CERGE
Center for Economics Research and Graduate Education
Charles University Prague

Essays on market competition, trading costs, and laboratory asset markets

Michal Ostatnický

Dissertation

Prague, June 2010

Michal Ostatnický

Essays on market competition, trading costs, and laboratory asset markets

Dissertation

Prague, June 2010

Dissertation Committee

Jan Hanousek (CERGE–EI; co-chair)

Andreas Ortmann (Australian School of Business; co-chair)

Avner Shaked (Bonn University)

Petr Zemčík (CERGE–EI)

Referees

Karel Janda (Charles University)

Karel Zimmermann (Charles University)

Table of Contents

Abstract

The thesis is motivated by the behavior of traders on stock markets, how they compete, and what the results of that competition are. I describe the effects of trading competition and costs empirically and theoretically, and the third chapter studies a related problem experimentally. In the first chapter, I study the intraday development of bid-ask spreads on stock exchanges. I show that the bid-ask spread decreases rapidly towards the end of the trading day on Czech markets. When a similar analysis is performed on NASDAQ-listed shares a similar phenomenon is found though it is much weaker. The second chapter deals with oligopolistic competition to reveal possible reasons for the declining bid-ask spread phenomenon. I show that when there is a mix of buyers informed and uninformed about the prices posted by a small group of sellers, all sellers mix their prices according to a specified formula, and this equilibrium is unique. The result can justify a positive bid-ask spread, its stochastic nature and the declining pattern. The third chapter (with Ondřej Rydval and Andreas Ortmann) is related to the bounded rationality of traders or the players of a game. We study three very simple dominance-solvable games and show that only one third of the subjects solve them correctly. Linking the performance in the experiment to their cognitive abilities and personality traits, we found that reasoning errors are more likely for subjects with lower working memory, intrinsic motivation and a premeditation attitude.

Disertační práce je motivována chováním obchodnídků na akciových trzích, jejich soutěžením a pozorovatelnými výsledky této soutěže. Mým cílem je popsat dopady soutěže a spojených nákladů empiricky a teoreticky a ve třetí kapitole studuji odvozený problém experimentálně. V první kapitole studuji vnitrodenní vývoj bid-ask spreadu na akciových trzích. Ukazuji, že bid-ask spread se na českých akciových trzích před ukončením obchodování dramaticky snižuje. Pokud je podobná analýza provedena na akciích obchodovaných v americkém systému NASDAQ, lze říci, že podobný, ač ne tak silný, trend lze zaznamenat i zde. Druhá kapitola se zabývá olygopolistickou soutěží. Motivací je možnost ukázat, zda snižující se bid-ask spread nelze vysvětlit strukturou trhu postaveném na souteži tvůrců trhu. V kapitole je popsán model s mixem zákazníků informovaných a neinformovaných o cenách nabízených malou skupinou prodejců. V rovnovážném stavu modelu se ukazuje, že všichni prodejci volí své ceny náhodně dle specifické netriviální distribuční funkce a že tento rovnovážný stav je unikátní. Výsledek

tak podporuje existenci bid-ask spreadu, jeho stochastické chování a klesající trend. Třetí kapitola (se spoluautory Ondřejem Rydvalem a Andreasem Ortmannem) se vztahuje k omezené racionalitě obchodníků nebo obecně hráčů specifické hry. Studujeme zde tři velmi jednoduché hry s dominantní strategií a ukazujeme, že pouze jedna třetina hráčů takovou hru dokáže správně vyřešit. Pokud vztáhneme správnost uvažování hráčů k jejich poznávacím schopnostem a povahovým rysům, ukazuje se, že chyby v úsudku jsou pravděpodobnější u subjektu s menší pracovní pamětí a u jedinců s menší vnitřní motivací a rozvahou.

Acknowledgments

I owe my deepest gratitude to Dr. Jan Hanousek and Dr. Andreas Ortmann, without whom I would not have been able to finish this dissertation. They gave me valuable advice and encouraged me from the preliminary to the concluding level.

I also thank Dr. Avner Shaked and Dr. Petr Zemčík for their hints at various stages of manuscript development as well as the referees of the dissertation and individual papers Dr. Karel Janda, Dr. Karel Zimmerman, Dr. Levent Celik and other anonymous referees. I am grateful to Dr. Ondřej Rydval and Dr. Andreas Ortmann for cooperation on the third chapter of the dissertation.

Finally, and most importantly, I would like to thank my wife Jana and my parents for their moral support and patience.

Czech Republic, Prague
June 2010

Michal Ostatnický

Introduction

The present thesis is motivated by issues pertaining to stock markets. Specifically, I am interested in the behavior of traders on stock markets, how they compete, and what the results of that competition are. In the first two essays I try to understand empirically and theoretically the costs of trading. The third essay reports an experiment about subjects' performance in the simplest dominance-solvable games.

Stock markets of various types have evolved over time around the world. The basic functions of such markets is to reveal information and provide liquidity. I focus on the latter function. Originally, the purpose of stock markets was to bring together investors willing to spend money on promising projects to get a return and entrepreneurs looking for cash to finance their projects. They met in one trading place to decrease the cost of finding their business counterpart.

Several types of transaction costs are associated with stock market trading, such as clearing fees, stock exchange membership fees or payments to a broker. The majority of these tend to be negligible. In contrast, the transaction costs resulting from bid-ask spreads tend to be significant. Stock exchanges publish share prices typically as averages of the last bid price and the last ask price. In other words, one has to pay more than the published price to buy a share. The difference between the published price and the best offer or ask is one-half of the bid-ask spread and it can be interpreted as a cost.

In the first essay, I study the intraday development of bid-ask spreads and show that they decrease relatively rapidly, especially on Czech markets.[1] I develop a testing procedure that confirms that the bid-ask spread was decreasing for Prague Stock Exchange blue-chips and for at least one half of RM-System (the less liquid Czech stock exchange) blue-chips. The Czech stock markets are quite young and the legal and business envi-

[1] 2003 data is used in the essay. It would be desirable to perform the same analysis on more recent data, but currently, I do not have access to the same data sources used in the essay.

ronment can be viewed as immature. I therefore also test a set of shares listed on the NASDAQ market using the same procedure. I show that for more than one-half of the shares in my sample the bid-ask spread decreases significantly. Since this phenomenon can be identified on a well established market, the intraday decline of the bid-ask spread seems not to be a consequence of market immaturity. A simple recommendation follows from the intraday trajectory of the bid-ask spread: if you believe that prices move randomly around a stable mean and you are not willing to speculate on the average price short-term increase/decrease it seems to make sense to trade in the hours before the market closes, when the bid-ask spread tends to be smallest.

An interesting question is why the bid-ask spread declines. I hypothesized in a working paper that market makers tacitly collude, thus behaving like a monopoly and extracting much of the market's surplus.[2] It turns out that the hypothesis seems not valid, as real trading patterns do not fit the theoretically developed trading pattern when monopoly is present on the market. The main reason is that in the presence of a monopolistic market maker the pattern of daily traded volume would be exponentially decreasing, while in reality it seems to be U-shaped.

In the second essay, I therefore study an oligopolistic setup rather than a monopoly setup, and I develop a model that explains the observed bid-ask spread pattern by oligopolistic price competition. The essay is framed in abstract terms: sellers, who compete on price, interact with a mix of informed and uninformed buyers. The purpose of this essay is to describe the properties of a one-shot game equilibrium that can be subsequently used to explain the intraday bid-ask spread development patterns. I show that the expected profit of all sellers is positive even if one of the sellers cannot expect any uninformed buyers to approach her, and that all sellers randomize their price in the equilibrium.[3] If we present one trading day by a series of time intervals, and assume that the above-mentioned game is played in each one, we can apply the results of the model.[4] Interpreting sellers as market makers, the positive profit (prices larger than production costs) reflects a positive bid-ask spread,[5] and the stochastic nature of the equilibrium

[2] Ostatnický, Michal, 2004. Coase's conjecture in finite horizon. CERGE-EI Working Paper Series, no. 241.

[3] Also, the equilibrium strategy is unique so it is unique also in a finitely repeated game.

[4] I show that the model has a unique equilibrium, and thus a finitely repeated game has the same equilibrium in every subgame.

[5] Though the positive profit can result also from adverse selection or uncertainty, the presented model can easily handle the volatility of the bid-ask spread. The volatility is caused by the stochastic nature of the equilibrium strategies and the minimum posted price.

reflects the stochasticity of a bid/ask spread.

The declining pattern of the bid-ask spread can be explained by the intraday changes of the group of buyers present on the market. During a trading day the mean and variance of buyers' reservation values decrease as traders with more extreme valuations trade and leave the market. In the model, the decreased mean and variance of buyers' reservation values imply a decreased expected minimum price. Interpreting the expected minimum price as the bid-ask spread, the model explains the declining pattern of the bid-ask spread.

Asset markets have relatively complicated structures, rules and payoff functions. In the third essay, co-authored with Ondrej Rydval and Andreas Ortmann and published as "Three Very Simple Games and What It Takes to Solve Them" in Journal of Economic Behavior and Organization, we show that even in experiments with very simple games, subjects tend to misinterpret the game rules and payoff functions. Specifically, we study very simple dominance-solvable games and the ability of subjects to detect the dominance in relation to their action choices. We ask subjects to select a strategy from a minimal strategy set (2 or 3 strategies) and to report the reasoning behind their choices. We found that the reported reasoning of only one third of the subjects is correct; the remaining two thirds of the subjects do not seem to realize that the games analyzed have a dominant strategy. We also identified a positive correlation between reasoning correctness and subjects' working memory (ability to maintain and allocate attention), intrinsic motivation to engage in demanding tasks, and propensity to deliberate while carrying out tasks. The fact that our experimental subjects have difficulties understanding and interpreting correctly a minimalistic game structure poses the interesting question of how many subjects are able to respond correctly to strategies, rules and payoff functions of more complicated laboratory asset markets,[6] never mind asset markets in the "real world".

Note: The third chapter has been published as a co-authored paper in an external journal. For this reason, the third chapter of this dissertation appears exactly as it did in that journal, and has not been edited ot modified for correctness or consistency with the rest of the dissertation.

[6]See also Ortmann and Ostatnicky. Proper experimental design and implementation are necessary conditions for a balanced social psychology. Behavioral and Brain Sciences, Vol. 27, No. 3, June 2004, 352-353, and Chou, McConnell, Nagel, and Plott. The control of game form recognition in experiments: understanding dominant strategy failures in a simple two person "guessing" game. Experimental Economics, Vol. 12, No. 2, June 2009, 133-252.

Chapter 1

Decreasing Bid-Ask Spread : The Czech Case and a Comparison to NASDAQ

Abstract

Czech stock markets developed rapidly from their re-opening in the year 1993, however, they are not generally considered to be fully developped due to lack of liquidity or private information exploiting phenomenon present. In this paper I look closely at the development of the bid-ask spread during an average trading day. I show that, on average, the bid-ask spread is sometimes decreasing and sometimes rather constant on Czech stock markets. To compare the results, especially the decreasing pattern, to well-established markets, a sample of NASDAQ shares is analyzed and shown that they behave similarly in several respects. This paper contributes to the discussion whether the pattern is "U-shaped" or "L-shaped" and enriches the used methods by fixed effects estimation of quoted spread intraday data.

Keywords: stock markets, market making, bid-ask spread, intraday behavior.

JEL classification: G14, L11, D40.

1.1 Introduction

In this paper we analyze the basic intraday trends of a quoted bid-ask spread observed on various stock markets. A bid-ask spread is defined as the difference between the best sell (bid) and the best buy (ask) offers on a stock exchange. It is an indicator of market activity, uncertainty, information currently revealed on the market, the existence of a trader with superior information (adverse selection), etc. Bid-ask spread behavior has been studied before extensively, and the existing literature can be split into two categories: inventory-based and information-based models.

Inventory-based models can be traced from Demsetz (1968), who investigated the relation between the bid-ask spread and volume traded. He found that the higher the trading activity the lower the bid-ask spread. His work was followed by Garman (1976), Stoll (1979), Ho and Stoll (1981), and O'Hara and Oldfield (1986). In these papers the main idea is to capture the stochastic nature of incoming bids and to derive the appropriate reactions of a market maker who has to optimize his inventories to be able to serve the demand. A risk-averse market maker is also assumed in the three later works.

Information-based models take into account an asymmetric information problem that naturally arises at stock markets as they are meant to serve as information-revealing institutions. It is assumed that there are liquidity traders[1] who do not speculate about the future prices, because they have no private information and trade for liquidity purposes. The second trader type is an informed trader who waits for the proper moment to benefit from the information other investors and market makers do not possess. These models can be represented by Bagehot (1971), Copelan and Galai (1983), and Glosten and Milgrom (1985). According to these studies the bid-ask spread increases as the fraction of informed traders gets higher, which implies that the higher the bid-ask spread the lower the trading activity, because the risk of facing an informed trader is relatively higher. Glosten and Milgrom mention in their study that the bid-ask spread should be decreasing in a competitive market. The reason is that the more information is revealed on the market, the smaller the bid-ask spread so if we do not expect new information inflow during the day, the bid-ask spread should, naturally, decrease during the day.

In a wide stream of empirical literature focused on the intraday behavior of stock markets authors analyse stocks or indices returns. They describe the main drivers of returns like the amount of insider trading. As this paper is focused on Czech markets,

[1] See e.g. Glosten (1985)

we can mention the work of Hanousek and Kocenda (2009).

The second stream of literature targets directly bid-ask spread analysis. In the analysis the bid-ask spread is either studied directly or it is decomposed to information asymmetry and order processing components - the basic decomposition model used most widely was developed by Huang and Stoll (1997).[2] The results of the model are threefold: inverse J-shaped, U-shaped, and L-shaped intraday patterns. The U-shaped pattern result is represented by McInish and Wood (1992). The U-shaped pattern result is probably the most widespread, e.g. in Abhyankar, Ghost, Levin, and Limmack (1997), Chung, Van Ness, and Van Ness (1999), Ding (1999), Huang (2004), Li, Van Ness, Van Ness (2005), and Vo (2007). The L-shaped intraday pattern means that the bid-ask spread is stable throughout the day and it declines fast at the end of the day. The first L-shaped result was presented by Chan, Christie and Schulz (1995), though only on a limited NASDAQ data sample. Later, evidence of an L-shaped pattern from other markets has been presented. See the summary given by Malinova and Park (2009), who also give a theoretical rationale for this behavior.

Naturally, there appear studies that compare different estimators and their performance. For example, McInish and Van Ness (2002) use a sample of NYSE data and three different models to decompose the bid-ask spread and to study the intraday patterns of the bid-ask spread components. The results of the three different models are mixed for some components of the bid-ask spread. A similar study on futures market data was done by Anand and Karagozoglu (2006). An interesting result was presented by Ap Gwilym and Thomas(2002): they conclude that "transaction based measures are found to be biased estimates of quoted and effective spreads". This means that the choice of spread definition may impact the result.

Currently, research focuses on smaller exchanges (Alabed and Al Khouri, 2008), markets where there is not much available data (Marcato and Ward, 2007), or new information is included in the model (Caglio and Kavajecz, 2006 include the size of the quotation into their model and Michayluk, Prather, Woo and Yip, 2009 include option markets indicators).

In this chapter I try to abstract from the bid-ask spread decomposition and analyze directly the intraday quoted bid-ask spread patterns, i.e. the prices a trader would get immediately trading a share. When analyzing the data it turns out that the daily average bid-ask spread sizes differ so fixed effects (with respect to daily levels of the bid-

[2]This model is used, for example, by Hanousek and Podpiera (2003) to analyze Czech markets.

ask spread) formulation is used to remove unnecessary variance in the data. The main result is that the bid-ask spread of (liquid) stocks on Czech stock exchanges show rather a decreasing intraday pattern than a U-shaped or J-shaped pattern. To compare the Czech markets with well-established ones the same methodology was used on a sample of NASDAQ data: 10+9 U.S. stocks and 7 Czech stocks traded on three Czech markets. I find various trends of the bid-ask spread including a steady series, a decreasing series with some breaks, and an increasing bid-ask spread at the end of trading. Also, we show that the bid-ask spread patterns may be significantly different on markets with different market structures (see also Huang, 2004) or market participants.

In the next section I describe the rules of the analyzed market, and in the third section the data used in the analysis. In the fourth section the basic analysis is performed and in the fifth section more detailed results are presented. The sixth section concludes.

1.2 Market Design

One of the leading stock U.S. markets - NASDAQ - and the only two organized stock markets in the Czech Republic - the PSE and the RM-System - are studied in this paper. Stock exchanges around the world vary a lot thus I describe the structure of these markets in this section. The NASDAQ market is generally well-known so I start with this market and then continue with the two Czech markets that have a generally similar design.

The main power of the NASDAQ Stock Market is the so-called 'supermontage'. There are many smaller stock exchanges (essentially electronic communication networks, ECNs, and regional stock exchanges) united under the NASDAQ-centralized system and the order books of these exchanges, i.e. the quantities offered and the corresponding limit prices, are sent to the centralized system where the best offers (bids and asks) are displayed. All brokers are obliged to route all unexecuted clients' orders to one of the exchanges so that every valid order can be displayed in the centralized system. In addition to standard orders routed through brokers and brokers' own orders, the majority of listed equities are quoted by market makers. Market makers in the NASDAQ environment are free to choose which equities they want to quote, and once they choose them they are obliged to publish both the bid and ask quotes.[3] Market makers bring liquidity to the market - if you want to trade there is always a market maker who concludes a trade with you.

[3]There are several situations when they are exempted from this obligation, i.e. during a short time interval after their quote is hit.

Also, when there is a lack of trading, market makers support price discovery. On the other hand, the market does not rely on market makers themselves; anybody is free to compete with them by submitting a limit order - a bid or ask with a specified amount of shares and a limit price.

In the Czech Republic the structure of the markets is a bit more complex, nevertheless, every market uses a mechanism similar to NASDAQ. Generally, there are two firms that run public stock markets: RM-System (RMS) and the Prague Stock Exchange (PSE). Both of them started to operate in the early nineties when a lot of shares were put on both markets simultaneously,[4] i.e. the majority of them were listed on both markets. Since that time, the markets have developed trading mechanisms, improved the rules and transparency and de-listed some titles. RMS is opened to the general public; anyone can submit a limit order directly to the main order book of the exchange. The PSE is a membership-based exchange; it is only possible to trade via brokers. In the description of the trading systems we focus on the main markets only - block trades, direct trades, etc. are not completely relevant for the purpose of this article.

There is one main market run by the RMS, a continuous double-side-auction computerized market. As mentioned above, anybody can send a limit order to the market, and these orders form an order book. In the morning the order book is cleared (the clearing price for the order book is found and relevant orders are cleared), and thereafter a continuous auction is run till the evening. As soon as a new order comes, a new auction is executed and the order book is cleared. The auction process starts at 9am and finishes at 5pm. There is some room to submit orders before 9am and after 5pm.

On the PSE it is possible to use two trading systems/floors, each with a different trading system.[5] From the beginning of its existence a continuous double-side market called KOBOS is used. KOBOS is similar to RMS: at 9am a clearing price from the active limit orders is computed and then a continuous double-side auction is run till 4pm when the trading floor is closed. In 1998 a new market type called SPAD was introduced. The SPAD trading system is based on market makers competition as well as the core of NASDAQ - market makers who choose to quote a share are obliged to publish both bid and ask quotes and trade a fixed lot[6] when the quote is hit. Trading hours are the same as for NASDAQ: 9:30am till 4pm. Czech bluechips are traded on this market and the

[4]See e.g. Hanousek and Podpiera (2004).

[5]There is also a possibility of using block/mass trade or direct trade within the PSE. I am not, however, interested in these trades as they are not price-discovering.

[6]The size of the lot is defined when the shares are accepted for trading in SPAD.

volume traded form the vast majority of all trades.

The main difference between the Czech markets and NASDAQ is that NASDAQ incorporates both trading systems in one. Sufficient liquidity is maintained by market makers whose role is to offer bids at every moment the market is open together with the possibility to submit individual limit orders for any trader interested. The PSE market system separates the two procedures into two seemingly distinct markets to offer cheaper transactions for those trading bigger volumes and to open the doors for smaller traders. However, due to existence of two distinct sub-markets there is fear for limited transparency because PSE supplies two distinct closing prices (though only one of them is said to be official). The solution found is that the SPAD closing price is published as the official one if shares are traded in both systems, because SPAD is considerably more liquid. The situation gives us an opportunity to show that the phenomenon indicated can be detected in all the trading mechanisms analyzed.

Another difference between the markets is the transaction fee paid to the exchange. This fee significantly impacts the absolute size of the bid-ask spread. Although we are interested in the relative daily pattern, and the time-flat fees do not impact the pattern, it is worthwhile to mention them. In the Czech Republic the fees are relatively high. The highest fees are collected on the public RMS, they range between 1.5% and 2.5% of the volume traded, depending on the trade size. A KOBOS trade costs around 1% and the cheapest is a SPAD trade as its volume must be relatively big; it costs 0.1%. The fees in the U.S. are based rather on a per-share-executed basis, therefore it is difficult to compare the U.S. and the Czech fees directly, however, the U.S. fees are negligible compared to the Czech environment.

1.3 Data Description

The data used for the description of bid-ask spread behavior is real time data from the four stock markets depicted in the previous section - NASDAQ, PSE-KOBOS, PSE-SPAD, and RMS. I study the intraday pattern of the bid-ask spread so we need real time data from the exchanges. The analysis was performed on 2003 real time data for all four markets to be directly comparable. The data from NASDAQ were taken from the Nastraq database, the Czech stock markets data were obtained from RM-System and the Prague Stock Exchange.

The Czech stock market is relatively small and there are only several titles traded often enough to allow us to perform a meaningful analysis. All shares traded on SPAD constitute the set of analyzed shares and none of the others have been added because of low trading activity. The sample includes seven titles, all of them traded on SPAD and KOBOS, six of them traded also on RMS. The sample represents 100% of volume traded on SPAD, 92% of shares volume traded on KOBOS and 62% of shares volume traded on RMS. It is possible to say that in the case of PSE the sample essentially *is* the market and in the case of RMS it is more than representative as it amounts to more than half of the total volume.

There are a lot of shares traded on the NASDAQ market compared to the Czech markets. It is not possible to analyze all the traded shares so I decided to prepare two groups of NASDAQ-listed shares - the shares of companies similar to the analyzed Czech ones and a sample of randomly chosen shares. The 'random' sample covers 10 NASDAQ titles. For the sample it was impossible to find a company similar to Unipetrol (ticker BAAUNIPE) and Philip Morris CR (ticker BAATABAK) so the 'similar' sample consists of a mix of telecommunications and financial companies with one energy company.

Table 1: Shares chosen to the sample. The shares listed in the right hand column correspond to the Czech shares listed in the left column.

Czech ticker	Industry	U.S. ticker
BAAERBAG BAAKOMB	banking / financial	BPOP FITB MRBK ZION
BAACRADI BAATELEC	telecommunications	LVLT MICC SHEN SURW
BAACEZ	electric energy	OTTR

The 'random' sample includes randomly picked shares. The following tickers were selected AMOT, CTCOB, CVSN, EBAY, GENBB, NIKU, NXTP, PRFT, TBCC, and WGRD. Detailed statistics of every presented share can be found in Appendix 1.

From the sample I dropped those trading days that were shortened for some reason. In the U.S. these days were July 3, November 28, December 24, 26, and 29, in the Czech Republic no trading day was shortened. Some shares were not traded for several days due to various reasons, so these days are missing in the data. As for the market maker-driven markets, NASDAQ and SPAD, the real time data for every stock consists of up to 247 daily series in the case of NASDAQ and up to 251 daily series in the case of SPAD. Every daily bid-ask spread series starts with opening inside quotes[7] followed by the values of inside quotes after every change during the day up to the market close. The difference between respective inside quotes directly defines the bid-ask spread. For each share and every trading day in our sample I get a time series of bid-ask spread data starting with the opening bid-ask spread followed by bid-ask spread value for every change recorded during the day. RMS and KOBOS use a different trading mechanism, a continuous double-side auction. In every instant of time the best buy and the best sale offers (of any size) are identified and the difference is recorded as the bid-ask spread. In reality, I record the initial bid-ask spread and then time and the new value if there is any change. I obtain a similar time series of the bid-ask spread as for the market maker-driven systems.

The time series of bid-ask spreads have been transformed into a format that enabled us to compare different stock exchanges that differ in the length of trading day. I split the trading day into 78 equal time periods[8] and compute the average of the bid-ask spread in every period. These numbers *(number of trading days* x *78)* make up our sample for every stock and stock exchange. One period represents 5 minutes on NASDAQ and PSE and 6.15 minutes on the RMS. In Appendix 1 are plots of the average relative bid-ask spread path in the 78 periods. Some analytically confirmed effects are visible by naked eye, for example, the decrease at the beginning of the trading, the decrease or increase at the end of trading.

1.4 Tests of decreasing pattern

In Appendix 1 the average intraday behavior of the bid-ask spread is shown. We would expect a turbulent bid-ask spread when there is fast information release or some uncertainty about the future; the bid-ask spread should be stable on average. Observing the time series pattern we notice different behavior of the analyzed markets. The bid-ask spread of the NASDAQ shares generally decreases at the beginning of trading. The regular everyday most intensive information release should be expected in the morning when

[7]The best bid and ask quotes at the trading opening.

[8]The number 78 comes from the PSE trading system, where the trading day lasts 6.5 hours, or 78 5-minute intervals.

traders come to the market with fuzzy expectations of the current day trading trends and new share-specific information. These expectations focus later during the day so that the bid-ask spread shape get stabilized in the 10th period (approximately the 50th minute), and the curve seems to be constant.

The Czech stock markets demonstrate slightly different behavior. The RMS markets' time series appear to be constant throughout the whole trading day. This shape differs completely from the bid-ask spread time series of the two Prague Stock Exchange markets, SPAD and KOBOS. The SPAD bid-ask spread time series declines till the 30th period (approximately 2.5 hours after trading-floor opening), stays constant around noon and again starts to decline in the 60th period (approximately an hour before trading closes). KOBOS follows the behavior of SPAD, the only difference is a faster decline - a very fast decline at the beginning of trading with a slowdown around noon and again a fast decline as the stock exchange approaches the end of regular trading hours.

Our aim in this section is to test the statistical validity the above-described effects, namely the time series decrease; the detailed shape of the decrease is to some extent analyzed in the next section. The statistical approach used is not very complicated as the effects are easy to identify. I manage with a sample mean comparison together with the fixed effects concept.

To identify the decreasing pattern of every individual share's bid-ask spread, I have chosen to compare three time intervals of equal length - the first 1/3 of the day, the period around noon and the end of the trading day. The simplest form of the estimation equation would be

$$BA_i^t = \beta_1 T_i^1 + \beta_2 T_i^2 + \beta_3 T_i^3 + \varepsilon_i^t,$$

where

BA_i^t denotes the bid-ask spread,

T_i^1 denotes the first third of the day, i.e. the first 26 periods,

T_i^2 denotes the second third of the day, i.e. periods 27 to 52, and

T_i^3 denotes the last third of the day, i.e. periods 53 to 78.

