately destroy the property of this special case. But also for the case where the possible utility functions $u_s(x)$ lead to different indirect utility functions the decision maker will take the explained effects into account.

Boundedly rational decision makers probably do not know their exact 'true' preference ordering precisely *and* moreover they err. We do not need any further analysis to see that these results still hold if one combines the random utility approach and the random choice approach.

## 4.4 Possible Extensions and Conclusion

Our analysis shows that one factor contributing to risk aversion is the fact that the decision maker must find his new consumption bundle after the outcome of the lottery has been realised. This requires that he incurs thinking costs *in* the realised state of the world. Evaluating a lottery is a much more difficult task because one does not only incur thinking costs in the realised state but in all possible states. Therefore the decision maker will apply a simpler heuristic in evaluating a lottery. It is then plausible that one will observe more misjudgements in decisions made about the acceptance of a lottery.

A related area of research examines the question of finding the *optimal planning horizon* in a world with uncertainty. Planning for distant future increases the number of states exponentially, which makes the maximisation problem much more complicated. Therefore, boundedly rational decision makers will apply a heuristic which is much more precise for short sighted problems. It remains to be shown that the optimal heuristic provides a fairly exact prediction for the near future and a rougher prediction for the distant future. It also seems plausible that increasing uncertainty levels makes people more short sighted, which can explain why high volatility in inflation rate, i.e. price uncertainty, hurts the economy. The optimal planning horizon solution also provides an explanation for why we observe incomplete contracts, and the demand for flexibility or liquidity.

The model can be extended to include uncertainty in both income and *prices*. It is a well known fact that in traditional microeconomic models where the decision maker's utility function is quasiconcave and exhibits constant marginal utility of income, the decision maker is risk loving with respect to price uncertainty. This is due to the fact that he chooses his optimal consumption bundle after the prices are realised. In an analysis with error possibilities similar to ours, this risk loving behaviour need not be true. More insight might be gained if we separate the income from the substitution effect by means of a Slutsky decomposition. Endogenous substitution costs can then be derived.

As we pointed out in Section 3 the results of the paper are also applicable to the production sector. Each firm has to "learn" its production function by gathering know-how. But gathering know-how incurs costs, so the firm's production function is only well-understood locally. If the output or input prices were to change suddenly, new information would have to be gathered and processed. A single shock in a one industry sector can change all relative prices and hence affect the whole economy, i.e. all firms have to adjust to their new production plan. If these adjustments cause costs (e.g. the new relevant part of the production function must be learnt) then it is very likely that adjustments do not take place for every single shock.

We have mentioned examples where bounded rationality effects do not average out, thereby affecting the aggregate economy. In these cases it is important to incorporate these boundedly rational aspects of agent behaviour into economic theory in trying to attain a better understanding of the real underlying economic relationships.

## 4.5 Appendix of Chapter 4

*Proof of Lemma 1:*

$$\frac{\partial Ev^E}{\partial m} = \frac{\partial f}{\partial v}\frac{\partial v}{\partial m}$$

$$\frac{\partial^2 Ev^E}{\partial m^2} = \left(\frac{\partial^2 f}{\partial v^2}\frac{\partial v}{\partial m}\right)\frac{\partial v}{\partial m} + \frac{\partial f}{\partial v}\frac{\partial^2 v}{\partial m^2}$$

$$\begin{aligned}
-\frac{\partial^2 Ev^E/\partial m^2}{\partial Ev^E/\partial m} &= -\frac{\partial^2 f/\partial v^2}{\partial f/\partial v}\frac{\partial v}{\partial m} - \frac{\partial^2 v/\partial m^2}{\partial v/\partial m} \\
RA^{Ev^E} &= RAC + RA^v \quad \blacksquare
\end{aligned}$$

*Proof of Lemma 2:*

(i) (1) for $m = 0$

Since $x_s^E \in \mathbb{R}_+^k$ and $px_s^E = m = 0 \; \forall s \; \forall x_s^E$

$x_s^E(0, m_r) = x^*(0, m_r) \; \forall s$

(2) for $m = m_r$

By A5(iii) $\underline{e}_s(x_i^{a*}(m_r, m_r), m_r) = 0 \; \forall s$.

Therefore $x^E(\hat{x}_i^*(m_r, m_r), m_r) \in X^*(m_r)$.

(ii) for each $m \in \mathbb{R}_+ \setminus \{0, m_r\}$

By definition D1, $v(m) \geq Ev^E(m)$.

By A5(iv) $\exists$ for each $\hat{x}$ at least one $s'$ with $\pi_{s'} > 0$.

such that $x_{s'}^E(\hat{x}, m_r) \notin X^*(m)$. Therefore $\exists$ for

each $m \in \mathbb{R}_+ \setminus \{0, m_r\}$ at least one $s'$ with $\pi_{s'}$ such that

$u(\hat{x}^* + e_{x,s'}) < u(x^*)$.

(iii) It follows immediately that (i) and (ii) is true not only for the upper bound $Ev^E(m)$, but also for any indirect utility function of the boundedly rational decision maker. $\blacksquare$

### 4.5.1 Proof of Proposition 1

(i) By Lemma 1 $RAC(m) = -\frac{\partial^2 f/\partial v^2}{\partial f/\partial v}\frac{\partial v}{\partial m}$

It is sufficient to derive the signs for the three factors.

(1) $\frac{\partial v}{\partial m} > 0$,

since $u(x)$ is strictly increasing in at least one argument.

Let $g(v(m)) := f(v(m)) - v(m)$. Since $f(\cdot)$ and $v(m) \in C^2$, $g(\cdot) \in C^2$. By Lemma 2 $g(\cdot)$ has a local maximum at $v(m_r)$, Therefore $\frac{\partial g}{\partial v} \mid_{v(m_r)} = 0$ and $\frac{\partial^2 g}{\partial v^2} \mid_{v(m_r)} < 0$, which yields

(2) $\frac{\partial f}{\partial v} \mid_{v(m_r)} = 1 > 0$,

(3) $\frac{\partial^2 f}{\partial v^2} \mid_{v(m_r)} < 0$.

(ii) Take any lottery $\tilde{m}$ (with distribution $F$) whose certainty equivalence for an rational decision maker is $m_r$, i.e. $E_F[v(\tilde{m})] = v(m_r)$. By Lemma 2 for any realisation $m_j$ of $\tilde{m}$, $Ev^E(m_j) < v(m_j)$. Thus, $E_F[Ev^E(\tilde{m})] < E_q[v(\tilde{m}] = v(m_r) = Ev^E(m_r)$. ■

### 4.5.2 Proof of Proposition 2

By Lemma 1 $RAC(m) = -\frac{\partial^2 f/\partial v^2}{\partial f/\partial v}\frac{\partial v}{\partial m}$

(1) $\frac{\partial v}{\partial m} > 0$ (see Proposition 1)

(2) $\frac{\partial f}{\partial v} > 0$,

since $\frac{\partial Ev^E}{\partial m} = \frac{\partial f}{\partial v}\frac{\partial v}{\partial m}$ and $v$ and $Ev^E$ are strictly increasing in $m$.

(3) $\exists$ $(\underline{m}, \overline{m}) \subset (c_1, c_2)$, s.t. $\frac{\partial^2 f}{\partial v^2} > 0$ on $(v(\underline{m}), v(\overline{m}))$.

This is shown in the following three steps:

(3.1) $\exists$ $a \subset (v(c_1), v(c_2))$ s.t. $\frac{\partial^2 f}{\partial v^2}|_a > 0$.

By Lemma 2 $f(v(c_1)) = v(c_1)$ and $f(v(c_1 + \epsilon)) < v(c_1 + \epsilon)$ for sufficiently small $\epsilon > 0$.

From this we can conclude that $\frac{\partial f}{\partial v}|_{v(c_1+\epsilon/2)} < 1$.

We also know from Proposition 1 that $\frac{\partial f}{\partial v}|_{v(c_2)} = 1$.

Applying the mean value theorem on $\frac{\partial f}{\partial v}(\cdot)$,

$\exists$ $a$ $\in (v(c_1 + \epsilon/2), v(c_2))$ such that

$$\frac{\partial^2 f}{\partial v^2}|_a = \frac{\overbrace{\partial f/\partial v\,|_{v(c_2)}}^{=1} - \overbrace{\partial f/\partial v\,|_{v(c_1+\epsilon/2)}}^{<1}}{\underbrace{v(c_2) - v(c_1 + \epsilon/2)}_{>0}} > 0.$$

(3.2) Since $f \in C^2$ and $\frac{\partial^2 f}{\partial v^2} > 0$ at $a$, this must be also true at $(a - \epsilon', a + \epsilon')$ for small $\epsilon' > 0$.

(3.3) Since $v(\cdot)$ is strictly increasing and continuous in $m$ there exists for each $\vartheta \in (a - \epsilon', a + \epsilon')$ a corresponding $m$ such that $\vartheta = v(m)$. ■

### 4.5.3 Proof of Proposition 3

It is sufficient to show that the random utility approach can be re-interpreted such that assumptions A1-A5 of the random choice approach are satisfied.

For this purpose let

$$e_x(m, m_r) := x^{Eu*}(m) - x_s^*(m)$$

where

$$\begin{aligned} x^{Eu*}(m) &\in X^{Eu*}(m) &&:= \arg\max\{E_s u_s(x) \text{ s.t. } px \le m\} \\ x_s^*(m) &\in X_s^*(m) &&:= \arg\max\{u_s(x) \text{ s.t. } px \le m\}. \end{aligned}$$

(1) A1-A4 are satisfied for the true $u_{s'}(x)$ by A2'.

(2) It remains to show that $\tilde{e}_x(\cdot)$ satisfies all restrictions of A5. (Prices are kept constant.)

(i) $x^{Eu*} \subset \mathbb{R}_+^k$ and $x_s^* \subset \mathbb{R}_+^k\ \forall s$

by A2' and A1.

(ii) $px^{Eu*} \le m\ \forall x^{Eu*} \in X^{Eu*}$,

which is satisfied by definition of $X^{Eu*}(m)$.

(iii) $\exists\ x^{Eu*}(m_r) \in X_s^*(m_r)\ \forall s$ s.t. $\tilde{e}_x(m_r, m_r) = 0$

By A3'(i) $x_s^*(m_r) = x_{m_r}^* \forall s$, thus $x^{Eu*}(m_r) = x_{m_r}^*$.

(iv) $\exists$ for each $m \in \mathbb{R}_+ \setminus \{0, m_r\}$ (incorporates Lemma 2) and for each $x^{Eu*}(m) \in X^{Eu*}(m)$ at least one $s'$ with $\pi_{s'} > 0$ such that $x_i^{Eu*}(m) \notin X_{s'}^*(m)$

This is satisfied by A3'(ii).

(v) $\exists\ x^{Eu*}(m)$ s.t. $e_{x,s}(m, m_r) \in C^2(m)\ \forall s$

By A2' and A4 all $u_s(x)$, $Eu_s(x)$ and all resulting indirect utility functions are twice continously differentiable. Thus there exist income expansion paths $x_s^*(m)\ \forall s$ and $x^{Eu*}(m)$ satisfying this property.

(vi) The decision maker always strictly prefers higher income since by A2' in conjunction with A3 all $u_s(x)$ are strictly increasing in at least one argument. ■

# Chapter 5

# Price Stickiness and Exchange Rate Volatility

## 5.1 Introduction

Monetary policy has real effects in a world with sluggish price adjustments. Furthermore, in an open economy setting, domestic monetary policy has an effect on economies abroad. The foreign economies are not only affected by exchange rate movements but also through other spillover effects. The effect of monetary policy on exchange rate volatility is of considerable interest. Macroeconomists have long discussed how domestic monetary policy affects foreign production, consumption and inflation rates. The violation of the uncovered interest rate parity and of the purchasing power parity (PPP) in the short run are other puzzles in international macroeconomics. In this paper we argue that exchange rate volatility, spillovers, interest rate parity and PPP depend strongly on the type of nominal rigidity that allows monetary shocks to have an effect on real variables. The importance of different forms of price stickiness can then be weighted given the existing empirical evidence.

Most arguments that were concerned with the effect of monetary policy in an open economy were based on a static Mundell Flemming type analysis. This has changed substantially over the last years. A growing body of recent research looks at these effects in a dynamic framework. Most of these papers use elements of the framework that has been developed by Obstfeld and Rogoff (1995) and Svensson and van Wijnbergen (1989). They combine the new intertemporal approach to the current account with rational expectations and the traditional Keynesian setting of sluggish price adjustment. In all these models, monetary policy has real effects by stimulating demand to which supply adjusts. While prices/wages are only sticky in the short run, asymmetric monetary policy leads to permanent effects through perfectly integrated international bond markets. Our model extends this framework in various ways.

The model describes a two country world, home and foreign, that is populated by workers that provide labour to firms. There is a complete home bias in the ownership of firms. Each firm produces in only one of the two countries and is in monopolistic competition with firms both abroad and at home. Unlike Obstfeld and Rogoff (1995) we assume that firms are able to price discriminate between countries. While we believe that this is realistic, we also need this assumption to study different sorts of price stickiness. There is a substantial amount of evidence that borders have a much bigger effect on price disparities than for example transport costs, (Engel and Rogers 1996). Furthermore, we assume that firms are monopsonists on the labour market. We need to introduce an imperfection to study wage stickiness following positive and negative monetary shocks. We believe that allocating the market power to the firm is more realistic than allocating it to the workers. This is in contrast to monopolistic competition between different trade unions which is assumed in a series of papers that use a similar framework, e.g. Obstfeld and Rogoff (1996) and Hau (1998). While monopsonistic market power of firms is certainly an extreme assumption, labour economists have previously argued that positive output effects following the introduction of minimum wages are a sign of firms with monopsonistic market power (Manning 1995). In comparison to the social optimum, prices are too high in our model due to monopolistic competition and wages are too low because of the monopsonistic market power of firms. If there is money expansion in an economy with sticky prices, nominal wages will adjust while real prices decrease. This leads a priori to more production in the country that expands its money supply and suggests a current account surplus. On the other hand, if wages are sticky, a money supply increase leads to higher prices and thus to lower real wages. This leads to lower production and a current account deficit. The monopsonistic market structure allows us to study both mirror images.

We distinguish between three different forms of price stickiness. We start with retail prices being fixed, by which we mean that prices are fixed in the consumers' currency. This is the sort of price stickiness that is traditionally assumed in Keynesian models such as Mundell (1961), (1963) and Dornbusch (1976). We then go on to analyse the implications of wholesale price stickiness, which we define as prices being fixed in the producers' currency. We show, this formulation is actually equivalent to the formulation in Obstfeld and Rogoff (1995) where producers can not price discriminate between countries. Finally, we compare these types of price stickiness with sticky wages.

