Warsaw University of Life Sciences – SGGW
Faculty of Applied Informatics and Mathematics

METODY ILOŚCIOWE W BADANIACH EKONOMICZNYCH

QUANTITATIVE METHODS IN ECONOMICS

Volume XI, No. 1

Warsaw 2010

ISBN 978-83-7583-236-5

Warsaw 2010, Volume XI, No. 1

Published by Warsaw University of Life Sciences Press

CONTENTS

QUANTITATIVE METHODS IN ECONOMICS
Vol. XI, No. 1, 2010, pp. 1-23

AN EXCURSION THROUGH TWO HUNDRED YEARS OF FINANCIAL CRISES

Yochanan Shachmurove
Department of Economics, The City College of the City University of New York,
and
Department of Economics, The University of Pennsylvania
e-mail: yshachmurove@ccny.cuny.edu

Abstract: This paper examines major U.S. financial crises from the beginning of the nineteenth century to the present. The financial crises of Poland are also discussed. Observing similarities and differences among various financial crises, including their causes and government responses, can shed light on the nature of financial crises. The current financial crisis is described and potential policies that may remedy the problem are discussed. Additionally, the paper briefly describes financial crises in Poland. The conclusion is that financial crises cannot be avoided.

Please address all correspondence to: Professor Yochanan Shachmurove, Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104-6297. Email address: yshachmurove@ccny.cuny.edu

I have greatly benefited from research assistance from Adel El Amir from The City College of The City University of New York, Michael Chirico, Gregory Kauffman, Aamani Mynampati, from the University of Pennsylvania and Zach Winston from Pennsylvania State University. I am grateful to many discussions about this topic with Professor Lawrence Klein, Emanuel, Amir and Tomer Shachmurove.

INTRODUCTION

Former British Prime Minister Tony Blair, in his recent memoir, devotes the last chapter to the recent financial crisis. Blair (2010) points at governments, regulations, politicians, and monetary policies as the key culprits of the crisis. He sees the crisis as a consequence of the perception that expansionary monetary policy and low inflation could co-exist in the long run. Although Blair (2010) blames political leadership, he mainly accuses regulators who, in his opinion, are responsible for alerting leaders that a serious crisis is about to break. He claims that if political leaders had a warning, they would have acted to prevent, or at least alleviate the impending crisis.

Nobel Laureate Joseph Stiglitz (2010) also notes that policymakers did not see the recent crisis coming and, once the bubble burst, thought that the consequences would be short lived. Addressing the issue of who should be blamed, Stiglitz (2010) emphasizes the roles of the financial institutions that did not understand the risks being taken, the regulators who failed to see what was coming, and not least, the economic profession. He observes that standard macroeconomic models did not incorporate adequate analyses of banks.

Similarly, Paul Krugman (2009), another Nobel Laureate, claims among other scholars that the current economic crisis parallels that of the Great Depression. Krugman (2009) traces the financial crisis -- the greatest since the 1930s -- to the failure of regulation to keep pace with what he calls an "out-of-control financial system."

Looking at financial crises through the lens of the last two centuries, this paper casts a doubt on the global rush to further empower the regulators who have failed to anticipate past panics and ballooning bubbles.

The paper is organized as follows: Section II provides a literature review. Section III discusses the main U.S. banking panics of the nineteenth century. Section IV explains the causes and effects of the current financial crisis. Section V briefly relays financial crises in Poland. Section VI is a conclusion.

LITERATURE REVIEW

Recent literature highlights a variety of theories with regards to the causes and conditions engendering financial panics. Reinhart and Rogoff (2008, 2009A, 2009B) present evidence that these crises are not limited to the U.S. They assert that any parallels drawn between crises within the United States can be extended to other countries. Bordo (2003) examines the economic history of the United Kingdom and the United States and concludes that stock market crashes are worsened by unstable financial conditions.

Kaminsky and Reinhardt (1999) claim that financial crises can be traced to deteriorating economic fundamentals, mainly increases in either private or national debt that at some point become impossible to refinance. Similarly, García-Herrero and Del Rio-Lopez (2003), studying episodes from seventy nine countries for the years 1979 until 1999, similarly conclude that when a central bank pursues an objective of financial stability, the likelihood of a crisis is reduced. Demirgüç-Kunt and Detragiache (1998) and Laeven and Valencia (2008) identify several conditions which encourage large-scale breakdowns, such as a ratio over ten percent of non-performing to total bank capital or a large cost of two percent of GDP as the cost of the government bailout.