Tests would then compare the statistical difference between $\beta_i s$. It turns out that every trading day has, on average, a different level of bid-ask spread so I use fixed effects estimation to increase the power of the test for the statistical difference of the bid-ask spread in different periods. The improved estimated equation can be written as

$$BA_i^t = \sum_{j=1}^{N} \delta_j D_j^t + \beta_1 T_i^1 + \beta_3 T_i^3 + \varepsilon_i^t,$$

where D_j^t are dummies for individual trading days and N is the number of trading days relevant for every share.

I am interested in the sign and statistical significance of the parameters β_1 and β_3 that show the difference between the first and the second periods, and the second and the last periods of the trading day. The results are summarized in Table 2.

Table 2: Estimated parameters of the trend.

Note: the listed Czech tickers exclude the standard prefix BAA, ♭ denotes a significant number on the level of 10%, ♮ denotes an insignificant number on the level of 10%, other values of β_i are significant on the 1% level.

Ticker	**# obs.**	**avg.BA**	β_1	β_3	β_1/ **avg.BA**	β_3/ **avg.BA**
SPAD						
CEZ	19578	.804	.200	-.0869	.249	-.108
CRADI	19578	5.19	.915	-.424	.176	-.0817
ERBAG	19578	13.3	.523	-.696	.039	-.0521
KOMB	19578	11.9	.914	-.530	.0768	-.0445
TABAK	19578	132	12.5	-10.9	.0951	-.0825
TELEC	19578	2.83	.859	-.386	.304	-.137
UNIPE	19578	.757	.141	-.0649	.186	-.0857
KOBOS						
CEZ	19578	1.05	.263	-.125	.250	-.119
CRADI	19422	5.67	2.71	-1.01	.477	-.178
ERBAG	6942	103	17.2	-16.0	.167	-.156
KOMB	19578	26.4	8.79	-3.55	.333	-.134
TABAK	19110	289	78.3	-65.2	.271	-.225
TELEC	19578	3.67	1.10	-.860	.299	-.235
UNIPE	19578	.752	.240	-.107	.319	-.143
RMS						
CEZ	19032	.948	.0212	-.0248	.0224	-.0261
CRADI	19266	3.67	.217	**-.0439**♮	.0592	-.0119
KOMB	19266	19.5	**.227**♭	-1.53	.0116	-.0784
TABAK	12948	188	.0497	-0.00868	2.65E-4	-4.63E-5
TELEC	19266	3.45	.0784	-.167	.0227	-.0484
UNIPE	19266	.635	.0373	**.00402**♭	.0587	-.00632

Table 2 (continued): Estimated parameters of the trend.

Note: the listed Czech tickers exclude the standard prefix BAA, ♭ denotes a significant number on the level of 10%, ♮ denotes an insignificant number on the level of 10%, other values of β_i are significant on the 1% level.

Ticker	**# obs.**	**avg.BA**	β_1	β_3	β_1/ **avg.BA**	β_3/ **avg.BA**
NASDAQ						
'similar'						
BPOP	19266	.0437	.0114	**5.12E-4**♮	.261	.0117
FITB	19266	.0265	.00592	-.00150	.224	-.0568
LVLT	19266	.0104	.00110	**4.92E-6**♮	.105	4.72E-4
MICC	17706	.178	.0607	**-.00163**♭	.340	-.00915
MRBK	19266	.0359	.00906	-.00123	.252	-.0341
OTTR	19266	.127	.0295	-.0128	.233	-.101
SHEN	19266	.985	.115	-.0994	.117	-.101
SURW	19266	.306	.0875	-.0311	.286	-.102
ZION	19266	.0296	.00835	-.00179	.282	-.0604
'random'						
AMOT	19266	.111	.00925	-.00437	.0833	-.0393
CTCOB	13104	1.23	.204	-.0828	.163	-.0660
CVSN	19266	.0333	.00291	-8.34E-4	.0874	-.0250
EBAY	19266	.0229	.00282	-4.09E-4	.123	-.0178
GENBB	19266	.171	.00875	**-.00158**♭	.0511	-.00924
NIKU	19266	.0831	.0201	**-.00148**♭	.242	-.0178
NXTP	19266	.0172	.00299	**-2.61E-4**♭	.714	-.0152
PRFT	19266	.0749	.00858	-.00127	.115	-.0169
TBCC	19266	.0914	.0330	**.00164**♭	.361	.0180
WGRD	19266	.0270	.00500	-.00139	.185	-.0514

It turns out that in the first period the bid-ask spread decreases for every share in all five samples with only one exception - KOMB traded in the RMS does not exhibit a significant decrease on a confidence level of 10%. Our first observation turns out to be confirmed statistically: there is a significant decrease of the bid-ask spread between the beginning of trading and the middle of the trading day. The decrease is relatively the smallest in the RMS where its average value is much less than 10%. In other markets the decrease, with just two exceptions, balances on the boundary of 10% or is much bigger even around 30 to 40%.

The second-to-third-period difference differs from the first-to-second-period one. The Prague Stock Exchange markets, the SPAD and the KOBOS, continue to decrease, but the decrease is smaller than at the beginning of the trading day. In the RMS two out of the six shares do not exhibit a significant decrease. The decrease of other shares remains similar to the first-to-second-period difference. In the NASDAQ case 7 shares exhibit an insignificant decrease, while 12 shares exhibit a significant decrease on the 1% level. The difference between the first and second periods is, however, bigger than that between the second and third periods. If the decrease is significant it amounts to a value between 1 and 10% for all markets with the exception of KOBOS where the decrease stays within the interval of 10 and 25%.

The comparison of average values between the three intervals demonstrates a decreasing trend throughout the day. I have decided to look into the behavior of the time series even closer and to estimate the decreasing trend of every share within each of the three intervals. Again, for every share and this time for each of the three time intervals I estimate the bid-ask spread trend in a fixed effects model:

$$BA_i^t = \sum_{j=1}^{N} \delta_j D_j^t + \theta t + \varepsilon_i^t.$$

In this equation I am interested in the variable β, especially if it is statistically significant for an individual share and time interval.

It turns out that during the first period the bid-ask spread always decreases, i.e. the variable θ is significant on a level much smaller than .1%, except for the three shares traded on RMS. The decrease variable *theta* is not significant on the level of 1% in the case of KOMB and on the level of 10% in the case of TELEC, UNIPE, and TABAK. Previously I found out that there is no significant difference between the first and second time period of the KOMB shares traded on the RMS, the first-to-second-period decrease of TELEC, UNIPE, and TABAK is realized during the second period. Quite a rare situation is revealed in the case of CEZ where there is detected a significant increase.

During the second period the decrease slows down and in many cases it is insignificant. On SPAD one out of seven shares does not reveal a significant decrease on the level of 1% (TABAK) and one on the level of 10% (KOMB). TABAK does not show any significant decrease on the level of 10% even on RMS. On KOBOS the bid-ask spread continues to decrease during the second period with one exception. KOMB do not seem to decrease on the level of 10% even on this market. In NASDAQ the score is 7-11-1, three shares show no significant decrease on the level of 1% (BPOP, EBAY, WGRD) and four show no significant decrease on the level of 10% (AMOT, MICC, MXTP, TBCC). LVLT shows a very slight significant increase.

In the third period all SPAD and KOBOS bid-ask spreads decrease in opposition to RMS where just one share decreases on the level of 1% (KOMB) and TABAK show a 10% significance level increase. In the case of NASDAQ the decrease of CVSN is not significant on the level of 1% and the decrease of OTTR is not significant on the level of 10%. Other shares split into two groups. In one of them the bid-ask spread significantly increases (BPOP, CTCOB, LVLT, MRBK, NIKU, NXTP, PRFT, TBCC, WGRD) and the other decreases (EBAY, FITB, GENBB, MICC, SHEN, SURW, ZION) on the significance level of 1%.

To summarize, SPAD and NASDAQ behave in a similar manner - a decline of bid-ask spreads in the first part of the trading day, stabilization around noon, and thereafter a further decrease - though some NASDAQ shares show a significant positive trend in the last period. The KOBOS bid-ask spread decreases constantly during the day and the pattern of the bid-ask spread on RMS reveals no regular pattern. The table below presents the number of shares for which the decrease is insignificant on the level of 1% in the given period.

Table 3: Summarized results of trends tests.

Market	Total number of shares	insignificant in 1st third	insignificant in 2nd third	insignificant in 3rd third
SPAD	7	0	2	0
KOBOS	7	0	1	0
RMS	6	3	2	4
NASDAQ	19	0	6	2

1.5 Decreasing pattern in detail

In this section I am going to present a detailed analysis of the bid-ask spread time structure within a trading day. In the previous section the analysis of the majority of the bid-ask spread time series revealed a decreasing bid-ask spread pattern. The three time intervals in the previous section were, however, set exogenously and I grouped the finest grid of 78 intervals per day into three time intervals so there is enough room for a more detailed preview of the time series pattern. Basically I perform a pairwise statistical comparison of the small time intervals, therefore I get a large number of statistical results and I have to count with a type I and type II error.

I am going to analyze 7+7+6 Czech markets time series and 10 NASDAQ time series.[9] Each time series consists of 78 time intervals, i.e. 78 random variables. For each variable I have usually around 250 observations, one observation per every 2003 trading day. I do not start our analysis with any prior belief about the time series - it may be constant, increasing or decreasing in any part of the trading day. The easiest concept to describe the behavior of such a series of random variables is a pairwise comparison of their means. Such a comparison gives us information about a relative increase or decrease between any two periods in our sample.

As I have a relatively large number of observations, I can use the central limit theorem and assume that the difference of bid-ask spread means is normally distributed.[10] The test procedure has been performed for every random variable series in the same manner; for the given series I performed a test for every pair of the 78 time intervals. As mentioned before, the test is a standard two-sample mean-equality test.

Using these procedures I did 3003 tests, where rejections of mean equality can be classified as a significant increase or a significant decrease of the bid-ask spread between the relevant periods. From the pattern of the set of rejections I have a *rough* idea of the intra-day bid-ask spread behavior. The tests are performed on the 95% confidence level, thus I can estimate that 2.5% of all rejections are wrong.[11] The standard way out

[9] I analyze the 'random' sample only as the behavior of the NASDAQ time series is similar and our primary focus is on the Czech market analysis.

[10] In reality, the bid-ask spread is constrained by stock exchange rules. These rules limit the difference between the bid and ask quotes a market maker is obliged to set, usually given as a percentage of the share price. To strengthen the power of the test I can use a censored normal instead of a normal distribution. In the tests I use just a normal distribution as it is sufficient for our purposes.

[11] The null hypotheses are rejected even on a much lower level of confidence, therefore much lower number of rejections is wrong (smaller number of type II errors). A similar idea holds for the type I error.

of such a situation is to use anova or a similar non-parametric estimation. My aim was not, however, to reject an all-time stability hypothesis. I wanted to observe the actual behavior of the time series. Stability was rejected already in the previous section. Sticking to the procedure I have to keep in mind that I need to judge just the overall picture and I cannot examine and comment on every test separately.

Looking closer at the procedure I intend to run on each series we observe that it is similar to ordinary least squares estimation in the limit. In such a least square estimation the explained variable would be the bid-ask spread and the explanatory variables set would consist of dummies for the 78 sample periods. The coefficients of the dummies represent the 78 random variables from the original series - the means are equal and the standard deviations are equal in the limit.[12] Albeit there is almost no difference between the two approaches, the least squares approach enables us to employ fixed effect estimation as in the previous section.

The bid-ask spread usually depends on the current day overall trading activity on the market, thus the level of the bid-ask spread may be different every trading day, or, more often, it is different during the 'high' season and summer, when there is much less frequent trading and wider bid-ask spreads. This portion of volatility can be swept out using fixed effects estimation instead of ordinary least squares as in the previous section. This approach leads to a day-to-day volatility reduction and a higher power of the means-equality tests.

The results plotted in Appendix 2 present the panel data (fixed effects) estimation. Every black spot represents a rejection of the means equality hypothesis in favor of the decreasing hypothesis in the top left part of a panel and in favor of the increasing hypothesis in the bottom right. I start the description with the NASDAQ market results, proceed with the Prague Stock Exchange - SPAD and KOBOS markets, and finish with the RMS.

The NASDAQ shares bid-ask spread pattern follows a downward trend at the beginning of trading. For some shares this interval lasts just ten periods or 50 minutes (GENBB, WGRD, CTCOB, CVSN, NXTP), for some of them 20 periods or even more (EBAY, NIKU, PRFT, TBCC). This initial period can be viewed as an initial information inflow period, when market participants learn what are going to be the current trading day overall and shape-specific trends. In mid-day (periods 20 to 50) the bid-ask spread stays relatively constant. The decrease of many shares' bid-ask spreads start thereafter.

[12]See Appendix 3 for the proof.

In period 50 the bid-ask spread of CTCOB, WGRD and to some extent NIKU start to fall down followed later by NXTP and WGRD. During the last several periods a rapid decline can be observed in the AMOT and EBAY bid-ask spread time series. The reason may be psychological in that many traders start to be more interested in trading or that traders usually try to close their position before the end of trading.

One more striking phenomenon can be observed in the results. During the last (several) periods I can notice a significant *increase* of the bid-ask spread (see SCCOB, WGRD, NIKU, NXTP, maybe also TBCC). I would expect the bid-ask spread to decrease due to information inflow. In these cases I have identified the completely opposite behavior.

The bid-ask spread behavior of the SPAD shares can be compared with NASDAQ share performance. The SPAD shares show a constant decline with the exception of the mid-day trading hours (ERBAG, KOMB, TABAK). The bid-ask spread stagnates from the 10th-20th to the 50th-60th periods that coincide with the NASDAQ stable period interval. The reason can be that automatic quotation systems are switched on and there is less interest in trading during lunch time. The KOBOS market can be thought as of inferior or complementary to SPAD so we would expect a similar trading pattern. I see a decline in the initial period, stability till the 50th period and a subsequent decline till the closing. There is one more effect identified - a rapid decline in the very last period or in the last 2-4 periods. I can justify this phenomenon by portfolio adjustments or closing at the end of the trading day. The SPAD market trades big lots only and on KOBOS it is possible to transact any number of shares. Some traders would be eager to get rid of small amounts of shares or to buy them at the very last moment if they are several shares short. These traders bring in non-negligible amount of bids and the bid-ask spread narrows. The bid-ask spread pattern observable on the RMS market completely differs from the NASDAQ and SPAD markets. I found some decline around the 50th to 60th period, however, it is a one-time decline. Also, I can notice a first-period bid-ask spread increase sometimes (TELEC, CEZ), but I do not have a rationale for this behavior.

The results have been checked by more robust data selection to ensure that 'non-standard' weeks in 2003 do not have a big influence on the result. I call non-standard such weeks when there is a stock exchange holiday or when new information is periodically presented. As for the U.S. the stock-exchange holidays coincide with national holidays. In addition, I defined information release days as the days when it is obligatory to publish new information about listed companies according to the Exchange Act Rule 11Ac1-5. The rule states that statistical data shall be published once per month, on NASDAQ it is the 25th day of the month. In the Czech Republic, stock exchange holidays coincide with national holidays and the whole December can be assumed as information release-intensive as shareholder meetings take place at the end of the year. From the sample I deleted all the weeks where there is a non-standard day and compared the restricted samples with the full ones. I found no noticeable difference in the pattern of the bid-ask spread decrease and I think the non-standard days do not cause non-standard behavior.

1.6 Conclusion

In this paper I analyzed the average intraday pattern of shares' bid-ask spreads. Our primary intention was to perform the analysis on the seven most liquid Czech shares and to compare them to NASDAQ behavior. As there are many shares listed on NASDAQ I prepared two samples: the sample 'similar' refers to nine companies that have similar industry and financial statistics as the Czech companies whose shares were analyzed, and the sample 'random' includes 10 randomly selected shares. In addition to a comparison of the bid-ask spread behavior of the selected shares I can compare the whole NASDAQ and Czech markets as I have a big enough sample from NASDAQ and the seven Czech shares in fact almost cover the market.

I used 2003 real time data to analyze the bid-ask spread intraday pattern. The real time data were reshaped to a more convenient format: they were transformed to 78 approximately 5 minute-long intervals. With this data two kinds of analysis was performed. The first showed that the bid-ask spread decreases during the day, the second revealed the detailed pattern of the decrease, although the result cannot be quantified.

In the first part of the analysis I split the trading day into thirds and compared the first interval with the second and the second with the third to show that the majority of shares exhibit a successive decrease of the bid-ask spread. The data have been analyzed further and the significance of the linear time trend in every third of the trading day was tested. Conclusions from this part of the analysis can be summarized into two points:

- Markets with restricted access, NASDAQ and the Prague Stock Exchange's KOBOS and SPAD, show similar bid-ask spread decreases. The publicly open RMS differs a lot as the bid-ask spread decreases slowly if the decrease can be detected.

- The general profile of the bid-ask spread decrease on the NASDAQ and PSE is a fast decrease in the first third of the trading day, a slow or no decrease during the middle of the trading day, and again a speed up in the last third.

A more detailed analysis was performed on the finest grid of 78 intervals. To see the actual decrease shape I made a pairwise comparison of the intervals for every individual share. The power of the tests was improved by employing a fixed effect model when the daily average bid-ask spread was 'fixed'. On the other hand it must be understood that there were performed 3003 tests for every share. This fact limits us in results interpretation as a given percentage of the test should suffer a type I or type II error. I try to conclude with 'general' results that should not be altered by a reasonable number of errors.

- The results from the first part are confirmed: the RMS shares' bid-ask spread seems to be stable and the bid-ask spread on other markets show a downward sloping trend.
- The bid-ask spread of NASDAQ and the PSE exhibit a decrease till the 10th to 20th period (50 minutes to 100 minutes from the beginning of trading), then there is a relatively stable period around noon and thereafter again a decrease from the 50th period (250 minutes from the start of trading, 140 minutes before the close).
- The bid-ask spread of several shares increases in the last several minutes of trading.

Generally I found that the SPAD and NASDAQ bid-ask spreads behave similarly; they decrease from the beginning of the trading day, around noon they are relatively stable and during the afternoon they start to decrease further. The KOBOS market follows closely SPAD (not surprisingly), and the decrease is more pronounced. The RMS is completely different. The bid-ask spread is quite stable or the shares do not show similar behavior. There can be various reasons for this fact: the market design, the openness to the general public or the absolute size of the bid-ask spread with added fees as it is the highest of all the markets presented.

1.7 References

Abhyankar, A., D. Ghosh, E. Levin, and R.J. Limmack, 1997. Bid-Ask Spreads, Trading Volume and Volatility: Intraday Evidence from the London Stock Exchange. Journal of Business Finance and Accounting 24(3), 343-362.

Anand, Amber, and Karagozoglu, Ahmet K., 2006. Relative Performance of Bid-Ask Spread Estimators: Futures Market Evidence. Journal of International Financial Markets, Institutions and Money 16(3), 231-245.

ap Gwilym, Owain, and Stephen Thomas, 2002. An Empirical Comparison of Quoted and Implied Bid-Ask Spreads on Futures Contracts. Journal of International Financial Markets, Institutions and Money 12(1), 81-99.

Bagehot, Walter, 1971. The Only Game in Town. Financial Analysis Journal March/April 1971.

Caglio, Cecilia, and Kenneth Kavajecz, 2006. A Specialist's Quoted Depth as a Strategic Choice Variable: An Application to Spread Decomposition Models. The Journal of Financial Research 29(3), 367-382.

Chan, K.C., William G. Christie, and Paul H. Schulz, 1995. Market Structure and the Intraday Pattern of Bid-Ask Spreads for NASDAQ Securities. The Journal of Business 68(1), 35-60.

Chung, Kee H., Bonnie F. Van Ness, and Robert A. Van Ness, 2001. Can the Treatment of Limit Orders Reconcile the Differences in Trading Casts between NYSE and Nasdaq Issues? The Journal of Financial and Quantitative Analysis 36(2), Special Issue on International Corporate Governance, 267-286.

Copeland, Thomas E., and Dan Galai, 1983. Information Effects on the Bid-Ask Spread. The Journal of Finance 38(5), 1457-1469.

Demsetz, Harold, 1968. The Cost of Transacting. Quarterly Journal of Economics February 1968, 33-53.

Ding, David K., 1999. The Determinant of Bid-Ask Spreads in the Foreign Exchange Futures Market: A Microstructure Analysis. The Journal of Futures Markets 19, 307-324.

Garman, Mark, 1976. Market Microstructure. The Journal of Financial Economics 3(3), 257-75.

Glosten, Lawrence, 1985. Bid, Ask, and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders. Journal of Financial Economics 14(1), 71-100.

Glosten, Lawrence R., and Paul R. Milgrom, 1985. Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders. Journal of Financial Economics 14, 71–100.

Hanousek, Jan, and Evzen Kocenda, 2009. Intraday Price Discovery in Emerging European Stock Markets. CERGE-EI Working Paper Series no.382.

Hanousek, Jan, and Richard Podpiera, 2004. Czech Experience with Market Maker Trading System. Economic Systems 28, 177-191.

Ho, Thomas, and Hans R. Stoll, 1981. Optimal Dealer Pricing under Transactions and Return Uncertainty. The Journal of Financial Economics 9(1), 47-74.

Huang, Roger D., and Stoll, Hans R., 1997. The Components of the Bid-Ask Spread: A General Approach. The Review of Financial Studies 10(4), 995-1034.

Huang, Yu Chuan, 2004. The Components of Bid-Ask Spread and Their Determinants: Taifex Versus SGX-DT. The Journal of Futures Markets 24(9), 835-860.

Li, Yang, Bonnie F. Van Ness, and Robert A. Van Ness, 2005. Daily and Intraday Patterns in Spread and Depth for Limit Orders and Specialists. Quarterly Journal of Business and Economics 44(3-4), 3-14.

Malinova, Katya, and Andreas Park, 2009. Intraday Trading Patterns: The Role of Timing. University of Toronto, Department of Economics, Working Papers.

Marcato, Gianluca, and Charles Ward, 2007. Back from Beyond the Bid-Ask Spread: Estimating Liquidity in International Markets. Real Estate Economics 35(4), 599-622.

Michalyuk, David, Laurie Prather, Li-Anne E Woo, and Henry Y.K. Yip, 2009. What do Options have to do with it?: Inclusion of Options Market Indicators in Bid-Ask Spread Decomposition. Asia-Pacific Journal of Financial Studies 38(3), 455-489.

McInish, Thomas H., and Bonnie F. Van Ness, 2002. An Intraday Examination of the Components of the Bid-Ask Spread. The Financial Review 37, 507-524.

McInish, Thomas H., and Robert A. Wood, 1992. An Analysis of Intraday Patterns in Bid/Ask Spreads for NYSE Stocks. The Journal of Finance 47(2), 753-764.

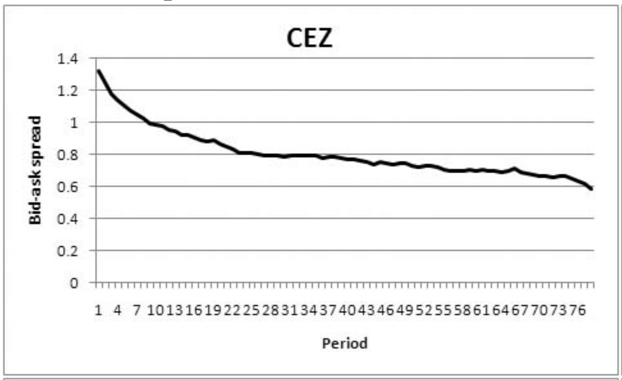
O'Hara, Maureen, and George S. Oldfield, 1986. The Microeconomics of Market Making. Journal of Financial and Quantitative Analysis 21(4), 361-76.

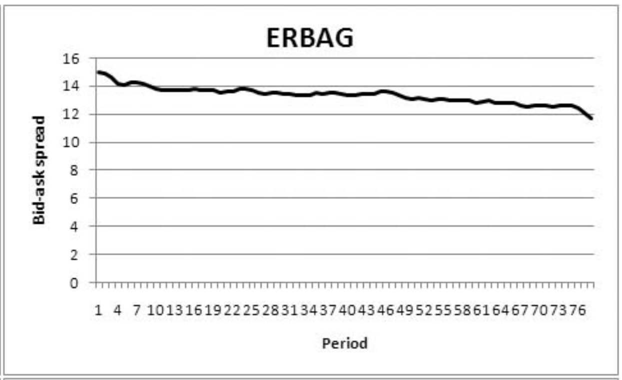
Stoll, Hans R., 1979. Commodity Futures and Spot Price Determination and Hedging in Capital Market Equilibrium. Journal of Financial and Quantitative Analysis 14(4), 873-94.

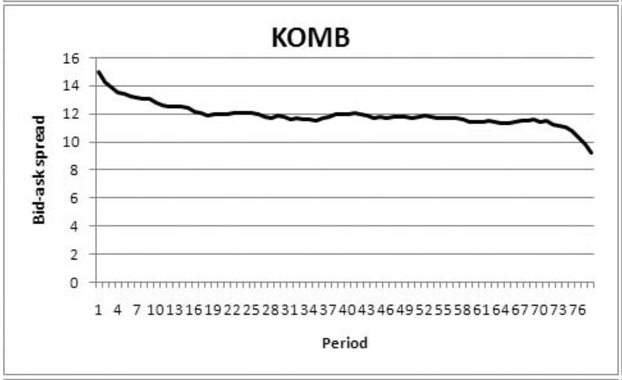
Vo, Minh T., 2007. Limit Orders and the Intraday Behavior of Market Liquidity: Evidence from the Toronto Stock Exchange. Global Finance Journal 17(3), 379-445.