The link between monetary policy and exchange rate volatility has drawn new attention. Mussa (1976) and (1986) first argued that the increased volatility of the real exchange rate in the post Bretton Wood period has to be explained by sluggish price adjustment and increased volatility of monetary disturbances. In contrast, Stockman (1988) proposed that the increased volatility is due to increased volatility of productivity shocks. Monetary models of exchange rates were further discredited when Meese and Rogoff (1983) showed that these models could not explain exchange rate movements. Recent research using VAR techniques has drawn attention back to monetary shocks. It has been shown that monetary disturbances explain a significant part of nominal and real exchange rate volatility, see e.g. Clarida and Gali (1994) and Eichenbaum and Evans (1995).

In our theoretical model the nominal exchange rate moves immediately no matter whether wages of whether wages, wholesale or retail prices are sticky. Under sticky wholesale prices, it jumps by less than the magnitude of the monetary expansion and immediately reaches its new steady state value. In contrast, under sticky retail prices the exchange rate jumps by more than the monetary expansion and returns to the old steady state level in the long-run. If wages are sticky, the exchange rate moves more than the money supply. The exchange rate immediately reaches its new steady state as in the case of sticky wholesale prices. Given sticky retail prices, the volatility of the real exchange rate, as measured by the relative price of a consumption basket in the two countries, displays the same volatility as the nominal exchange rate. This is in line with the empirical findings of Rogoff (1996). Under sticky wholesale prices and under sticky wages, the real exchange rate does not move at all because the law of one price always holds.

Empirical evidence about the spillover effects appears to be inconclusive. McKibbin and Sachs (1991) argue that the spillover effects of monetary policy on real variables are small while Canzoneri and Minford (1986) claim that they are reasonably big and negative. It is important to understand the size and direction of spillover effects before one can discuss the need for international monetary coordination . Traditional Keynesian models predict a negative response of foreign output to domestic monetary expansions (e.g. Mussa (1979)). A depreciation in the home currency raises the price of foreign goods. This leads to a substitution away from foreign goods and to a reduction in production abroad. This is not necessarily true in the Obstfeld and Rogoff (1995) model because the income effect can potentially dominate the substitution effect. However, in their model welfare always increases in both countries no matter which country expands its money supply.

We show in this paper that both the size of spillover effects on foreign consumption and production and their direction depends crucially on the type of nominal stickiness assumed. Under sluggish wholesale prices Obstfeld and Rogoff's (1995) result is confirmed even though we do not assume the law of one price. On the other hand, if retail prices are sticky, the foreign country's welfare is unambiguously negatively affected by monetary expansions at home. The traditional Keynesian notion of "beggar thy neighbour" policies is reinstated. Foreign consumption is negatively correlated with a money expansions at home whilst the equilibrium labour input is positively correlated. Under sticky wholesale, prices the correlations of both consumption and production change from the short to the long-run. While consumption is initially positively affected by a foreign money expansion, it is negatively correlated in the long run. The opposite is true for production. Spillover effects under sticky wages are very different from the effects under sticky prices. The effect is almost the mirror image of what happens under sticky wholesale prices. Foreign production is negatively correlated in the long-run to home money expansions. Consumption abroad declines in the short-run but increases in the long-run.

The empirically established J-curve effect shows that the trade balance is negatively correlated with current and future exchange rates while it is positively correlated with past exchange rates. In our model, the current account is initially positive if either of the two prices are sticky but turns out to be negative under wage stickiness. In the long run the sign of the current account is reversed turning negative under sticky prices and positive under sticky wages. It is worthwhile noting that while the cross-correlation of the trade balance with the current exchange rate has different signs under sticky wages and sticky prices, the cross-correlation of the terms of trade and the trade balance is always positive. Even under sticky wages, where the exchange rate is negatively correlated the prices move far enough to allow the terms of trade to be positively correlated with the trade balance. Our findings extends the findings of Backus, Kehoe, and Kydland (1994) to monetary shocks. They found that while the J-curve effect can be reconciled with permanent productivity shocks, it is not possible to reconcile the negative correlation with fiscal shocks. In our model the efficiency gain of monetary disturbances is also only short-term even though they lead to permanent effects due to international lending.

The remaining paper is organised as follows. Section 2 introduces the model and Section 3 analyses the steady state. Section 4 discusses the effects of monetary disturbances under different kinds of price stickiness. Section 5 summarises the results and compares the effects of different types of price stickiness and real imperfections. Conclusions are presented in Section 6.

## 5.2 The Model

### 5.2.1 Consumers' Problem

The world is a $1 \times 1$ square in our model. A fraction $n$ of the population lives in the home country and a fraction $(1-n)$ abroad. There is also a continuum of firms on the interval $[0,1]$. A measure of $n$ firms produce at home and a measure $(1-n)$ in the foreign country. Home firms are symmetrically owned by home citizens and foreign firms by foreign citizens. Each inhabitant works in one firm located in his country but consumes the whole range of home and foreign produced goods. The group of potential workers for each firm is of measure one. All citizens maximise an additively separable utility function with a common discount rate $\delta$,

$$U = \sum_{t=1}^{\infty} \left(\frac{1}{1+\delta}\right)^t u(C_t^h, \frac{M_t^h}{P_t^h}, l_t^h).$$

As in Obstfeld and Rogoff (1996), the flow utility is Cobb Douglas in money and in the composite consumption good. The marginal disutility of labour is constant $\kappa$.

$$u(C_t^h, \frac{M_t^h}{p_t^h}, L_t^h) = \ln C_t^h(z) + \chi \ln \frac{M_t^h(z)}{p_t^h} - \frac{\kappa}{2} L_t^h(z)^2$$

The citizens derive positive utility from holding real money in their own currency. Holding more cash saves them trips to their bank. The flow utility exhibits constant elasticity of substitution (CES) of $\rho$ among the different commodities. The composite consumption good is, therefore, given by

$$C_t^h(z) = \left[\int_0^1 c_t^h(k,z)^{\frac{\rho-1}{\rho}} dk\right]^{\frac{\rho}{\rho-1}}$$

and the price index is defined as

$$p_t^h = \left[\int_0^1 p_t^h(k)^{1-\rho}\right]^{\frac{1}{1-\rho}}.$$

The superscript $h$ refers to the home country and $f$ to the foreign country.

The budget constraint for an individual agent of type $z$ is given by

$$p_t^h C_t^h + p_t^h B_{t+1}^h + M_t^h = L_t^h(z) w_t^h(z) + \pi_t^h + M_{t-1}^h + p_t^h(1+r_t) B_t^h - p_t^h \tau_t^h,$$

where $\tau_t^h$ are real government transfers, $B_{t+1}$ denotes the real bond holdings of internationally traded bonds from time $t$ to $t+1$ and $r_t$ is the interest rate which is determined at the time of its payment at $t$. $w$ is the nominal wage and $\pi_t^h$ is the share of profits from home firms that the agent holds stocks of.

As in Obstfeld and Rogoff (1995) (1996) citizens are not allowed to trade their shares of the firms. However they can trade real bonds in order to smooth their consumption. The international real bond market clears if $B_t^h = -B_t^f =: B_t$. Agents choose their labour supply, their consumption stream, their money holdings and their bond holdings.

The government's revenue results from seigniorage. We will assume throughout this analysis that the government balances its budget in each period.[1]

$$M_t^h - M_{t-1}^h = p_t^h \tau_t^h$$

The consumption side is identical to the formulation used by Obstfeld and Rogoff (1996).

### 5.2.2 Firms' Problem

As in the standard framework, we assume that companies are monopolistic competitors in the goods market. Each good $k$ is produced by firm $k$ only. Furthermore we assume that

[1] We do not really have to assume this. As long as the government spends all its revenue on transfers or buys the same consumption baskets as the economy's agents, Ricardian equivalence in the model ensures that a temporary deficit or surplus has no effect.

each company is a monopsonist in the labour market. This is one crucial assumption that leads to very different dynamics in our model under sticky wages compared to the standard framework. The dynamics under sticky prices is largely unaffected by this assumption. We believe that there is empirical evidence suggesting that this is a reasonable assumption. The market power is typically with the employers rather than with the employees (Manning 1995). Therefore, it can be misleading to shift the market power to the workers for modelling purposes.

For the price setting, we assume that producers can differentiate between foreign and home markets. The production function for an individual home firm $k$ takes the simple constant returns form

$y^{hh}(k) = L^{hh}(k)$ for the home market $h$ and

$y^{hf}(k) = L^{hf}(k)$ for the foreign (export) market $f$.

The firm $k$ maximises its profit $\pi^h(k)$, which depends not only on the prices it sets but also on the exchange rate $E$,

$$\max_{L^{hh},L^{hf}} \pi^h(k) = p^h(k)L^{hh}(k) + Ep^f(k)L^{hf}(k) - w^h\left(L^{hh}(k) + L^{hf}(k)\right),$$

subject to

| | |
|---|---|
| home goods demand: | $p^h(k) = p^h(k; L^{hh}(k))$, |
| foreign goods demand: | $Ep^f(k) = Ep^f(k; L^{hf}(k))$, |
| labour supply: | $w^h = w^h(L^{hh}(k) + L^{hf}(k))$. |

In the next section we solve the consumers' and producers' optimisation problem by assuming that both prices and wages are flexible.

## 5.3 Steady State Analysis

We analyse the steady state by assuming that all prices are flexible. Maximising the consumers' utility and the entrepreneurs' profits in this setting leads us to a system of equations that determines the equilibrium.

**Proposition 1** *The symmetric equilibrium of the economy is fully determined by the following eight equations and their foreign counterparts. (all variables besides the bond holdings are per capita)*

1. $C^h_{t+1}(z) = \left(\frac{1+r_{t+1}}{1+\delta}\right) C^h_t(z)$ *(consumption Euler equation),*

2. $\frac{M_t^h(z)}{p_t^h} = \chi C_t^h \frac{1+i_{t+1}^h}{i_{t+1}^h}$ , *where* $1 + i_{t+1}^h = \frac{p_{t+1}^h}{p_t^h}(1 + r_{t+1})$ *(money demand),*

3. $L_t^h = \frac{1}{\kappa}\frac{1}{C_t^h}\frac{w_t^h}{p_t^h}$ *(labour supply),*

4. $p_t^h = \left[np_t^h(h)^{1-\rho} + (1-n)p_t^h(f)^{1-\rho}\right]^{\frac{1}{1-\rho}}$ *(price index),*

5. $C_t^h = \frac{p_t^h(h)}{p_t^h}L^{hh} + \frac{Ep_t^f(h)}{p_t^h}L^{hf} + (1+r_t)\frac{B_t}{n} - \frac{B_{t+1}}{n}$ *(budget constraint),*

6. $L_t^{hh} = \left[\frac{p_t(h)}{p_t^h}\right]^{-\rho} nC_t^h$, $L_t^{hf} = \left(\frac{p_t^f(h)}{p_t^f}\right)^{-\rho}(1-n)C_t^f$
*(goods demand for home and export goods market),*

7. $L_t^h = L_t^{hh} + L_t^{hf}$ *(total labour demand),*

8. $L_t^{hh} = \left(2\frac{\rho}{\rho-1}\frac{w_t^h}{p_t^h}\right)^{-\rho} nC_t^h$, $L_t^{hf} = \left(2\frac{\rho}{\rho-1}\frac{w_t^h}{E_t p_t^f}\right)^{-\rho}(1-n)C_t^f$
*(labour demand for home and export goods market).*

This system of equations is almost identical to the system in Obstfeld and Rogoff (1996). The only differences occur in the labour supply and demand equation as well as in the goods supply equation. We give entrepreneurs monopsonistic power in the labour market, thereby reducing the labour demand by a factor of $2^\rho$. The reduced supply enables the entrepreneurs to charge a mark up that is double the one that Obstfeld and Rogoff (1996) find. Additionally we allow firms to discriminate in prices between home and foreign markets, i.e. they can choose the labour input that serves the domestic and export markets separately. The consumers' CES utility function leads to a simple mark up pricing by firms. A comparison of the goods and the labour demand functions (equation 6 and 8) shows that entrepreneurs always set prices that are higher by a factor of $(2\frac{\rho-1}{\rho})$ than the production costs. Since the costs of serving the two markets are determined by the home wage, the price firms charge in the two countries is the same. Effectively a Purchasing Power Parity (PPP) or a no arbitrage condition holds even though it has not been assumed ($Ep^f(h) = p^h(h)$). This fact is proven formally in the next lemma.

**Lemma 1** *Purchasing Power Parity ($p^h = Ep^f$) holds when prices and wages are flexible, even though firms could price discriminate.*

**Proof.** The firm's profit maximisation problem is given by
$\max_{L^h, L^{hh}} L^{hh}p^h(h) + (L^h - L^{hh})(p^f(h)E) - wL^h$

subject to

(1) inverse goods demands in both countries

$p^h(h) = \left(\frac{nC^h}{L^{hh}}\right)^{\frac{1}{\rho}} p^h$ and $p^f(h) = \left(\frac{nC^h}{L^{hh}}\right)^{\frac{1}{\rho}} p^f$ and

(2) labour supply function

$w^h = \frac{1}{k}\frac{C^h}{L^h}p^h.$

The first order conditions (FOC) are given by

$\left(p^h(h) - p^f(h)E\right) + L^{hh}\frac{\partial p^h(h)}{\partial L^{hh}} - L^{hf}E\frac{\partial p^f(h)}{\partial L^{hf}} = 0$

and

$p^f(h)E - w - L^h\frac{\partial w^h}{\partial L^h} = 0$

The assumption of the constant elasticity utility function ensures that the demand functions are isoelastic.

$\frac{\partial p^h(h)}{\partial L^{hh}}\frac{L^{hh}}{p^h(h)} = L^{hf}E\frac{\partial p^f(h)}{\partial L^{hf}}\frac{L^{hf}}{Ep^f(h)} = -\frac{1}{\rho}$

Substituting these relations into the second and third terms of the first FOC shows that the relative price that ensures the optimal allocation between foreign and home market, is given by

$p^h(h) = Ep^f(h).$

As long as the first FOC holds, firms set the same price in both markets. Since this holds for all individual prices it is also valid for the price indices. Hence, as long as prices are flexible, PPP holds even though it is not assumed.■

However, we will see in Lemma 5 presented in the next section that purchasing power parity need not hold if certain prices are sticky.

It is difficult to determine the steady state of the economy unless we assume that bond holdings are internationally balanced. Hence, we adopt the strategy of determining the symmetric steady state and later on log-linearise the system of equations of Proposition 1 around this steady state.