Boyd, Nicolò, and Loukoianova (2009) postulate that banking crises are due to responses of the financial institutions to government intervention. Klomp (2010) examines 132 banking crises across 110 countries and concludes that the GDP growth rate and the real interest rate can indicate the onslaught of a financial crisis, but that not a single factor is significant in causing more than 60 percent of crises.

NINETEENTH AND TWENTIETH CENTURY FINANCIAL CRISES IN THE UNITED STATES

As a newly developed country, the United States experienced a period of frequent banking panics in the nineteenth century with eight major crises. These eight episodes include the Panics of 1819, 1837, 1839, 1857, 1873, 1884, 1893 and 1896. The major crises of that century, which occurred in the years 1819, 1837, 1857, 1873, and 1893, are described below. Then the two major crises from the twentieth century, taking place in the years 1907 and 1929, are recounted.

The Panic of 1819

The Panic of 1819 was the first national financial crisis in the United States. The panic began with a sharp expansion of the money supply, leading to a bubble that burst and caused a sharp downturn. Prior to the War of 1812, regulation of banks was minimal with only two main constraints. During the War of 1812 the government of the United States had limited means of raising revenue for the war effort. Following the expiration of the First Bank of the United States' Charter in 1811, the constraints on bank lending were loosened, and many banks extended credit beyond their reserves, thus creating a large expansion of the money supply. As this increased lending continued, banks saw their reserves flowing to other banks with safer lending practices. For example, since New England opposed the war, its banks did not extend credit to the government, causing reserves to flow from the rest of the nation to New England.

As these practices continued, the government tried to prevent further withdrawals from banks, and thus suspended convertibility of banknotes into specie, the main venue for extending the money supply. Following the war, it was apparent that convertibility could not be restored immediately. The Second Bank of the United States was created in 1817 to allow for the transition back to the bimetallic standard. However, while the creation of this bank allowed banks to resume convertibility, it facilitated further lending that expanded the money supply even more. For example, by 1818, the bank had lent out $41 million and transferred the risk from bank balance sheets to its own balance sheet. The bank also increased the total banknote issue by $23 million, with only $2.5 million in reserves.

The policies of the new bank encouraged other banks to further expand their issuance of banknotes. This expansion caused a sharp increase in the price level, speculation in real estate and the founding of many banks. For example, the number of banks increased from 208 to 245 in the year of 1815 alone. However, with the repayment of the "Louisiana debt" that was used for the Louisiana Purchase in 1818 and 1819, required in specie, The Second Bank of the United States had to engage in a series of policies to reverse the expansion in the money supply.

The actions of the bank facilitated the Panic of 1819 by creating a deflationary environment where debtors were no longer able to repay the banks, and the value of banknotes depreciated. By 1821, the panic and depression began to clear and the credit contraction ended. Though it had been painful, the United States had survived the first of the eight banking panics in the nineteenth century (Rothbard, 1962).

The Panic of 1837

The next banking crisis was the Panic of 1837. In the five years following the suspension of specie payments, 194 of the 729 chartered banks failed. The book assets of state owned and controlled banks fell 45 percent. Five-year depression ensued. Rousseau (2002) enumerates two main causes of the crisis. The first was the Deposit Act of 1836, which forced the government to distribute $28 million of the $34 million in Federal surplus to state banks, other than New York City banks. The other was the "Specie Circular," also passed in 1836, which required public lands to be paid for in specie. These two policies led to a drain in the reserves of deposit banks from New York City and other commercial centers to the rest of the country. The reserves in New York fell from $7.2 million on September 1, 1836 to $1.5 million on May 1, 1837. This drain in reserves from the commercial centers of the United States led to a loss of confidence in New York banks, making the panic inevitable. The policy of enacting regulations intended to respond to one problem simultaneously and inevitably creates other impediments, not less severe. This is the hallmark of financial regulations since these essential liquid instruments, just like water, will eventually find their way out to the unregulated high seas. This phenomenon repeats itself in many past, present and future crises.

Rolnick, Smith and Weber (1997) emphasize two different causes of the Panic of 1837. One view sees the culprit as President Andrew Jackson for not renewing the charter for The Second Bank of the United States, which had disciplined weaker banks by returning their paper. The other highlights the fall in cotton prices as conducive to the crisis.

Following the implementation of the Deposit Act and the Specie Circular, banks began suspension of converting banknotes into specie in May of 1837. Almost a year later, in April of 1838, banks began to resume convertibility. In the following year, a further economic slowdown occurred, causing banks to again suspend payments, leading to many additional bankruptcies. Two more years of recession followed (Rolnick, Smith and Weber, 1997).