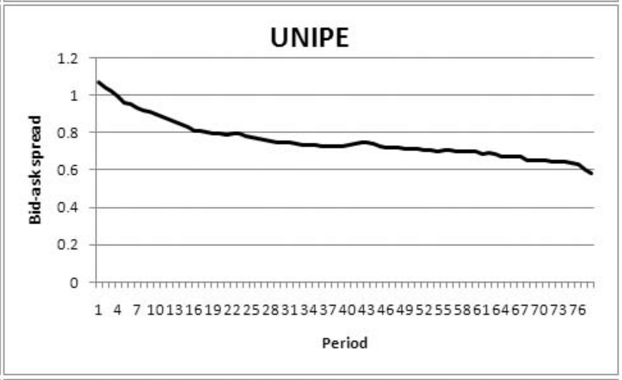
1.A Appendix 1: Basic characteristics of analyzed shares

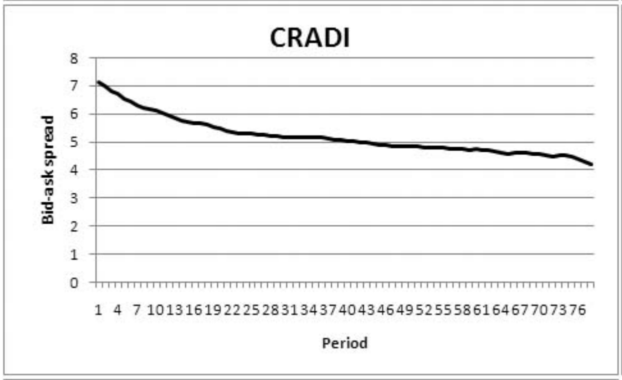
Bid-ask spread - SPAD

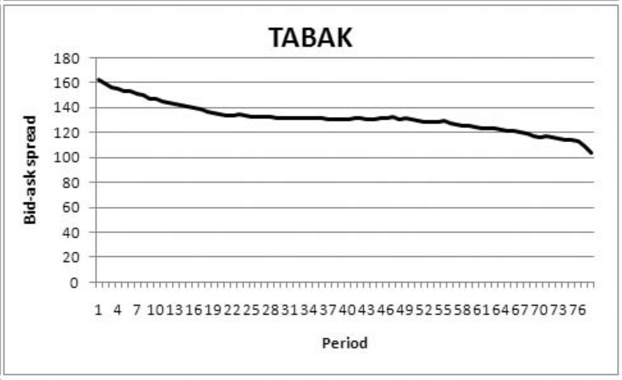


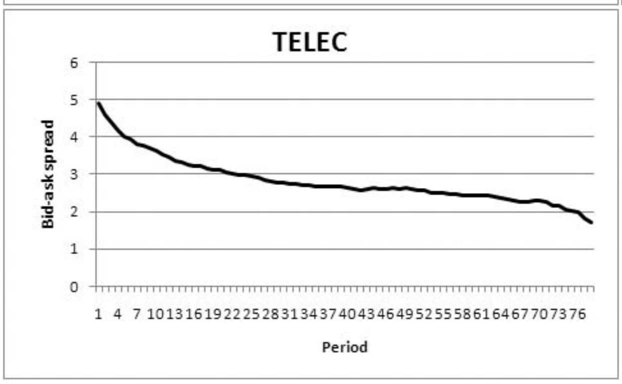




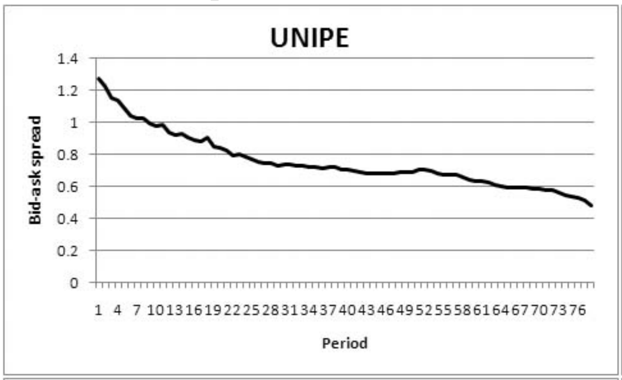


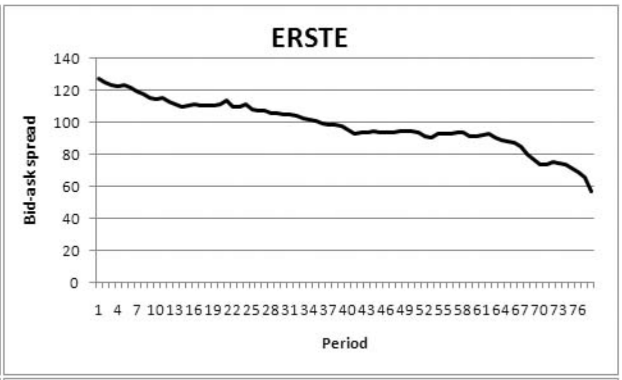


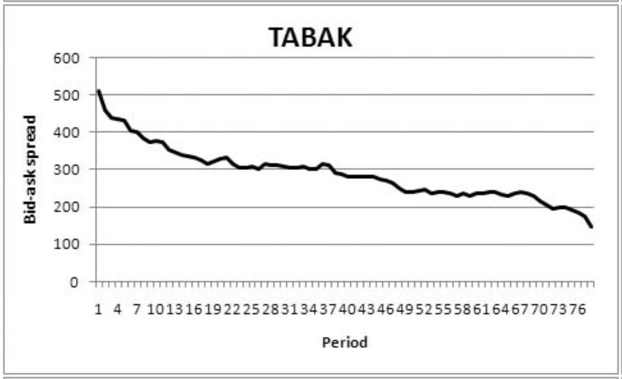


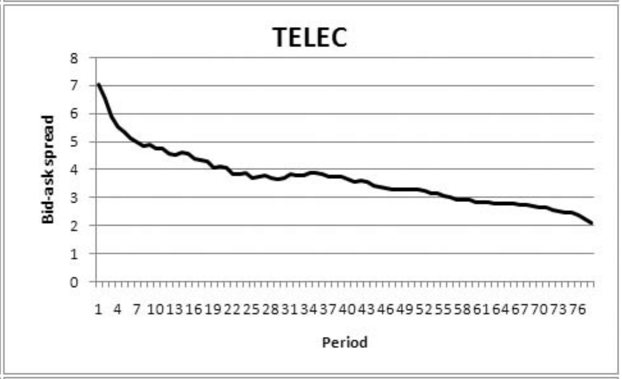


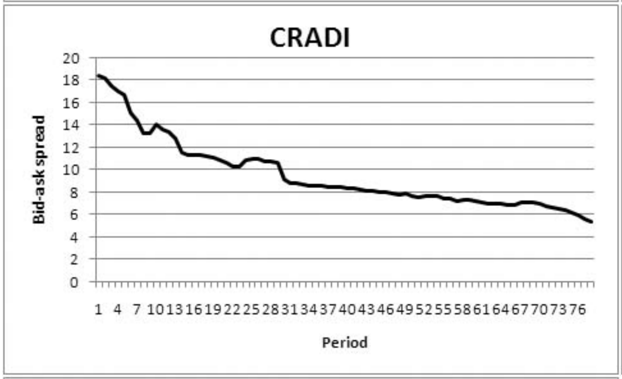
Bid-ask spread - KOBOS

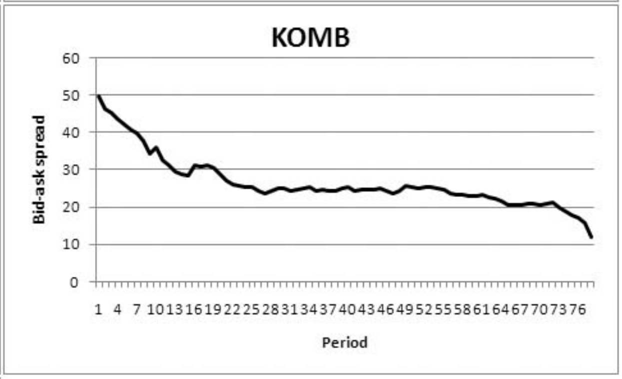


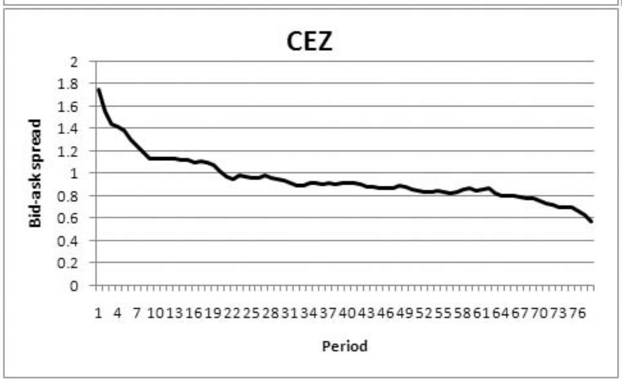




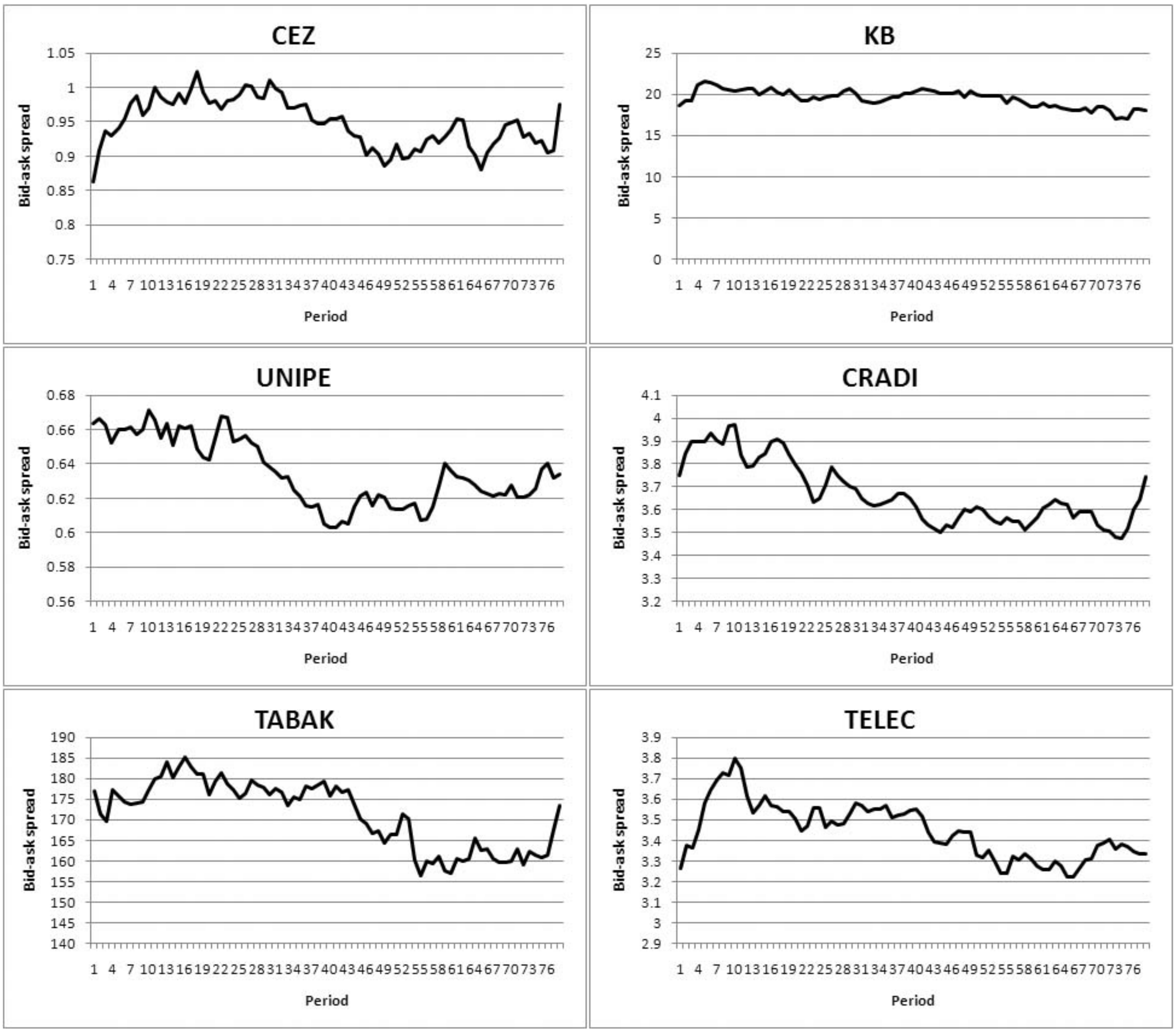




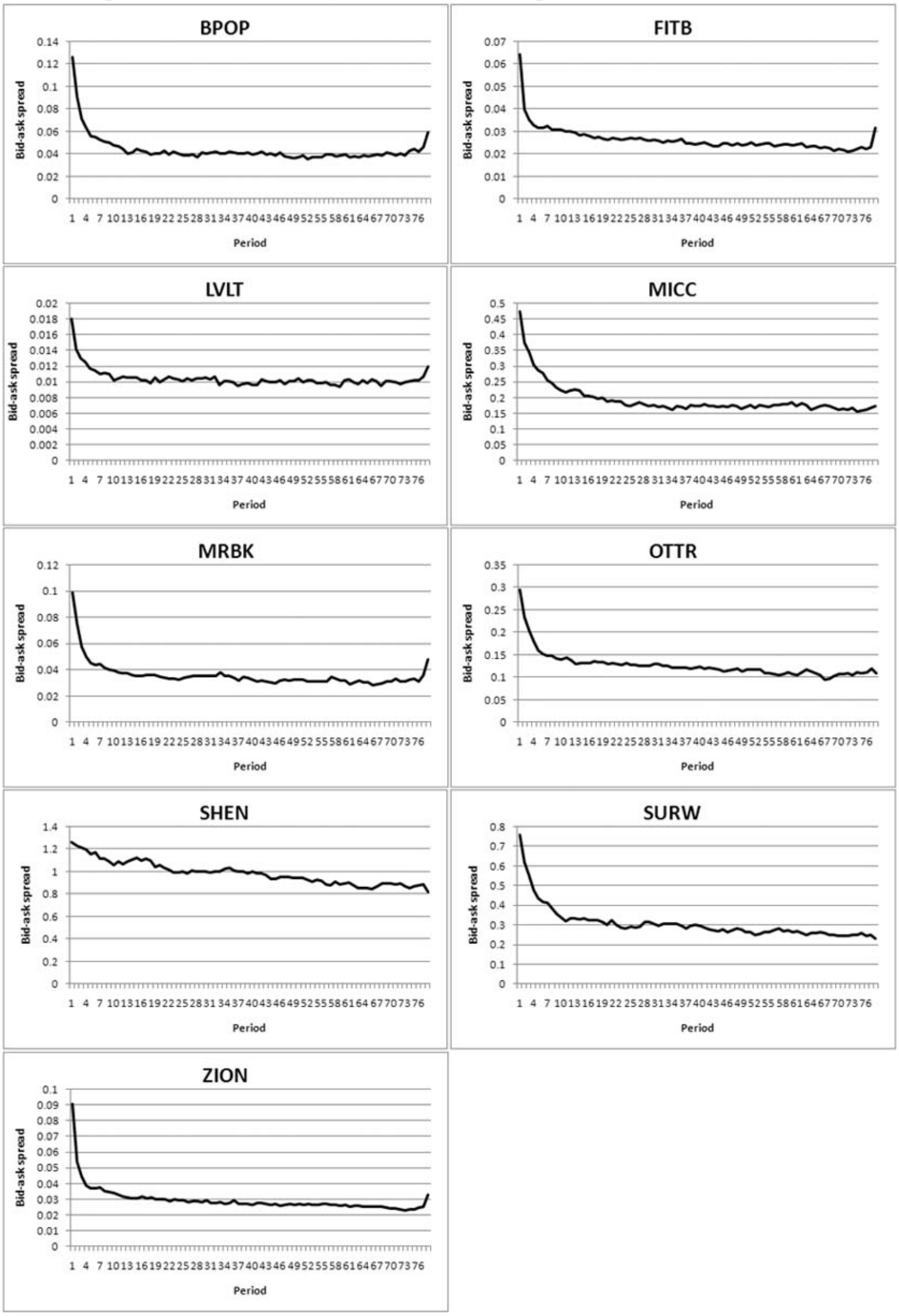




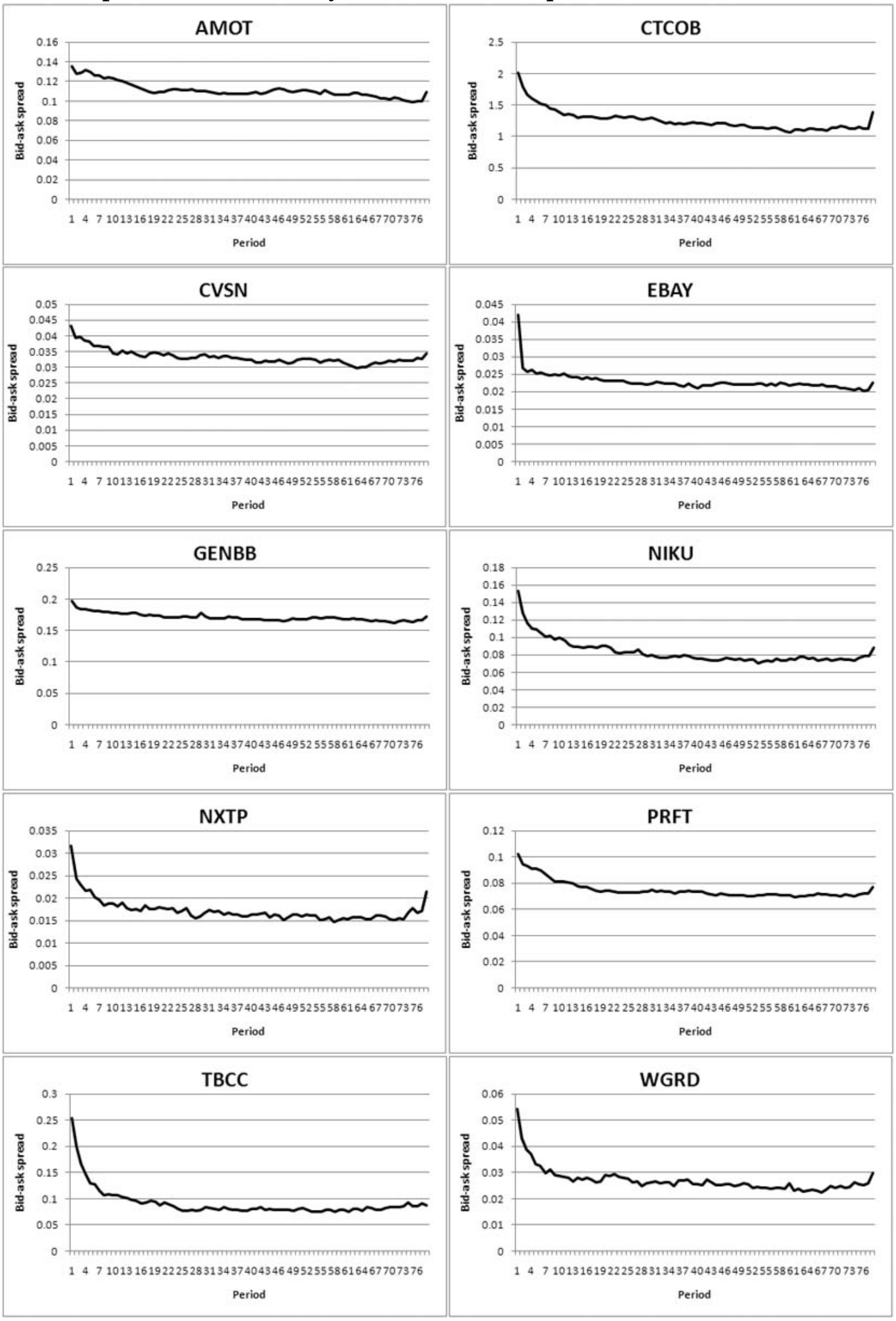
Bid-ask spread - RMS



Bid-ask spread - NASDAQ - 'similar' sample



Bid-ask spread - NASDAQ - 'random' sample



Quantitative data

Ticker	Days	BA changes	Min price	Max price	Avg BA	Avg BA/ Avg price
SPAD						
CEZ	251	98494	87.55	146.3	.804	.00687
CRADI	251	31906	181.5	345	5.19	.0197
ERBAG	251	67898	1860	3210	13.3	.00527
KOMB	251	147784	1824	2669	11.9	.00530
TABAK	251	57270	10500	15700	132	.0101
TELEC	251	90393	242	379.2	2.83	.00910
UNIPE	251	52393	34	66.5	.757	.0151
KOBOS						
CEZ	251	98724	83.4	147	1.05	.00915
CRADI	249	40887	180	345	5.67	.0688
ERBAG	89	17016	1861	3200	103	.0406
KOMB	251	144812	1578	2800	26.4	.0121
TABAK	245	87933	10490	15785	289	.109
TELEC	251	70796	219.6	400	3.67	.0407
UNIPE	251	40280	34	70	.752	.0145
RMS						
CEZ	244	72366	84.4	147.4	.948	.00818
CRADI	247	33878	178	353.8	3.67	.0138
KOMB	247	69012	1806.4	2679	19.5	.00871
TABAK	166	48824	10560.5	15766.8	188	.0143
TELEC	247	48787	237.6	390	3.45	0.0110
UNIPE	247	33045	33.4	68.3	.635	.0125

The Czech tickers include the prefix BAA.

Ticker	Days	BA changes	Min price	Max price	Avg BA	Avg BA/ Avg price
NASDAQ						
'similar'						
BPOP	247	294280	31.7	37.9	.0296	.000851
FITB	247	1742341	27.05	62.12	.0265	.000485
LVLT	247	259834	4.34	7.9	.0296	.00484
MICC	227	193644	1.31	81	.178	.00433
MRBK	247	391202	30.15	45.96	.0359	.000944
OTTR	247	158646	23.75	29.95	.127	.00481
SHEN	247	110035	26.2	55	.985	.0243
SURW	247	199488	22.1	41.48	.306	.00964
ZION	247	921488	39.31	63.86	.0296	.000574
'random'						
AMOT	247	24833	1.49	4.55	.111	.0368
CTCOB	168	45895	22	52	1.25	.0339
CVSN	247	23781	.66	4.39	.0333	.0132
EBAY	247	4487388	50.45	117.86	.0229	.000272
GENBB	247	34911	3.32	10.21	.171	.0253
NIKU	247	46758	2.97	9.82	.0831	.0130
NXTP	247	230584	3.9	13.5	.0172	.00198
PRFT	247	27539	.41	3.95	.0749	.0343
TBCC	247	185130	11.8	30.34	.0914	.00434
WGRD	247	154601	3.85	9.0	.0270	.00421

1.B Appendix 2

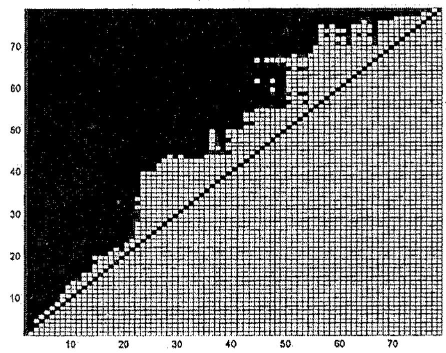
In this appendix we present the results of the detailed analysis of the bid-ask spread intraday pattern in a graphical form. For every share there is one pane on which the result of 3003 tests is presented. The null hypothesis of the mean equality was tested on the 5% (two tail) confidence level.

Every pane is split by a secondary diagonal to the top-left and bottom-right triangles. The top-left triangle presents the rejections of the null hypothesis in favor of a decreasing pattern (i.e. the earlier period bid-ask spread is bigger than the later period). The bottom-right triangle presents the rejections of the null hypothesis in favor of an increasing pattern.

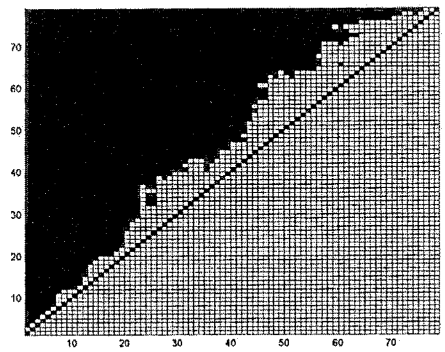
On the axes there are the periods from period 1 to period 78. If the appropriate test is rejected the cell of the intersection of the respective periods is black. E.g. in the case of TELEC traded on RMS the hypothesis of an equal bid-ask spread in periods 1 and 10 can be rejected in favor of an increasing pattern, and between periods 10 and 60 it can be rejected in favor of a decreasing pattern.

On every page shares traded on one stock exchange or in one system are presented. We begin with SPAD, followed by KOBOS and RMS and we finish with NASDAQ.