**Proposition 2** *The symmetric steady state in which the bond holdings are internationally balanced is given by*

1. $\bar{L}_0^h = \bar{L}_0^f = \bar{C}_0^h = \bar{C}_0^f = \sqrt{\frac{1}{k}\frac{1}{2}\frac{\rho-1}{\rho}},$

2. $\bar{r}_0 = \delta,$

3. $\bar{p}_0^h = \frac{\bar{M}_0^h}{\chi}\frac{1}{\bar{L}^h}\frac{\delta}{1+\delta} = \frac{\bar{M}_0^h}{\bar{M}_0^f}\bar{p}_0^f$,

4. $\bar{w}_0^h = \frac{1}{2}\frac{\rho-1}{\rho}\bar{p}_0^h = \frac{\bar{p}_0^h}{\bar{p}_0^f}\bar{w}_0^f$,

5. $\bar{E}_0 = \frac{\bar{p}_0^h}{\bar{p}_0^f} = \frac{\bar{M}_0^h}{\bar{M}_0^f}$.

Please consult the appendix of Chapter 5 for the proof of all major propositions.

The scale of production is reduced and the real wage is depressed due to the market imperfections inherent in monopolistic goods market and monopsonistic labour markets. The real interest rate is entirely determined by the exogenous time preference of the agents and the exchange rate solely depends on the relative money supply. Money is neutral in this economy and does not have any effect on real variables.

The mark up $\bar{p}_0^h = 2\frac{\rho}{\rho-1}\bar{w}_0^h$ in our model is twice as high as in Obstfeld and Rogoff (1996). Because companies are able to use their market power to set wages, they set them too low. This in turn leads to a lower scale of production by a factor of $\sqrt{2}$.

As mentioned earlier we log-linearise the model around the symmetric steady state. $\widehat{\overline{x}}$ approximates the percentage change from the symmetric steady state. We drop the subscript $t$ from all equations which apply only within a period.

**Lemma 2** *The log-linearized system of equations around the symmetric steady state with $B = 0$ is given by*

1. $\widehat{\overline{C}}_{t+1}^h = \widehat{\overline{C}}_t^h + \frac{\delta}{1+\delta}\hat{r}_{t+1}$ *(consumption Euler equation),*

2. $\hat{M}_t^h - \widehat{\overline{p}}_t^h = \widehat{\overline{C}}_t^h - \frac{\hat{r}_{t+1}}{1+\delta} - \frac{\hat{p}_{t+1}^h - \hat{p}_t^h}{\delta}$ *(money demand),*

3. $\widehat{\overline{L}}^h = -\widehat{\overline{C}}^h + \hat{w}^h - \hat{p}^h$ *(labour supply),*

4. $\widehat{\overline{p}}^h = n\widehat{\overline{p}}^h(h) + (1-n)\widehat{\overline{p}}^h(f)$ *(price index),*

5. $\widehat{\overline{C}}^h + \widehat{\overline{p}}^h = \widehat{\overline{L}}^h + n\widehat{\overline{p}}^h(h) + (1-n)\left(\widehat{\overline{p}}^h(f) + \widehat{\overline{E}}\right) + \frac{1}{n}\frac{\delta dB}{\overline{C}_0^h}$ *(budget constraint),*

6. $\widehat{\overline{L}}^{hh} = -\rho\left(\widehat{\overline{p}}^h(h) - \widehat{\overline{p}}^h\right) + \widehat{\overline{C}}^h, \widehat{\overline{L}}^{hh} = -\rho\left(\widehat{\overline{p}}^f(f) - \widehat{\overline{p}}^f\right) + \widehat{\overline{C}}^f$
*(goods demand for home and export market),*

7. $\widehat{\overline{L}}^h = n\widehat{\overline{L}}^{hh} + (1-n)\widehat{\overline{L}}^{hf}$ *(total labour demand),*

8. $\widehat{L}^{hh} = -\rho\left(\widehat{w}^h - \widehat{p}^h\right) + \widehat{C}^h$, $\widehat{L}^{hf} = -\rho\left(\widehat{w}^h - \widehat{p}^f - \widehat{E}\right) + \widehat{C}^f$
*(labour demand for home and export market).*

The log-linearisation allows us to understand the reaction of the economy to exogenous wealth and money shocks. We will use the equations later in order to determine the long run effects of monetary expansions if either wages or prices are sticky in the short term. For convenience we first determine the difference in the growth rates of domestic and foreign variables and only later determine the growth rates of individual countries' consumption and production.

**Proposition 3** *A one time redistribution of the bond holdings by dB does not affect aggregate world consumption or production but leads to the following permanent changes in home consumption, home employment, exchange rate and terms of trade.*

1. $\widehat{L}^w = \widehat{C}^w = 0$,
2. $\widehat{C}^h = \widehat{C}^w + (1-n)(\widehat{C}^h - \widehat{C}^f) = \frac{1+\rho}{2\rho}\frac{1}{n}\frac{\delta dB}{\overline{C}_0^h}$,
3. $\widehat{L}^h = \widehat{L}^w + (1-n)(\widehat{L}^h - \widehat{L}^f) = -\frac{1}{2}\frac{1}{n}\frac{\delta dB}{\overline{C}_0^h}$,
4. $\widehat{E} = \left[\hat{M}^h - \hat{M}^f\right] - \frac{1+\rho}{2\rho}\frac{1}{n(1-n)}\frac{\delta dB}{\overline{C}_0^h}$,
5. $\widehat{p}^h - \widehat{E} - \widehat{p}^f = \widehat{w}^h - \widehat{E} - \widehat{w}^f = \frac{1}{2\rho}\frac{1}{n(1-n)}\frac{\delta dB}{\overline{C}_0^h}$.

Home agents consume more as a reaction to an exogenous wealth transfer towards the home country. The extent of the increase in consumption depends positively on the substitutability of home and foreign goods. Consumption does not change as much as the income from bond holdings since agents also choose to work less. The home wage rises relative to the foreign wage and the exchange rate falls to lower the price of foreign goods at home and to increase the price of home goods abroad. Thus the foreign country is able to repay its debt. Not surprisingly, an exogenous change in the money supply does not affect any real variables. The exchange rate moves according to the relative money supply in the two countries.

## 5.4 Nominal Rigidities

So far we have kept prices and wages flexible and have found that a money supply shock has no real effect. It only alters the nominal prices, wages and the exchange rate. In other words, with flexible prices and wages, money is "neutral" and since a money shock does not change the dynamics, it is even "super-neutral".

This changes fundamentally if we assume a sluggish price adjustment. With sticky prices a money shock will not only affect the short-run real variables but will also cause the economy to settle in a different steady state. We will look at a situation where in period zero the economy is in the symmetric steady state as described by Proposition 2. In period one a monetary supply shock occurs but nominal wages/prices are held fixed for that period. In period two all nominal prices and wages adjust and the economy reaches its new steady state. The new steady state can be characterised by the new levels of bond holdings and money supplies ($B$, $M^h$, $M^f$).

We distinguish between three different types of price stickiness:

- nominal retail price stickiness,
- nominal wholesale price stickiness and
- nominal wage stickiness.

Retail prices are the prices that are paid by the consumers in the two countries. By wholesale prices we mean the prices the producers charge in their own currency.

We follow the methodology developed in Obstfeld and Rogoff (1995) in deriving the dynamic equilibrium with nominal rigidities. We log-linearise the system around the symmetric steady state to find out the short term dynamics and take into account the fact that certain prices are fixed between period zero and one. We denote the first order percentage change of a variable $x$ in the shock period by $\hat{x}$.

The economy reaches its new steady state in period two. As in the previous section we denote the percentage deviation between the new steady state and the original symmetric steady state by $\hat{\hat{x}}$. After the money shock at the beginning of period one, agents adjust their net international bond holdings $B$ immediately. From period two onwards all variables stay constant. Bond holdings do not change from period one to period two. Any steady state of the economy is fully characterised by the money supply and the international bond holdings (the only real state variables). Therefore, the steady state from period two onwards is the same as the steady state under flexible prices if

(1) the money supply changes in the same way, and

(2) the bond holdings are exogenously changed to the levels that endogenously arise under price stickiness.

If one knows the money shock and the endogenous redistribution of bonds, the change in period two can be fully characterised by the long run relationships in Proposition 3.

Because of intertemporal nature of the model, solving for the short-run involves also the long-run changes in the variables consumption $\hat{c}$, the price index $\hat{p}^h$ and the interest rate $\hat{r}$. The money demand depends on future price levels and agents want to smooth their consumption path. To determine the short-run changes we will hence need in addition to the equations in Lemma 3 the long-run budget constraint and the linearised long-run money demand equation from Lemma 2.

**Lemma 3** *For a given form of price/wage stickiness the log-linearized system of equations around the symmetric steady state with $B = 0$ is given by*

1. $\widehat{\overline{C}}^h = \hat{C}^h + \frac{\delta}{1+\delta}\hat{r}$ *(consumption Euler equation),*

2. $\hat{M}^h - \hat{p}^h = \hat{C}^h - \frac{\hat{r}}{1+\delta} - \frac{\widehat{\overline{p}}^h - \hat{p}^h}{\delta}$ *(money demand),*

3. $\hat{L}^h = -\hat{C}^h + \hat{w}^h - \hat{p}^h$ *(labour supply),*

4. $\hat{p}^h = n\hat{p}^h(h) + (1-n)\hat{p}^h(f)$ *(price index),*

5. $\hat{C}^h + \hat{p}^h = \hat{L}^h + n\hat{p}^h(h) + (1-n)\left(\hat{E} + \hat{p}^f(h)\right) - \frac{1}{n}\frac{dB}{C_0^h}$ *(budget constraint),*

6. $\hat{L}^{hh} = -\rho\left(\hat{p}^h(h) - \hat{p}^h\right) + \hat{C}^h$, $\hat{L}^{hf} = -\rho\left(\hat{p}^f(h) - \hat{p}^f\right) + \hat{C}^f$

   *(goods demand for home and foreign market),*

7. $\hat{L}^h = n\hat{L}^{hh} + (1-n)\hat{L}^{hf}$ *(total labour demand),*

8. *(labour demand equations are replaced by equations which* ***vary with the form of price stickiness****).*

The labour demand equation in lemma 2 is replaced by $\hat{p}^h(h) = \hat{p}^h(f) = 0$ in the case of sticky retail prices. Under sluggish wholesale prices, i.e. when prices are sticky in the producers' currency, the additional equation is given by $\hat{p}^h(h) = \hat{p}^f(f) = 0$. Similarly, if wages are sticky, it is given by $\hat{w}^h = \hat{w}^f = 0$.

The labour demand equation also varies depending on the form of price stickiness. With both forms of price stickiness, the monopolists always serve the goods demand as long as they earn a positive mark up. The monopolists need not be concerned that additional supply reduces the price. The labour demand, therefore, results directly from the goods demand equation. In the case of sticky prices, the labour demand is determined by the labour supply at this fixed wage.

Note that the budget constraint in the short-run differs from the long-run budget constraint. Fixing the prices or wages leads to a temporary change in real income which agents smooth by saving or dissaving in the international bond market.

Before we go on to Sections 4.1 - 4.3 to explicitly analyse the effect of monetary shocks under the three forms of price stickiness, we derive some qualitative result.

The nominal interest rates are the same in period one regardless of the form of price stickiness. Lemma 4 also shows that the inflation rate from period one to period two has to be the same in both countries.

**Lemma 4** *Both countries always face the same ex ante nominal interest rate $i^h = i^f$. Furthermore, they experience the same inflation rates between period one and period two. Thus*

$$\left(\hat{\overline{p}}^h - \hat{p}^h\right) = \left(\hat{\overline{p}}^f - \hat{p}^f\right).$$

**Proof.** In the steady state, the nominal interest rate coincides with the real interest rate. Both countries always face the same real interest rate. This is also true in the shock period. Hence, taking the difference between the home and foreign consumption Euler equations, we conclude that the consumption differentials are constant in time. Thus it is

$$\hat{\overline{C}}^h - \hat{\overline{C}}^f = \hat{C}^h - \hat{C}^f.$$

Subtracting the difference of the home and foreign long run money demands from the short term money demand differential, we find that the short run differential of nominal interest rates is given by

$$\hat{\imath}^h - \hat{\imath}^f = (1+\delta)\left[\left(\hat{p}^h - \hat{p}^f\right) - \left(\widehat{\overline{p}}^h - \widehat{\overline{p}}^f\right)\right].$$

Given the definition of the nominal interest rate, the differential of interest rate changes is given by

$$\hat{\imath}^h - \hat{\imath}^f = -\tfrac{(1+\delta)}{\delta}\left[\left(\hat{p}^h - \hat{p}^f\right) - \left(\widehat{\overline{p}}^h - \widehat{\overline{p}}^f\right)\right].$$

A comparison of the last two equations gives the result.■

The next lemma analyses whether PPP, which holds under flexible prices, still applies when price stickiness is assumed.

**Lemma 5** *In the long run, purchasing power parity ($p^h = Ep^f$) holds under any form of price stickiness. In the short run, it still holds under sticky wholesale prices and under sticky wages but not under sticky retail prices.*

**Proof.** In the long run, firms can adjust their prices and the result that PPP holds under flexible prices applies (Lemma 1). If prices are not flexible, the first order condition becomes irrelevant in the short term. Nevertheless, it is true that PPP holds under sticky wholesale prices. The argument is as follows. PPP holds in the initial steady state because prices and wages are flexible. In the shock period, the relative price of the same goods in the home and the foreign market moves only with the exchange rate. Hence, the no arbitrage condition continues to hold for each good and, therefore, also for the price levels.

This is obviously not true under fixed retail prices because the exchange rate moves in the shock period ($\hat{E} \neq 0$). It is intuitively easy to understand why the exchange jumps under sticky retail prices. Under sticky retail prices, the price of consumption stays constant in the shock period. There is no substitution between home and foreign goods. Hence, production is the same in both countries. Now, suppose the exchange rate would not move. This would imply that home and foreign agents have the same real income and, therefore, there is no international borrowing. Consequently, they both consume the same amount. Both also face the same nominal interest rates (Lemma 4). Given all these symmetries, they would demand the same amount of real money. This cannot be an equilibrium because the money supply differs. (For an explicit proof see Proposition 6).■

These two lemmas allow us to show that whether interest rate parity holds and the exchange rate overshoots depends on the price stickiness assumed.

**Proposition 4** *Uncovered nominal interest rate parity holds under sticky wholesale prices and sticky wages but is violated under sticky retail prices.*

**Proof.** The linearised interest rate parity $(1 + i_{t+1}^h = \frac{E_{t+1}}{E_t}\left(1 + i_{t+1}^f\right))$ in the shock period is given by

$\hat{\imath}^h - \hat{\imath}^f = \frac{(1+\delta)}{\delta}\left[\widehat{\overline{E}} - \hat{E}\right].$

Given the definition of the nominal interest rate, the differential of interest rate changes is given by

$\hat{\imath}^h - \hat{\imath}^f = \frac{(1+\delta)}{\delta}\left[\left(\widehat{\overline{p}}^h - \widehat{\overline{p}}^f\right) - \left(\hat{p}^h - \hat{p}^f\right)\right].$

Under sticky whole sale and sticky wages, we replace the price differentials by the exchange rate changes because PPP holds (Lemma 1 and 5). This proves the first part of the proposition.