The Panic of 1857

The Panic of 1857 was characterized by the closure of the Ohio Life and Trust Company, a collapse of the stock and bond markets and a sharp recession. The United States was on a bi-metallic (gold and silver) standard at the time. Thus the discovery of new gold deposits sharply increased the money supply. Consequently, a speculative bubble emerged, primarily in railroads and the land required to build them. Nearly $700 million was spent over a nine year period to construct about 18,000 miles of rails, which at the time accounted for over 77 percent of all railroad mileage in the United States (Conant, 1915, p. 637).

When the bubble burst, speculators were unable to repay their debts causing some banks to fail. This increased the probability that other banks would go bankrupt. Bank runs resulted in the collapse of stock and bond markets, which further exacerbated the situation. As banks failed, many cities and states suspended the convertibility of bank deposits into gold. However, lacking a central monetary authority, nation-wide coordination among banks was impossible. Thus, when Philadelphia suspended convertibility in September of 1857 and New York City did not, fear–induced bank runs ensued in New York City. Consequently, on October 13, 1857, New York City suspended convertibility. Only on November 20 did New York resume partial convertibility, thus averting a major crisis (Calomiris and Schweikart, 1991, p. 822-828).

The Panic of 1873

The Panic of 1873 was caused by rapid overexpansion following the American Civil War. From the mid-1860s to the early 1870s, the United States railroad industry thrived. As tracks were laid, railroad companies benefited from government grants and subsidies. These grants and subsidies, in turn, spurred private investment in the industry (Oberholtzer, 1926). The railroad industry grew at a fast pace, eventually creating an excess supply of railroads and dramatically reducing the returns to investment in new rails.

Significant debt accumulation led to failure of many large entities, including Jay Cooke & Company. The stock exchange closed for ten days starting on September 20, 1873, setting off a chain reaction. Bank runs ensued as people panicked due to losses from railroad investments, causing further failures of banks. The U.S. GDP decreased for the following six years.

As in previous panics, the Panic of 1873 was worsened by government policies. The Coinage Act of 1873, issued earlier in the year, moved the United States to a pure gold standard, so that silver was dethroned as a form of exchange. The amount of currency available decreased, which led to sudden and severe deflation. Furthermore, the monetary policy of President Ulysses Grant consisted of further contracting the money supply (Wheeler, 1973). The resultant sudden and unexpected high interest rates hindered the repayment of debt. Although the U.S.

government tried to mitigate the deflation by buying bonds, the attempt was insufficient to counter the decrease in GDP.

The Panic of 1893

The Panic of 1893 is considered the worst financial crisis in the United States prior to the Great Depression (Timberlake, 1997). Similar to the Panic of 1873, it was caused by an array of factors, the most important being the overexpansion of the railroad industry and an expansionary monetary policy.

The previous decade of the 1880s was a time of high economic growth and optimism. However, the main engine of this growth was risky speculation. In the 1860s, citizens invested heavily in American railroads creating excess capacity in the industry, leading to bankruptcies. The dissolution of the Philadelphia and Reading Railroad was the first sign of the panic (Holton, 1990). As more railroads companies and their suppliers filed for bankruptcy, concern among the general public grew, and bank runs began. More entities failed, leading to widespread layoffs and causing unemployment to increase to 19 percent at the peak of the panic (Hoffman, 1970).

Analogous to the previous panics, the Panic of 1893 was worsened by a deflationary monetary policy. The Sherman Silver Purchase Act of 1890 required the U.S. government to purchase silver using currency backed by metallic specie. Soon there were not enough reserves in banks to exchange silver for gold. The value of silver became practically negligible, and deflation persisted for years, making it difficult for firms and individuals to repay debt.

The Panic of 1907

The Panic of 1907, another of the most severe financial crises before the Great Depression, was triggered by a sudden downturn in the New York Stock Exchange. Stock prices fell almost fifty percent from the previous year (Braunstein, 2009). One of the most significant causes of the downturn was a failed bid to corner the stock of the United Copper Company. When the venture failed, entities that had banked on its success incurred substantial losses.

Bank runs began as concerns about decreased liquidity surfaced (Bruner and Carr, 2007). The runs only compounded the shortage of funds, and firms accrued significant amounts of debt. The lack of a U.S. central bank to inject liquidity meant that the crisis could only be temporarily ameliorated by the lending of emergency funds by wealthy magnates and large corporations.