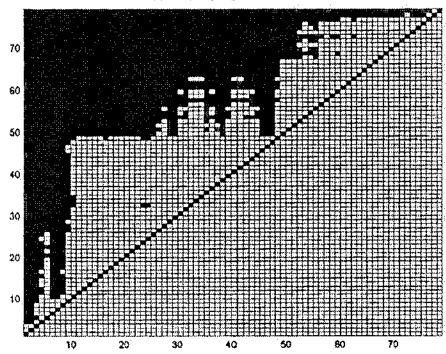
CEZ SPAD



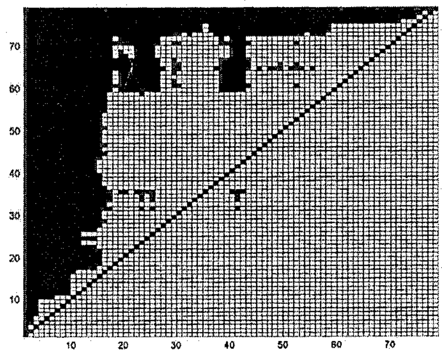
CRADI SPAD



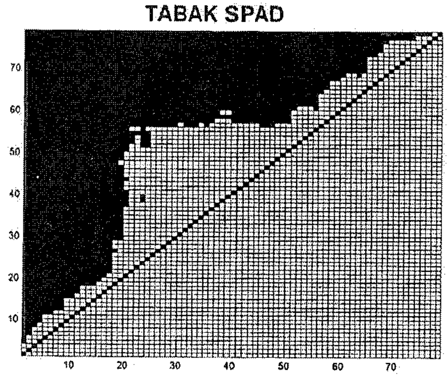
ERBAG SPAD



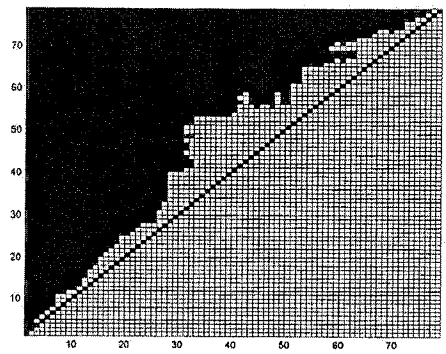
KOMB SPAD



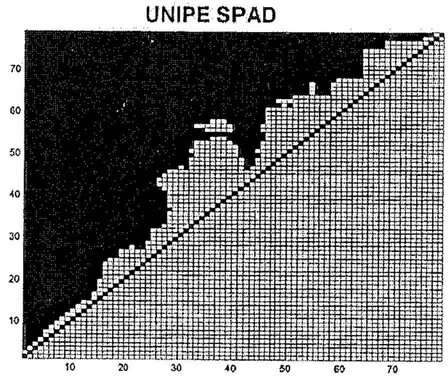
TABAK SPAD



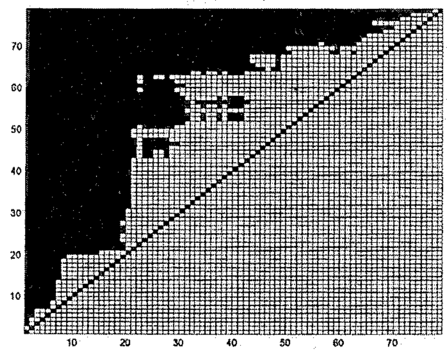
TELEC SPAD



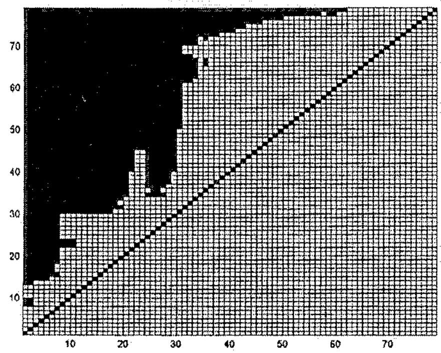
UNIPE SPAD



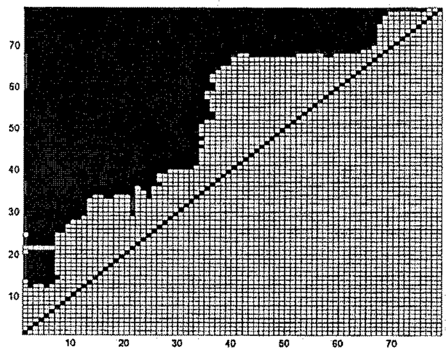
CEZ KOBOS



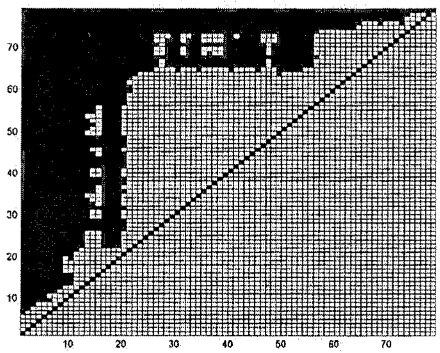
CRADI KOBOS



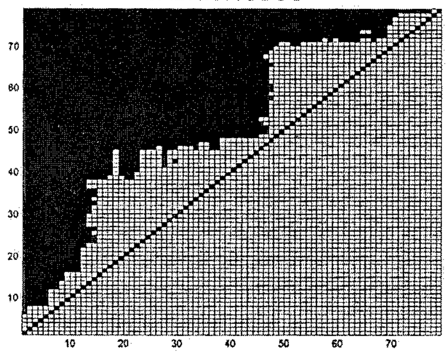
ERBAG KOBOS



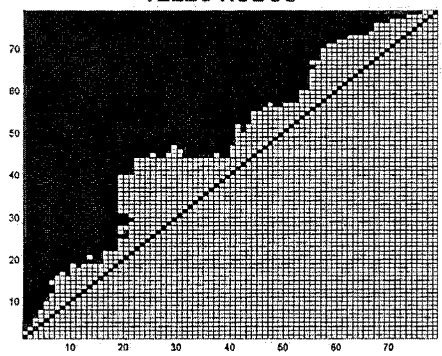
KOMB KOBOS



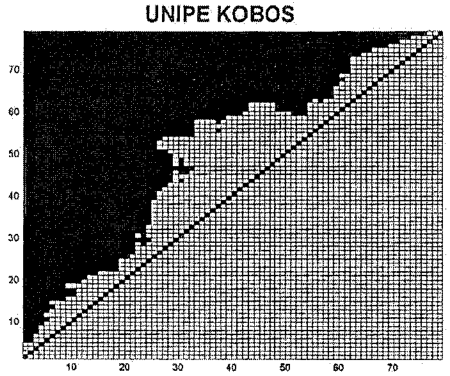
TABAK KOBOS



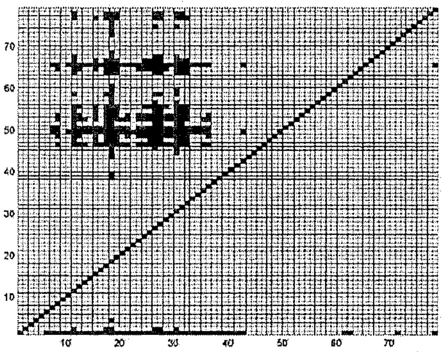
TELEC KOBOS



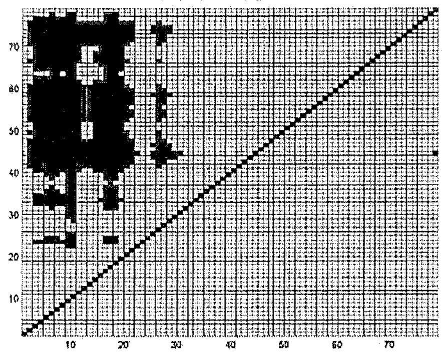
UNIPE KOBOS



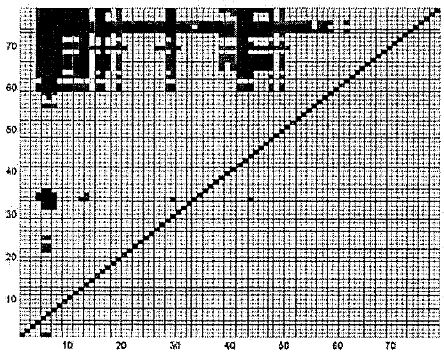
CEZ RMS



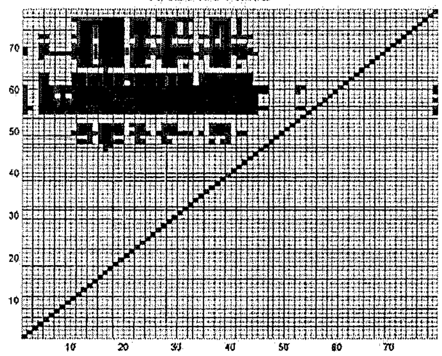
CRADI RMS



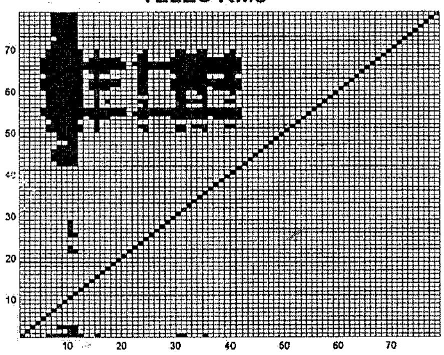
KOMB RMS



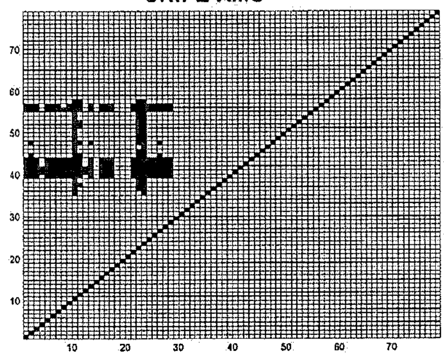
TABAK RMS

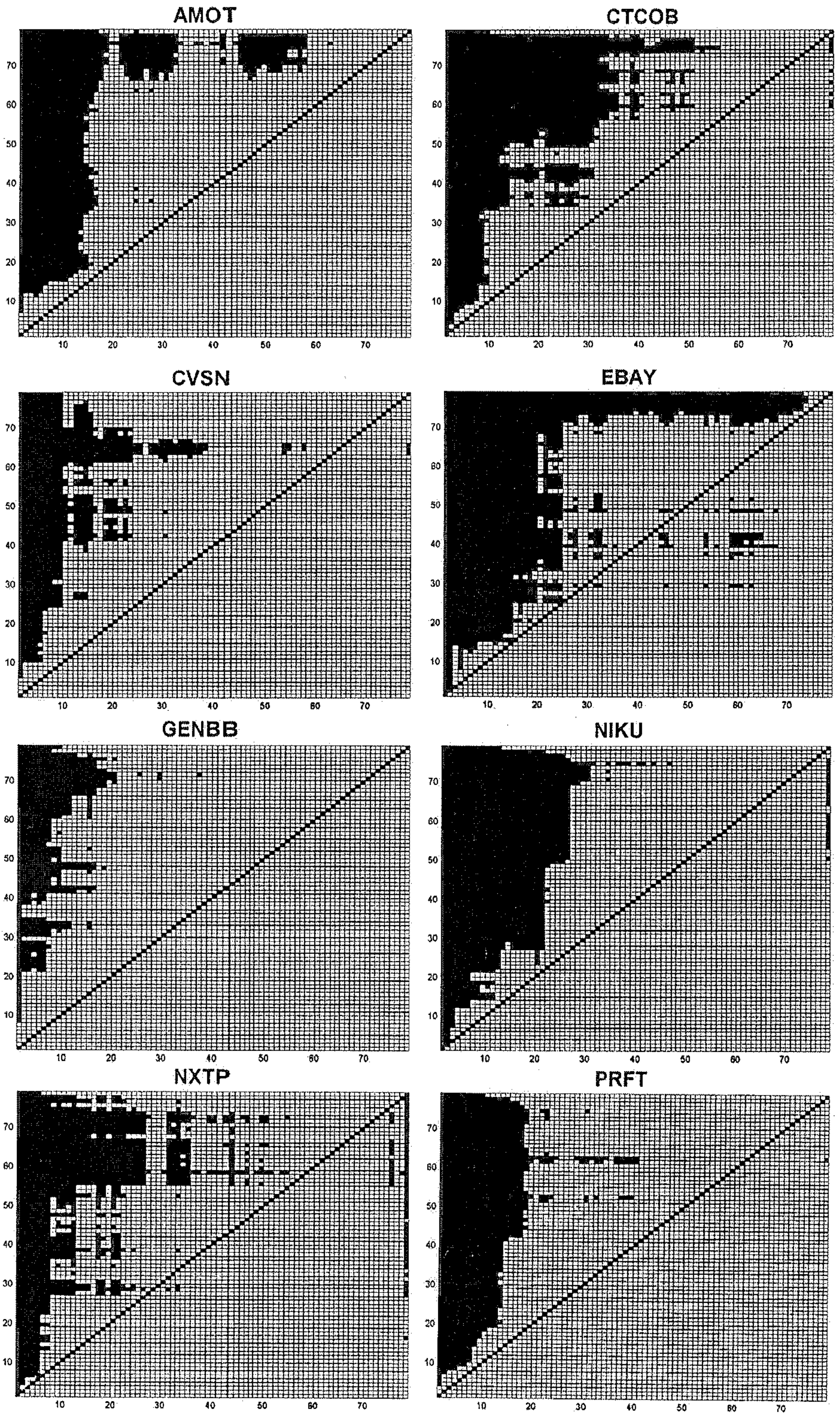


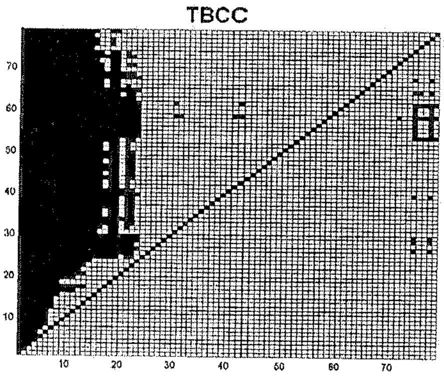
TELEC RMS

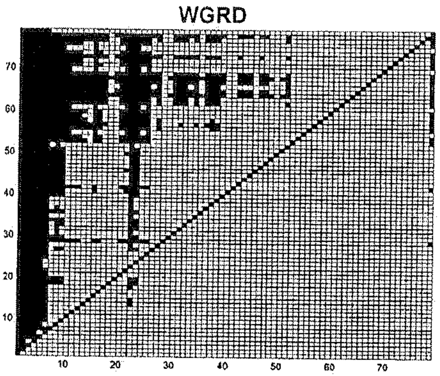


UNIPE RMS




AMOT
CTCOB
CVSN
EBAY
GENBB
NIKU
NXTP
PRFT


TBCC
10
20
30
40
50
60
70


WGRD
10
20
30
40
50
60
70

1.C Appendix 3: Relationship of OLS and standard statistics sample mean comparison

In the model we analyze data y_n^i in 78 groups of the same size N. The main statistics for each group i are

$$\begin{aligned}\mu_i &= \frac{\sum_{n=1}^{N} y_n^i}{N} \\ \sigma_i^2 &= \frac{\sum_{n=1}^{N}(y_n^i - \mu_i)^2}{N-1}.\end{aligned}$$

Instead of analyzing the data directly we can use a model estimated using least squares estimation with White heteroscedasticity correction. The model can be defined as

$$y_n^i = \sum_{j=1}^{78} \alpha_j D^j + \varepsilon_n^i,$$

where D^j are dummies for every group, i.e. the trivial fixed effects model with no other explanatory variables than the group-specific dummies. The hypothesis $\mu_i = \mu_j$ is equivalent to the hypothesis $\alpha_i = \alpha_j$, so the test statistics should be equal for both hypotheses. The confidence interval for the difference $\mu_i - \mu_j$ is $\mu_i - \mu_j \pm t(p, 78N)\sqrt{\frac{\sigma_i+\sigma+j}{N}}$, where $t(\cdot,\cdot)$ is a relevant statistics. We would like to show that this interval is equal to the similar one for $\alpha_i - \alpha_j$.

Let y_n^i be the data sorted by groups, D^i be the dummy for the group i, i.e. the vector of zeros and ones on the $(i-1)N+1$ to iN-th place, and the explanatory variables matrix X consists of the dummies. The coefficients α_i are defined to be

$$\alpha_i = (X'X)^{-1}X'Y.$$

The inverse is a diagonal matrix $(X' * X)^{-1} = 1/N\, I$, and the product $X'Y$ is the sum of the group values, thus $\alpha_i = \mu_i$. The variance-covariance matrix $var(\alpha)$ can be expressed as

$$var(\alpha) = (X'X)^{-1}X'ee'X(X'X)^{-1} = \frac{1}{N^2}IX'ee'XI = \frac{1}{N^2}X'(ee')X,$$

where e is the vector of residuals $Y - X\alpha$, implying $e_n^i = y_n^i - \mu_i$. When we assume heteroscedasticity and no autocorrelation the errors variance-covariance matrix is a diagonal matrix, and its estimated diagonal values are $(e_n^i)^2$ (White estimation). The term $X'ee'X$

equals a diagonal matrix with terms $\sum_{n=1}^{N}(y_n^i - \mu_i)^2$ on the diagonal. The mentioned confidence interval equals $\mu_i - \mu_j \pm t(p, 78N)\sqrt{\frac{\sigma_i + \sigma + j}{N\frac{N}{N-1}}}$. The difference between the two confidence intervals is the term $\frac{N}{N-1}$. The two confidence intervals are equal in the limit and there is just a tiny difference for high values of N; in the limit they are equal. For the purposes of our analysis the difference seems to be negligible.

Chapter 2

Oligopolistic price competition with informed and uninformed buyers

Abstract

The standard price competition of two or more players leads to Bertrand equilibrium in basic economic theory (if complete information is assumed, there are no capacity constraints, etc.). In reality, even on highly competitive Internet-based markets, the prices of seemingly undifferentiated goods (e.g. books and CDs on Amazon and similar e-shops) vary, although competition seems *prima facie* based on prices. I follow the literature that originated with Varian's (1980) model, especially Kocas and Kiyak (2006), and analyze oligopolistic markets where buyers have reservation values drawn from a common distribution function rather than a single value (inelastic demand), as typically assumed in the models of Varian's or Kocas and Kiyak's type. The model presented in this paper is developed from the simplest symmetric set-up (uninformed buyers are assigned to sellers evenly) to the most complex asymmetric set-up with many competing sellers (uninformed buyers are distributed over sellers unevenly). The most complex set-up theoretically rationalizes the empirical findings of Kocas and Kiyak. In the equilibrium of my model, all sellers randomly choose prices from a non-trivial interval for (almost) every seller, while in Kocas and Kiyak's theoretical model only two sellers randomize while others always offer the same price.

Keywords: oligopoly, price competition, price dispersion.

JEL classification: L11, D43.

2.1 Introduction

In this paper, I follow the literature on price dispersion. Using the concept of informed and uninformed buyers, I show that price competition with no capacity constraint can result in pricing that differs from the Bertrand equilibrium pricing and fits empirically observed behavior. Specifically, I derive a model that fits Kocas and Kiyak's (2006) empirical findings better than their own model.

The basic features of the model I develop in this paper have been applied before in a simpler, slightly different set-up. The informed buyers[1] know the prices posted by all sellers and choose the seller with the lowest price, and the uninformed buyers[2] choose sellers randomly. Sellers choose prices so that they benefit maximally from uninformed buyers (which entices sellers to choose high prices) and informed buyers (which entices sellers to choose low prices).

The literature on this topic was initiated by Salop and Stiglitz (1977), and followed by Varian (1980).[3] These articles present models without sequential search in their articles. Consumers are either fully informed or never search and stay uninformed. Other examples of such models can be found in Braverman (1980), Stiglitz (1979), and Narasimhan (1988). In the present paper, I follow this precedent.[4]

Non-sequential, Varian-type search models have been applied to explain promotional strategies (Raju, Srinivasan, and Lal, 1990), international trade (Neven, Norman, and Thisse, 1991 or Baye and De Vries, 1992) and other applications where buyers may have an exogenously given tendency to prefer one product (seller) over another. All these models (with one laudable exception to be discussed presently) assume either a symmetry of sellers, meaning uninformed consumers are distributed evenly over sellers, or a duopoly configuration.[5]

Recently, informed by Varian-type models, empirical researchers have used data collected from Internet e-shops (e.g. Clay, Ramayya, Wolff, and Fernandes, 2002; Clemons, Hann, and Hitt, 2002; or Iyer and Pazgal, 2003).

[1]In the literature they are also called switchers, or searchers, even though search may not be modelled in the paper. I stick to the term "informed buyers" in this paper.

[2]In the literature they are also called non-searchers or loyal buyers.

[3]The model I derive in this paper is mainly based on Varian (1980) and Kocas and Kiyak (2006), so I refer to these models as Varian-type models.

[4]A related stream of literature does feature sequential search. Burdett and Judd (1983), Reinganum (1979), Stahl (1989), Stiglitz (1987), and Wilde and Schwartz (1979) are prominent examples.

[5]Additionally, Sinitsyn (2008) theoretically characterizes the basic features of the Varian-type model class. He shows that for a Varian-type model with heterogeneous tastes, a mixed-strategy equilibrium exists and identifies the basic features of the support of the equilibrium strategies.

The most recent and most relevant research in the present context is Kocas and Kiyak (2006). These authors take the Varian (1980) model, where uninformed buyers choose a seller randomly (i.e. the probability an uniformed buyer chooses a specific seller is equal for all sellers) and add an asymmetry introduced by Narasimhan (1988): Uninformed buyers always go to "their" seller. The amount of uninformed buyers (which can be thought of as loyal) can be different for every seller. Kocas and Kiyak show that at most two sellers offer discounts,[6] while others choose the monopoly price. This theoretical result does not correspond well to the reality of the online-book retailer market in which all sellers offer discounts.[7] The authors enrich the model by seller-specific reservation prices to increase price dispersion. They fail, however, to get mixed strategies for all sellers as observed in reality. The result seems to be caused by the assumption of fixed reservation values. In this paper, I show that by relaxing this assumption and, instead, assuming that the reservation prices of all buyers are drawn from a commonly known distribution function, all sellers apply indeed mixed strategies. I also show that the lower bound of the support of the mixed-strategy distribution function is increasing with the increasing share of uninformed buyers. This result is in line with the empirical findings that Kocas and Kiyak provide.

I develop my model from the simplest set-up to the most complex, analyzing gradually symmetric duopoly and oligopoly (section 2) and then asymmetric duopoly and oligopoly (section 3).[8] Section 4 concludes.

2.2 Symmetric model

I start the analysis with the simplest possible model: symmetric duopoly. Two sellers offer undifferentiated products. Both sellers have to post a price so that everyone can buy the product at the given price, and they cannot discriminate between individual buyers. There is a mass of potential buyers willing to buy one unit of the product. Each

[6]Offering a discount implies that they follow a mixed strategy in equilibrium. The mixed strategy nature of the equilibrium is standard in non-sequential, Varian-type models. The mixed strategy equilibrium results from optimal price-setting to capture the informed buyers (setting a low enough price to beat competitors) and to gain from the uniformed ones (setting a high enough price). Kocas and Kiyak show that in their set-up "only those with the least to lose from deep price cuts will offer discounts" (Kocas and Kiyak, 2006, page 90), and others do not want to compete and sell only to the uninformed buyers.

[7]The fact that all sellers offer discounts is empirically documented in their article; see Figure 3 of Kocas and Kiyak (2006).

[8]The last, and most complex, model includes also the simpler cases. Due to its complexity, I think it is better to understand the mechanics of the simpler models first.

buyer has a reservation value drawn from a commonly known distribution function; the reservation value itself is the private information of each buyer.

A buyer can either be informed about prices posted by the sellers or be uninformed. An informed buyer either buys one unit of the product from the seller with the lower posted price (if it is lower than the reservation value) and leaves the market or does not buy anything. An uninformed buyer chooses randomly, with equal probabilities, one seller and then compares the posted price with his reservation value. If the reservation value is higher than the posted price, the buyer gets one unit of the product and leaves the market; otherwise, he does not buy the product and, without checking the price of the other seller, he leaves the market.[9]

The game is one-shot. There is no repetition. Search is not possible.[10]

The model is symmetric because the probability that an uninformed buyer comes to one of the sellers equals the probability he comes to the other. From a seller's perspective, there is a fixed probability that a specific buyer comes to her whatever price she posts. It is the probability that the buyer is uninformed and he (randomly, with probability 1/2) chooses to come to the seller. The probability that an informed buyer comes to the seller is equal to the probability that the seller posts a price that is lower than her competitor's (and lower than the buyer's reservation value).[11]

This model is similar to Varian's model, with two exceptions: I do not assume perfect competition, and the reservation price is specified by a distribution function rather than by a single value identical for all buyers. Perfect competition, as in Varian's model, can be approximated by zero expected profit in the symmetric model. In the asymmetric model, a zero expected profit equilibrium cannot be reached, and asymmetry rules out perfect competition. Varian's single value reservation price is a degenerate case of the distribution function assumption.

I first show that, as in other similar models, a pure strategy equilibrium does not exist. Thereafter, I derive a mixed strategy equilibrium. All propositions in this paper have a common part of the assumptions that are formulated in the following paragraph.

[9]The difference between informed and uninformed buyers is usually their search cost. A general search cost would make the model too complicated to be solvable, so it is assumed that some buyers have a zero search cost and some of them have too high a search cost to perform any search.

[10]It will be shown that the game has a unique equilibrium in mixed strategies, so a finitely repeated game has unique equilibrium with the same properties as the presently analyzed one-round game.

[11]In the set-up with a continuum of buyers, both sellers have a share of buyers secured, and they compete on price for the remaining buyers.

Common Assumptions

There is a mass of buyers whose distribution of reservation prices, $\hat{D}(\cdot)$, generates a profit function with a single finite local maximum (monopoly price and profit). For notation simplicity, denote $D(p) = 1 - \hat{D}(p)$. Sellers produce identical units of a good at no cost. Buyers are either informed or uninformed about the prices posted by sellers.

Proposition 1

Suppose the Common Assumptions hold, and there are two sellers. Then no pure-strategy equilibrium exists when the probability that a seller sells the good at a higher price than a competitor is positive, and there are some informed buyers on the market.

Proof

Assume the sellers post prices p_1 and p_2.. If the prices are unequal, I can, without loss of generality, assume $p_1 < p_2 \leq p_M$, where p_M is the monopoly price (by the definition of a monopoly price, no seller chooses a price bigger than the monopoly price, although they generally can). Seller 1 can increase the expected profit by marginally increasing her price. If $0 < p_1 = p_2 a$ any seller can increase the expected profit by decreasing her price. The case of $0 = p_1 < p_2$ cannot be an equilibrium either because the expected profit is zero, and both sellers can earn strictly positive expected profit by posting p_M or, in fact, any positive price within the reservation price distribution support. •

The logic behind the mixed-strategy equilibrium results from an optimal weighting of two considerations: making money on volume and competing with the other seller for informed buyers through a low price on the one hand and profiting from a higher price on uninformed buyers on the other. The resulting equilibrium is derived from the distribution function of the buyer's reservation price.

Proposition 2

Suppose the Common Assumptions hold. Assume that there are two sellers, and the probability that a buyer is uninformed (chooses randomly between the sellers) is $0 < \alpha < 1$;[12] he is informed and comes to the seller with the best price.

Then there is a unique (symmetric) mixed-strategy equilibrium with a price-mixing cumulative distribution function (cdf):

$$\begin{aligned}
\beta(p) &= 0, \quad p \leq a, \\
&= \frac{1-\alpha/2}{1-\alpha}\left(1 - \frac{\alpha}{2-\alpha}\frac{M}{pD(p)}\right), \quad p \in [a,b], \\
&= 1, \quad p \geq b,
\end{aligned}$$

where M is the monopoly profit:

$$M = \max{}_q qD(q);$$

a solves the implicit equation

$$aD(a) = \frac{\alpha}{2-\alpha}M;$$

and $b = p_M = \mathtt{argmax}_q \alpha qD(q)$ is the monopoly price.

The expected profit is $\frac{\alpha}{2}\max_q qD(q)$, i.e. the profit a seller would secure by setting a monopoly price and selling to 'her' at a fraction of the market.

Proof

The beginning of the proof is standard, similar to the proofs done by, for example, Varian (1980) or Baye and De Vries (1992). We find the equilibrium mixing strategy represented by a distribution function of price choice. By the definition of mixed-strategy equilibrium a the distribution function makes the competitor indifferent between price choices over a specified interval (the support of the mixing function).

Let β_i be the mixed strategy distribution function of seller i, $i = 1, 2$ (I assume that even an asymmetric distribution may exist though I will show later that the distribution function $\beta()$ is unique and therefore, the symmetric equilibrium is unique). Assume first that the support of the distribution function is an interval (a_i, b_i).[13] The lower bounds

[12] For example, each seller has 'in expectation' secured by $\alpha/2$ fraction of the market, i.e. the seller sells on average to this amount of buyers. Of course, the offered price must be smaller than the buyer's reservation price for the seller to be able to sell the product to him.

[13] The interval is open, i.e. the probability that a player chooses a price equal to a or b is zero. It is shown later that there are no mass points of the distributions.

of the supports of both β_i must be identical because setting a lower $a_i < a$ would be a mistake, increasing it by half of the difference $(a_{-i} - a_i)$ would increase the profit due to the single local maximum of the profit function property (the probability of winning informed buyers would be 1 in both cases, and the profit would be higher when the price is set to $(a_{-i} - a_i)/2$).

The expected profit function can be expressed as

$$\begin{aligned}
\pi_i(p) &= \frac{\alpha}{2} pD(p) + (1-\alpha)pD(p)\left(1 - \beta_{-i}(p)\right) \\
&= \left(1 - \frac{\alpha}{2}\right)\left(\frac{\alpha/2}{1-\alpha/2} pD(p) + \frac{1-\alpha}{1-\alpha/2} pD(p)(1 - \beta_{-i}(p))\right) \\
&= \left(1 - \frac{\alpha}{2}\right) pD(p) \left(1 - \frac{1-\alpha}{1-\alpha/2}\beta_{-i}(p)\right) \\
&= \left(1 - \frac{\alpha}{2}\right) pD(p)(1 - c\beta_{-i}(p)),
\end{aligned}$$

where $c = \frac{1-\alpha}{1-\alpha/2}$.

From this point on, I simplify the notation of the profit and the distribution function indices: $\pi() = \pi_i()$ and $\beta() = \beta_{-i}()$.