Under sticky retail prices, the nominal interest rate differential can be written as

$\hat{\imath}^h - \hat{\imath}^f = \frac{(1+\delta)}{\delta}\left[\widehat{\overline{E}}\right]$, since PPP is valid in the long run.

This shows that interest parity would only hold if the exchange rate does not change in the first period ($\hat{E} = 0$). This is not the case as the proof of the previous lemma shows.■

**Proposition 5** *While the exchange rate overshoots its long run value under sticky retail prices, it immediately reaches its new steady state value under sticky wholesale prices as well as under sticky wages.*

**Proof.** Since the interest rate parity holds under sticky wholesale prices and wages (Proposition 4) and the nominal interest rates are the same (Lemma 4), it must be true that in these cases $\widehat{\overline{E}} = \hat{E}$.

From the proof of Lemma 5 for sticky retail prices, we know that the exchange rate jumps in the shock period ($\hat{E} \neq 0$). Additionally, we know that the nominal interest rate differential and the long run exchange rate is given by (Lemma 4, Proposition 4)

$0 = \hat{\imath}^h - \hat{\imath}^f = \frac{(1+\delta)}{\delta}\left[\widehat{\overline{E}}\right].$

The long run exchange rate coincides with the initial exchange rate. This completes the proof.■

Intuitively, the exchange rate has to return to its original level under sticky retail prices since in both steady states PPP holds and inflation from period zero to period two is the same in both countries. From period zero to period one, inflation is zero due to retail price stickiness. Furthermore, both countries experience the same inflation rate from period one to period two as shown in Lemma 4.

In the following three subsections we analyse the dynamics of the model in more detail assuming in turn one of the three prices to be sticky.

### 5.4.1 Sticky Retail Prices

In this subsection we assume that prices are sticky in the consumers' currency ($\hat{p}^h(h) = \hat{p}^f(h) = \hat{p}^f(f) = \hat{p}^h(f) = 0$). These four equation together with Lemma 3 allow us to calculate explicitly the dynamics of the two countries' economies if one or both of them expand their money supply. Specifically, we can analyse spillovers of one country's money expansion on production and consumption abroad.

**Proposition 6** *Under sticky retail prices, money supply shocks give rise to an endogenous change in international net bond holdings given by*

$\frac{dB}{C_0^h} = \frac{2\rho}{(1+\rho)\delta} n (1-n) \left[\hat{M}^h - \hat{M}^f\right].$

*Changes in each country's consumption, production, exchange rates and terms of trade are given by*

- *in the short-run*

  $\hat{C}^h = \hat{M}^h,$

  $\hat{L}^h = \hat{M}^w = n\hat{M}^h + (1-n)\hat{M}^f,$

  $\hat{E} = \left(1 + \frac{2\rho}{(1+\rho)\delta}\right)\left[\hat{M}^h - \hat{M}^f\right],$

  $\hat{w}^h - \hat{E} - \hat{w}^f = -\frac{2\rho}{(1+\rho)\delta}\left[\hat{M}^h - \hat{M}^f\right],$

  $\hat{r} = -\left(\frac{1+\delta}{\delta}\right)\hat{M}^w,$

- *in the long-run*

  $\widehat{\overline{C}}^h = (1-n)\left[\hat{M}^h - \hat{M}^f\right],$

  $\widehat{\overline{L}}^h = -\frac{\rho}{(1+\rho)}(1-n)\left[\hat{M}^h - \hat{M}^f\right],$

  $\widehat{\overline{E}} = 0,$

$$\widehat{\overline{p}}^h(h) - \widehat{\overline{E}} - \widehat{\overline{p}}^f(f) = \widehat{\overline{p}}^h(h) - \widehat{\overline{p}}^f(f) = \tfrac{1}{1+\rho}\left[\hat{M}^h - \hat{M}^f\right],$$
$$\widehat{\overline{w}}^h - \hat{E} - \widehat{\overline{w}}^f = \tfrac{1}{1+\rho}\left[\hat{M}^h - \hat{M}^f\right],$$
$$\hat{p}^h = n\hat{M}^h + (1-n)\hat{M}^f.$$

To grasp the intuition more easily, let us consider the special case that there is a money expansion at home while the foreigners keep their money supply constant.

Money holding and consumption at home increase by the same degree since consumers' preferences are homothetic between real money holding and consumption. Note the relative price between real money and consumption is equal to one by definition. Consumers do not substitute between different products since the retail prices stay the same. The additional income which is necessary to afford the higher consumption comes from two sources. First, a positive money shock reduces real prices. At lower real prices, consumers demand more goods and producers, having lost their price setting power, are willing to meet the demand. This leads to lower deadweight losses and higher consumer surplus. Second, the exchange rate jump allows domestic exporters to earn more from their exports. They sell their products at the same foreign retail price and convert the revenues into the home currency at a more favourable exchange rate. Their income increases in real terms since the domestic consumer prices do not change.

A money shock at home affects the foreign economy as well. Whereas the reduction of monopolistic distortions generates some additional consumer surplus, the second source of income is just a redistribution from foreign consumers to home consumers. For foreigners, who export to the home country, an increase in the exchange rate reduces their returns in the foreign currency. Consumers in the home country do not only demand more home-produced goods but also more foreign made products (by the same degree). They do not substitute between home-made goods and imported goods since the retail prices are fixed. Higher demand for foreign goods combined with sticky prices reduces the monopolistic deadweight loss abroad as well. Consequently, production increases abroad too. The percentage increase in production is the same in both countries. This is due to the absence of substitution between the goods. More production at home and abroad might suggest higher income for foreigner too. However, as indicated above, the large jump in the exchange rate diminishes their real revenues from exporting to the home country. Their exports measured in terms of the number of goods increases but their revenues in their own foreign currency decline.

This redistributional effect makes the foreigners worse off. In equilibrium they have to work harder in order to export more goods but their real revenues decline. In summary, an unexpected money expansion at home is beneficial for home citizens but it reduces welfare for foreigners. Therefore, in a world with sticky retail prices, a central bank always has an incentive to increase money supply.[2] This explains the well known strategy "beggar thy neighbour" conducted by many industrialised countries in the beginning of this century. All countries increased their money base in order to profit from the others (Nurkse 1944).

The better off home citizens try to smooth their additional income and, therefore, buy bonds from foreigners at a low real interest rate. This allows foreigners to keep their consumption constant. From period two onwards the prices adjust and, hence, each monopolist will restrain its output in order to achieve higher prices. The trade balance surplus of the home country in period one leads to a trade balance deficit from period two onwards since foreigners have to pay interest for the borrowed amount. Consequently, foreigners have to produce more and consume less in the long run, whereas home citizen enjoy lower production and higher consumption.

Reduced production at home makes home-produced goods relatively more scarce and, thus, improves the terms of trade for the home country in the long run. The exchange rate displays very strong short term volatility. It jumps up in the short-run but comes back to its original level in period two. After a money expansion at home, the change in the exchange rate exceeds the change in the money supply. In other words, the exchange rate increase is larger than in the case of flexible prices. This overshooting is in line with the seminal work of Dornbusch (1976). And indeed Dornbusch (1976) also assumes sticky retail prices. In period two the exchange rate bounces back to its original level. This seems surprising given the fact that the home money supply is higher in the new steady state. Sticky prices, therefore, explain the excess volatility observed for exchange rates.

### 5.4.2 Sticky Wholesale Prices

Whereas in the former section the prices were fixed for the consumers, with sticky wholesale prices the prices are fixed in the producer's currency. Just like in the last section, entrepreneurs, knowing that they have no influence on the price, are happy to meet the

[2]Note that this effect is mitigated if agents expect the central bank to increase the money supply.

additional demand. Nevertheless, the implications of monetary shocks are fundamentally different. Similar to the case of sticky retail prices, the labour demand is replaced by a fixed nominal prices in producers' currency ($\hat{p}^h(h) = \hat{p}^f(f) = 0$). As the proof of Lemma 5 shows, PPP holds both for the price indices ($\hat{p}^h = \hat{E} + \hat{p}^f$) and for individual goods prices ($\hat{p}^h(h) = \hat{E} + \hat{p}^f(h)$, $\hat{p}^h(f) = \hat{E} + \hat{p}^f(f)$). This implies a very different reaction of the economy to a monetary shock.

**Proposition 7** *Under sticky wholesale prices money supply shocks give rise to an endogenous change in international net bond holdings given by*

$\frac{dB}{C_0^h} = \frac{2(\rho-1)}{(\rho+1)\delta+2} n(1-n) \left[\hat{M}^h - \hat{M}^f\right].$

*Changes in each country's consumption, production, exchange rates and terms of trade are given by*

- *in the short- run*

$$\hat{C}^h = \underbrace{\frac{\rho\left[\delta(\rho+1)+2n\right]-(1-n)\left[(1+\rho)\delta\right]}{\rho\left[(\rho+1)\delta+2\right]}}_{<1} \hat{M}^h + (1-n)\underbrace{\frac{2\rho+(\rho+1)\delta}{\rho\left[(\rho+1)\delta+2\right]}}_{<1} \hat{M}^f,$$

$$\hat{L}^h = \underbrace{\frac{\rho\left[\delta(\rho+1)+2n\right]+(1-n)\left[2\rho^2\right]}{\rho\left[(\rho+1)\delta+2\right]}}_{>1} \hat{M}^h \underbrace{-(1-n)\frac{2(\rho-1)}{(\rho+1)\delta+2}}_{<0} \hat{M}^f,$$

$$\hat{E} = \frac{\delta(\rho+1)+2\rho}{\rho((1+\rho)\delta+2)}\left[\hat{M}^h - \hat{M}^f\right] = \hat{\overline{E}},$$

$$\hat{p}^h(h) - \hat{E} - \hat{p}^f(f) = -\hat{E},$$

$$\hat{w}^h - \hat{E} - \hat{w}^f = \frac{2\rho^2(\delta+1)+(\rho-1)\delta}{\rho((\rho+1)\delta+2)}\left[\hat{M}^h - \hat{M}^f\right],$$

$$\hat{C}^w = \hat{L}^w = \hat{M}^w,$$

$$\hat{r} = -\left(\frac{1+\delta}{\delta}\right)\hat{M}^w,$$

- *in the long-run*

$$\hat{\overline{C}}^h = \frac{1+\rho}{\rho}\frac{(\rho-1)\delta}{(\rho+1)\delta+2}(1-n)\left[\hat{M}^h - \hat{M}^f\right],$$

$$\hat{\overline{L}}^h = -\frac{(\rho-1)\delta}{(\rho+1)\delta+2}(1-n)\left[\hat{M}^h - \hat{M}^f\right],$$

$$\left[\hat{\overline{p}}^h(h) - \hat{\overline{E}} - \hat{\overline{p}}^f(f)\right] = \left[\hat{\overline{w}}^h - \hat{\overline{E}} - \hat{\overline{w}}^f\right] = \frac{(\rho-1)\delta}{\rho((\rho+1)\delta+2)}\left[\hat{M}^h - \hat{M}^f\right],$$

$$\hat{\overline{E}} = \left(\frac{(\rho+1)\delta+2\rho}{\rho((\rho+1)\delta+2)}\right)\left[\hat{M}^h - \hat{M}^f\right] = \hat{E},$$

$$\hat{\overline{C}}^w = \hat{\overline{L}}^w = 0.$$

Providing home country consumers with more money stimulates their demand for home-produced and foreign-produced goods. Whereas the consumer price for the home-made products is fixed for one period, the retail price for imported good changes with the exchange rate. The exchange rate goes up because the increased demand of foreign products raises the demand of foreign currency as well. This makes imported foreign products for home consumers more expensive and, thus, they will substitute them partly with home-made products.

An increase in home money supply affects home consumers' income in three ways. First, the higher demand for home-produced goods combined with fixed prices reduces the monopoly distortions and, thus, increases production and real income for consumers. Second, the increase in the exchange rate leads to higher export revenues. For given fixed wholesale prices, it makes home-produced goods relatively cheaper for foreigners. This boosts the number of exported goods. Third, the increased exchange rate also makes imported goods more expensive which not only leads to the above described substitution effect but also to a negative real income effect. The overall income effect on home consumption is positive.

Abroad, an increase in the exchange rate makes products from the home country cheaper as well. Therefore, even abroad foreign-produced commodities become less popular. This combined with the decline in export explains why production goes down. The higher the elasticity of substitution $\rho$, the larger the impact on foreign production. A lower level of production reduces their real income. On the other hand, foreigners' profit from lower import prices, resulting in lower inflation. In period one foreigners want to enjoy the low import prices reflected by the favourable terms of trade. Therefore, they sell real bonds to the home citizens, consume more, and work less in period one. In the long-run they have to pay interest to the home citizens. Therefore, in the new steady state they have to produce more and consume less in comparison to the original steady state. The terms of trade increase since home-produced goods are more scarce in the long-run. Nominal prices increase at home in period two and the exchange rate increase is smaller than in the case of flexible prices. Note that the higher the elasticity of substitution between the goods, the smaller is the adjustment in the terms of trade through the exchange rate.

### 5.4.3 Sticky Wages

The labour demand equation is replaced by an assumption of fixed wages ($\hat{w}^h = \hat{w}^f = 0$). These equations together with Lemma 3 allow us to determine the dynamics explicitly. As under sticky wholesale prices, PPP holds again. In contrast to the fixed prices scenarios the scale of production is determined by the labour supply rather than by the goods demand. This has important implications specifically for the current account dynamics.