In just a few months, however, the stock value of the Tennessee Coal, Iron, and Railroad Company (TCI) plummeted. Crisis was once again closely averted by another emergency purchase—J.P. Morgan bought TCI to bolster the value of its shares (Bruner and Carr, 2007).

Unlike in Panics of 1853, 1873, and 1893, economic overexpansion did not play a significant role in the Panic of 1907. Instead, the cause of the crisis was a lack of funds, exacerbated by the absence of a central banking system, which was disbanded when President Andrew Jackson did not renew a charter for a central bank in 1823. However, the crisis was among a series of factors that led to the establishment of the Federal Reserve in 1913 (Caporale and McKiernan, 1998).

The Great Depression

The most severe financial crisis of the twentieth century in the United States began after a dramatic stock market crash in October 1929. In just a few months, the Dow Jones Industrial Average fell from the high 300s to below 100 (PBS, 2008). The causes of the crash were similar to those of previous financial panics. There was a rise in speculation during the “Roaring Twenties” which led to a financial bubble. The bubble burst in 1929, and was followed by widespread panic. As many attempted to retrieve their bank deposits and to sell stocks, financial institutions incurred more debt and stock prices fell further. Other factors that led to the depression include the unbalanced world economy in the wake of World War I and the resulting inflation, left unchecked by the newly-established Federal Reserve.

The Great Depression lasted for several years, despite efforts at resolution by the U.S. government. The Hoover Administration directly lent capital to banks and individuals with the passing in 1932 of the Reconstruction Finance Corporation Act and the Federal Home Loan Bank Act. In the years from 1933 to 1938, President Franklin Roosevelt actuated the New Deal, a series of economic programs which generated millions of jobs for the unemployed. Ultimately, though, it was the build-up for World War II that bolstered spending and Gross Domestic Product, ending the depression (Klein, 1947).

The Great Depression led to the establishment of the Glass – Steagall Act in 1933, which regulated the economy for years to come. It divided banks into separate categories, and created the Federal Deposit Insurance Corporation (FDIC) to insure citizens in case of bank runs.

The Savings and Loan Crisis

The Savings and Loan Crisis was rooted in the failure of many Savings and Loans (S&L) associations. These associations traditionally offer interest on savings deposits and use the deposits to make additional loans. Until 1980, the thrift industry was regulated by the U.S. government. However, with the passage of the Depository Institutions Deregulation and Monetary Control Act in 1980 and the Garn – St. Germain Depository Institutions Act of 1982, thrift institutions became financial intermediaries with the power of banks, but without the associated regulations (Mason, 1993). Risky speculation began, especially concerning real estate. With time, as yields on speculation declined, many Savings and Loans institutions failed.

Other factors which led to the crisis include the Tax Reform Act of 1986, which removed tax shelters and decreased values of many investments.

Government deposit insurance of thrift institutions led S&Ls to invest in riskier ventures. The collapse was ineluctable (Strunk and Case, 1988).

As a result of The Savings and Loan Crisis, the government repealed the Glass-Steagall Act. Furthermore, the federal government provided funds and bailed out many S&L institutions. Many point to these bailouts as contributing to the large budget deficit in the 1990s and beyond.

THE CURRENT FINANCIAL CRISIS

The current financial crisis, which began in 2007, is said to be the worst economic crisis to strike the U.S. since the Great Depression. It was caused by a variety of factors, but once again, its roots can be traced to speculation and the bursting of a financial bubble. Before the crisis struck, in 2006, the value of real estate peaked. However, much of the rise in prices was hinged on speculation.

In the early- and mid-2000s, as many foreign and domestic citizens invested in mortgage-related companies and products, liquidity increased in the real estate industry. Companies took this availability of funds for granted and began to issue mortgages at very low rates. Financial innovations such as collateralized debt obligations (CDOs) and mortgage-backed securities (MBSs) surfaced. People who could not necessarily afford some homes in the long run began to purchase them. The effects of these new deals were not strictly internalized or regulated.

When mortgage rates began to rise due to declining prices and lagging investment in the housing industry around late 2006, new homeowners could not pay the amounts that were demanded by the subprime mortgage industry. A wave of home foreclosures began. Even when foreclosing, however, companies could not retrieve the full worth of the house that had been sold. This is because the U.S. government does not allow companies to claim assets besides the house itself to retrieve lost value. Though a landlord has the right to evict a tenant with five days notice if rent payments are late, banks must foreclose on a house and have a court order issued before a tenant can be evicted. Also, if the value of a house is lower than it was at the time of purchase, the owner loses money by having to pay more than the house is currently worth. Thus, the owner is technically permitted to walk away from paying the mortgage. The U.S. does not allow banks to seize assets in order to obtain repayment of loans.