For the competitor to be indifferent between choosing prices $p_1, p_2 \in (a, b)$, the expected profits related to the two prices must be equal.

$$\begin{aligned}
\pi(p_1) &= \pi(p_2) \\
p_1 D(p_1)(1 - c\beta(p_1)) &= p_2 D(p_2)(1 - c\beta(p_2)) \\
\frac{1 - c\beta(p_1)}{1 - c\beta(p_2)} &= \frac{p_2 D(p_2)}{p_1 D(p_1)}.
\end{aligned}$$

Since $\beta(a) = 0$, I can rewrite the equation to

$$1 - c\beta(p_1) = \frac{aD(a)}{p_1 D(p_1)} \tag{2.1}$$

that gives me the mixed strategy function formula

$$\beta(p) = \frac{1}{c} - \frac{aD(a)}{cpD(p)}.$$

The expected profit is equal to $\pi = (1 - \alpha/2)aD(a)$ because setting the price equal to the lower bound a we would sell the good to all searchers almost sure,[14] and the

[14] As the support of the mixed strategy is open, a is never chosen. A completely correct proof would

expected profit of all choices of price must be equal. From the equality and the profit when the monopoly price is chosen ($\max \alpha/2\, pD(p)$), it is possible to derive the following inequality:[15]

$$\pi = (1 - \alpha/2)aD(a) \geq \max \alpha/2\, pD(p). \tag{2.2}$$

I assume for now that the mixing function $\beta(b)$ is continuous everywhere, and that there are no "gaps", i.e. the mixing function is increasing. The discontinuity case and "gaps" are discussed at the end of the proof. The continuity and equality $\beta(b) = 1$ imply, using formula (2.1),

$$1 - c = \frac{aD(a)}{bD(b)}. \tag{2.3}$$

The upper bound of support b must be lower than or equal to the monopoly price[16] and a must be strictly smaller than the monopoly price. That gives me the implicit inequality:

$$aD(a) \leq (1 - c) \max pD(p),\ a < \texttt{argmax} pD(p),$$

and substituting for $aD(a)$ in the equation $\pi = (1 - \alpha/2)aD(a)$ derived above, I get the formula for maximum expected profit:

$$\pi \leq (1 - \alpha/2)(1 - c) \max pD(p).$$

Substituting for c, I obtain

$$\pi \leq \alpha/2 \max pD(p).$$

Using inequality (2.2), the equality holds, i.e.

$$\pi = (1 - \alpha/2)aD(a) = \alpha/2 \max pD(p).$$

use limits and is straightforward.

[15] When the monopoly price is chosen, the seller sells only to uninformed customers. Therefore the expected profit π must be at least as big as if the monopoly price is chosen.

[16] A formal proof of the equality $b = p_M$ follows. Informally, if $b < p_M$, when choosing the price b the probability of selling to informed buyers is zero; therefore, $\pi(b) < \pi(p_M)$, which contradicts the equilibrium assumptions.

The lower bound of the mixing interval a is the solution of the implicit equation

$$aD(a) = \frac{\alpha}{2-\alpha} \max pD(p),[17]$$

and using equation (2.3), the upper bound b is equal to the monopoly price.

Gaps in β

A gap in a distribution is defined to be an interval where the mixed strategy function is constant (different from 0 and 1). We will show there are no gaps in the functions β_i. Assume the mixing functions of the two competitors are not equal, $\beta_1 \neq \beta_2$. Assume that β_1 is constant in the interval $(c, d) \subset (a, b)$. If player 2 chooses a price in this interval with a positive probability, she would make a mistake. She would increase her expected profit when she moves the whole mass from the interval (c, d) to the point d. Therefore, gaps must be symmetric if there are any.

Assume there is a symmetric gap $(c, d) \subset (a, b)$, and it is the lowest one (if there are many gaps (c_i, d_i), c and d are the smallest limits). Note that $c > a \geq 0$ because a is the lower bound of the mixed strategy function. $\epsilon > 0$ exists such that a player benefits from moving the probability from the interval $(c - \epsilon, c)$ to the point d. For ϵ, it suffices to satisfy the equation $\beta(c - \epsilon)d > \beta(c)c$, i.e. the benefits outweigh the losses. Thus, there are no gaps in the distribution.

Discontinuity of β_i

The discontinuity of β_i means that there are atoms in the mixed strategy cumulative distribution function, or that a particular price is chosen with a strictly positive probability. I will show that such a price does not exist in equilibrium.

First, note that there can be only a symmetric equilibrium. We have shown before that the lower bounds of the mixing intervals are identical for all sellers. The formula for the mixing function $\beta()$ is dependent only on the share of uninformed buyers α, mixing interval lower bound a and the distribution function of the reservation prices. Therefore, the distribution function must be identical for all sellers.

This implies that, if there were a mass point, it would be identical for both players. In that case, an undercutting procedure would start - a move of the mass point to a

[17] Again, solution existence and uniqueness follow from the single local maximum property of the profit function.

slightly smaller price would be beneficial (again, recall that $a > 0$). Therefore, the mixed strategy cumulative distribution functions β_i are continuous.

•

The solution can be evaluated even in the extreme case of $\alpha = 1$ when each seller owns half of the market. In that case, the mixing interval shrinks to one point $a = b = p_{monopoly}$. In the other extreme, $\alpha = 0$, when all buyers are informed, and the lower bound of the interval as well as the expected profit are equal to 0, the upper bound should be also 0. Therefore, even the limits in extreme cases are in line with the existing theory.

The above model has been developed for price competition between two sellers. Naturally, we may ask what happens when a higher number of sellers enters the market. I keep the basic set-up unchanged and solve it for an arbitrary number of N competing sellers. The characteristics of the buyers remain the same. The probability of a buyer choosing a seller randomly (and not searching for the lowest price) is α. It means that every seller has secured a α/N share of the market. The model equilibrium is characterized in the following proposition.

Proposition 3

Assume that the Common Assumptions hold, and there are N sellers. Assume that the probability a buyer is uninformed and chooses randomly between the sellers is $0 < \alpha < 1$, and the probability a buyer chooses one particular seller is identical for all sellers and is equal to α/N. The probability that a buyer is informed and knows the posted prices of all sellers in advance is therefore $1 - \alpha$.

Then, there is no pure-strategy equilibrium, and a unique symmetric mixed-strategy equilibrium with price-mixing cdf is

$$\begin{aligned}
\beta(p) &= 0, \quad p \leq a \\
&= 1 - \left(\frac{\alpha}{N(1-\alpha)} \frac{M}{pD(p)} - \frac{\alpha}{N(1-\alpha)} \right)^{\frac{1}{N-1}}, \quad p \in [a, b] \\
&= 1, \quad p \geq b,
\end{aligned}$$

where M is the monopoly profit, $b = p_M$ the monopoly price, and a solves the implicit equation $aD(a) = \frac{\alpha}{N-(N-1)\alpha} M$.

The expected profit is $\frac{\alpha}{N} \max_q qD(q)$, i.e. the profit a seller would secure setting a monopoly price and selling to 'her' fraction of market.

Proof

The proof of the non-existence of the pure-strategy equilibrium is analogous to the proof of Proposition 1.

The logic behind the derivation of the mixed-strategy equilibrium copies the proof of Proposition 2; thus I indicate here just the key points of the proof.

I assume the symmetry of the equilibrium. Let $\beta(p)$ label the mixed strategy cdf and the interval (a,b) its support. The associated cdf of minimum price set by $N-1$ competitors is $B(p) = 1 - (1 - \beta(p))^{N-1}$. I can express the expected profit as

$$\pi(p) = \frac{\alpha}{N} p + (1 - \alpha) p (1 - B(p)).$$

Using the argument that $\pi(p_1) = \pi(p_2), \quad \forall p_1, p_2 \in (a, b)$ I get the counterpart of the equation (2.1):

$$1 - cB(p) = \frac{aD(a)}{pD(p)}, \tag{2.4}$$

where $c = \frac{1-\alpha}{1-\frac{N-1}{N}\alpha}$. Substituting for $B(\cdot)$ and rearranging the equation I obtain the formula for the mixing function:

$$\beta(p) = 1 - \left(\frac{1}{c} \frac{aD(a)}{pD(p)} - \frac{1-c}{c} \right)^{\frac{1}{N-1}}.$$

Setting the price at the lower bound of the interval (a,b), the seller sells to all informed buyers *almost sure,* and I have to have at least the profit of selling only to uninformed buyers at the monopoly price, so[18]

$$\pi = \left(1 - \frac{N-1}{N} \alpha \right) aD(a) \geq \max \frac{\alpha}{N} qD(q). \tag{2.5}$$

Equation (2.4) expressed for $p = b$ gives me the inequality

$$aD(a) = (1 - c) bD(b) \leq (1 - c) \max qD(q), a < \texttt{argmax} qD(q).$$

[18] As in the previous proof, the precise derivation of the first equality is the limit $\pi =$ (Probability of winning)*(amount won), where the first term approaches 1 when $p \to a$, and the second term approaches $\max \frac{\alpha}{N} qD(q)$.

Together with (2.5) and substituting for c I get

$$aD(a) = \frac{\alpha}{N-(N-1)\alpha} \max qD(q),$$

and

$$\pi = \frac{\alpha}{N} \max qD(q).$$

Substituting for c and $aD(a)$ in the equation defining $\beta(p)$, I obtain the final form of the mixing function.

The proof of the continuity and strict monotonicity (no gaps property) of the distribution function $\beta(\cdot)$ is completely analogous to the proof of Proposition 2.

Uniqueness of the upper bound

If there were two different upper bounds $b_1 \neq b_2$, then there would need to be a gap or a mass point in the mixing function, which is ruled out. If $b < p_M$, then the expected profit would need to be smaller than α/Np_M; therefore, sellers would tend to choose $p = p_M$, and the upper bound is therefore unique.

Uniqueness of the lower bound

Assume sellers choose different lower bounds $a_1 \leq a_2 \leq a_3 \leq ...$ The lowest of all bounds has to be equal to $a_1 = a$. To set a lower bound is unprofitable, and someone has to choose it because the equilibrium profit is equal to the profit gained from choosing a. Also, at least one other player has to have the lower bound equal to a; otherwise, seller 1 would tend to increase her lower bound. Assume that $M < N$ players have the lower bound of the mixing interval equal to a and the seller $M+1$ has the lower bound equal to $\hat{a} > a$. When I set the number of sellers N to be a parameter of the function $\beta = \beta(p, N)$, then the mixing function of the first M sellers β_1 in the first part of the interval would be $\beta_1(p) = \beta(p, M), p \in (a, \hat{a})$. As $\frac{\partial}{\partial N}\beta(p, N) < 0$, the inequality $\beta_1(p) < \beta(p, M), p \in (a, \hat{a})$ holds, and the player would get a higher-than-equilibrium profit choosing a price from the interval $(a, \hat{a})$ that is, however, outside her mixing interval. Therefore, her mixing interval lower bound has to be equal to a, and the equilibrium is unique.

Because the bounds of the supports of the mixed strategy cumulative distribution functions are equal for all sellers, there are no gaps or mass points, and since the mixed strategies follow the above derived formulae, the equilibrium is unique.

•

Equilibrium prices can be influenced by a variation of two basic parameters, the fraction of informed buyers $1-\alpha$ and the number of competitors N. The increasing number of informed buyers (or the increasing probability of the buyer being informed) directly lowers the (expected) profit of sellers and changes the division of the surplus in favor of the buyers.[19] The lower bound of the mixing interval decreases towards zero faster than linearly as the fraction of searchers increases. The shape of the mixing function itself stretches out proportionally over the mixing interval.

The effect of increased competition differs from the effect of the increased proportion of informed buyers. The expected buyers' and sellers' surpluses do not change in the case of increased competition, only the sellers' surplus is spread across a larger number of sellers. The lower bound of the mixing interval (the interval of prices that can be posted by sellers in equilibrium) is inversely proportional to the number of sellers. The shape of the mixing function thus changes as the number of competitors increases; the probability of choosing a smaller price close to the lower bound decreases and the probability of choosing a price close to the monopoly price increases. There are two countervailing phenomena that influence the expected minimum price on the market: a decreased lower bound of the mixing interval that decreases the expected minimum price on the market and the shifted weight of the mixing function towards the monopoly price, which increases the expected minimum price. I now analyze the minimum in more detail.

Keeping the notation of Proposition 3, the expected minimum price for N sellers is[20]

$$Ep_{min} = a + \int_a^b \left(\frac{\alpha}{N(1-\alpha)} \left(\frac{M}{pD(p)} - 1 \right) \right)^{\frac{N}{N-1}} dp.$$

The effect of increased competition can be determined from the derivative of the expected minimum with respect to N :

$$\frac{dEp_{min}}{dN} = \int_a^b \frac{d}{dN} \left(\frac{\alpha}{N(1-\alpha)} \left(\frac{M}{pD(p)} - 1 \right) \right)^{\frac{N}{N-1}} dp.$$

The expression in the outer brackets gets a value of 1 for $p = a$ and a value of 0 for

[19] I assume there are no search costs in the model. A natural extension of the model would be the introduction of the endogenous choice of the fractions of informed buyers. In such a model, I would observe not only a transfer of surplus from sellers to buyers, but also a decrease of total surplus due to increased search costs.

[20] The expected minimum price is equal to $\int_a^b p d\hat{B}(p)$, where $\hat{B}(p) = 1 - (1-\beta(p))^N$. Integrating by parts, I get the formula for the expected minimum price.

$p = b$; thus, its values in the integration limits belong to the interval $(0, 1)$. The exponent is always bigger than 1, and it is decreasing in N so the whole function is decreasing in N, and the derivative is always negative (the derivative exists). It implies that the expected minimum price decreases with increasing competition.

This result implies that informed buyers benefit from increased competition, sellers maintain the expected sum of profits unchanged as competition increases, and uninformed buyers fully pay the benefit of informed buyers.

To summarize, an increased number of informed buyers depletes the profit of sellers. An increased number of sellers leads to the total volume of profits being split across a larger number of competing sellers, while the absolute value remains unchanged. Informed buyers benefit from increased competition while the uninformed lose.

2.3 Asymmetric model

I now add asymmetry to the model. In the symmetric set-up presented so far, uninformed buyers randomly selected a seller they decided to visit. Randomness was symmetric with respect to sellers: The probability that a specific uninformed buyer would visit a specific seller was equal across all sellers and buyers. In an asymmetric set-up, the probability that a buyer visits a specific seller may differ across sellers.

An asymmetric set-up can capture the behavior of uninformed buyers that for various reasons may not be randomizing. A real-life situation captured by such a model would be brand sensitivity (or even loyalty). A fraction α_1 of potential buyers prefers a particular seller, another fraction α_2 of buyers prefers a different seller, and the remaining buyers are searchers. If we randomly select a buyer in this brand sensitivity model, the probability that he prefers the first seller is α_1, and the probability that he prefers the second seller is α_2. Thus, in the brand sensitivity model, I can think of buyers that prefer a seller as uninformed buyers and of searchers as informed buyers.[21]

The formulation of the model remains essentially the same as in the symmetric case with the exception that uninformed buyers are positioned asymmetrically. For every seller, I specify her share of buyers who are not informed of the lowest price and/or who do not go directly to other sellers. These shares may be different for every seller.[22] As in

[21] The model can be used in the field of promotion if I assume brand sensitivity, or loyalty, can be endogenized assuming it can be changed by sellers' investments.

[22] Or, in other words, I specify the probabilities that a specific buyer visits each seller even if the seller does not post the lowest price on the market.

the previous section, I first analyze the simplest possible model - the competition of two sellers.

Proposition 4

Assume that the Common Assumptions hold, and there are two sellers. Assume that the probability a buyer is uninformed and visits the first seller is α_1, and the probability the buyer is uninformed and visits the second seller is α_2.[23] Assume, without loss of generality, that $\alpha_1 > \alpha_2$. The remaining probability $1 - \alpha_1 - \alpha_2$ represents informed buyers.

Then there is a unique mixed-strategy equilibrium with price-mixing cdf of the first seller:[24]

$$\begin{aligned}\beta_1(p) &= 0, \quad p < a, \\ &= \frac{1-\alpha_1}{1-\alpha_1-\alpha_2}\left(1 - \frac{aD(a)}{pD(p)}\right), \quad p \in [a, p_M), \\ &= 1, \quad p \geq p_M,\end{aligned}$$

and the price-mixing cdf of the second seller

$$\begin{aligned}\beta_2(p) &= 0, \quad p < a, \\ &= \frac{1-\alpha_2}{1-\alpha_1-\alpha_2}\left(1 - \frac{aD(a)}{pD(p)}\right), \quad p \in [a, p_M], \\ &= 1, \quad p \geq p_M,\end{aligned}$$

where M is the monopoly profit, p_M the monopoly price, and a solves the implicit equation

$$aD(a) = \frac{\alpha_1}{1-\alpha_2} M.$$

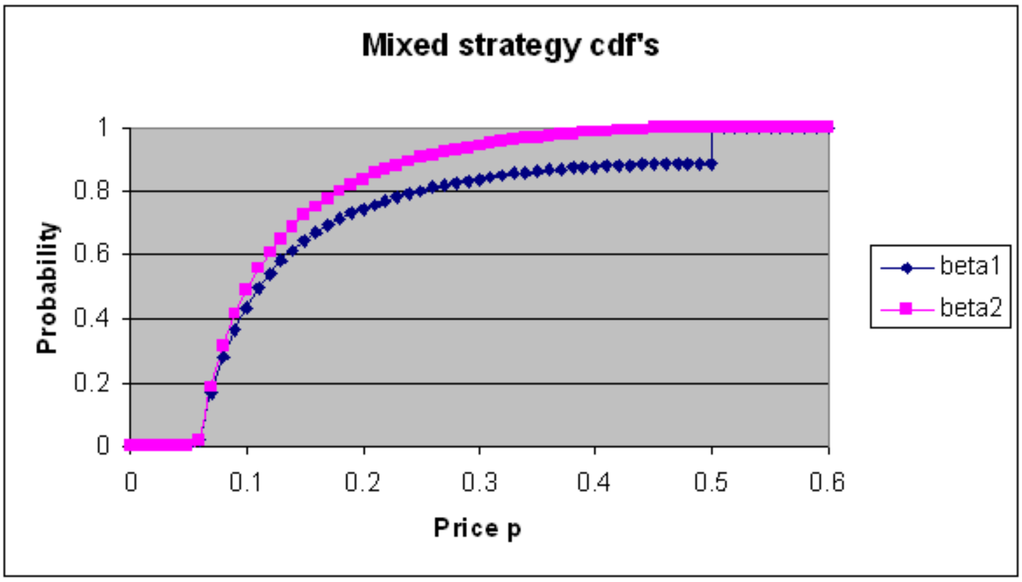
The expected profit of the stronger seller (the seller with a higher share of uninformed buyers) is $\alpha_1 \max_q qD(q)$, and the profit of the weaker seller is $\alpha_1 \max_q qD(q) - (\alpha_1 - \alpha_2)aD(a)$. •

In equilibrium an interesting behavior can be seen close to the monopoly price. On the graph below, the equilibrium of the set-up $\alpha_1 = 0.2, \quad \alpha_2 = 0.1$, and $D(p) = 1 - p$,

[23]This means the first seller has secured a fraction α_1 of the market and the second seller has secured α_2 of the market.

[24]Note there is a positive probability that seller 1 chooses the monopoly price, and the mixed strategy distribution function has a mass point in p_M. The result and the mathematic properties of the equilibrium are discussed between this Proposition and the Proof.

(uniformly distributed reservation values of buyers over the interval $(0, 1)$) is plotted. The monopoly price is $1/4$, and the lower bound of the mixed strategies is approximately 0.06 .



We can see that there is a probability of more than 10% that the stronger seller sets the monopoly price. The behavior of the weaker and the stronger seller just below the monopoly price is clearly visible there. The weaker player sets a price from any interval $(p_M - \epsilon, p_M), \quad \epsilon > 0$ with positive probability, but the price p_M is never selected. This behavior prevents the stronger player from shifting the probability mass to p_M so that a gap would appear under the monopoly price. On the contrary, the stronger player selects the prices below the monopoly price with probabilities much lower than 1 so that the weaker player is motivated to assign some probability to these prices. The mass point in the monopoly price is formed by the "remaining" probability.

Mathematically, this situation is described by the supports of the associated probability functions. The support of the weaker seller is (a, p_M), i.e. any price can be selected but p_M. The support of the stronger seller is $(a, p_M]$, i.e. any price can be selected including p_M, which is selected with a positive probability.

Proof of Proposition 4

The proof follows the same logic as the proofs of Propositions 2 and 3. First, I develop the expected profit function of the two competing sellers. In the previous section, the expected profit functions coincided for both sellers, and the equilibrium strategies were equivalent. Because the set-up is asymmetric, the solution cannot be symmetric either. I have to develop profit functions and strategies for each seller individually. I label the first seller's expected profit function and strategy (the cdf of the price choice) $\pi_1(p)$ and $\beta_1(p)$, respectively, and $\pi_2(p)$ and $\beta_2(p)$ analogously label the corresponding variables for the second seller. The expected profit functions of the two sellers are:

$$\begin{aligned}
\pi_1(p) &= \alpha_1 pD(p) + (1-(\alpha_1+\alpha_2))pD(p)\,(1-\beta_2(p)) \\
&= (1-\alpha_2)pD(p)(\frac{\alpha_1}{1-\alpha_2} + \frac{1-\alpha_1-\alpha_2}{1-\alpha_2} - \frac{1-\alpha_1-\alpha_2}{1-\alpha_2}\beta_2(p) \\
&= (1-\alpha_2)pD(p)(1-c_1\beta_2(p)), \\
\pi_2(p) &= (1-\alpha_1)pD(p)(1-c_2\beta_1(p)),
\end{aligned}$$

where $c_1 = \frac{1-\alpha_1-\alpha_2}{1-\alpha_2}$ and $c_2 = \frac{1-\alpha_1-\alpha_2}{1-\alpha_1}$.

I continue to evaluate the strategy of the 'weaker' second seller (the seller with a smaller amount of uninformed buyers). The formulae relevant for the first seller apply analogously for a couple of the following steps.

The first seller has to be indifferent between prices in the mixing interval, therefore

$$\begin{aligned}
\pi_1(p_1) &= \pi_1(p_2) \\
p_1 D(p_1)(1-c_1\beta_2(p_1)) &= p_2 D(p_2)(1-c_1\beta_2(p_2)) \\
\frac{1-c_1\beta_2(p_1)}{1-c_1\beta_2(p_2)} &= \frac{p_2 D(p_2)}{p_1 D(p_1)}.
\end{aligned}$$

I label the lower bounds for the first and second player a_1 and a_2 and the upper bounds b_1 and b_2. Note that I do not assume there is not a mass point in b_i. In fact, I will show that seller 1 chooses price b_1 with a positive probability. I evaluate the previous formula at $p_1 = b_2, p_2 = a_2$ and find $(1-c_1)b_2 D(b_2) = a_2 D(a_2)$ or

$$\frac{\alpha_1}{1-\alpha_2} b_2 D(b_2) = a_2 D(a_2).$$

If the second seller sets the price $p = a_2$[25] and wins the informed buyers, she gets a profit of at least $\alpha_2 M$, where M denotes the monopoly profit with respect to the demand function $D(\cdot)$. Otherwise she would prefer to sell to her fraction of the market α_2 at the monopoly price p_M. This argument can be expressed as the equality $(1-\alpha_1)a_2 D(a_2) \geq \alpha_2 M$. Rearranging terms, I get $\frac{(1-\alpha_1)\alpha_1}{(1-\alpha_2)\alpha_2} b_2 D(b_2) \geq M$. Using an analogical procedure, I get $\frac{(1-\alpha_2)\alpha_2}{(1-\alpha_1)\alpha_1} b_1 D(b_1) \geq M$. Note that $b_1 = b_2$. If not, there would be a gap between b_1

[25] The seller chooses this price, in fact, with zero probability; therefore, it would be an invalid argument. I use it to simplify the full argumentation of setting price $p = a_2 + \epsilon$ and evaluating the limit $\epsilon \to 0$.

and b_2, and subsequently, it would be beneficial to switch some mass of the distribution function of the strategy from the lower bound of the gap to the upper bound of the gap for one player. Thus, in equilibrium, $b_1 = b_2 = b$. The previous two equations imply that $bD(b) = M$. Then, I can express the lower bounds as

$$a_2 D(a_2) = \frac{\alpha_1}{1-\alpha_2} M,$$

and

$$a_1 D(a_1) = \frac{\alpha_2}{1-\alpha_1} M.$$

However, I know that $a_1 = a_2$ otherwise the strategy of one seller would be irrational. Namely, as $\alpha_1 > \alpha_2$ and thus $a_2 > a_1$, the first player would prefer to move the mass of the function $\beta_1(\cdot)$ from the interval (a_1, a_2) to the point $a_2 - \epsilon$ and increase the expected profit. In fact, the first seller would not be willing to select a price lower than a_2 – even if the probability of winning is 1, she would not get more than $\alpha_1 M$ in expected value. Therefore, I let the second player's strategy remain the same as derived above and derive the first player's strategy that satisfies the equality $a_1 = a_2$ and the requirement that the second player is indifferent between choosing any price from the interval $[a_2, b_2]$. As before,

$$\begin{aligned}
\pi_2(p_1) &= \pi_2(p_2) \\
p_1 D(p_1)(1 - c_2\beta_1(p_1)) &= p_2 D(p_2)(1 - c_2\beta_1(p_2)) \\
\frac{1 - c_2\beta_1(p_1)}{1 - c_2\beta_1(p_2)} &= \frac{p_2 D(p_2)}{p_1 D(p_1)}.
\end{aligned}$$

As defined earlier, $\beta_1(a_2) = 0$. Setting $p_2 = a_2$ and rearranging, I get

$$\beta_1(p) = \frac{1}{c_2}\left(1 - \frac{a_2 D(a_2)}{pD(p)}\right). \tag{2.6}$$

Because the following equalities hold,[26] $\beta_1(p_M-) = \frac{1}{c_2}\left(1 - \frac{a_2 D(a_2)}{p_M D(p_M)}\right) < 1$ and $\beta_1(p_M) = 1$, the stronger player sets the monopoly price with a positive probability. This function leaves the second player indifferent between prices within the interval $[a_2, p_M)$, and the above derived strategy leaves the first player indifferent between prices within the interval $[a_2, p_M]$.

[26] The notation $\beta_1(p_M-)$ assigns the left limit of the function β_1 in the point p_M.

Therefore, the system

$$\begin{aligned}
\beta_2(p) &= \frac{1}{c_1}\left(1-\frac{aD(a)}{pD(p)}\right), \quad p\in[a,p_M] \\
\beta_1(p) &= \frac{1}{c_2}\left(1-\frac{aD(a)}{pD(p)}\right), \quad p\in[a,p_M) \\
\beta_1(p_M) &= 1, \\
aD(a) &= \frac{\alpha_1}{1-\alpha_2}
\end{aligned}$$

constitutes the equilibrium.