**Proposition 8** *Under sticky wages money supply shocks give rise to an endogenous change in international net bond holdings given by*

$$\frac{dB}{C_0^h} = -\frac{2(\rho-1)}{(1+\rho)\delta+2}n(1-n)\left[\hat{M}^h - \hat{M}^f\right].$$

*Changes in each country's consumption, production, exchange rates and terms of trade are given by*

- *in the short- run*

$$\hat{C}^h = \underbrace{-\left(n + (1-n)\frac{\rho-1}{\rho}\frac{(\rho+1)\delta}{(\rho+1)\delta+2}\right)}_{<0}\hat{M}^h \underbrace{-(1-n)\left(1-\frac{\rho-1}{\rho}\frac{(\rho+1)\delta}{(\rho+1)\delta+2}\right)}_{<0}\hat{M}^f,$$

$$\hat{L}^h = -\hat{M}^h,$$

$$\hat{E} = \left(1+\frac{\rho-1}{\rho}\frac{(\rho+1)\delta}{(\rho+1)\delta+2}\right)\left[\hat{M}^h - \hat{M}^f\right] = \hat{\hat{E}},$$

$$\hat{p}^h(h) - \hat{E} - \hat{p}^f(f) = \frac{1}{\rho}\left[\hat{M}^h - \hat{M}^f\right],$$

$$\hat{w}^h - \hat{E} - \hat{w}^f = \left(1+\frac{\rho-1}{\rho}\frac{(\rho+1)\delta}{(\rho+1)\delta+2}\right)\left[\hat{M}^h - \hat{M}^f\right],$$

$$\hat{L}^w = \hat{c}^w = -\hat{M}^w,$$

$$\hat{r} = \left(\frac{1+\delta}{\delta}\right)\hat{M}^w,$$

- *in the long run*

$$\hat{\hat{C}}^h = -\frac{\rho-1}{\rho}\frac{(\rho+1)\delta}{(\rho+1)\delta+2}(1-n)\left[\hat{M}^h - \hat{M}^f\right],$$

$$\hat{\hat{L}}^h = \frac{(\rho-1)\delta}{(\rho+1)\delta+2}(1-n)\left[\hat{M}^h - \hat{M}^f\right],$$

$$\hat{\hat{E}} = \left(1+\frac{\rho-1}{\rho}\frac{(\rho+1)\delta}{(\rho+1)\delta+2}\right)\left[\hat{M}^h - \hat{M}^f\right] = \hat{E},$$

$$\left[\hat{\hat{p}}^h(h) - \hat{\hat{E}} - \hat{\hat{p}}^f(f)\right] = \left[\hat{\hat{w}}^h - \hat{\hat{E}} - \hat{\hat{w}}^f\right] = -\frac{(\rho-1)}{\rho(\rho+1)\delta+2\rho}\left[\hat{M}^h - \hat{M}^f\right],$$

$$\hat{\hat{L}}^w = \hat{\hat{C}}^w = 0.$$

Increasing home money supply causes upward price pressure at home. Due to the stickiness of nominal wages, higher consumer prices result in lower real wages. Workers, therefore, substitute consumption for leisure and work fewer hours. The resulting contraction in the production of home-made products has at least two effects. First, it reduces the income for home citizens. In expectation of higher future income, they try to borrow from abroad and, therefore, push up the interest rate. Second, home-produced goods become more expensive. Consumers substitute them for imported foreign products. More demand for foreign products and, hence, foreign currency results in a higher exchange rate.

Though a high exchange rate should make imported home-produced goods cheaper abroad, the opposite happens because the price $p^h(h)$ skyrockets. The calculations of the terms of trade highlight this. Consequently, foreign consumers also substitute home-produced goods with foreign goods. Higher demand for their foreign products and higher prices for the imported goods increases their price index too. Foreigners reduce their consumption in favour of more savings. They lend a larger amount to the home citizens. The high real interest rates in period one makes it worthwhile for them to reduce their consumption but to keep their production constant, even though the real wages decline abroad too.

From period two onwards, foreigners will receive interest payments in the form of home-produced goods. Therefore, in the long run production at home has to increase whereas consumption declines. The opposite is true abroad. Note that the setting with sticky wages replicates the empirical regularity known as the J-curve effect. It is often claimed that after an exchange rate appreciation the trade balance becomes negative for a while before bouncing back and leading to a long-run trade balance surplus. In period one the exchange rate and the trade balance are negatively correlated. However, the terms of trade and the trade balance are positively correlated.

## 5.5 Comparing Different Forms of Stickiness

The formal analysis in Section 4 demonstrates that different forms of price stickiness lead to strikingly different economic outcomes. In this section we compare the implication of a monetary shock for the case of sticky retail prices, sticky wholesale prices and sticky wages. Empirical evidence might then suggest which form of stickiness seems most plausible.

We restrict ourselves to the case of a positive money expansion in the home country. Due to the symmetry between both countries, the effects of a positive money supply shock abroad would mirror the effects. Similarly, a contraction of the money supply leads to the opposite effects. Given different price stickiness, a monetary shock does not only affect nominal variables differently but also affects real variables and the whole dynamics of the economy. The following figure illustrates the impulse response functions triggered by an unexpected money expansion at home.

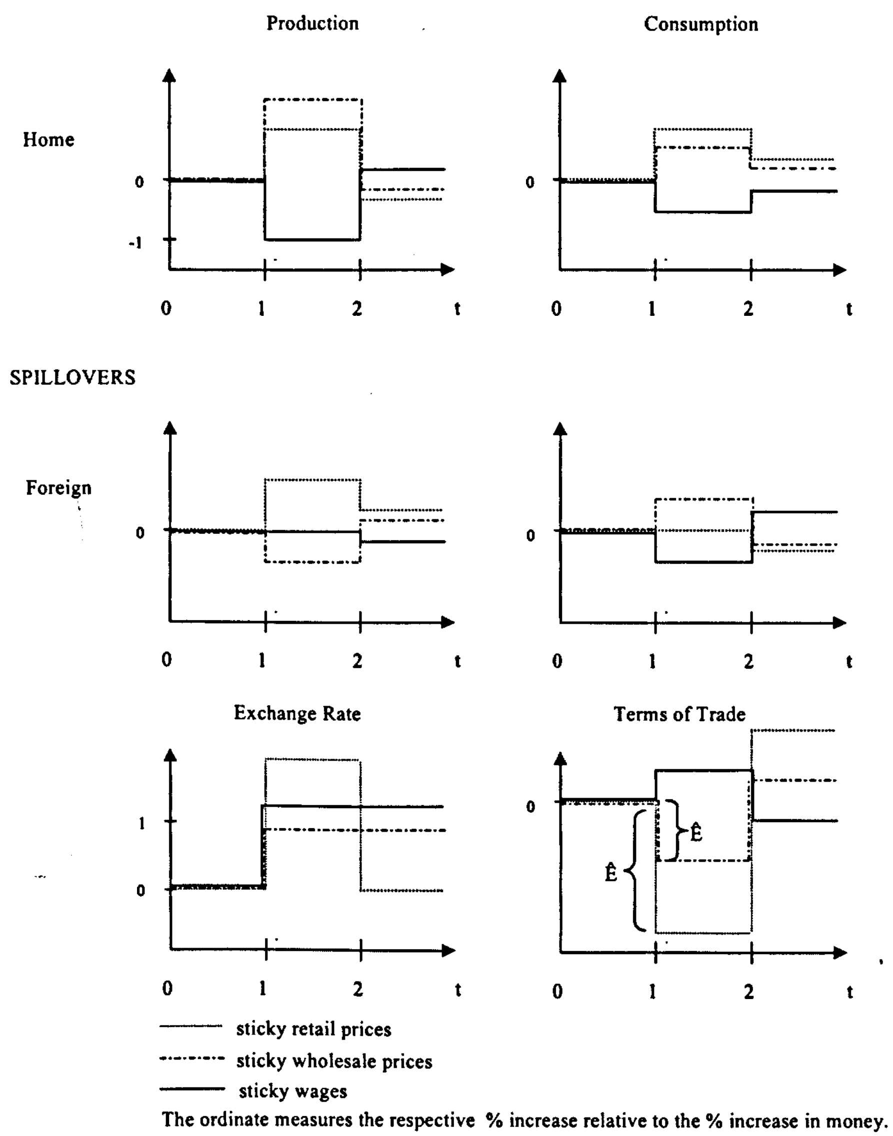



Figure 5.1: Impulse Response Functions

A monetary expansion under sticky retail prices leads to more production at home as well as to an equal increase in production abroad. This is the case since sticky retail prices prevent any substitution between home and foreign goods. With sticky wholesale prices home-produced goods become more expensive relative to foreign-made ones in both countries. The resulting substitution leads to a sharp increase in production at home and a reduction abroad.

The absolute level of production is not only affected by the substitution effect but also by the real income effects due to the unexpected money expansion at home. In both cases, consumers profit from a reduction in the monopolistic distortion due to price stickiness. The crucial distinction between both settings is that with sticky retail prices the large spike in the exchange rate boosts home citizens' revenue from exporting goods, while the consumer prices stay constant. This increases the real income of home citizens at the expense of the foreigners. With sticky wholesale prices nominal export revenues increase too, but so do the consumer prices for the important goods. This explains why in period one consumption by the home citizens increases more when retail prices are sluggish. They are also able to build up a larger trade balance surplus under sticky retail prices. This guarantees that the long-run consumption rise also exceeds the increase in the case of sticky wholesale prices.

The second row of graphs illustrates how the monetary expansion spills over to the foreign country. In the case of sticky retail prices, foreign producers meet the increased home demand. Nevertheless, their export revenue declines in their own foreign currency since the exchange rate increases. The money expansion has a negative wealth impact to foreigners. Since the interest rate is lower due to higher world production, foreigners sell international bonds in order to keep up with the consumption level that they are used to. In short, foreigners have to work harder, become debtors and consequently consume less in the long-run. The 'beggar thy neighbour' strategy is surely optimal in a setting with sluggish retail prices. On the other hand, sticky wholesale prices allow the foreigners to work less and consume more in period one. They can enjoy part of the additional consumer surplus due to reduced monopolistic distortions.

Whereas the effects for the two different forms of price stickiness are demand driven, the effects due to sticky wages are governed by the supply side. The economic implications of sticky wages are in sharp contrast to the outcomes of the other two settings. As outlined in Section 4.3, a money expansion and its resulting price increase leads to lower real wages. Workers work less and production declines. Consequently, they have to reduce their consumption in the short-run and in the long-run. Fewer home-made products and, thus, higher consumer prices reduce consumption abroad too. Nevertheless, foreigners keep up with their production stimulated by higher interest rates. They achieve a current account surplus which leads to a long run increase in consumption. Interestingly, the current account surplus for the foreigners is exactly the same size as in the case of sticky wholesale prices where foreigners suffer a current account deficit.

In all three forms of price sluggishness, the size of the spillovers effects depends on the size of the home country. The model predicts that smaller countries are more vulnerable to money supply shocks of neighbouring large countries than larger countries.

However, the size of the countries has however no impact on the dynamics of the exchange rate or the terms of trade. The third row of graphs shows that with sticky retail prices, the exchange rate skyrockets in period one and surprisingly comes back to its original level in period two. For the other two forms of stickiness, the exchange rate changes only once. Under flexible prices and wages, the exchange rate moves by the same degree as the money supply and the terms of trade are not affected. The terms of trade $\left[\widehat{\overline{p}}^h(h) - \widehat{\overline{E}} - \widehat{\overline{p}}^f(f)\right]$ represent the number of foreign goods one would receive in exchange for one home-produced good. Sticky wages lead to a larger exchange rate movement than sticky wholesale prices. The reason is that the exchange rate does not only accommodate the relative increase in the money supplies but, since money is not neutral, it also helps the terms of trade to adjust. In the case of sticky wholesale prices the terms of trade increase in the long run, whereas they decrease under sticky wages in the long-run. Under sticky wholesale prices, the home country becomes the net creditor. Its terms of trade have to deteriorate and the exchange rate jumps by less than the money supply. Under sticky wages, the terms of trade have to move in the home country's favour. The nominal exchange rate jumps more than the money supply. An alternative definition of the terms of trade - which measure the competitiveness of domestic products abroad - is best understood by the relative scarcity of the products. Hence, this definition follows immediately from the production activities in both countries. Both definitions coincide only as long as PPP holds. With fixed retail prices, this is not the case in period one.

## 5.6 Conclusion

The main message of this paper is that the form of price stickiness matters. Given the empirical regularities like the violation of PPP in the short run and of the uncovered interest rate parity etc., it seems plausible that the stickiness of retail prices is very important. Retail price stickiness leads to the large spillover effects and reinstates the "beggar thy neighbour" policy. The sticky retail price analysis also suggests that there should be an international coordination of monetary policy.

Some further extensions are left for future research. It would be interesting to extend the analysis to a setting where monetary shocks occur with positive probabilities. An analysis along the lines of Obstfeld and Rogoff (1999) seems promising. We did not cover the case of asymmetric forms of price stickiness, such as when whole sale prices are sticky in the home country while abroad retail prices do not adjust. Some interesting insights might emerge from such an analysis. Introducing productivity shocks bundled with a certain form of price stickiness might lead to slightly different results, especially when the monetary policy cannot adjust immediately and lags the productivity shocks. Another worthwhile extension would be to find an appropriate empirical test that allows us to discriminate between different forms of price stickiness and to empirically estimate their relative importance.

## 5.7 Appendix of Chapter 5

### 5.7.1 Proof of Proposition 2

Let us assume that labour and consumption are identical in the two countries. The consumption Euler equation as usual determines the real interest rate

$$r = \frac{1-\beta}{\beta} =: \delta.$$

The budget constraint in the symmetric steady state is given by

$$C_t^h = \frac{p_t^h(h)}{p_t^h} L^{hh} + \frac{E p_t^f(h)}{p_t^h} L^{hf}.$$

Since the no arbitrage condition holds, it simplifies to

$$C_t^h = \frac{p_t^h(h)}{p_t^h} L^h.$$

The labour market equilibrium and the world goods market equilibrium imply

$$L^{hh} + L^{hf} = L^h = L^f = L^{ff} + L^{fh}$$

and

$$L^{hh} + \frac{(1-n)}{n}L^{fh} = C^h = C^f = L^{ff} + \frac{n}{(1-n)}L^{hf}.$$

The last two equations imply that

$$nL^{hf} = (1-n)L^{fh}.$$

Since the capital account is balanced by assumption the current account has to be balanced

$$nL^{hf}Ep^f(h) - (1-n)L^{fh}p^h(f) = 0.$$

which implies that the terms of trade are zero

$$p^h(h) - p^f(f)E = 0.$$

This implies for the price index that

$$p^h = p^h(h).$$

The labour supply equation together with the mark up formula and the budget constraint implies the scale of production

$$L^h = \sqrt{\frac{1}{k}\frac{1}{2}\frac{\rho-1}{\rho}} = L^f.$$

The money demand equation is given by

$$p^h = \frac{M^h}{\chi}\frac{1}{L^h}\frac{\delta}{1+\delta}$$

Dividing this by the foreign equivalent leads to

$$E = \frac{p^h}{p^f} = \frac{p^h(h)}{p^f(f)} = \frac{M^h}{M^f}.$$

### 5.7.2 Proof of Proposition 3

Taking the differences of the linearised equations of home and foreign variables allows us to write these as a function of the exogenous wealth transfer $dB$.