As housing prices began to fall, buyers preferred to renege on payments of less-valued houses than to make an effort to pay for them. Thus, the value of the housing industry and its stocks declined further. Figure I shows the median and average sales prices of new homes sold in United States from 1963 until 2009. Whereas the median and average home prices were $245,300 and $301,200 in February 2008, the corresponding figures for July 2010 are $204,000 and $235,300, respectively. These figures represent a drop of about 17 percent in the median and a dive of about 22 percent in the price of the average house sold in the United States.

Figure II presents annual data of the number of housing units in thousands for sale in the United States and in its sub-regions from 1975 until 2009. Whereas in the year 2006 the numbers were 537, 54, 97, 267, and 119 for the U.S., North-East, Mid-West, South and West, respectively, these numbers dropped in 2009 to 232, 27, 38, 118, and 48. Percentage-wise these figures represent a drop of about 57 percent for the United States as a whole and decreases of 50, 61, 56 and 60 percent, for the North-East, Mid-West, South and West, respectively. Figure III presents annual data of the lots sold for future construction of housing units in the United States in what are called "permit-issuing areas that will never have a permit authorization." This figure is useful because it indicates expectations for future development, since at the time of purchase people do not have permits to build on these lots. According to these data, whereas in 2005 the figures for the United States, the Northeast, the Midwest, the South and the West regions were 1,283, 81, 205, 638 and 358, respectively, these figures dropped in 2009 to 375, 31, 54, 202, and 87. Figure IV shows the current state of the U.S. housing market. The regions in the blue are the housing markets that were the most stable in August 2010. However, even in these "stable" locations, prices are dramatically lower than would be expected. For some houses, prices are currently more than 80% lower than they were in 2006.

The result of falling sales and prices of houses was manifested in devaluation of companies with mortgage-backed securities. Bear Sterns was taken over by J.P. Morgan Chase in 2008. Later that same year, Lehman Brothers filed for bankruptcy. Also in 2008, Fannie Mae and Freddie Mac, both companies that sold many risky deals and assets, were essentially taken over by the U.S. government. By October 2010, Fannie Mae and Freddie Mac have cost taxpayers an estimated $400 billion, and there is currently no path to resolution. Merrill Lynch was sold at a price far below its market value. Goldman Sachs and Morgan Stanley, though avoiding failure, yielded to more government regulation (Labaton, 2008).

The sudden decline of large, well-known financial entities led to general concern and panic. As in other financial crises, bank runs began, and there was a sharp drop in investor confidence. Companies and industries that relied on credit from the financial services industry began to suffer as well. For example, the purchase of automobiles generally requires substantial credit, and consequently General Motors and Chrysler required government assistance to be bailed out in 2009. As companies tried to cut back on spending, unemployment rose and general spending declined, bringing the United States to a full-fledged recession.

Figure V shows the unemployment rate for U.S. citizens 16 years of age and over since 1948. A period of unemployment similar to the present state occurred in 1982-1983. For 18 months, from March 1982 to November 1983, the unemployment rate was above 9 percent. The U.S. actually had fifty-six months of unemployment at 7 percent or above, lasting from May 1980 until January 1986. In this most recent recession, the U.S. unemployment rate has been above 9 percent for 16 months from May 2009 through August 2010, and above 7 percent for 21 months since December 2008 to August 2010. History again provides evidence of the difficulties in reducing unemployment. Recently released data from the U.S. Labor Department indicates that unemployment from month-to-month and year-to-year is still stagnant and remains high. Eleven of the fifty states showed no change in unemployment, and twenty-seven states showed only a slight increase (Izzo, 2010).

Such long-term high unemployment rates are intimately related to both demand and supply of labor. Unemployed laborers lose skills, making the process of rehiring slow and difficult. The effects on general standard of living are also detrimental. Figure VI and Table 1 show the poverty rate from 1959 to 2009. The poverty rate is currently at 14.3 percent - the highest it has been in fifteen years. The number is even higher among certain minorities. For example, more than one-fourth of both the African American and Hispanic American populations are estimated to be living “in poverty.”

There have been significant government attempts to quell the recession beginning in 2008. During the Bush Administration, the U.S. government purchased troubled assets that were causing lack of funds. These policies, however, were not very successful. The Obama Administration introduced large stimulus packages, in order to increase employment and spending. However, the effect of the stimulus has not yet been particularly pronounced. In fact, some argue that the Keynesian multiplier, which determines how effective government spending is in combating recession, is much lower than it was in previous decade (Cogan, Cwik, Taylor, and Wieland, 2009). Significant public spending, monetarists argue, contributes to national debt more than it supports the economy.