Uniqueness

I have derived a formula for a mixing function that leaves the opponent indifferent between choosing prices from the interval $[a_2, p_M)$ in the previous part of the proof. This formula defines a unique solution if:

1. the lower and upper bounds of the support are unique and

2. the functions are strictly increasing in the support.[27]

The uniqueness of the upper bound has been shown before. It needs to be shown that the functions have to be strictly increasing in the support (or that there are no gaps in the support) and that the lower bound is unique.

There are no gaps in the distribution functions β_i

If there is a gap in a distribution function there must be the same in the second one; otherwise, the second seller would not behave rationally. As in the proof of Proposition 2, it is possible to show that shifting some mass from the lower bound of the gap to the upper bound, the seller would benefit; therefore, there can be no gap in the distribution function.

There are no mass points except the one in p_M for β_1

The situation of a symmetric mass point, i.e. the situation where both β_1 and β_2 have a mass point for the same price, would initiate an undercutting procedure. Therefore, there can be no symmetric mass point. If there is a mass point in $x \in (a, p_M)$ in the function β_i and $\beta_{-i} < 1$, then it it easy to show that shifting the mass of the probability

[27]Theoretically, if my opponent never chooses prices from a specific interval I do not need to 'make him indifferent' in that interval and I would not need to follow the prescribed formula. In that case there would exist infinitely many strategies.

function β_{-i} from the interval $(x, x + \epsilon_1)$ to the point $x - \epsilon_2$ for some $\epsilon_1, \epsilon_2 > 0$ would increase the expected profit of the seller $-i$. Therefore, there can be no mass point in any of the two distribution functions with the mentioned exception.

The lower bound a is unique

As the functions β_i must be increasing in (a, p_M) and there are no mass points, the functions can be represented by the formulae derived above for a given a. If I assume that $\hat{a} < \frac{\alpha_1}{1-\alpha_2}$, then the stronger seller has no incentive to set a positive probability to the prices in the interval $(\hat{a}, \frac{\alpha_1}{1-\alpha_2})$, which violates the no gaps property. If, in contrast, $\hat{a} > \frac{\alpha_1}{1-\alpha_2}$, the values of mixing functions at monopoly prices are $\beta_i(p_M) < 1$ for both functions, which violates the no mass points property. Therefore, the lower bound $a = \frac{\alpha_1}{1-\alpha_2}$ is unique. •

I have found an interesting result. In an asymmetric duopoly, the weaker seller (the one with the smaller share of uninformed buyers) gains from the position of the stronger seller. If the stronger seller's market share increases, the weaker seller's expected profit rises as fast as the stronger seller's even though the weaker seller does not invest in capturing a larger share of the market. Also, it is much more probable that the weaker seller sets a lower price than the stronger one and wins the searchers. The wider the gap between the weaker and the stronger seller, i.e. the bigger difference between the market shares of the two sellers, the higher the probability that the stronger seller sets the monopoly price and sells to 'her' market share.

The above-mentioned facts have consequences for buyers. If we start from a symmetric position and let one seller capture a larger market share, then the expected profit of the sellers increases as if both players' market shares increase in the same way. The increase of the sellers' profit is not paid by the buyers who caused the market share to increase but in the majority by previously uninformed buyers and partially by informed buyers. When, in contrast, I start from an asymmetric position and the gap between the sellers' market shares moves towards the symmetric position so that more buyers become uninformed, the expected profit of the sellers does not change. The buyers who switch from informed to uninformed necessarily lose (recall I do not have search cost in the model), therefore informed buyers benefit from the increased market share of the weaker seller.

The last step is to develop the most general model. It is an asymmetric model of competition among N sellers when every individual seller has secured her own market share α_i, or the probability that a specific uninformed buyer visits her. This model, naturally, covers all the already-analyzed specific cases. It will help us understand the logic and nature of the equilibrium.

Proposition 5

Assume that the Common Assumptions hold, and there are N sellers. Assume that the probability a buyer chooses to visit seller i is α_i.[28] The sellers are, without loss of generality, sorted according to ascending market share: $\alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_N$. The remaining probability $1 - \sum \alpha_i$ represents informed buyers.

Then there is a unique mixed-strategy equilibrium, where the price-mixing cdfs β_i can be expressed as[29]

$$\begin{aligned}
\beta_i(p) &= 0, \quad p \leq a_i, \\
&= 1 - \frac{\prod_{j, a_j < p} k_j^{\frac{1}{m-1}}}{k_i} \quad p \in [a_i, p_M), \\
&= 1, \quad p \geq p_M \\
k_i &= (1 + c_i((\prod_{j<i}(1 - \beta_j(a_i)) - 1)))\frac{a_i D(a_i)}{c_i p D(p)} - \frac{1 - c_i}{c_i} \\
c_i &= \frac{1 - \mathcal{A}}{1 - \mathcal{A}_{-i}}, \quad \mathcal{A} = \sum \alpha_i, \quad \mathcal{A}_{-i} = \mathcal{A} - \alpha_i, \\
m &= max_l(a_l < p),
\end{aligned}$$

where M is the monopoly profit and p_M the monopoly price. The lower bound of every distribution function a_i is chosen as the solution of the implicit equation valid $\forall i > 1$ and for $i = 1$ if $\alpha_1 = \alpha_2$

$$(\alpha_i + (1 - \mathcal{A})(\prod_{j<i}(1 - \beta_j(a_i))))a_i D(a_i) = \alpha_i M,$$

where M is the monopoly profit. If $\alpha_1 < \alpha_2$, then $a_1 = a_2$.

The expected profits of sellers are $\alpha_i M, \forall i > 1$, and $i = 1$ if $\alpha_1 = \alpha_2$, i.e. the profit the first seller would secure setting a monopoly price and selling to 'her' fraction of market. If $\alpha_1 < \alpha_2$, then the first seller's expected profit is $\alpha_2 M - (\alpha_2 - \alpha_1)a_2 D(a_2)$, i.e. the profit of the second-weakest seller minus the difference of the profit from the uninformed buyers at the lower bound of the mixing interval.

[28]This means that each seller has "in expectation" secured a fraction α_i of the market.

[29]For the weakest seller 1, the function is continuous; for other sellers, there can be a discontinuity in p_M, i.e. they can choose the price p_M with a positive probability.

Proof

The proof follows the standard line used in the previous proofs. First, the profit function of each seller is defined. I use the fact that each seller has to have identical expected profits for all prices she is willing to set to obtain a system of equations that defines the mixed strategy cumulative distribution functions of all players.

Assume that $\alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_N$. Label the respective mixing functions for sellers $i = 1, \ldots, N$ $\beta_i(p)$ and the supports of the mixing functions $(a_i, b_i]$.[30] To keep the notation simple, I use the fact that $a_1 \leq a_2 \leq \cdots \leq a_N$, and will be proven later, and no step in the proof depends on this fact. I denote $\mathcal{A} = \sum_{j=1}^{N} \alpha_j$ and $\mathcal{A}_{-i} = \mathcal{A} - \alpha_i$.

The expected profit of seller i is her expected profit from uninformed buyers who come directly to her and the expected profit from the searchers if she selects the lowest price on the market.

$$\begin{aligned}\pi_i(p) &= \alpha_i p D(p) + (1-\mathcal{A}) p D(p) \prod_{j\neq i}(1-\beta_j(p)) \\ &= (1-\mathcal{A}_{-i}) p D(p)(1 + c_i(\prod_{j\neq i}(1-\beta_j(p)) - 1)),\end{aligned}$$

where $c_i = \frac{1-\mathcal{A}}{1-\mathcal{A}_{-i}}$.

I compare the seller's expected profits for two different price choices. We know that they have to be equal if we want her to randomize between them (or other prices with an identical outcome): $\pi_i(p_1) = \pi_i(p_2), \forall p_1, p_2 \in (a_i, b_i]$. This equation gives us the relationship

$$\frac{p_1 D(p_1)}{p_2 D(p_2)} = \frac{1 + c_i(\prod_{j\neq i}(1-\beta_j(p_2)) - 1)}{1 + c_i(\prod_{j\neq i}(1-\beta_j(p_1)) - 1)}.$$

I use a_i as the reference value,[31] and the fact that $\beta_j(a_i) = 0\, \forall j > i$ and $\beta_j(p) = 0\, \forall j, a_j \geq p$ to get the equality

$$\frac{a_i D(a_i)}{p D(p)} = \frac{1 + c_i(\prod_{j, a_j < p \,\&\, j\neq i}(1-\beta_j(p)) - 1)}{1 + c_i(\prod_{j<i}(1-\beta_j(a_i)) - 1)}.$$

Rearranging the terms, I get one equation from the system defining the mixing func-

[30] Later I will show that the weakest seller chooses the upper bound with zero probability and others may choose it with a positive probability.

[31] Though a_i lies outside the interval (a_i, b_i), I can use the limit $p_1 \to a_i+$ to verify the correctness of this equation. The proof that a_i cannot be chosen with a positive probability is straightforward.

tion

$$\prod_{j,a_j<p\&j\neq i}(1-\beta_j(p)) = (1+c_i(\prod_{j<i}(1-\beta_j(a_i))-1))\frac{a_iD(a_i)}{c_ipD(p)} - \frac{1-c_i}{c_i} =: k_i(p).$$

If I use a logarithm for every equation from the system,

$$\prod_{j,a_j<p\&j\neq i}(1-\beta_j(p)) = k_i(p), i=1\ldots m, m = max_l(a_l<p), \tag{2.7}$$

I get a system of linear equations:

$$\sum_{j,a_j<p\&j\neq i} log(1-\beta_j(p)) = log(k_i(p)), i=1\ldots m, m = max_l(a_l<p).$$

I think of the expressions $x_i(p) = log(1-\beta_i(p))$ as unknowns. Then assigning $y_i(p) = log(k_i(p))$, vectors $X(p) = x_i(p), \quad Y(p) = y_i(p)$ the system of equations can be rewritten as AX=Y, where A is a matrix of ones with zeroes on the diagonal. The solution of the system is $x_i = \sum_{j,a_j<p} y_j/(m-1) - y_i$. Solving the substitution I get the formula for $\beta_i(p)$:

$$\beta_i(p) = 1 - \frac{\prod_{j,a_j<p} k_j^{\frac{1}{m-1}}}{k_i}.$$

I need to determine the values of a_i to finish the analysis. Observe that the distribution function β_i of at least one player cannot have a mass point in the monopoly price. If every distribution function had one, the undercutting process would start. Also, $\beta_i(p_M-\epsilon) < 1, \forall i, \epsilon > 0$. If some player sets $\beta_i(p_M-\epsilon) = 1$, a gap in the distribution function of other players would emerge between $p_M-\epsilon$ and p_M and thus, the strategy of player i would be sub-optimal.

Assume that the mixing function of player i satisfies $\lim_{p\to p_M}\beta_i(p) = 1$.[32] Then, using equation (2.7), $\forall l \neq i, \lim_{p\to p_M} k_l(p) = 0$. The equation can be translated to

$$(\alpha_l + (1-\mathcal{A})(\prod_{j<l}(1-\beta_j(a_l))))a_lD(a_l) = \alpha_l p_M D(p_M). \tag{2.8}$$

This equation says that a_l is chosen so that the expected profit setting the price to a_l equals the expected profit when the price is set equal to the monopoly price and selling only to uninformed buyers. The assumed properties of the function $D(\cdot)$ (or $pD(p)$)

[32] I know there is at least one - the one whose distribution function does not have a mass point in p_M.

ensure the uniqueness of the solution. It means the equation has a unique solution for $a_l, l \neq i$. If seller l's mixing interval lower boundary would be higher, $\hat{a}_l > a_l$, then $\hat{\beta}_i(p) < \beta_i(p), i < l, p \in (a_l, \hat{a}_l)$.[33] This would mean that seller l would get a higher profit choosing $p \in (a_l, \hat{a}_l)$ than what she would get in equilibrium, $\alpha_l p_M D(p_M)$. Therefore choosing any lower bound other than a_l is not an equilibrium strategy.

The above derived formula also gives us the reason why $a_l, l \neq i$ are sorted increasingly $a_1 \leq a_2 \leq \cdots \leq a_N$. The proof by contradiction is apparent. I show later that even the value a_i itself is sorted in the chain.

I need to analyze the two following cases. In the first case, the two least powerful sellers have the same market share, or $\alpha_1 = \alpha_2$. In the second case, the least powerful seller has a strictly smaller market share than all other sellers, or $\alpha_1 < \alpha_2$. I continue the analysis with the more complicated second case.

From the previous analysis, I know that the lower bounds of the distribution functions a_i are specified by the above-mentioned formula except, possibly, for one. Assume that all a_i are given by formula (2.8), then inequality $a_1 < a_2 \leq \cdots \leq a_N$ holds. However in that case, seller 1 would choose a sub-optimal strategy, and she would prefer to shift the lower bound a_1 upwards. In other words, there must be at least two players choosing the smallest value of the lower bound of their mixing probability function support. This implies that if a_1 is derived from equality (2.8), there must be a_i such that $a_i = a_1$. Seller i would, however, make a mistake doing this because she would lower her expected profit and prefer to choose the price p_M. Therefore, player 1 must be the one who does not set the lower bound a_1 according to formula (2.8). The value of $k_1(p_M) > 0$ and due to equality (2.7), all other players choose the monopoly price with a positive probability.

I still need to find the value of a_1. The selection of $a_1 = a_2$ is an equilibrium choice. I need to show that it is a unique choice that leads to an equilibrium. If seller 1 chooses a smaller value of the lower bound $a_1 < a_2$, she would not gain the maximum expected profit. If seller 1 chooses a lower bound bigger than a_2, other players will adapt to her choice and select strategies such that player 1 gets a higher profit than choosing $a_1 = a_2$. There is no reason, however, why seller 2 would not mimic player 1's strategy and obtain a higher expected profit. That would not be an equilibrium, however, because it has been proven that seller 2 chooses the lower bound a_2 with respect to equation (2.8). Therefore, the choice $a_1 = a_2$ is the unique equilibrium choice. Also, this is the final fragment from

[33]This follows directly the formula for $\beta_i(p)$ and the fact that $k_i(p) < 1, p > a_i$.

the proof of the inequality chain $a_1 \leq a_2 \leq \cdots \leq a_N$.[34]

The first case, when there are at least two sellers with the same, smallest market share is easier to analyze. There is an obvious equilibrium when every seller chooses the lower bound according to formula (2.8), and all mixing functions are continuous everywhere. I need to review the situation when one player chooses a different lower bound. When one player chooses a different lower bound a_i than the one derived from equation (2.8), she must get a higher expected profit (she does not want to get a lower expected profit). There is always a player who would be willing to mimic player i's strategy and grab the same profit. Then we would have two players not following equation (2.8), which is ruled out in an equilibrium. •

We see that the equilibrium of the most complex model corresponds to, and includes, all three specific models presented before. The model describes symmetric competition and asymmetric competition of two or more sellers. The main feature, the uniqueness of the solution, remains preserved even in the most complicated set-up. The equilibrium uniqueness implies that even in a finitely repeated competition model, the equilibrium remains unique and fully defined by the proposed general-model solution. The general model applies even to finitely repeated competition with variable market shares if current actions do not change future market share.

Compared to welfare implications derived from the three previously analyzed models, there are no qualitatively new results in this general model. I can, however, summarize and generalize those presented earlier.

I have distinguished three groups of agents in the analysis: sellers, informed buyers, and uninformed buyers. The expected profit of sellers is always equal to the profit they would gain from selling to their market share at the monopoly price with one exception - the seller with the smallest market share (the "weakest" seller), who obtains almost the same expected profit as the second-weakest seller (with the difference of the profit from uninformed buyers at the lower bound). Every seller has an expected profit proportional to her own market share; only the uniquely weakest seller's expected profit is driven, apart from luck, by the expected profit of the second-weakest seller. This exception has more impact as the gap widens between the uniquely weakest seller and the second-weakest seller.

The buyers are assumed to be informed or uninformed, and uninformed buyers rely

[34] In fact, the first inequality holds with equality.

on the sellers to which they go. Their individual expected profit is higher if their seller is weaker (again, with the exception of the weakest seller). Informed buyers benefit from market fragmentation the most. For them, it is beneficial to face many weak sellers, and strong sellers do not have a significant impact on their profit because they do not push the prices down. In the exceptional case, with the existence of one (uniquely) weakest seller, the shift of the lower bound of the weakest seller's mixing interval subsequently increasing the seller's welfare is absorbed by the corresponding uninformed buyers (and there are few of them as the seller is the weakest one) and, more importantly, by the informed buyers. Informed buyers are, therefore, heavily dependent on the market structure, especially on the market shares of the weakest sellers.

2.4 Conclusion

In this paper, I have developed an oligopolistic model of sellers who compete on price for informed buyers, and have secured their share of uninformed buyers. Informed buyers know which seller sets the lowest price, and uninformed, or "loyal", buyers do not search for the lowest price but go to one seller only without knowing ex-ante his posted price. The basic, and for the result crucial, difference from previously published models is the reservation values of buyers. In the model presented in this paper, the reservation values are randomly drawn from a common distribution function rather than a pre-defined single value that has been used previously in the existing literature.

I have found that there is a unique mixed-strategy, Nash equilibrium of this model. In the equilibrium, every seller chooses randomly a price between an endogenously determined lower bound and the monopoly price associated with the demand function generated by the reservation value distribution function. Compared to the theoretical and empirical results of Kocas and Kyiak (2006), the model presented in this paper, namely the equilibrium strategy of sellers, explains the empirical findings of Kocas and Kyiak better than their own theoretical model. The main difference is that in my models all sellers, not just the two weakest, choose their prices randomly from a non-trivial interval.

There are several welfare implications in the result. As derived in the text, the expected profit, or the seller's surplus, is equal to the profit that every seller would get setting the monopoly price and selling to her share of the market (to the uninformed buyers who visit this seller) plus the difference between the 'market share value' of the two weakest sellers.[35] As for buyers, the non-searchers' surplus increases with increased competition between sellers and a constant amount of searchers. Searchers benefit from increased competition in every situation with one exception.

The interesting exception from the rule "bigger competition = bigger buyer surplus" is the entry of a new seller who becomes the uniquely weakest one on the market. Imagine there are several sellers, and all with a positive market share. If a seller with no market share enters, she gets the expected profit close to the profit of the second weakest seller and thus steals a share of the total surplus. As the expected profit of the incumbent sellers does not change, their share on the total surplus remains constant. Therefore, the entrant steals the surplus fully from buyers.

The model applies to many areas. It naturally follows and generalizes the literature originated in Varian's article and also the vast literature on the Hotelling lemma: the size and distribution of the network of stores influences the size of the seller's market share. In fact, the present paper integrates the two concepts into one model: oligopolistic competition is introduced into Varian's model. Another straightforward application, apart from the analysis of concentration/location of stores, would be a study of advertisement from a competitive (not demand enhancing) point of view.

As a theoretical extension of the model, I would like to study further the introduction of production costs, the introduction of costs to increase market share (the model would approach the literature on promotions), and the introduction of search costs (the model would approach the literature on search mentioned in the introduction). It would also be interesting to introduce bounded rationality or incomplete information into the model.[36] Sellers, for example, might not know the competitors' market shares and may just have estimates. It might be interesting to develop some kind of learning model that would fit reality better than the existing models do, the present one included.

[35] A seller becomes weaker as a smaller amount of uninformed buyers come to her. The lower bound of the mixing strategy support of the two weakest sellers is identical – the same result as in the findings of Kocas and Kyiak.

[36] This extension would target the field of experimental economics; see, for example, Huck, Muller, and Vriend (2002).

2.5 References

Baye, Michael R., and Casper G. De Vries, 1992. Mixed Strategy Trade Equilibria. The Canadian Journal of Economics 25(2), 281-93.

Braverman, Avishay, 1980. Consumer Search and Alternative Market Equilibria. Review of Economic Studies 47, 487-502.

Burdett, Kenneth, and Kenneth L. Judd, 1983. Equilibrium Price dispersion. Econometrica 51, 955-70.

Clay, Karen, Krishnan Ramayya, Eric Wolff, and Danny Fernandes, 2002. Retail Strategies on the Web: Evidence from the Online Book Industry. The Journal of Industrial Economics 50(3), 351-367.

Clemons, Eric K., Il-Horn Hann, and Lorin M. Hitt, 2002. Price Dispersion and Differentiation in Online Travel: An Empirical Investigation. Management Science 48(4), 534-549.

Huck, Steffen, Wieland Muller, and Nicolaas J. Vriend, 2002. The East End, the West End, and King's Cross: On Clustering in the Four-Player Hotelling Game. Economic Inquiry 40(2), 231-240.

Iyer, Ganesh, and Amit Pazgal, 2003. Internet Shopping Agents: Virtual Co-Location and Competition. Marketing Science 22(1), 85-106.

Kocas, Cenk, and Tunga Kiyak, 2006. Theory and Evidence on Pricing by Asymmetric Oligopolies. International Journal of Industrial Organization 24, 83-105.

Narasimhan, Chakravarthi, 1988. Competitive Promotional Strategies. The Journal of Business 61(4), 427-449.

Neven, D., G.Norman, and J.F.Thisse, 1991. Attitudes Towards Foreign Products and International Price Competition. The Canadian Journal of Economics 24(1), 1-11.

Raju, J.S., V.Srinivasan, and R.Lal, 1990. The Effects of Brand Loyalty on Competitive Price Promotional Strategies. Management Science 276-304.

Reinganum, Jennifer F., 1979. A Simple Model of Equilibrium Price Dispersion. Journal of Political Economy 87, 851-58.

Salop, Steven C., and Joseph E. Stiglitz, 1979. Bargains and Ripoffs: A Model of Monopolistically Competitive Price Dispersion. Review of Economic Studies 44, 493-510.

Sinitsyn, Maxim, 2008. Characterization of the Support of the Mixed Strategy Price Equilibria in Oligopolies with Heterogeneous Consumers. Economics Letters 99(2), 242-245.

Stahl, Dale O., 1989. Oligopolistic Pricing with Sequential Consumer Search. The American Economic Review 79(4), 700-12.

Stiglitz, Joseph E., 1987. Competition and the Number of Firms in the Market. Journal of Political Economy 95, 1041-61.

Stiglitz, Joseph E., 1979. Equilibrium in Product Markets with Imperfect Information. The American Economic Review 69, 339-45.

Varian, Hal R., 1980. A Model of Sales. The American Economic Review 70(4), 651-59.

Wilde, Louis L., and Allan Schwartz, 1979. Equilibrium Comparison Shopping. Review of Economic Studies 46, 543-53.

Chapter 3

Three Very Simple Games and What It Takes to Solve Them

Note: This chapter has been published as a co-authored paper in an external journal. For this reason, this chapter of the dissertation appears exactly as it did in that journal, and has not been edited or modified for correctness or consistency with the rest of the dissertation.

Abstract

We study experimentally the nature of dominance violations in three minimalist dominance-solvable guessing games. Only about a third of our subjects report reasoning consistent with dominance; they all make dominant choices and almost all expect others to do so. Nearly two-thirds of subjects report reasoning inconsistent with dominance, yet a quarter of them actually make dominant choices and half of those expect others to do so. Reasoning errors are more likely for subjects with lower working memory, intrinsic motivation and premeditation attitude. Dominance-incompatible reasoning arises mainly from subjects misrepresenting the strategic nature (payoff structure) of the guessing games.

Keywords: cognition, bounded rationality, beliefs, guessing games, experiment

JEL classification: C72, C92, D83

Joined paper with Ondřej Rydval and Andreas Ortmann. This research project was supported by a grant from Bank Austria and by CAPER grant No. LC542 of the Ministry of Education of the Czech Republic implemented at CERGE-EI (a joint workplace of the Center for Economic Research and Graduate Education, Charles University, and the Economics Institute of the Academy of Sciences of the Czech Republic). We thank Vince Crawford, Jan Kmenta, Laura Mentz and two anonymous referees for helpful comments; Vitezslav Babicky, Marian Krajc and Jana Richmanova for helping us conduct the experiment; Randall Engle and Richard Heitz of the Attention and Working Memory Lab, Georgia Institute of Technology for providing us with their "automated working memory span" tests; and Brian MacWhinney of the Department of Psychology, Carnegie Mellon University for providing us with his "audio digit span" test. All remaining errors are our own.

3.1 Introduction

Experimental studies extensively document deviations of initial responses from equilibrium predictions in iterated-dominance-solvable games, including matrix and other normal-form games, extensive-form bargaining games and guessing games (see Costa-Gomes, Crawford and Broseta, 2001, and Costa-Gomes and Crawford, 2006, for overviews). Non-equilibrium behavior is typically attributed to subjects' non-equilibrium beliefs about others' irrationality rather than their own irrationality. Especially Costa-Gomes and Crawford (2006, hereafter CGC), through joint analysis of initial responses and information search patterns in iterated-dominance-solvable guessing games, convincingly show that many subjects' deviations from equilibrium "can be confidently attributed to non-equilibrium beliefs rather than irrationality, risk aversion, altruism, spite, or confusion." (p. 1740) CGC conclude that the findings "affirm subjects' rationality and ability to comprehend complex games and reason about others' responses to them..." (p. 1767).

However, evidence on individual rationality from simpler, dominance-solvable games seems much less conclusive. In Grosskopf and Nagel's (2008) two-player dominance-solvable guessing game, 90% of subjects violate simple dominance. Moreover, Devetag and Warglien (2007) show that nearly a quarter of their subjects cannot even correctly represent the relational structure of preferences in a two-player game similar to dominance-solvable guessing games. On the other hand, in Bone et al.'s (2008) extensive-form game against nature, only 5% of subjects violate simple dominance, which suggests that most people are in principle capable of applying dominance when it is transparent.[1]

Similar to Charness and Levin (2008), who simplify common value auctions to study the origin of the winner's curse, we examine the nature of dominance violations in three "minimalist" dominance-solvable guessing games featuring two or three players choosing among two or three strategies. Also called beauty contest games, guessing games are *ex ante* well-suited for studying individual rationality bounds without the potentially confounding effects of other-regarding and risk preferences. Guessing games of the dominance-solvable nature have the additional appeal of making a player's optimal choice independent of her beliefs about others' choices (and hence others' rationality).

To better understand the decision-making errors our subjects commit, we ask them to report (alongside making choices) their detailed reasoning underlying the choices, and also to state their beliefs about their anonymous partners' choices (i.e., about others' rationality). We then study

[1]We refer to the second decision node of Bone et al.'s game where almost all of 152 subjects choose a four-payoff distribution that first-order stochastically dominates another such distribution, which is far above random choice (50%). However, the whole game is a two-stage game, and only about a third of the subjects detect dominance at the first (prior) decision node (using, for example, backward induction or the strategy method).

how the reported reasoning – classified according to dominance-compatibility by two independent examiners – translates into the subjects' stated choices and beliefs. Following the lead from psychology (e.g., Simon, 1978 and 1989; Stanovich and West, 2000) and recently experimental economics (e.g., Ballinger et al., 2008; Rydval, 2007), we also examine how subjects' reasoning classes and choices relate to their measured cognitive abilities and personality traits (see Sections 4 and 6.2).