1. $\widehat{p}^h - \widehat{E} - \widehat{p}^f = \widehat{w}^h - \widehat{E} - \widehat{w}^f = \frac{1}{2\rho}\frac{1}{n(1-n)}\frac{\delta dB}{\overline{C}_0^h},$

2. $\widehat{C}^h - \widehat{C}^f = \frac{1+\rho}{2\rho}\frac{1}{n(1-n)}\frac{\delta dB}{\overline{C}_0^h},$

3. $\widehat{L}^h - \widehat{L}^f = -\frac{1}{2}\frac{1}{n(1-n)}\frac{\delta dB}{\overline{C}_0^h},$

4. $\widehat{E} = \left[\hat{M}^h - \hat{M}^f\right] - \frac{1+\rho}{2\rho}\frac{1}{n(1-n)}\frac{\delta dB}{\overline{C}_0^h}.$

Adding the labour supply functions weighted by the country size and using the price levels leads to

$$\widehat{L}^w := n\widehat{L}^h + (1-n)\widehat{L}^f = -n\widehat{C}^h - (1-n)\widehat{C}^f = -\widehat{C}^w.$$

Since world production and world consumption has to be equal it follows that

$\widehat{\overline{L}}^w = \widehat{\overline{C}}^w = 0.$

The changes of consumption and labour are derived from

$\widehat{\overline{C}}^h = \widehat{\overline{C}}^w + (1-n)(\widehat{\overline{C}}^h - \widehat{\overline{C}}^f) = \frac{1+\rho}{2\rho}\frac{1}{n}\frac{\delta dB}{\overline{C}_0^h},$

$\widehat{\overline{L}}^h = \widehat{\overline{L}}^w + (1-n)(\widehat{\overline{L}}^h - \widehat{\overline{L}}^f) = -\frac{1}{2}\frac{1}{n}\frac{\delta dB}{\overline{C}_0^h}.$

### 5.7.3 Proof for Short Term World Changes

Adding the consumption Euler equations weighted by the country size leads to

$\hat{C}^w = -\frac{\delta}{1+\delta}\hat{r}.$

Calculate the world long term and short term money demand functions

$\hat{M}^w := n\hat{M}^h + (1-n)\hat{M}^f = \widehat{\overline{C}}^w + n\widehat{\overline{p}}^h + (1-n)\widehat{\overline{p}}^f$ (long term),

$\hat{M}^w + \frac{1}{\delta}\left(n\widehat{\overline{p}}^h + (1-n)\widehat{\overline{p}}^f\right) - \hat{C}^w = \left(\frac{\delta+1}{\delta}\right)\left(n\hat{p}^h + (1-n)\hat{p}^f\right) - \frac{\hat{r}}{1+\delta}$(short term).

Substituting the long term relationship into the short term one leads to

$\left(\frac{\delta+1}{\delta}\right)\hat{M}^w - \hat{c}^w = \left(\frac{\delta+1}{\delta}\right)\left(n\hat{p}^h + (1-n)\hat{p}^f\right) - \frac{\hat{r}}{1+\delta}.$

This relationship can be used to determine the short term growth rates of world consumption in the three cases.

- sticky wages

  Use the labour supply to replace the short term price changes

  $\left(\frac{\delta+1}{\delta}\right)\hat{M}^w - \hat{C}^w = \left(\frac{\delta+1}{\delta}\right)\left(-\hat{C}^w - \hat{L}^w\right) + \frac{\hat{C}^w}{\delta},$

  and finally since $\hat{C}^w = \hat{L}^w$,

  $\hat{C}^w = -\hat{M}^w.$

- sticky retail prices

  retail prices do not change in the short term, hence

  $\left(\frac{\delta+1}{\delta}\right)\hat{M}^w - \hat{C}^w = \frac{\hat{C}^w}{\delta}$ or

  $\hat{C}^w = \hat{M}^w.$

- sticky wholesale prices

  $\left(\frac{\delta+1}{\delta}\right)\hat{M}^w - \hat{C}^w = \left(\frac{\delta+1}{\delta}\right)\left(n\hat{p}^h + (1-n)\hat{p}^f\right) + \frac{\hat{C}^w}{\delta},$

  and, hence, again

  $\hat{C}^w = \hat{M}^w.$

### 5.7.4 Proof of Proposition 6

We first subtract the foreign short term equilibrium equations from their home counterparts using Lemma 3. We do not impose sticky retail prices at this stage because we will use these equations in the proofs for sticky wholesale prices and sticky wages. Therefore, we have

$$\left(\hat{L}^h - \hat{L}^f\right) = -\rho\left[\left(\hat{p}^h - \hat{p}^f\right) + \left(\hat{p}^f(h) - \hat{p}^h(f)\right)\right] \text{ (demand)},$$

$$\left(\hat{C}^h - \hat{C}^f\right) - \left(\hat{L}^h - \hat{L}^f\right) + \frac{1}{n(1-n)}\frac{dB}{C_0^h} = \left[-\hat{p}^h(f) + \hat{p}^f(h) + \hat{E}\right] \text{ (budget constraint)},$$

$$\left(\hat{M}^h - \hat{M}^f\right) - \left(\hat{p}^h - \hat{p}^f\right) = \left(\hat{C}^h - \hat{C}^f\right) - \tfrac{1}{\delta}\left(\widehat{\overline{p}}^h - \widehat{\overline{p}}^f\right) + \tfrac{1}{\delta}\left(\hat{p}^h - \hat{p}^f\right) \text{ (money demand)},$$

$$\left(\hat{C}^h - \hat{C}^f\right) = \left(\widehat{\overline{C}}^h - \widehat{\overline{C}}^f\right) \text{(consumption Euler equation)},$$

$$\left(\hat{p}^h - \hat{p}^f\right) = -\left(\hat{C}^h - \hat{C}^f\right) - \left(\hat{L}^h - \hat{L}^f\right) + \left(\hat{w}^h - \hat{w}^f\right) \text{ (labour supply)}.$$

Additionally we need the difference between the long term budget constraints and the long term money demand equations for the reasons outlined in section 4. We use the fact that PPP always holds in the long run (Lemma 1). Thus,

$$\left(\widehat{\overline{C}}^h - \widehat{\overline{C}}^f\right) - \left(\widehat{\overline{L}}^h - \widehat{\overline{L}}^f\right) - \frac{1}{n(1-n)}\frac{\delta dB}{C_0^h} = \left[-\widehat{\overline{p}}^h(f) + \widehat{\overline{E}} + \widehat{\overline{p}}^f(h)\right] \text{ (budget constraint)},$$

$$\left(\hat{M}^h - \hat{M}^f\right) - \widehat{\overline{E}} = \left(\widehat{\overline{C}}^h - \widehat{\overline{C}}^f\right) \text{(money demand)},$$

$$\left(\widehat{\overline{L}}^h - \widehat{\overline{L}}^f\right) = -\rho\left[-\widehat{\overline{p}}^h(f) + \widehat{\overline{E}} + \widehat{\overline{p}}^f(h)\right] \text{ (long term demand)}.$$

Under the sticky retail price scenario, we know from the proof of Proposition 5 that the exchange rate does not change in the long run ($\widehat{\overline{E}} = 0$). From the long run money demand equation and the consumption Euler equation, we conclude that the change in both periods consumption is proportional to the change in the money supply

$$\left(\widehat{\overline{C}}^h - \widehat{\overline{C}}^f\right) = \left(\hat{C}^h - \hat{C}^f\right) = \left(\hat{M}^h - \hat{M}^f\right).$$

Substituting this last equation and the long run demand equation into the long run budget constraint we arrive at

$$\left(\hat{M}^h - \hat{M}^f\right) - \frac{1}{n(1-n)}\frac{\delta dB}{C_0^h} = (1-\rho)\left[-\widehat{\overline{p}}^h(f) + \widehat{\overline{E}} + \widehat{\overline{p}}^f(h)\right].$$

Using the expression for the long term change in the terms of trade that is given in Proposition 3, we can derive the change in net international bond holdings..

$$\frac{dB}{C_0^h} = \frac{2\rho}{(1+\rho)\delta} n(1-n)\left(\hat{M}^h - \hat{M}^f\right).$$

Substituting this equation into the equations of Proposition 3 we can calculate all the long run changes of the variables.

For the differences in the short run, we see from the short term demand function that under sticky retail prices their is no substitution between foreign and home goods. Thus,

$\left(\hat{L}^h - \hat{L}^f\right) = 0.$

Using the relative short term changes in consumption, price levels and production it is easy to see from the labour supply that

$(\hat{w}^h - \hat{w}^f) = \left(\hat{M}^h - \hat{M}^f\right).$

We can now derive the short term change in the exchange rate given the short term budget constraint .

Having derived the differences in short run changes abroad and at home we use the change in world aggregates, given by appendix 5.7.3 to calculate the changes in the individual countries. The methodology is the same as in the proof of Proposition 3.■

### 5.7.5 Proof of Proposition 7

We again use the differences of the short and long run changes derived at the beginning of the proof for sticky retail prices. Under sticky wholesale prices, we can make use of the results that PPP also holds in the short run and that the exchange rate immediately reaches its long term value ($\hat{E} = \hat{\overline{E}}$).

Substituting the goods and money demand equation into the budget constraint, both for the long and short run we derive

$\left(\hat{M}^h - \hat{M}^f\right) - \hat{E} = (\rho - 1)\hat{E} - \frac{1}{n(1-n)}\frac{dB}{C_0^h}$ (short term budget),
$\left(\hat{M}^h - \hat{M}^f\right) - \hat{E} = \frac{1+\rho}{2\rho}\frac{1}{n(1-n)}\frac{\delta dB}{C_0^h}$ (long term).

From these two equations we derive the change in the international bond holdings and the change in the exchange rate.

$\hat{E} = \left(\frac{\delta(\rho+1)+2\rho}{\rho((1+\rho)\delta+2)}\right)\left[\hat{M}^h - \hat{M}^f\right],$
$\frac{dB}{C_0^h} = \frac{2(\rho-1)}{(\rho+1)\delta+2}n(1-n)\left[\hat{M}^h - \hat{M}^f\right].$

Just like in the sticky retail price scenario we can derive all the long run changes using Proposition 3.

We can derive the short term difference in production from the short term demand equation using the expression for the exchange rate. Thus,

$\hat{L}^h - \hat{L}^f = \left(\frac{\delta(\rho+1)+2\rho}{(1+\rho)\delta+2}\right)\left[\hat{M}^h - \hat{M}^f\right].$

The short term difference in consumption can then be read from the short term budget constraint.

$\hat{C}^h - \hat{C}^f = \left(\frac{\rho^2-1}{\rho((1+\rho)\delta+2)}\delta\right)\left[\hat{M}^h - \hat{M}^f\right].$

Finally, the relative change in wages can be calculated using the labour supply equation. $\hat{w}^h - \hat{E} - \hat{w}^f = \left(\frac{2\rho^2(\delta+1)+\delta(\rho-1)}{\rho((1+\rho)\delta+2)}\right)\left[\hat{M}^h - \hat{M}^f\right]$.

Having derived the differences in short run changes abroad and at home, we use the change in world aggregates, given by Appendix 5.7.3 to calculate the changes in the individual countries. The methodology is the same as in the proof of Proposition 3.■

### 5.7.6 Proof of Proposition 8

We again use the differences in short term changes that have been derived at the beginning of the proof for changes under sticky retail prices. Just like under sticky wholesale prices, we can make use of the facts that PPP holds in the short run and that the exchange rate does not overshoot (Proposition 5). The crucial difference under sticky wages is that the scale of production is determined by the labour supply rather than by the demand.

Using the differences in the long run money demand equation and the short run labour supply equations, we can derive the short term change in labour. Thus,

$\left(\hat{M}^h - \hat{M}^f\right) = \left[\hat{c}^h - \hat{c}^f\right] + \hat{E} = -\left(\hat{L}^h - \hat{L}^f\right)$.

The short run terms of trade change can be read from the difference in the short term goods demand equation. Thus,

$\left(\hat{M}^h - \hat{M}^f\right) = \rho\left(\hat{p}^h(h) - \widehat{E} - \hat{p}^f(f)\right)$.

The difference between the two short term budget constraints leads to

$\left(\hat{c}^h - \hat{c}^f\right) = -\frac{\rho-1}{\rho}\left(\hat{M}^h - \hat{M}^f\right) - \frac{1}{n(1-n)}\frac{dB}{\bar{c}_0^h}$.

The difference between the long run budget constraints can be written as

$\left(\hat{c}^h - \hat{c}^f\right) = \frac{1+\rho}{2\rho}\frac{1}{n(1-n)}\bar{r}_0\frac{dB}{\bar{c}_0^h}$.

We derive the change in the bond holdings and the change in consumption, by substituting the last two equations into each other. Thus,

$\frac{dB}{\bar{C}_0^h} = -\frac{2(\rho-1)}{(1+\rho)\delta+2}n(1-n)\left[\hat{M}^h - \hat{M}^f\right]$,

$\widehat{\bar{C}}^h - \widehat{\bar{C}}^f = -\left(\frac{\rho-1}{\rho}\frac{1}{\delta+2\frac{\rho}{\rho+1}}\delta\right)\left[\hat{M}^h - \hat{M}^f\right]$.

Just like in the sticky price scenarios, the long term changes can now be calculated using Proposition 3.

The change in the exchange rate can be read from the long run money demand equation using the change in consumption. It is

$\widehat{\bar{E}} = \left(1 + \frac{\rho-1}{\rho}\frac{1}{\delta+2\frac{\rho}{\rho+1}}\delta\right)\left[\hat{M}^h - \hat{M}^f\right]$.

Having derived the differences in short run changes abroad and at home, we use the change in world aggregates, given by Appendix 5.7.3 to calculate the changes in the individual countries. The methodology is the same as in the proof of Proposition 3.■

# Bibliography


ADMATI, A. R. (1985): "A Noisy Rational Expectations Equilibrium for Multi-Asset Securities Markets," *Econometrica*, 53, 629–657.

ADMATI, A. R. (1989): "Information in Financial Markets: The Rational Expectations Approach," in *Financial Markets and Incomplete Information*, ed. by S. Bhattacharya, and G. M. Constantinides, pp. 139–152. Rowman and Littlefield Publishers, Inc., Totowa, NJ.

ADMATI, A. R., and P. PFLEIDERER (1986): "A Monopolisitic Market for Information," *Journal of Economic Theory*, 39, 400–438.

ADMATI, A. R., and P. PFLEIDERER (1988): "A Theory of Intraday Patterns: Volume and Price Variability," *Review of Financial Studies*, 1(1), 3–40.

ADMATI, A. R., and P. PFLEIDERER (1990): "Direct and Indirect Sale of Information," *Econometrica*, 58, 901–928.

AKERLOF, G. A. (1970): "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, 84, 488–500.

ALLEN, B. (1981): "A Class of Monotone Economies in which Rational Expectations Equilibria exist but Prices do not Reveal all Information," *Economics Letters*, 7, 227–232.