As Figure VII and Table 2 show, federal outlays as a percent of GDP are currently very high—about 24.7 percent of GDP can be attributed to government spending. The only time the government spending comprised a larger percent of GDP was during WWII (41.88 percent in 1945). The comparable number for 1920, two years after the end of World War I was only 7.33 percent. Even during the American Civil War of 1861–1865, the number peaked at a mere 14.2 percent. Meanwhile, the Federal Reserve Bank is currently cutting and maintaining low interest rates. While there are concerns that inflation may result from the latter action, proponents claim that interest rates can and will be stabilized once the economy begins to recover.

What else can be done to help the U.S. out of the current recession? Shultz, Boskin, Cogan, Meltzer, and Taylor (2010) claim that it was deviating from sound economic principles which caused the recession. Policies, they argue, should take into account incentives and disincentives. Tax cuts are effective, but they only evince the intended incentives when they are permanent. Government intervention should be less than it is at the moment, given the current U.S. deficit and large federal budget, and the potentially stifling crowding out effect that it has on industries. To prevent future panic, the Fed should have a set of concrete rules to follow in hypothetical situations. Rules which are simple and transparent, rather than discretionary economic policies, should be the norm. Any discretionary government interventions should be discouraged.

Additionally, in an increasingly globalized world, it is important to consider the role of world economies in the onset and the alleviation of recessions. There is implicit competition among countries in terms of regulation. Economies with fewer regulations inherently tend to attract more investment and resources.

FINANCIAL CRISES IN POLAND

Poland experienced a major crisis in 1991. The Bank for Food Economy and the cooperative banking system, together with seven of the nine state-owned banks, which controlled ninety percent of the national credit market, all experienced solvency problems. It is estimated that the losses in the crisis amounted to seven percent of the Polish GDP (Klytchnikova, 2000).

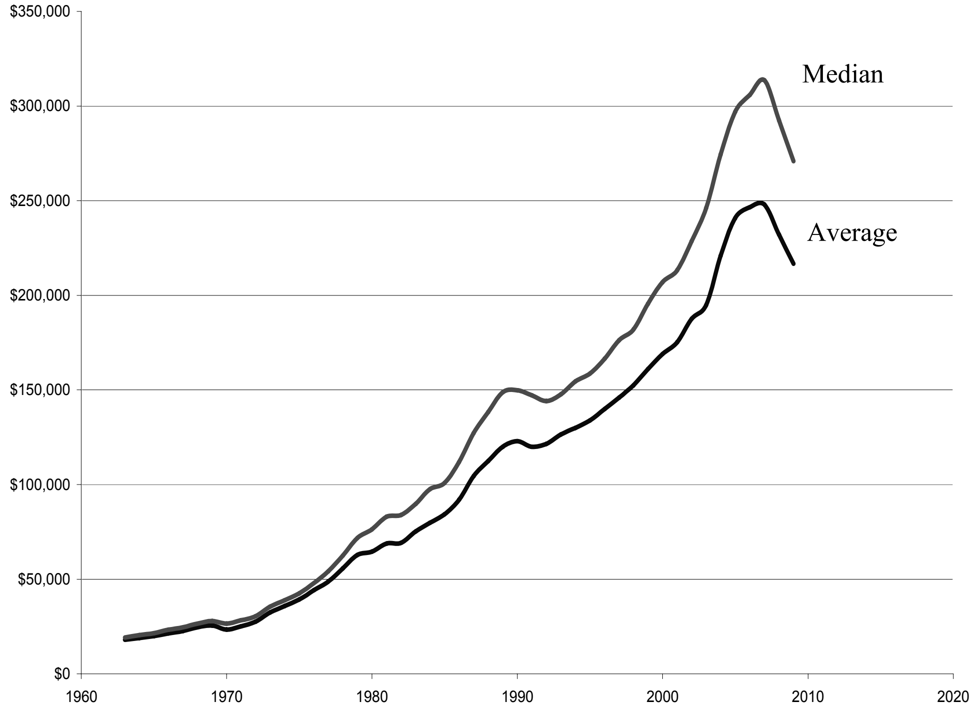
Despite the banking crisis in the early 90s, Poland has fared well in the most recent crisis. As the International Monetary Fund 9IMF0 Managing Director Dominique Strauss-Kahn observes, Poland employs sound macroeconomic policies and financial management and thus has remained relatively unscathed. With Poland being one of the few European Union countries to register positive economic growth for 2009, his claims are not unfounded. Table 3 shows that Poland has a higher GDP growth rate compared with the average figure for the emerging markets in eastern and central Europe. Figure VIII shows the Warsaw Stock Exchange Index (WIG) from 1991 to 2009, as well as a narrower view to provide detail on recent fluctuations.