Only about a third of our subjects reason in line with dominance; they all make dominant choices and almost all expect others to do so. By contrast, nearly two-thirds of subjects report reasoning processes incompatible with dominance, yet a quarter of them actually make dominant choices and half of those expect others to do so. Reasoning errors are more likely for subjects with lower ability to maintain and allocate attention (measured by a working memory test), and for subjects with lower intrinsic motivation to engage in cognitively demanding tasks (measured by the need-for-cognition personality scale) and lower propensity to deliberate while carrying out tasks (measured by the premeditation personality scale). In section 6.3, we further explore origins of dominance violations and find that dominance-incompatible reasoning arises mainly from subjects misrepresenting (to themselves) the strategic nature (payoff structure) of the guessing games.

3.2 The guessing games

We study behavior in three symmetric dominance-solvable guessing games depicted in normal-form representation in Figure 1. A pair or a triplet of players simultaneously choose (or guess) among two (0, 1) or three (0, 1, 2) numbers. A fixed monetary prize, M, is won by the player whose choice is closest to one-half of the pair's or triplet's average choice; multiple winners divide the prize equally. Under complete information – an assumption justified by publicly announcing the games' structure – our games have a unique equilibrium in which all players choose 0. Games 2p2n and 3p2n are strict-dominance-solvable, i.e., choosing 0 yields a strictly higher payoff compared to choosing 1, for any choice(s) of the other player(s). Game 2p3n is weak-dominance-solvable, i.e., choosing 0 yields a higher or equal payoff compared to choosing 1 or 2, for any choice of the other player.

Figure 1: The guessing games in normal-form representation

Game 2p2n: 2 players, 2 numbers

		Player 2	
		0	1
Player 1	0	M/2,M/2	M,0
	1	0,M	M/2,M/2

Game 2p3n: 2 players, 3 numbers

		Player 2		
		0	1	2
Player 1	0	M/2,M/2	M,0	M,0
	1	0,M	M/2,M/2	M,0
	2	0,M	0,M	M/2,M/2

Game 3p2n: 3 players, 2 numbers

Player 3's choice = 0

		Player 2	
		0	1
Player 1	0	M/3,M/3,M/3	M/2,0,M/2
	1	0,M/2,M/2	0,0,M

Player 3's choice = 1

		Player 2	
		0	1
Player 1	0	M/2,M/2,0	M,0,0
	1	0,M,0	0,0,M/3,M/3,M/3

Previous studies predominantly focus on iterated-dominance-solvable guessing games, with two or more players facing various (a)symmetric (across players) strategy spaces larger than ours, and with the "winning guess" determined by various sample statistics of players' guesses (e.g., mean, median or maximum) multiplied by various (a)symmetric (across players) target numbers smaller or greater than one. These features jointly determine how one's own guess influences the winning guess, and the number of rounds of iterated elimination of dominated guesses necessary to identify one's iteratively undominated guess(es).[2] Iterated-dominance-solvable guessing games require a simultaneous assessment of players' individual rationality and their beliefs about others' rationality. In dominance-solvable guessing games, by contrast, equilibrium predictions and players' best responses rely only on individual rationality in the game-theoretic sense of obeying simple dominance. This allows us to focus on the basic limits of cognition as revealed by dominance violations.

Behavior in a two-player weak-dominance-solvable guessing game is studied in Grosskopf and Nagel (2008 and 2007, hereafter GNa and GNb, respectively, or GN).[3] A fixed monetary prize is won by the player(s) whose guess is closest to two-thirds of the pair's average guess. Guesses can range from 0 to 100 (inclusive), so the strategy space is much larger than in our games. Although guessing 0 is a weakly dominant strategy, 90% of subjects (132 undergraduates with no formal training in game theory) initially make dominated guesses above 0, which is close to random guesses (99%). The strikingly frequent dominance violations appear robust to increasing stakes or implementing more detailed explanation of the guessing game,[4] and they only partly vanish with more expertise or on-task experience.[5]

GN further offer a comparison of behavior in the two-player weak-dominance-solvable game with behavior in an otherwise identical eighteen-player iterated-dominance-solvable game, played by another 36 subjects from the same population. In the eighteen-player game, only about 10% of subjects initially violate dominance, far below random guesses (33%) and typical for iterated-dominance-solvable games including guessing games.[6] Surprisingly, initial guesses in the two- and eighteen-player games are similar and even marginally higher in the former game with the unique undominated guess of 0. GNb also observe what appears as a lack of knowledge transfer between the two games: In a treatment where subjects switch after four rounds from the eighteen-player to the two-player game, most of them make a higher guess in the fifth round compared to the fourth round.[7]

[2]See CGC and Grosskopf and Nagel (2008) for an overview of iterated-dominance-solvable guessing games. Two-player guessing games are normally dominance-solvable but CGC introduce a new class that are iterated-dominance-solvable.

[3]We occasionally use GN to refer to both GNa and GNb which use the same experimental dataset, with GNa analyzing first-round behavior and GNb behavior over time.

[4]This result is based on correspondence with Brit Grosskopf.

[5]Experts (economic researchers at conferences) do better than students but their dominance compliance is still only 37%. GNb further observe that dominance violations persist even after ten rounds of playing the game in fixed pairs: depending on the extent of feedback provided during the game, one- to three-quarters of student subjects guess above 0 in the tenth round.

[6]About 90% of subjects guess below *100, thus seemingly respecting at least one round of iterated dominance. This is typical for initial responses in iterated-dominance-solvable games (see, e.g., CGC's Table 6).

[7]These (and other) comparisons should of course be viewed in light of the different sets of dominated and undominated strategies across the games. Moreover, since subjects played GN's two-player game repeatedly in fixed pairings and knew about it a priori, they might have viewed their first-round guesses

In the spirit of Charness and Levin's (2008) exploration of the origins of the winner's curse, we implement minimalist dominance-solvable guessing games to look closer at the potential sources of dominance violations. Compared to GN, our two-player games 2p2n and 2p3n constrain the strategy space to only two and three numbers, respectively. In principle, especially game 2p2n permits mentally or visually listing all contingencies – i.e., all combinations of both players' possible choices and their payoff consequences – so one can "gradually recognize" the dominance of choosing 0 even without being a priori aware of the notion of dominance. As illustrated in Figure 1, listing contingencies may be (cognitively) hardest in our three-player game 3p2n. In a game-theoretic sense, however, game 3p2n is not harder than game 2p3n since the former is strict-dominance-solvable while the latter is weak-dominance-solvable.

Although our guessing games are simple, they are neither trivial nor necessarily easier to represent mentally than guessing games with larger strategy spaces. The latter games may facilitate (or necessitate) subjects translating them into stories that are easier to represent mentally, although the high rate of dominance violation in GN's two-player game suggests otherwise. Given GN's findings, we do not expect everyone to solve our games, especially not the arguably more complex games 2p3n and 3p2n. The variation in cognitive and game-theoretic complexity among our games is meant to aid our understanding of the sources of dominance violation. Particularly, the three games may differ in subjects' reasoning processes and reasoning errors.

Since our focus is on cognition rather than learning, we collect only initial choices in a between-subjects design. Each of our subjects makes a single choice for one of the games depicted in Figure 1, which suppresses any form of learning (including introspective one), repeated-game effects, and experimentation.[8] Nevertheless, the single choice results from a relatively lengthy deliberation process undertaken by subjects when reporting their reasoning. While having only one choice per subject could still undermine the reliability of across-game comparisons of choice behavior, our primary focus is rather on the relationship between subjects' reasoning processes, choices, beliefs, and cognitive and personality characteristics, as detailed below.

---

as influential for subsequent game play; hence the first-round guesses might not represent true initial responses free of repeated-game effects and experimentation. Also, most of GN's subjects obtained some degree of outcome and payoff feedback which is uncommon in studies of initial responses, though GN document that the distributions of first-round guesses do not differ across their feedback treatments.

[8]In our opinion, our games are too similar to each other to warrant their implementation in a within-subjects design, especially given our parallel elicitation of reasoning processes, choices and beliefs. Even with new partners for each game and no feedback, or even with the games embedded within a set of other games, we would risk considerable introspective learning, making it difficult to disentangle learning from cognition. See CGC for a detailed argument for studying truly initial responses.

3.3 Reasoning classes, decision-making errors and stated beliefs

In an answer protocol appended to the experiment's instructions, we prompted subjects to report their complete reasoning leading them to their choice, and then to state their choice and beliefs about the choice(s) of the other player(s) in their pair or triplet (see Appendix 1).[9] Subjects were told to report their reasoning in as much detail as possible in order to get paid.

Two examiners from outside the research team – CERGE-EI third-year Ph.D. students with advanced training in game theory – independently classified subjects' reasoning processes based on inspecting copies of the answer protocols, without observing the stated choices and beliefs which we deleted from the protocols. This was to ensure that the examiners focus on classifying subjects' reasoning processes rather than inferring the classification from the stated choices and beliefs, with the ultimate aim of detecting any differences between reasoning processes and choices.[10]

We gave the examiners classification instructions (see Appendix 2 for details) asking them to assign each subject's reasoning process into one of the following three reasoning classes:

Reasoning class A

Wrong reasoning – e.g., due to misrepresenting the strategic nature of the guessing game or making a numerical mistake, or irrelevant belief-based reasoning.

Reasoning class B

Reasoning based on listing contingencies involving own dominant choice of 0, but without explicitly explaining why 0 is the dominant choice.

Reasoning class C

Reasoning explicitly recognizing and explaining why 0 is the dominant choice, with or without listing contingencies.

Class A includes a variety of wrong reasoning processes discussed in detail in Section 6.3. Class A, for example, includes irrelevant belief-based reasoning such as "I believe the other player chooses 1, so I will choose 1 and we will split the prize."[11] By contrast, belief-based

[9] Subjects in game 3p2n were reminded that they could state different beliefs about the choices of the other two players in their triplet, but none of them actually did so.

[10] The examiners were of course not completely blind with respect to choices and beliefs since subjects often indirectly stated them as part of their reasoning. However, as will become clear below, such indirect statements could still be part of various (correct or incorrect) reasoning processes and had to be carefully interpreted by the examiners in the context of a particular reported reasoning process.

[11] We argue in Section 6.3 that irrelevant belief-based reasoning might in fact stem from misrepresenting the strategic nature of the guessing games. There is no indication that irrelevant belief-based reasoning could be induced by our belief elicitation procedure or result from fairness considerations.

explanations of dominance are included in reasoning class C – e.g., "I believe the other player chooses 0 because that's the best for her, so I will choose 0 not to lose the game," or "I expect the other player to choose between 0 and 1 randomly or with some probabilities, but no matter what she chooses, my best choice is 0."

For class B, listing contingencies means listing the combinations of the pair's or triplet's possible choices and their consequences, in any plausible mathematical, verbal or graphical form. However, since it would have been impossible for the examiners to distinguish between intentional and unintentional omission of (irrelevant) contingencies involving own dominated choices, class B requires listing only the contingencies involving own dominant choice of 0:

Game 2p2n: contingencies involving choice pairs (0, 0) and (0, 1)

Game 2p3n: contingencies involving choice pairs (0, 0), (0, 1) and (0, 2)

Game 3p2n: contingencies involving choice triplets (0, 0, 0), (0, 0, 1) and (0, 1, 1)

Class B therefore includes subjects who used the correct (if not most efficient) approach which in principle allowed them to recognize the dominance of choosing 0, but who apparently did not recognize it. By contrast, class C includes subjects who explicitly recognized and explained the dominance of choosing 0. In addition to the aforementioned belief-based explanations of dominance, class C subjects used reasoning such as "Choosing 0 is an always-winning choice," or "If I choose 1, I can lose, whereas if I choose 0, I always win or at worst tie," or "If I choose 0, I don't need to take the choice(s) of the other player(s) into account."

To the extent that subjects did not always report their reasoning clearly and completely, we cannot rule out classification errors. If uncertain whether a subject falls into class A (class C), the examiners assigned the subject into a "borderline" class A/B (class B/C). If uncertain whether a subject used erroneous belief-based reasoning or rather a belief-based explanation of dominance, the examiners assigned the subject into a "borderline" class A/C. Appendix 2 outlines further steps taken to minimize classification errors.

We repeatedly reminded the examiners that our primary classification goal was to maximize the accuracy of assignment into reasoning classes A and C. This assignment turns out to be robust in that, except for four and three subjects, respectively, the examiners' independent assignments into class A and class C coincide. The robustness is much lower for class B and the borderline classes where the assignments mostly arise from an initial disagreement between the examiners and subsequent re-classification.[12] In the discussion of results below, we therefore mainly concentrate on the robust classes A and C.

Our classification procedure improves upon previously implemented classifications of reason-

[12]See Appendix 2 for details of the re-classification procedure.

ing processes in iterated-dominance-solvable guessing games (e.g., Bosch-Domenech et al., 2002; CGC) which relied on correctly disentangling individual (ir)rationality from beliefs about others' (ir)rationality.[13] Our classification, by contrast, focuses solely on whether subjects are rational in terms of obeying simple dominance. Reasoning processes in a dominance-solvable guessing game were also collected by GN but were used only to illustrate specific cases of dominance violation. Our advantage over GN lies in our games having constrained strategy spaces. As a result, our subjects mostly report their reasoning in an easily interpretable manner, which reduces the potential scope for classification errors.

This has important implications for interpreting the relationship between reasoning classes and choices (Section 6.1). In particular, provided that classification errors are minimal, class C subjects should make the dominant choice of 0, unless they slip up during the ultimate decision-making stage of actually stating their (dominant) choice. By inspecting the choice distribution of class C subjects, we can assess the extent of such choice errors. On the other hand, class A subjects most likely make errors during an earlier reasoning stage of the decision-making process, and we explore the nature of such reasoning errors (Section 6.3) and trace them back to subjects' cognitive and personality characteristics (Section 6.2). We also check the extent to which class A subjects make dominated or (accidentally) dominant choices.

Beliefs about others' choices play no strategic role in our dominance-solvable guessing games as they are irrelevant for own optimal behavior. For that reason, we do not elicit beliefs in an incentive-compatible manner, and our aim is not to assess whether subjects act on their beliefs (see, e.g., Costa-Gomes and Weizsäcker, 2008). We merely interpret subjects' stated beliefs as an interesting indicator of their view of others' rationality, and we report the beliefs conditional on subjects' own dominance compliance as revealed by their reasoning class and choice (Section 6.1). This usefully complements the evidence from iterated-dominance-solvable games where individual rationality and beliefs are necessarily assessed concurrently.[14]

[13]Bosch-Domenech et al. classify mostly optionally reported reasoning processes from lab, classroom, and field experiments. As an implementation caveat, the classification is done by the authors themselves. The authors use the classification to conclude that guess distributions visually differ across reasoning classes broadly as predicted by iterated belief types. CGC collect reasoning processes only ex-post through a debriefing questionnaire, and use them to diagnose reasoning errors (of the kind we discuss in Section 6.3) or exotic decision rules not discernible from subjects' guesses alone.

[14]We acknowledge that the assessment of beliefs is more informative for class C subjects who apparently understood that others' choices are irrelevant for their own best response, but perhaps less informative for class A subjects who rarely understood the strategic nature of the guessing games (see Section 6.3).

3.4 Cognitive, personality and demographic characteristics

Our subjects completed several tests of cognitive abilities, several scales measuring personality traits, and a demographic questionnaire. Because of no strong priors regarding which individual characteristics might predict behavior in our games, measuring a broader set of potentially relevant characteristics seemed desirable in order to explore and compare their effect.[15] Below we briefly outline the measured cognitive, personality and demographic characteristics, of which working memory (cognitive ability), intrinsic motivation and premeditation attitude (personality traits) turn out to be important predictors of subjects' behavior. We refer the reader to Rydval (2007) and Ballinger et al. (2008) for further details of the cognitive tests and personality scales.

Working memory is viewed by psychologists as the ability to keep relevant information accessible in memory when facing information interference and to allocate attention among competing uses when executing cognitively complex tasks. Working memory tests proxy general cognitive abilities in that they robustly predict general "fluid intelligence" and performance in a broad range of cognitive tasks requiring controlled (as opposed to automated) information processing (e.g., Feldman-Barrett et al., 2004; Kane et al., 2004). Working memory also positively affects economic performance, such as precautionary saving behavior (Ballinger et al., 2008) or forecasting performance (Rydval, 2007). We measure working memory by a computerized version of the "operation span" test (Turner and Engle, 1989) that requires memorizing sequences (of various lengths) of briefly presented letters interrupted by solving simple arithmetic problems. At the end of each sequence, subjects are asked to recall as many letters as possible in the correct position in the sequence, which in turn determines the test score.

Short-term memory reflects information storage capacity as well as information coding and rehearsal skills that make the stored information more memorable (e.g., Engle et al., 1999). We measure short-term memory by a computerized auditory "digit span" test similar to the Wechsler digit span test (e.g., Devetag and Warglien, 2007). Our test requires memorizing pseudo-random sequences (of various lengths) of briefly presented digits and recalling them immediately after hearing each sequence.[16] The test score is based on the number of digits recalled in the correct position in the sequences.

We measure subjects' personality traits using several item-response personality scales described below. Personality traits could predict guessing game behavior but could also correlate

[15]Some of the cognitive tests and personality scales were primarily implemented for the purpose of an unrelated follow-up experiment completed by the subjects (see Section 5 for details).

[16]What distinguishes the short-term and working memory tests (and cognitive constructs) is an "attention interference" task in the latter tests, such as the simple arithmetic problems in the operation span test.

with measured cognitive abilities, so we measure both to disentangle their effect. Each personality scale consists of a collection of statements (worded positively or negatively) for which subjects indicate their agreement or disagreement on a scale from 1 to 4. The personality scales were included in a single item-response survey in a randomized order identical across subjects.

The need for cognition scale measures intrinsic motivation to engage in cognitively demanding tasks (e.g., Cacioppo et al., 1996). There is an extensive (inconclusive) literature in economics and psychology on the channels through which intrinsic motivation could interact with financial incentives in stimulating mental or physical effort and performance (e.g., Deci et al., 1999; Gneezy and Rustichini, 2000; McDaniel and Rutström, 2001; Ariely et al., 2008). Not addressing the complex interactions, we measure intrinsic motivation to account for the possibility that subjects are *ex ante* differentially motivated to solve the guessing games or that intrinsic motivation correlates with subjects' measured cognitive abilities.

The premeditation scale captures the propensity to pause and think carefully while carrying out (cognitive) tasks, which might be relevant for forming sound reasoning processes in our games. The sensation-seeking scale is a general proxy for risk-taking attitude which might affect subjects' willingness to experiment with alternative approaches to solving the guessing games. The perseverance scale measure subjects' determination and perseverance in solving lengthy and demanding tasks.[17] The math anxiety scale is a proxy for feelings of tension when manipulating numbers and solving math problems (e.g., Pajares and Urdan, 1996).

We further elicit risk preferences using a hypothetical "multiple price list" procedure (e.g., Holt and Laury, 2002). While risk preferences theoretically should not affect subjects' ability to reason in line with dominance, subjects who do not recognize the dominance of choosing 0 might view their own choice as risky, and risk-taking attitude might also matter for the reasons hypothesized above for sensation-seeking. Finally, we administer a demographic questionnaire to collect data on subjects' age, gender, field of study, and socioeconomic status such as (family and personal) car ownership.

3.5 Implementation details

The experiment was conducted at the Bank Austria Portable Experimental Laboratory at CERGE-EI in November 2005 and January 2006, as displayed in Table 1.[18] The subjects

[17]The premeditation, sensation-seeking and perseverance scales capture various aspects of impulsive behavior (Whiteside and Lynam, 2001). See Ballinger et al. (2008) for further details.

[18]Due to concerns that subjects in successive experimental sessions might share information relevant for performance in the guessing games and some of the cognitive tests, we ensured to the extent possible that successive sessions overlapped or that subjects in non-overlapping sessions were recruited from different universities or university campuses. Judging from the experiment following the guessing game, subjects' behavior suggests little or no degree of social learning (see Rydval, 2007).

were 112 full-time students (Czech natives, with a couple of exceptions permitted based on proficiency in Czech) from Prague universities and colleges, namely the University of Economics, Czech Technical University, Charles University, and Anglo-American College, with the majority of subjects from the first two universities.[19] None of the subjects had prior formal training in game theory.

Table 1: Order of experimental sessions and number of participants

Session	1	2	3	4	5	6	7	8
Game	3p2n	2p2n	2p3n	3p2n	2p2n	2p3n	2p3n	3p2n
# participants	15	14	14	12	14	14	13	16

Each experimental session started with conducting the cognitive tests and personality scales, followed by the guessing game and the demographic questionnaire.[20] The guessing game itself lasted about 20-30 minutes. We read the instructions aloud (see Appendix 1) and then gave subjects virtually unlimited time to re-read the instructions, to ask any questions, and to fill out the answer protocol. We did not explicitly check subjects' understanding of the instructions. While experimentalists often implement prior understanding tests or unpaid practice rounds to ensure that subjects understand the potential consequences of their and others' decisions, doing so in our simple guessing games would almost inevitably induce undesirable experimenter demand effects or suggest strategies to the subjects.[21] As an alternative to an understanding test, GN implement more elaborate instructions in several sessions of their dominance-solvable guessing game, but find no impact on behavior.[22] In Section 6.3, we report on additional sessions aimed at gauging the nature and extent of our subjects' misunderstanding.

The experiment lasted 1.5-2 hours and subjects earned 150 CZK (≅ PPP$12) for its completion. In addition, the guessing games featured the fixed prize of M=1500CZK (≅ PPP$117) for

[19] Czech Technical University is a relatively non-selective university mostly offering education in various branches of engineering, while the University of Economics is a more selective university mostly offering education in economics, management and accounting. We do not detect any differences in subjects' behavior related to their field of study, though the sample sizes involved in those comparisons are too small to draw any firm conclusions.

[20] After a short break the sessions continued with an individual decision-making experiment unrelated to the guessing game (a time-series forecasting task; see Rydval, 2007).

[21] Understanding tests of course strive hard to avoid such adverse effects – usually by checking solely that subjects understand how their and others' decisions determine payoffs – but even that may have behavioral consequences. For example, in Bosch-Domenech et al.'s (2002) iterated-dominance-solvable guessing game, subjects who a priori observed an example outlining the consequence of guessing a low number violated dominance less frequently than other subjects not observing that example.

[22] GN explain how the average of the pair's guesses is computed and then multiplied by the target number to determine the winning guess. This explanation has no effect on the distribution of guesses, though one should note that the change in instructions coincided with an increase in stakes as well as a minor change in the subject population.

the winner(s) originating from one pair or triplet selected at random in each session.[23] All parts of the experiment were anonymous (subjects were assigned a unique ID that they kept throughout the session) and earnings were paid out privately in cash after the experiment. The order of cognitive tests and personality scales was the same across sessions, with the former generally preceding the latter. The working memory and short-term memory tests were computerized using E-prime (Schneider et al., 2002) while the remainder of the experiment was administered in a paper-and-pencil format.

3.6 Results

3.6.1 Relationship between reasoning classes, choices and beliefs

Table 2 displays the number of subjects in the reasoning classes defined earlier, aggregated across the three guessing games. The first row shows that 66 subjects (59%) used wrong reasoning processes (class A), whereas 30 subjects (27%) reasoned consistently with dominance (class C). The remaining 16 subjects are scattered among class B and the borderline classes. Thus while class B contains only 3 subjects, it can in principle contain up to 13 subjects depending on how one interprets the borderline classes A/B and B/C. Similarly, class A can contain up to 72 (64%) subjects if adding the borderline classes A/B and A/C, and class C can contain up to 40 (36%) subjects if adding the borderline classes A/C and B/C.

[23]The guessing games were announced as a bonus task. Subjects knew about the existence of a bonus task (and the potential prize) from initial instructions. Subjects also knew they could earn an additional 900CZK (PPP$70) in the experiment following the guessing games.

Table 2: Frequency of subjects sorted by reasoning classes, choices and beliefs

	Total	Class					
		A	A/B	A/C	B	B/C	C
All subjects	112	66	3	3	3	7	30
Choice=0	62	17	2	3	3	7	30
Choice=1	50	49	1	0	0	0	0
Belief=0	49	12	1	2	3	4	27
Belief=1	63	54	2	1	0	3	3
Choice=0 & Belief=0	46	9	1	2	3	4	27
Choice=0 & Belief=1	16	8	1	1	0	3	3
Choice=1 & Belief=0	3	3	0	0	0	0	0
Choice=1 & Belief=1	47	46	1	0	0	0	0

The second and third rows of Table 2 display the frequencies of dominant and dominated choices. The first column shows that 62 subjects (55%) made the dominant choice of 0 while the remaining 50 subjects (45%) violated dominance.[24] This can be compared to a rate of dominance violation of 56% for random guesses, 10% typically reported for iterated-dominance-solvable games, and 90% for GN's weak-dominance-solvable game (see Section 2). In GN's unpublished treatment closest to ours – because of collecting guesses, beliefs and reasoning processes – 78% of subjects violated dominance.[25]

The second and third rows of Table 2 further show that subjects in class A/C or higher all made the dominant choice of 0: Hence they apparently did not commit any errors at the ultimate decision-making stage of stating their dominant choice. On the other hand, the high frequency of class A and class A/B subjects suggests a prevalence of reasoning errors made at an earlier stage of the decision-making process. Note, however, that little over a quarter of class A and class A/B subjects made the *dominant* choice. Thus to the extent that our classification is correct, the observed frequency of dominated choices in fact understates the actual frequency of dominance violations as revealed by dominance-incompatible reasoning.

The remaining rows of Table 2 display subjects' beliefs, first conditional on reasoning classes and then also conditional on choices. The last column indicates that all but three class C

[24]For ease of exposition, the Choice=1 category includes the two subjects who in fact chose 2 in game 2p3n. Similarly, the Belief=1 category includes the subject who stated a belief of 2. The prevalence of the dominated choice of 1 over the dominated choice of 2 might signal a focal-number effect of unity.

[25]We are grateful to Brit Grosskopf for providing us with the data for this unpublished treatment conducted with 18 student subjects. Dominance violation rates do not differ between this treatment and other treatments without belief elicitation. We again note design differences between our and GN's games, such as the multi-round nature of their experiment and their payoff function rewarding the winner(s) in each fixed pair in every round.

subjects believed that the other player(s) would likewise make the dominant choice. The second column shows that out of the 49 class A subjects who made dominated choices, all but three believed that the other player(s) would likewise do so. On the other hand, out of the 17 class A subjects who made the dominant choice, little over half believed that the other player(s) would likewise do so. This further illustrates that seemingly dominance-compatible choices and beliefs can sometimes be based on dominance-incompatible reasoning processes.

Table 3 disaggregates the percentages of reasoning classes and choices for each game. While our primary interest is not in across-game comparisons, we note that the percentage of class C subjects is highest in game 3p2n – even higher than in the simplest game 2p2n – and lowest in game 2p3n, and vice versa for the percentages of class A subjects. Accordingly, the frequency of dominant choices is highest in game 3p2n and lowest in game 2p3n (which is weak-dominance-solvable and in that respect closest to the GN's unpublished treatment mentioned above) but is always higher than random dominance compliance.[26] Thus we might have somewhat "less smart" subjects playing game 2p3n – an issue addressed in the next section – or game 2p3n might be generally harder to solve, perhaps due to its weak-dominance-solvable nature or its larger set of dominated strategies (and hence its lower random dominance compliance).