ALLEN, B. (1982): "Strict Rational Expectations Equilibria with Diffuseness," *Journal of Economic Theory*, 27, 20–46.

ALLEN, B. (1985): "The Existence of Fully Rational Expectations Approximate Equilibria with Noisy Price Observations," *Journal of Economic Theory*, 37, 213–253.

ALLEN, F., and D. GALE (1992): "Stock-Price Manipulation," *Review of Financial Studies*, 5(3), 503–529.

ALLEN, F., and D. GALE (1994): *Financial Innovation and Risk Sharing.* MIT Press, Cambridge, MA.

ALLEN, F., and G. GORTON (1992): "Stock Price Manipulation, Market Microstructure and Asymmetric Information," *European Economic Review*, 36(2-3), 624–630.

ALLEN, F., S. MORRIS, and A. POSTLEWAITE (1993): "Finite Bubbles with Short Sale Constraints and Asymmetric Information," *Journal of Economic Theory*, 61, 206–229.

ANDERSON, R. M., and H. SONNENSCHEIN (1982): "On the Existence of Rational Expectations Equilibrium," *Journal of Economic Theory*, 26, 261–278.

ANDERSON, T. W. (1984): *An Introduction to Multivariate Statistical Analysis.* John Wiley & Sons, New York, Chapter 2.

AUMANN, R. J. (1976): "Agreeing to Disagree," *Annals of Statistics*, 4, 1236–1239.

AUMANN, R. J. (1987): "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica*, 55, 1–18.

AUSUBEL, L. M. (1990): "Partially-Revealing Rational Expectations Equilibria in a Competitive Economy," *Journal of Economic Theory*, 50, 93–126.

AVERY, C., and P. ZEMSKY (1995): "Multi-dimensional Uncertainty and Herd Behavior in Financial Markets," *Working Paper.*

BACK, K. (1992): "Insider Trading in Continuous Time," *Review of Financial Studies*, 5, 387–409.

BACK, K., H. H. CAO, and G. A. WILLARD (1997): "Imperfect Competition among Informed Traders," *Journal of Finance*, (forthcoming).

BACKUS, D. K., P. J. KEHOE, and F. E. KYDLAND (1994): "Dynamics of the Trade Balance and the Terms of Trade: The J-Curve?," *American Economic Review*, 84(1), 84–103.

BAGNOLI, M., S. VISWANATHAN, and C. HOLDEN (1994): "On the Existence of Linear Equilibria in Models of Market Making," *Discussion Paper 9421, Duke University.*

BANERJEE, A. V. (1992): "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 107, 797–817.

BENABOU, R., and G. LAROQUE (1992): "Using Privileged Information to Manipulate Markets: Insiders, Gurus, and Credibility," *Quarterly Journal of Economics*, 107(3), 921–958.

BERGEMANN, D., and J. VÄLIMÄKI (1996): "Market Experimentation and Pricing," *Cowles Foundation Discussion Paper, Yale University.*

BHATTACHARYA, S., and B. L. LIPMAN (1995): "Ex ante versus Interim Rationality and the Existence of Bubbles," *Economic Theory*, 6(3), 469–494.

BHATTACHARYA, U., and M. SPIEGEL (1991): "Insider, Outsider, and Market Breakdown," *Review of Financial Studies*, 4, 255–282.

BIAIS, B., D. MARTIMORT, and J.-C. ROCHET (1997): "Competition Between Strategic Market Makers," *Mimeo, University of Toulouse.*

BIAIS, B., and J.-C. ROCHET (1997): "Risk Sharing, Adverse Selection and Market Structure," in *Financial Mathematics*, ed. by B. Biais, T. Björk, J. Cvitanić, N. El Karoui, and M. Quenez, pp. 1–51. Springer Verlag, Berlin.

BIKHCHANDANI, S., D. HIRSHLEIFER, and I. WELCH (1992): "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy*, 100, 992–1026.

BLUME, L., and D. EASLEY (1990): "Implementation of Walrasian Expectations Equilibria," *Journal of Economic Theory*, 51, 207–227.

BLUME, L., D. EASLEY, and M. O'HARA (1994): "Market Statistics and Technical Analysis: The Role of Volume," *Journal of Finance*, 49(1), 153–181.

BOLTON, P., and C. HARRIS (1993): "Strategic Experimentation," *STICERD TE 93/261, London School of Economics.*

BRANDENBRUGER, A., and B. POLAK (1996): "When Managers Cover their Posteriors: Making the Decisions the Market Wants to See," *Rand Journal of Economics*, 27, 523–541.

BRANDENBURGER, A. (1992): "Knowledge and Equilibrium in Games," *Journal of Economic Perspectives*, 6, 83–101.

BRAY, M. M. (1982): "Learning, Estimation, and the Stability of Rational Expectations," *Journal of Economic Theory*, 26, 318–339.

BRAY, M. M., and D. KREPS (1987): "Rational Learning and Rational Expectations," in *Arrow and the Ascent of Modern Economic Theory*, ed. by G. R. Feiwel, pp. 597–625. New York University Press, New York.

BRENNAN, M. J. (1990): "Latent Assets," *Journal of Finance*, 45, 709–730.

BROWN, D. P., and R. H. JENNINGS (1989): "On Technical Analysis," *Review of Financial Studies*, 2, 527–552.

BRUNNERMEIER, M. K. (1997): *Prices, Price Processes, Volume and Their Information - A Survey of the Market Microstructure Literature -*. forthcoming Oxford University Press.

BRUNNERMEIER, M. K. (1998): "Buy on Rumors - Sell on News: A Manipulative Trading Strategy," *Mimeo, London School of Economics.*

CAMERER, C. (1995): "Individual Decision Making," in *The Handbook of Experimental Economics*, ed. by J. H. Kagel, and A. E. Roth, pp. 587–703. Princeton University Press, Princeton, New Jersey.

CAMPBELL, J. Y., A. W. LO, and A. C. MACKINLAY (1997): *The Econometrics of Financial Markets.* Princeton University Press, Princeton, New Jersey, USA.

CANZONERI, M. B., and P. MINFORD (1986): "When International Policy Coordination Matters: An Empirical Analysis," *CEPR Discussion Paper 119.*

CHAKRABORTY, A. (1997): "Information and Manipulation in Financial Markets," *Mimeo, Carnegie Mellon University.*

CHAMLEY, C., and D. GALE (1994): "Information Revelation and Strategic Delay in a Model of Investment," *Econometrica*, 62, 1065–1085.

CHO, I. K., and D. M. KREPS (1987): "Signalling Games and Stable Equilibria," *Quarterly Journal of Economics*, 102, 179–221.

CLARIDA, R. H., and J. GALI (1994): "Sources of Real Exchange Rate Fluctuations: How Important are Nominal Shocks?," *Carnegie-Rochester Conference Series on Public Policy*, 41(December), 1–56.

COPELAND, T. E. (1976): "A Model of Asset Trading under the Assumption of Sequential Information Arrival," *Journal of Finance*, 31, 1149–1168.

DE LONG, J. B., A. SHLEIFER, L. H. SUMMERS, and R. J. WALDMANN (1990a): "Noise Trader Risk in Financial Markets," *Journal of Political Economy*, 91, 929–956.

DE LONG, J. B., A. SHLEIFER, L. H. SUMMERS, and R. J. WALDMANN (1990b): "Positive Feedback Investment Strategies and Destabilizing Rational Speculation," *Journal of Finance*, 45, 379–395.

DEBONDT, W. F. M., and R. H. THALER (1995): "Financial Decision-Making in Markets and Firms: A Behavioral Perspective," in *Handbooks in Operations Research and Management Science*, ed. by R. A. Jarrow, V. Maksimovic, and W. T. Ziemba, pp. 385–410. North Holland, Amsterdam, Volume 9, Chapter 13.

DEBREU, G. (1959): *Theory of Value.* Wiley, New York, Cowles Foundation Monograph, Vol. 17.

DEKEL, E., and F. GUL (1997): "Rationality and Knowledge in Game Theory," in *Advances in Economics and Econometrics: Theory and Applications, Seventh World Congress*, ed. by D. M. Kreps, and K. F. Wallis, vol. 1, pp. 87–172. Cambridge University Press, Cambridge.

DENNERT, J. (1994): "Price Competition between Market Makers," *Review of Economic Studies*, 60, 735–751.

DIAMOND, D. W., and R. E. VERRECCHIA (1981): "Information Aggregation in a Noisy Rational Expectations Economy," *Journal of Financial Economics*, 9, 221–235.

DORNBUSCH, R. (1976): "Expectations and Exchange Rate Dynamics," *Journal of Political Economy*, 84(February), 141–153.

DUBEY, P., J. GEANAKOPLOS, and M. SHUBIK (1987): "The Revelation of Information in Strategic Market Games: A Critique of Rational Expectations Equilibrium," *Journal of Mathematical Economics*, 16, 105–137.

DUFFIE, D., and R. RAHI (1995): "Financial Market Innovation and Security Design: An Introduction," *Journal of Economic Theory*, 65, 1–42.

EASLEY, D., and M. O'HARA (1987): "Price, Trade Size, and Information in Securities Markets," *Journal of Financial Economics*, 19, 69–90.

EASLEY, D., and M. O'HARA (1992): "Time and the Process of Security Price Adjustment," *Journal of Finance*, 47, 577–605.

EICHENBAUM, M., and C. EVANS (1995): "Some Empirical Evidence on the Effects of Monetary Policy Shocks on Real Exchange Rates," *Quarterly Journal of Economics*, 110(November), 975–1009.

ENGEL, C., and J. H. ROGERS (1996): "How Wide is the Border?," *American Economic Review*, 86(5), 1112–1125.

FAMA, E. F. (1970): "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance*, 25, 383–417.

FAMA, E. F. (1976): *Foundations of Finance.* Basic Books, Inc., New York.

FELDMAN, M., and C. GILLES (1985): "An Expository Note on Individual Risk without Aggregate Uncertainty," *Journal of Economic Theory*, 35, 26–32.

FISHMAN, M. J., and K. M. HAGERTY (1995): "The Mandatory Disclosure of Trades and Market Liquidty," *Review of Financial Studies*, 8(3), 637–676.

FLEMING, W. H., and R. W. RISHEL (1975): *Deterministic and Stochastic Optimal Control.* Springer-Verlag, New York.

FOSTER, D. F., and S. VISWANATHAN (1990): "A Theory of Interday Variations in Volume, Variance, and Trading Costs in Security Markets," *Review of Financial Studies*, 3, 593–624.

FOSTER, D. F., and S. VISWANATHAN (1994): "Strategic Trading with Asymmetrically Informed Traders and Long-Lived Information," *Journal of Financial and Quantitative Analysis*, 29(4), 499–518.

FOSTER, D. F., and S. VISWANATHAN (1996): "Strategic Trading When Agents Forecast the Forecasts of Others," *Journal of Finance*, 51(4), 1437–1478.

FRIEDMAN, J. W., and J.-F. THISSE (1994): "Boundedly Rational Nash Equilibrium: A Probabilistic Choice Approach," *Mimeo.*

FROOT, K. A., D. S. SCHARFSTEIN, and J. C. STEIN (1992): "Herd on the Street: Informational Inefficiencies in a Market with Short-Term Speculation," *Journal of Finance*, 47, 1461–1484.

FUDENBERG, D., and J. TIROLE (1991): *Game Theory.* MIT Press, Cambridge, Massachusetts, USA.

GALE, D. (1996): "What Have We Learned from Social Learning?," *European Economic Review*, 40, 617–628.

GARBER, P. (1990): "Famous First Bubbles," *Journal of Economic Perspectives*, 4, 35–54.

GEANAKOPLOS, J. (1994): "Common Knowledge," in *Handbook of Game Theory II*, ed. by R. J. Aumann, and S. Hart, pp. 1437–1496. Elsevier Science B.V.

GENNOTTE, G., and H. LELAND (1990): "Market Liquidity, Hedging, and Crashes," *American Economic Review*, 80(5), 999–1021.

GERVAIS, S. (1995): "Market Microstructure With Uncertain Information: A Multi-Period Analysis," *Working Paper.*

GLOSTEN, L. (1989): "Insider Trading, Liquidity, and the Role of the Monopolist Specialist," *Journal of Business*, 62, 211–235.

GLOSTEN, L. R. (1994): "Is the Electronic Open Limit Order Book Inevitable," *Journal of Finance*, 69, 1127–1161.

GLOSTEN, L. R., and P. R. MILGROM (1985): "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders," *Journal of Financial Economics*, 14, 71–100.

GROSSMAN, S. J. (1976): "On the Efficiency of Competitive Stock Markets Where Traders Have Diverse Information," *Journal of Finance*, 31, 573–585.

GROSSMAN, S. J. (1978): "Further Results on the Informational Efficiency of Competitive Stock Markets," *Journal of Economic Theory*, 18, 81–101.

GROSSMAN, S. J. (1981): "An Introduction to the Theory of Rational Expectations Under Asymmetric Information," *Review of Economic Studies*, 48, 541–559.

GROSSMAN, S. J. (1988): "An Analysis of the Implications for Stock and Futures price Volatility of Program Tading and Dynamic Hedging Strategies," *Journal of Business*, 61, 275–298.

GROSSMAN, S. J. (1995): "Dynamic Asset Allocation and the Informational Efficiency of Markets," *Journal of Finance*, 50(3), 773–787.

GROSSMAN, S. J., and J. E. STIGLITZ (1980): "On the Impossibility of Informationally Efficient Markets," *American Economic Review*, 70, 393–408.

GROSSMAN, S. J., and Z. ZHOU (1996): "Equilibrium Analysis of Portfolio Insurance," *Journal of Finance*, 51(4), 1379–1403.

GRUNDY, B. D., and M. MCNICHOLS (1989): "Trade and Revelation of Information through Prices and Direct Disclosure," *Review of Financial Studies*, 2, 495–526.

GUL, F., and R. LUNDHOLM (1995): "Endogenous Timing and the Clustering of Agents' Decisions," *Journal of Political Economy*, 103, 1039–1066.

HARRIS, C., M. STINCHCOMBE, and W. ZAME (1997): "Resolving the Closure Problem in Normal Form Games," *Mimeo*.

HARRISON, M., and D. KREPS (1978): "Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations," *Quarterly Journal of Economics*, 89, 519–542.

HAU, H. (1998): "Real Exchange Rate Volatility and Economic Openess: Theory and Evidence," *Mimeo, Princeton University.*

HAYEK, F. (1945): "The Use of Knowledge in Society," *American Economic Review*, 35, 519–530.

HE, H., and J. WANG (1995): "Differential Information and Dynamic Behavior of Stock Trading Volume," *Review of Financial Studies*, 8(4), 919–972.

HELLWIG, M. F. (1980): "On the Aggregation of Information in Competitive Markets," *Journal of Economic Theory*, 22, 477–498.