CONCLUSION

Economic recession can often trace its roots back to unregulated expansion, as shown in many crises in the last three centuries. Almost every past panic has begun with rapid expansion in some industry. The rapidly expanding industry often falls into the pattern of acquiring assets and funding from new or risky financial institutions and instruments. The expanding bubble ultimately bursts when the growth slows. Millions lose money and assets, and the resulting failure of banks, companies and even government entities leads to panic, deflation and unemployment.

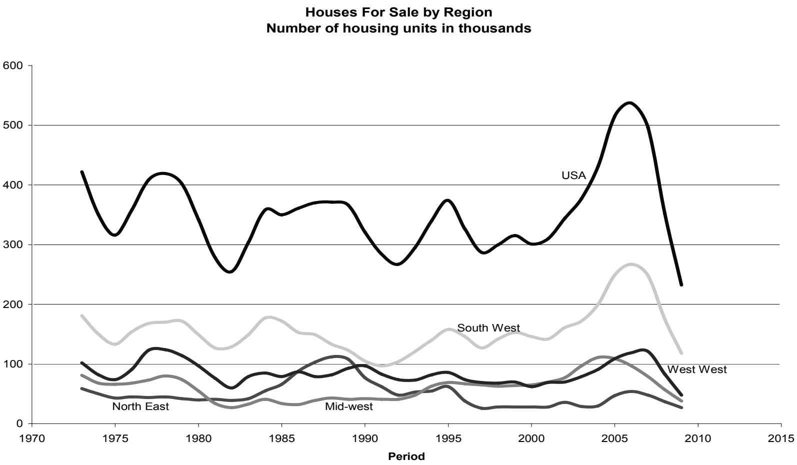
If expansion of an industry were to be kept in check, so that bubbles and artificial growth could be avoided, perhaps panics would be less pronounced. However, examining past crises shows one very important pattern. No matter what regulations a government attempts to impose, people find ways to circumvent them. The boom-bust cycle that is typical of a capitalist economy is seemingly unavoidable. Further regulations are no panacea because regulators systematically fail to foresee the extent of human ingenuity in circumventing their restrictions.

Figure I
The Median and Average Sales Prices of New Homes Sold in United States
Note that the median is above the average price



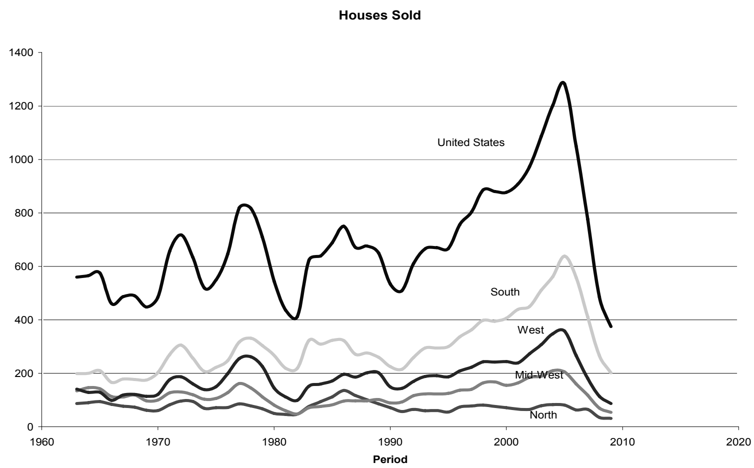
Source : www.census.gov/const/uspricemon.pdf

Figure II
Houses for Sale by Region



Source: www.census.gov/const/fsalmon.pdf

Figure III
The Annual Data of the Number of Housing Units in Thousands of Houses Sold in the United States by Region in Permit-issuing Areas That Will Never Have a Permit Authorization