[26] Nevertheless, the three games are very similar in terms of the proportions of dominant and dominated choices made by class A subjects.

Table 3: Percentages (rounded to integers) of reasoning classes and choices for each game

	Total	Class					
		A	A/B	A/C	B	B/C	C
Game 2p2n (28 subj.)	100	57	4	0	4	11	25
Choice=0	57	14	4	0	4	11	25
Choice=1	43	43	0	0	0	0	0
Game 2p3n (41 subj.)	100	76	2	5	2	2	12
Choice=0	39	17	0	5	2	2	12
Choice=1	61	59	2	0	0	0	0
Game 3p2n (43 subj.)	100	44	2	2	2	7	42
Choice=0	70	14	2	2	2	7	42
Choice=1	30	30	0	0	0	0	0

3.6.2 Cognitive and personality predictors of reasoning classes and choices

Here we assume in a very simple manner that reasoning errors and choice errors have a logistic structure. In particular, Table 4 reports logit estimates of the effect of statistically relevant cognitive and personality characteristics on reasoning classes and choices. Other cognitive, personality and demographic characteristics are individually and jointly insignificant at the 10% level.[27] The dummies for game 2p2n and game 3p2n capture any remaining differences with respect to game 2p3n. In all estimations, we drop the three class A/C subjects, leaving us with 109 subjects.

[27]As an exception, arithmetic ability affects both reasoning classes and choices when included instead of or besides working memory, but its impact is generally weaker than that of working memory. We measure arithmetic ability using a "two-digit addition and subtraction" test under time pressure (see Rydval et al., 2008, for more details). Since we lack arithmetic ability scores for the first experimental session, we focus on the effect of working memory in the full sample, noting that working memory is correlated with arithmetic ability at the 10% significance level (Spearman correlation of 0.19) and hence that part of the explanatory power of working memory may reflect the impact of arithmetic ability. One could in principle separate the impact of working memory, arithmetic ability, and short-term memory on behavior (see Rydval, 2007), but this seems undesirable here due to the limited sample of subjects with arithmetic ability and short-term memory scores (additional ten observations are missing for logistical reasons).

Table 4: Logistic regressions of reasoning classes (*Model 1* and *Model 2*) and choices (*Model 3*) on cognitive and personality characteristics

REGRESSOR	***Model 1***			***Model 2***	***Model 3***
	marg.eff.	**marg.eff.**	**marg.eff.**	**marg.eff.**	**marg.eff.**
	(std.err.)	**(std.err.)**	**(std.err.)**	**(std.err.)**	**(std.err.)**
game 2p2n	-0.303^{**}	0.043^{*}	0.261^{**}	0.334^{**}	-0.241^{**}
	(0.133)	(0.022)	(0.123)	(0.136)	(0.119)
game 3p2n	-0.346^{***}	0.055^{**}	0.291^{***}	0.354^{***}	-0.320^{***}
	(0.114)	(0.026)	(0.104)	(0.120)	(0.108)
working memory	-0.084^{*}	0.018	0.066^{*}	0.106^{**}	-0.089^{*}
	(0.049)	(0.013)	(0.038)	(0.054)	(0.052)
need for cognition	-0.108^{**}	0.023	0.085^{**}	0.115^{**}	-0.044
	(0.051)	(0.015)	(0.039)	(0.054)	(0.054)
premeditation	-0.118^{**}	0.025^{**}	0.093^{**}	0.129^{**}	-0.010^{*}
	(0.046)	(0.012)	(0.038)	(0.056)	(0.056)
Joint significance		***		***	
Number of subjects	69	10	30	109	109
	LR chi-square(5)=5.37, p=0.372			**% correctly predicted**	
	Hausman chi-square(5)=2.07, p=0.840			74.31	64.22

Notes: Marginal effects are evaluated at the means of the regressors. Marginal effects in the ordered logit *Model 1* are for classes A, B and C. Working memory, need for cognition, and premeditation are z-standardized using their sample means and sample standard deviations. Standard errors and tests are based on the heteroskedasticity-robust "sandwich" estimator. *, ** and *** indicate significance of estimates at the 10%, 5% and 1% level, respectively. "Joint significance" stands for a chi-square test of joint significance of working memory, need for cognition, and premeditation. "LR" stands for an approximate likelihood ratio test of the null hypothesis that the coefficients are equal across classes. "Hausman" stands for a Hausman-type specification test of the null hypothesis that classes B and C can be merged.

Model 1 reports marginal effects for ordered logit estimation with the reasoning classes as the dependent variable: We conservatively re-assign subjects from the borderline classes A/B and B/C to classes A and B, respectively. The estimates for the game dummies confirm the overall higher likelihood of sounder reasoning processes in games 2p2n and 3p2n compared to game 2p3n. The remaining estimates suggest that higher working memory, need for cognition, and premeditation are associated with a lower likelihood of reasoning inconsistently with dominance (class A) and with a higher likelihood of reasoning consistently with dominance (class C).[28]

Since a Hausman-type specification test for *Model 1* suggests that treating reasoning classes B and C separately is unnecessary, we merge the classes in *Model 2*. The resulting logit estimates reaffirm the results of *Model 1*, namely the positive predictive power of measured working memory, need for cognition, and premeditation for our subjects' ability to reason consistently with dominance. A one-standard-deviation increase in any of the three variables is associated with an increase in the likelihood of using a sound reasoning process (here class B or higher) by over 10 percentage points.[29]

Model 3 reports marginal effects for logit estimation with choices as the dependent variable. The negative estimates reflect that higher working memory, need for cognition, and premeditation are associated with a higher likelihood of making the dominant choice of 0. However, comparing *Model 2* and *Model 3* reveals that the predictive power and magnitude of impact of the three cognitive and personality characteristics is much higher in the former model.[30] This in our view further confirms the value-added of attending to reasoning processes, besides choices, when assessing our subjects' ability to reason consistently with dominance.

We note once again the marked across-game differences in reasoning processes and choices which prevail even after accounting for the effect of cognitive and personality characteristics.[31] Though not reported in Table 4, there are no differences in the effect of working memory, need for cognition, and premeditation across the games. Therefore, dominance seems harder to understand or apply in the weak-dominance-solvable game 2p3n (with lower random dominance compliance) compared to the strict-dominance-solvable games 2p2n and 3p2n.

3.6.3 Further exploring origins of dominance-incompatible reasoning

Our classification procedure reveals further insights about the nature of reasoning errors. First, the examiners indicated that misunderstanding the experiment's instructions appeared rare: only

[28]In all estimations, the working memory score is the total number of correctly recalled letters only in letter sequences recalled entirely correctly. An alternative score, based on the total number of correctly recalled letters, has less predictive power in our estimations. See Conway et al. (2005) for a comparison of the two valid working memory scoring procedures.

[29]These effects are independent to the extent that working memory, need for cognition, and premeditation are not correlated in our subject sample at the 10% significance level.

[30]We include need for cognition in Model 3 for the purpose of a direct comparison with Model 2. Although the preferred model of choice behavior does not feature need for cognition, including it does not affect the significance of the other regressors.

[31]Although need for cognition is on average significantly lower for subjects in game 2p3n than in game 3p2n at the 5% level (using a two-sided rank-sum test and t-test), Table 4 shows that this cannot explain the lower performance of subjects in game 2p3n.

up to three class A subjects seemingly misunderstood that they played an iterated-dominance-solvable guessing game against everyone else in their session.[32] Second, nearly a quarter of class A subjects apparently failed to incorporate the target number, ˝, in their reasoning process, yet this was unlikely due to their misunderstanding of the instructions *per se* (since we stressed the target number when reading the instructions aloud) but rather due to a reasoning or computational error. Third, nearly half of class A subjects reported irrelevant belief-based reasoning such as "I believe the other player chooses 1, so I will choose 1 and we will split the prize," or irrelevant focal-number reasoning such as "I like number 1 more than number 0 and hence choose 1." These kinds of irrelevant reasoning might likewise stem from failing to incorporate the target number, i.e., interpreting it as unity. In the 2p2n and 2p3n games, this would imply a game where both players win regardless of their choices. In game 3p2n, this would imply a game where one's own choice has a pivotal influence on the winning choice only if the other two players choose 0 and 1. Thus choosing 0 would no longer be a dominant strategy in any of our games.

This suggests that many class A subjects might have misrepresented the strategic nature of our guessing games and hence played a wrong game. Indeed, in Devetag and Warglien's (2007) two-player game similar to our guessing games, nearly a quarter of subjects misrepresented the relational structure between own and other's preferences.[33] Strategic misrepresentations could be even more widespread in our games since Devetag and Warglien explicitly display the to-be-represented preferences, whereas we rely on subjects inferring the preference (payoff) structure from the verbal instructions.

To explore the nature and extent of strategic misrepresentations, we conducted four additional sessions for the simplest game 2p2n. We changed our experimental design in that subjects first filled out all contingencies, i.e., the four combinations of both players' possible choices and their payoff consequences (see Appendix 1). We also asked subjects to rank the contingencies according to their preferences, had they been able to choose among them (subjects could express indifference between any contingencies by using ranking such as "1, 2, 2, 4" or "1, 3, 3, 3"). Only then were subjects prompted to report their complete reasoning leading them to their choice and to state their choice and beliefs. The additional sessions otherwise resembled the original

[32]Here and hereafter, we quantify the maximum extent of specific types of reasoning errors, as indicated by either of the examiners in the nine-class classification scheme (see Appendix 2).

[33]Devetag and Warglien categorize two-player games by type of bi-ordered preference structures varying in relational complexity. Their subjects select four out of 16 possible squares simultaneously representing two order relations, one represented by the size and the other by the color of the squares. Our guessing games fall into a category of preference bi-orders found in games of conflict where players' preference relations are the reverse of one another (though our games feature non-strict payoff relations unlike Devetag and Warglien's games). Harder relational structures found in chicken games and prisoner's dilemma games were misrepresented by 34% and 52% of Devetag and Warglien's subjects, respectively (138 undergraduate and MBA students in two related experiments).

ones, including the payoff function and the subject population.[34]

Before answering our key question, we note that out of the 64 additional subjects, 50% reasoned in line with dominance (class C) while the rest did not (class A), and 31% of subjects violated dominance by making dominated choices. Hence compared to our findings in Tables 2 and 3, asking subjects to represent game 2p2n in terms of its contingencies seems to slightly reduce but certainly not eliminate dominance violations.[35] This finding is in the spirit of GNa's observation that implementing more elaborate instructions in their dominance-solvable guessing game does little to improve subjects' understanding of dominance.

As to our key question, all but one class C subject filled out the four contingencies correctly.[36] By contrast, all but two class A subjects were unable to do so, despite always correctly listing all combinations of choices. A third of class A subjects assigned identical payoff, M/2, to both players in all contingencies (and their reasoning reveals a failure to incorporate the target number, 1/2); another third assigned the prize, M, to the dominated rather than the dominant choice; and the remaining third filled out partly or entirely wrong payoffs – seemingly illogical fractions of M or payoffs not summing to M (for a given contingency).

3.7 Discussion and conclusion

To understand the nature of dominance violations in dominance-solvable guessing games, we study the relationship among subjects' reasoning processes, choices, beliefs, and cognitive and personality characteristics. Our classification of reasoning processes suggests that only 27-36% of subjects reason in line with dominance; they all make the dominant choice and almost all expect others to do so. On the other hand, 59-64% of subjects reason inconsistently with dominance, of which about three-quarters make dominated choices (and almost all of those expect others to do so) but the remaining quarter perhaps accidentally make the dominant choice (and half of those expect others to do so).

Our additional findings in Section 6.3 reveal that dominance-incompatible reasoning stems primarily from subjects misrepresenting the strategic nature of the games. Specifically, half

[34]Subjects' earnings did not depend on how they filled out and ranked the contingencies, but completing these tasks was a precondition for receiving the participation fee. Subjects were undergraduates from the University of Economics and Czech University of Life Sciences in Prague.

[35]These observations rest on classification done by the authors. Similar to our findings in Tables 2 and 3, our additional class C subjects all make the dominant choice and almost all believe that others would do so; most of the additional class A subjects make dominated choices and expect others to do so, yet nearly a third of class A subjects make the dominant choice and nearly half of those expect others to do so.

[36]Interestingly, a quarter of class C subjects indicated in their ranking (and sometimes also in their reasoning) a preference for splitting the prize with the other player – by ranking highest the contingencies with choice pairs (0, 0) and (1, 1) – or an indifference between winning and splitting the prize. The remaining class C subjects ranked the four contingencies according to their own payoff.

of our additional subjects, and likely a similar fraction of our original subjects, were unable to connect their own and others' choices with the payoff consequences, despite the minimalist nature of our games and despite having virtually unlimited time for clarification questions and for making decisions. This kind of bounded rationality, observed in a similar form by Devetag and Warglien (2007), underlies the game-theoretic notion of "sampling equilibria" (e.g., Osborne and Rubinstein, 1998). We note, however, that this kind of bounded rationality may conceivably be less widespread in more naturalistic settings than ours.

In our original subject sample, the likelihood of reasoning errors – most likely misrepresentation errors – is higher for subjects' with lower ability to maintain and allocate attention, as measured by working memory. This is in line with Devetag and Warglien's (2007) finding of a positive link between short-term memory and the ability to represent preference structures similar to our guessing games – though short-term and working memory are quite distinct cognitive constructs (see Section 4) – and also in line with Burnham et al.'s (2007) finding of a positive link between a short test of general cognitive ability and performance in an iterated-dominance-solvable guessing game.[37] We acknowledge that the observed impact of working memory may partly reflect the influence of arithmetic ability (see footnote 27). Also, the effect of working memory might be a combination of a direct effect on behavior and an indirect one activated by requiring subjects to report their reasoning processes.[38]

In our original subject sample, reasoning (misrepresentation) errors are also more likely for subjects with *ex ante* lower intrinsic motivation and premeditation attitude, presumably due to their lower willingness to engage in solving the guessing games or to carefully think through the solution. In our view, this does not contradict CGC's conclusion that deviations from theoretical predictions in their iterated-dominance-solvable guessing games are mainly driven by cognitive errors rather than insufficient motivation. Our findings suggest that insufficiently (intrinsically) motivated subjects were most likely excluded from CGC's subject sample after failing an understanding test.[39]

Put differently, our findings suggest that had we implemented the listing of contingencies as an understanding test and dismissed subjects failing it – which would have been mostly those with low working memory, intrinsic motivation or premeditation attitude – we would have

[37]Neither of the two studies account for other potential sources of individual heterogeneity in cognitive abilities and personality traits (as Devetag and Warglien acknowledge), though Burnham et al. control for individual differences in gender, education, and age.

[38]Cognitive scientists, especially proponents of Protocol Analysis, usually take more care than we did to train subjects in verbalizing thought processes in a manner not interfering with solving the task itself (e.g., Ericsson and Simon, 1993; Ericsson, 2002). Describing thought processes – especially aloud, i.e., not in our case – may require additional cognitive resources, divert task-specific cognitive processes and hence generate invalid descriptions of thoughts, particularly in insight tasks requiring creative thinking (e.g., Schooler et al., 1993).

[39]CGC dismiss about 20% of subjects based on failing a detailed understanding test.

observed a much lower rate of dominance violation. From this perspective, our findings do not contradict the much lower dominance violation rates observed in Bone et al.'s (2008) extensive-form dominance-solvable game (where presenting the payoff structure visually presumably makes dominance transparent) or in iterated-dominance-solvable games (where subjects usually have to pass a "payoff structure understanding" test). A similar qualification probably also applies to the dauntingly high dominance violation rate reported in GN's dominance-solvable guessing game.

3.8 References

Ariely, D., Gneezy, U., Loewenstein G., Mazar, N., 2008. Large stakes and big mistakes. Review of Economic Studies, forthcoming.

Ballinger, P., Hudson, E., Karkoviata, L., Wilcox, N., 2008. Saving performance and cognitive abilities. University of Houston manuscript.

Bone, J., Hey, J. D., Suckling, J., 2008. Do people plan? Experimental Economics, forthcoming.

Bosch-Domenech, A., Montalvo, J. G., Nagel, R., Satorra, A., 2002. One, two, (three), infinity, ...: newspaper and lab beauty-contest games. American Economic Review 92, 1687-1701.

Burnham, T. C., Cesarini, D., Wallace, B., Johannesson, M., Lichtenstein, P., 2007. Billiards and brains – cognitive ability and behavior in a p-beauty contest. SSE/EFI Working Paper Series in Economics and Finance No. 684.

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., Jarvis, W. B. G., 1996. Dispositional differences in cognitive motivation: the life and times of individuals varying in need for cognition. Psychological Bulletin 119, 197-253.

Charness, G., Levin, D., 2008. The origin of the winner's curse: a laboratory study. American Economic Journal: Microeconomics, forthcoming.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., Engle, R. W., 2005. Working memory span tasks: a methodological review and user's guide. Psychonomic Bulletin and Review 12, 769-786.

Costa-Gomes, M. A., Crawford, V. P., 2006. Cognition and behavior in two-person guessing games: an experimental study. American Economic Review 96, 1737-1768.

Costa-Gomes, M. A., Crawford, V. P., Broseta, B., 2001. Cognition and behavior in normal-form games: an experimental study. Econometrica 69, 1193-1235.

Costa-Gomes, M. A., Weizsäcker, G., 2008. Stated beliefs and play in normal-form games. Review of Economic Studies, forthcoming.

Deci, E., Koestner, R., Ryan, R., 1999. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. Psychological Bulletin 125, 627–668.

Devetag, G., Warglien, M., 2007. Playing the wrong game: an experimental analysis of relational complexity and strategic misrepresentation. Games and Economic Behavior, forthcoming.

Ekstrom, R. B., French, J. W., Harman, H., Derman, D., 1976. Kit of Factor-Referenced Cognitive Tests (revised edition). Princeton, NJ: Educational Testing Service.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., Conway, A. R. A., 1999. Working memory, short-term memory, and general fluid intelligence: a latent variable approach. Journal of Experimental Psychology: General 128, 309–331.

Ericsson, K. A., 2002. Toward a procedure for eliciting verbal expression of nonverbal experience without reactivity: interpreting the verbal overshadowing effect within the theoretical framework for protocol analysis. Applied Cognitive Psychology 16, 981-987.

Ericsson, K. A., Simon, H. A., 1993. Protocol Analysis: Verbal Reports as Data (revised edition). Cambridge, MA: Bradford books/MIT Press.

Feldman-Barrett, L., Tugade, M. M., Engle, R. W., 2004. "Individual differences in working memory capacity and dual-process theories of the mind. Psychological Bulletin 130, 553-573.

Gneezy, U., Rustichini, A., 2000. Pay enough or don't pay at all. Quarterly Journal of Economics 115, 791-811.

Grosskopf, B., Nagel, R., 2008. The two-person beauty contest. Games and Economic Behavior 62, 93-99.

Grosskopf, B., Nagel, R., 2007. Rational reasoning or adaptive behavior? Evidence from two-person beauty-contest games. Harvard NOM Research Paper No. 01-09 (revised version).

Holt, C. A., Laury, S. K., 2002. Risk aversion and incentive effects. American Economic Review 92, 1644–55.

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., Engle, R. W., 2004. The generality of working memory capacity: a latent variable approach to verbal and visuospatial memory span and reasoning. Journal of Experimental Psychology: General 133, 189-217.

McDaniel, T, M., Rutström, E. E., 2001. Decision making costs and problem solving performance. Experimental Economics 4, 145-161.

Osborne, M. J., Rubinstein, A., 1998. Games with proceduraly rational players. American Economic Review 88, 834-847.

Pajares, F., Urdan, T., 1996. An exploratory factor analysis of the mathematics anxiety scale. Measurement and Evaluation in Counseling and Development 29, 35-47.

Rydval, O., 2007. Financial incentives and cognitive abilities: evidence from a forecasting task with varying cognitive load. Jena Economic Research Paper No. 2007-040 and CERGE-EI Discussion Paper No. 173.

Rydval, O., Ortmann, A., Ostatnicky, M., 2008. Three very simple games and what it takes to solve them. CERGE-EI Working Paper No. 347.

Schooler, J. W., Ohlsson, S., Brooks, K., 1993. Thoughts beyond words – when language overshadows insight. Journal of Experimental Psychology: General 122, 166-183.

Schneider, W., Eschman, A., Zuccolotto, A., 2002. E-prime User's Guide. Pittsburgh: Psychology Software Tools.

Simon, H. A., 1978. Models of Thought, vol. 1. New Haven: Yale University Press.

Simon, H. A., 1989. Models of Thought, vol. 2. New Haven: Yale University Press.

Stanovich, K. E., West, R. F., 2000. Individual differences in reasoning: implications for the rationality debate? Behavioral and Brain Sciences 23, 645-665.

Turner, M. L., Engle, R. W., 1989. Is working memory capacity task-dependent? Journal of Memory and Language 28, 127-154.

Whiteside, S. P., Lynam, D. R., 2001. The five factor model and impulsivity: using a structural model of personality to understand impulsivity. Personality and Individual Differences 30, 669-689.

3.A APPENDIX 1: Instructions and answer protocol

[The instructions below were presented to subjects for game 2p2n. Italics denote alterations for games 2p3n and 3p2n. The instructions were in Czech and were preceded by general instructions explaining, among other things, the anonymity of the experiment and the privacy of the paying-out procedure. Explanatory notes in square brackets do not appear in the instructions. The bold face appears in the instructions.]

BONUS TASK! BONUS TASK! BONUS TASK!

ID: _________________

In this task, you will be randomly matched with one *(two)* other participant*(s)* in this room who will be solving the same task as you will. The task will be explained below.

From now on, the two *(three)* of you will be called a 'group'.

After everyone has finished the task, the winner from one randomly selected group will earn a prize of 1,500 CZK. If the group has more than one winner, the prize of 1,500 CZK will be split evenly between the winners.

The task:

Each member of the group chooses a number: 0 or 1 *(0, 1 or 2)*.

The winner is the group member whose choice is closest to **1/2 of the average of the numbers chosen by all group members**. [The experimenter read the instructions aloud, stressing the "1/2" to ensure the target number was not overlooked.]

[In the additional sessions (see Section 6.3), we inserted here instructions asking subjects to fill out the four combinations of both players' possible choices and their payoff consequences, in four consecutive tables of the following format:

Your choice:		Your payoff:	
His/her choice:		His/her payoff:	

Then subjects were asked to rank the tables (i.e., the combinations of choices and resulting payoffs) according to their preferences, had they been able to choose among them.]

Below, please write down the complete reasoning leading you to your choice and then answer the questions at the bottom of the page. **Write while you think! (If you need more space, please turn over and continue.)**

[Here subjects were given much more space to report their reasoning.]

Your choice:
Number (please circle) 0 1 *(2)*

Question: What choice do you expect from the other member*(s)* of your group?

Number (please circle) 0 1 *(2)*

3.B APPENDIX 2: Details of the classification procedure

In the classification instructions, we first presented the examiners with the three guessing games through a condensed version of the experiment's instructions accompanied by Figure 1. We reminded the examiners that they may encounter different reasoning processes across the three guessing games but that it is important to classify them consistently across the games.

The classification instructions further stressed that "[i]t is extremely important for us that you are consistent in your classification, from the very first to the very last subject. It may well happen during the classification that you change your mind about how you classified a previous subject. This is not an error on your part but please do tell us about such cases before proceeding with further classification." The examiners were encouraged to independently contact the second author in case of any questions or ambiguities, preferably before starting (or restarting) their classification.

To minimize the potential scope for subjective classification errors, we initially asked the examiners to independently classify subjects' reasoning processes according to a more detailed nine-class classification scheme. Being based on our evaluation of reasoning processes in previous pilot experiments, the nine narrower reasoning classes corresponded to the various subtle distinctions among potentially reported reasoning processes (see the above discussion in Section 3).

After the nine-class classification scheme, we were able to clarify those distinctions to the examiners (through examples unrelated to their specific nine-class classification results), to explain them how to classify reasoning processes within the three-class classification scheme and what types of classification errors to attend to. Judging from the examiners' feedback and their classification adjustments between the two classification schemes, we were successful in tackling these issues. Being based on classifying reasoning processes using pre-specified, narrowly-defined nine classes of potentially reported reasoning processes, our classification procedure meets the standards of the Protocol Analysis (see, e.g., Ericsson, 2002).

The three-class classification scheme yielded about 20% of classification disagreements between the examiners, half of which they subsequently jointly resolved (only if they deemed appropriate). This final re-classification procedure therefore left us with 10% (11 out of 112) of classification disagreements, which the examiners jointly assigned into the borderline classes in accordance with the nature of their disagreement. In similar fashion, the examiners also jointly re-examined the remaining subjects in the borderline classes and re-classified them if they deemed appropriate. For all the above cases, we revealed to the examiners the subjects' stated choices and beliefs to which they were a priori blind.

Chapter 4

Conclusion

The reader is right if he notices that the three chapters in this dissertation are loosely connected. As mentioned in the introduction, while all three chapters in this dissertation are motivated by stock market phenomena, the individual studies are only loosely connected to these phenomena and to each other.

The first chapter directly analyses the bid-ask spread on stock markets. The phenomenon of a decreasing bid-ask spread motivated further research that resulted in the second chapter. Due to the techniques used and the link to the price dispersion literature, the second chapter was written in the language of price dispersion literature. The last chapter is motivated by experimental research on stock markets.

In the first chapter I show, using a simple econometric model, that many shares exhibit a decreasing bid-ask spread. A decreasing bid-ask spread can be easily exploited. Those traders who wait for the end of the trading day pay lower costs associated with the bid-ask spread. The phenomenon can be studied further by utilizing models detecting information release and comparing them to the trend in the bid-ask spread. Also, the information exchange can be modeled across several exchanges using the unique situation of stock exchanges in the Czech Republic (two unrelated stock exchanges running together 3-4 individual markets).

In the second chapter I theoretically study a specific case of the price competition of an oligopoly. I show that prices posted by oligopolistic sellers are strictly positive even if the production costs are assumed to be zero. A positive bid-ask spread may relate to a positive price so that the spread may equal twice the lowest price of competing sellers. The trend in the bid-ask spread may be explained by change in the characteristics of buyers during the day, e.g. the reservation prices. In future research I think it would be possible to link the theoretical model to real data and prepare a testing procedure to see if the model is theoretically applicable on stock markets. As for the theoretical model itself, it can be extended in many directions - from asymmetric costs, through costs associated with increasing seller's share of the market, to more complex behaviors of buyers.

The third chapter is focused on experimental studies of competition. It is shown that our laboratory participants often have trouble to "solve" (i.e., understand correctly by the measuring rod of the most fundamental solution concept: dominance) even very simple games correctly. It seems likely that such failure would be even more pronounced in experiments dealing with more complicated institutions such as a stock market. But it may also be that our laboratory environment was simply too simple and that more context, even if more difficult to read, might have triggered subjects' recognition patterns more successfully. The question of the abstractness of the laboratory remains an interesting one.