HELLWIG, M. F. (1982): "Rational Expectations Equilibrium with Conditioning on Past Prices: A Mean-Variance Example," *Journal of Economic Theory*, 26, 279–312.

HELLWIG, M. F. (1992): "Fully Revealing Outcomes in Signalling Models: An Example of Nonexistence When the Type Space is Unbounded," *Journal of Economic Theory*, 58, 93–104.

HEY, J. D. (1997): "Experiments and the Economics of Individual Decision Making under Risk and Uncertainty," in *Advances in Economics and Econometrics: Theory and Applications*, ed. by D. M. Kreps, and K. F. Wallis, pp. 173–205. Cambridge University Press, Cambridge.

HIRSHLEIFER, D., A. SUBRAHMANYAM, and S. TITMAN (1994): "Security Analysis and Trading Patterns When Some Investors Receive Information Before Others," *Journal of Finance*, 49, 1665–1698.

HIRSHLEIFER, J. (1971): "The Private and Social Value of Information and the Reward to Inventive Activity," *American Economic Review*, 61, 561–574.

HO, T., and H. R. STOLL (1981): "Optimal Dealer Pricing Under Transactions and Return Uncertainty," *Journal of Financial Economics*, 9, 47–73.

HOLDEN, C. W., and A. SUBRAHMANYAM (1992): "Long-Lived Private Information and Imperfect Competition," *Journal of Finance*, 47, 247–270.

HOLDEN, C. W., and A. SUBRAHMANYAM (1994): "Risk Aversion, Imperfect Competition and Long-Lived Information," *Economics Letters*, 44, 181–190.

HOLDEN, C. W., and A. SUBRAHMANYAM (1996): "Risk Aversion, Liquidity, and Endogenous Short Horizons," *Review of Financial Studies*, 9, 691–722.

HUDDART, S., J. S. HUGHES, and C. LEVINE (1998): "Public Disclosure of Insider Trades, Trading Costs and Price Discovery," *mimeo, Duke University.*

JACKSON, M. O. (1991): "Equilibrium, Price Formation and the Value of Private Information," *Review of Financial Studies*, 4, 1–16.

JENNINGS, R. H., L. T. STARKS, and J. C. FELLINGHAM (1981): "An Equilibrium Model of Asset Trading with Sequential Information Arrival," *Journal of Finance*, 36, 143–161.

JORDAN, J. S. (1983): "On the Efficient Markets Hypothesis," *Econometrica*, 51, 1325–1343.

JORDAN, J. S., and R. RADNER (1982): "Rational Expectations in Microeconomic Models: An Overview," *Journal of Economic Theory*, 26, 201–223.

KAHNEMAN, D., and A. TVERSKY (1979): "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, XLVII, 263–291.

KEYNES, J. M. (1936): *The General Theory of Employment, Interest and Money.* Macmillan, London.

KREPS, D. (1977): "A Note on Fulfilled Expectations Equilibria," *Journal of Economic Theory*, 14, 32–43.

KUMAR, P., and D. J. SEPPI (1992): "Futures Manipulation with "Cash Settlement"," *Journal of Finance*, 47(4), 1485–1502.

KYLE, A. S. (1985): "Continuous Auctions and Insider Trading," *Econometrica*, 53, 1315–1335.

KYLE, A. S. (1989): "Informed Speculation with Imperfect Competition," *Review of Economic Studies*, 56, 317–356.

LAFFONT, J.-J. (1985): "On the Welfare Analysis of Rational Expectations Equilibria with Asymmetric Information," *Econometrica*, 53(1), 1–29.

LAFFONT, J.-J., and E. S. MASKIN (1990): "The Efficient Market Hypothesis and Insider Trading on the Stock Market," *Journal of Political Economy*, 98, 70–93.

LEACH, J. C., and A. MADHAVAN (1993): "Price Experimentation and Security Market Structure," *Review of Financial Studies*, 6, 375–404.

LEE, I. H. (1993): "On the Convergence of Informational Cascades," *Journal of Economic Theory*, 61, 395–411.

LEROY, S. F., and R. D. PORTER (1981): "The Present Value Relation: Tests Based on Implied Variance Bounds," *Econometrica*, 64, 555–574.

LIPSTER, R. S., and A. N. SHIRYAYEV (1977): *Statisitcs of Random Processes I, II.* Springer-Verlag, Berlin.

MADHAVAN, A. (1992): "Trading Mechanisms in Securities Markets," *Journal of Finance*, 47, 607–641.

MANNING, A. (1995): "How Do We Know That Real Wages Are Too High," *Quarterly Journal of Economics*, 110(4), 1111–1125.

MARCET, A., and T. J. SARGENT (1989a): "Convergence of Least Squares Learning in Environments with Hidden State Variables and Private Information," *Journal of Political Economy*, 83, 1113–1144.

MARCET, A., and T. J. SARGENT (1989b): "Convergence of Least Squares Learning Mechanism in Self Referential Linear Stochastic Models," *Journal of Economic Theory*, 48, 337–368.

MARIN, J. M., and R. RAHI (1996): "Information Revelation and Market Incompleteness," *Working Paper, University of London.*

MAS-COLELL, A. (1980): "Noncooperative Approaches to the Theory of Perfect Competition: Presentation," *Journal of Economic Theory*, 22, 121–135.

MCALLISTER, P. H. (1990): "Rational Behavior and Rational Expectations," *Journal of Economic Theory*, 52, 332–363.

MCKIBBIN, W. J., and J. D. SACHS (1991): *Global Linkages: Macroeconomic Interdependence and Cooperations in the World Economy.* Brookings Institution, Washington, DC.

MEESE, R. A., and K. ROGOFF (1983): "Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample," *Journal of International Economics*, 14(February), 3–24.

MILGROM, P., and N. STOKEY (1982): "Information, Trade and Common Knowledge," *Journal of Economic Theory*, 26, 17–27.

MORRIS, S. (1994): "Trade with Heterogeneous Prior Beliefs and Assymmetric Information," *Econometrica*, 62, 1327–1347.

MORRIS, S., A. POSTLEWAITE, and H. S. SHIN (1995): "Depth of Knowledge and the Effect of Hihger Order Uncertainty," *Economic Theory*, 6, 453–467.

MUNDELL, R. A. (1961): "A Theory of Optimum Currency Area," *American Economic Review*, 51(September), 657–665.

MUNDELL, R. A. (1963): "Capital Mobility and Stabilization Policy Under Fixed and Flexible Exchange Rates," *Canadian Journal of Economics and Political Science*, 29(May), 421–431.

MUSSA, M. (1976): "The Exchange Rate, The Balance of Payments, and Monetary and Fiscal Policy under a Regime of Controlled Floating," *Scandinavian Journal of Economics*, 78(May), 229–248.

MUSSA, M. (1979): "Macroeconomic Interdependence and the Exchange Rate Regime," in *International Economic Policy*, ed. by R. Dornbusch, and J. A. Frenkel. The Johns Hopkins University Press, Baltimore, Maryland.

MUSSA, M. (1986): "Nominal Exchange Rate Regimes and the Behavior of Real Exchange Rates: Evidence and Implications," *Carnegie-Rochester Series on Public Policy*, 25(Autumn), 117–214.

MUTH, J. R. (1960): "Optimal Properties of Exponentially Weighted Forecasts," *Journal of the American Statistical Association*, 55, 229–305.

MUTH, J. R. (1961): "Rational Expectations and the Theory of Price Movements," *Econometrica*, 24, 315–335.

NEEMAN, Z., and G. O. OROSEL (1998): "Herding and the Winner's Curse in Markets with Sequential Bids," *Mimeo, University of Vienna.*

NÖLDEKE, G. (1993): *Lecture Notes in Financial Economics.* Mimeo, Princeton University.

NURKSE, R. (1944): *International Currency Experience: Lessons of the Interwar Period.* League of Nations, Geneva, Switzerland.

OBSTFELD, M., and K. ROGOFF (1995): "Exchange Rate Dynamics Redux," *Journal of Political Economy*, 103(31), 624–660.

OBSTFELD, M., and K. ROGOFF (1996): *Foundations of International Macroeconomics.* MIT Press, Cambridge, Massachusetts, USA.

OBSTFELD, M., and K. ROGOFF (1999): "Exchange Rate Risk," *Mimeo, Princeton University.*

O'HARA, M. (1995): *Market Microstructure Theory.* Blackwell Publishers, Cambridge, Massachusetts.

PAGANO, M. (1989a): "Endogenous Market Thinness and Stock Price Volatility," *Review of Economic Studies*, 56, 269–288.

PAGANO, M. (1989b): "Trading Volume and Asset Liquidity," *Quarterly Journal of Economics*, 104, 255–274.

PAGANO, M., and A. RÖELL (1992): "Auction and Dealership Markets - What is the Difference?," *European Economic Review*, 36(2-3), 613–623.

PAGANO, M., and A. RÖELL (1993): "Front-running and Stock Market Liquidity," in *Financial Markets' Liberalisation and the Role of Banks*, ed. by V. Conti, and R. Hamaui, pp. 178–199. Cambridge University Press.

PFLEIDERER, P. (1984): "The Volume of Trade and the Variability of Prices: A Framework for Analysis in Noisy Rational Expectations Equilibria," *Working Paper, Stanford University.*

PRATT, J. (1964): "Risk Aversion in the Small and in the Large," *Econometrica*, 32, 122–136.

RABIN, M. (1998): "Psychology and Economics," *Journal of Economic Literature*, XXXVI, 11–46.

RADNER, R. (1979): "Rational Expectations Equilibrium Generic Existence and the Information Revealed by Prices," *Econometrica*, 47, 655–678.

RAO, K. P. S. B., and M. B. RAO (1983): *Theory of Charges.* Academic Press, New York.

ROBERTS, K. (1996): "The Le Châtelier Principle can be 'Expected' to Hold Globally: Implications of Rationality in Constrained and Unconstrained Situations," *STICERD Working Paper TE 96/298, London School of Economics.*

ROCHET, J. C., and J. L. VILA (1994): "Insider Trading without Normality," *Review of Economic Studies*, 61, 131–152.

ROGOFF, K. (1996): "The Purchasing Power Parity Puzzle," *Journal of Economic Literature*, 34(June), 647–668.

ROMER, D. (1993): "Rational Asset-Price Movements Without News," *American Economic Review*, 83, 1112–1130.

ROTHSCHILD, M. (1974): "A Two-Armed Bandit Theory of Market Pricing," *Journal of Economic Theory*, 9, 185–202.

ROTHSCHILD, M., and J. E. STIGLITZ (1976): "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics*, 90, 629–649.

RUBINSTEIN, M. (1974): "An Aggregation Theorem for Securities Markets," *Journal of Financial Economics*, 1, 201–224.

SARGENT, T. J. (1979): *Macroeconomic Theory.* Academic Press, New York.

SARGENT, T. J. (1991): "Equilibrium with Signal Extraction from Endogenous Variables," *Journal of Economic Dynamics and Control*, 15, 245–273.

SAVAGE, L. J. (1954): *The Foundation of Statistics.* Wiley, New York.

SCHARFSTEIN, D. S., and J. C. STEIN (1990): "Herd Behavior and Investment," *American Economic Review*, 80, 465–479.

SHILLER, R. J. (1989): *Market Volatility.* MIT Press, Cambridge, MA.

SHLEIFER, A., and R. W. VISHNY (1990): "Equilibrium Short Horizons of Investors and Firms," *American Economic Review Proceedings*, 80, 148–153.

SINGLETON, K. J. (1987): "Asset Prices in a Time-Series Model with Disparately Informed, Competitive Traders," in *New Approach to Monetary Economics*, ed. by W. A. Barnett, and K. J. Singleton. Cambridge University Press, Cambridge.

SMITH, L., and P. SØRENSEN (1994): "Pathological Outcomes of Observational Learning," *Working Paper 94-24, Massachusetts Institute of Technology.*

SMITH, L., and P. SØRENSEN (1997): "Informational Herding and Optimal Experimentation," *Mimeo, Massachusetts Institute of Technology.*

STOCKMAN, A. C. (1988): "Sectoral and National Aggregate Disurbances to Industrial Output in Seven European Countries," *Journal of Monetary Economics*, 21(March/May), 387–409.

SUBRAHMANYAM, A. (1991): "A Theory of Trading in Stock Index Futures," *Review of Financial Studies*, 4, 17–51.

SVENSSON, L. E. O., and S. VAN WIJNBERGEN (1989): "Excess Capacity, Monopolisitic Competition, and International Transmission of Monetary Disturbances," *Economic Journal*, 99(September), 785–805.

TIROLE, J. (1982): "On the Possibility of Speculation under Rational Expectations," *Econometrica*, 50, 1163–1182.

TOWNSEND, R. M. (1983): "Forecasting the Forecasts of Others," *Journal of Political Economy*, 91(4), 546–588.

TREYNOR, J. L., and R. FERGUSON (1985): "In Defense of Technical Analysis," *Journal of Finance*, 40, 757–775.

TRUEMAN, B. (1994): "Analyst Forecasts and Herding Behavior," *Review of Financial Studies*, 7, 97–124.

TVERSKY, A., and D. KAHNEMAN (1986): "Rational Choice and the Framing of Decisions," in *Rational Choice: The Contrast between Economics and Psychology*, ed. by R. Hogarth, and M. Reder, pp. 67–94. University of Chicago Press, Chicago and London.

TVERSKY, A., and D. KAHNEMAN (1991): "Loss Aversion in Riskless Choice: A Reference-Dependent Model," *The Quaterly Journal of Economics*, CVI.

VAYANOS, D. (1996): "Strategic Trading and Welfare in a Dynamic Market," *Research Paper 1416, Stanford University.*

VILA, J.-L. (1989): "Simple Games of Market Manipulation," *Economics Letters*, 29, 21–26.

VITALE, P. (1997): "Speculative Noise Trading and Manipulation in the Foreign Exchange Market," *Mimeo, London School of Economics.*

VIVES, X. (1993): "How Fast Do Rational Agents Learn?," *Review of Economic Studies*, 60(2), 329–347.

VIVES, X. (1995): "Short-Term Investment and the Informational Efficiency of the Market," *Review of Financial Studies*, 8, 125–160.

WANG, J. (1993): "A Model of Intertemporal Asset Prices Under Asymmetric Information," *Review of Economic Studies*, 60, 249–282.

WANG, J. (1994): "A Model of Competitive Stock Trading Volume," *Journal of Political Economy*, 102(1), 127–168.

ZHANG, J. (1997): "Strategic Delay and the Onset of Investment Cascades," *Rand Journal of Economics*, 28, 188–205.

ZWIEBEL, J. (1995): "Corporate Conservatism and Relative Compensation," *Journal of Political Economy*, 103, 1–25.