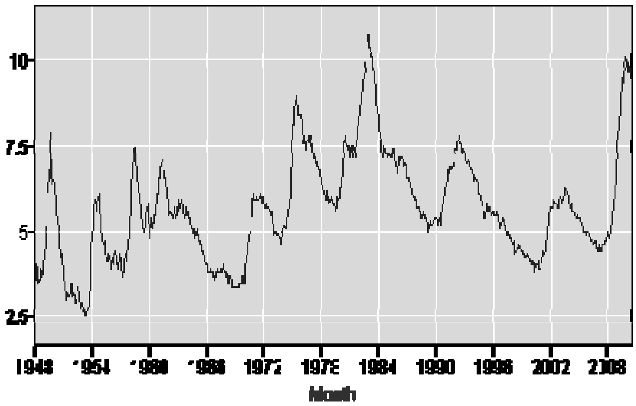


Source: www.census.gov/const/soldann.pdf

Figure IV
The Best and Worst Markets in the U.S. for Single-Family Real-Estate Investment Property

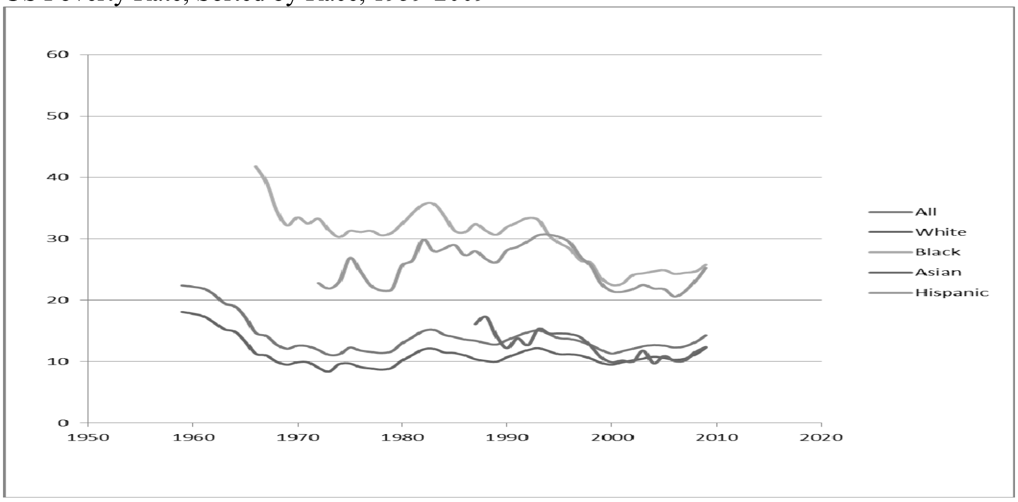
Source: Wall Street Journal-McQueen, M.P, August 2010

Figure V
Unemployment Rate, 16+ Years of Age, 1948-Present



Source: Bureau of Labor Statistics (BLS), 2010

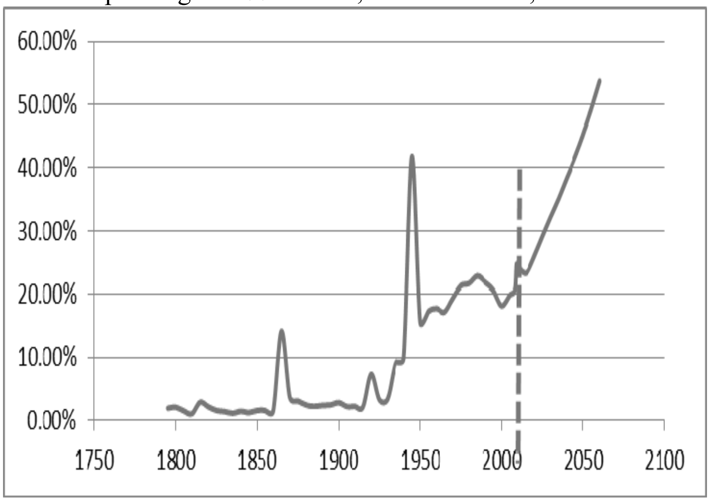
Figure VI
US Poverty Rate, Sorted by Race, 1959-2009



Source: DeNavas-Walt Carmen, Bernadette D. Proctor and Jessica C. Smith (2010) U.S. Census Bureau, September 2010; Table B1

Figure VII

Federal Spending as a % of GDP, 1795 – Present; Present – 2050 (Projected)



Source: Wall Street Journal- Shultz, Taylor, Cogan, Meltzer, 2010

Figure VIII
Warsaw Stock Exchange Index (WIG) 1991 – Present

Source : http://www.bloomberg.com/markets/

Figure VIII
Warsaw Stock Exchange Index (WIG) 1991 – Present (continued)

This more close up view shows that though there was a rather consistent increase in the index until 2007, from July 2007 onwards, the index has been declining. The index is currently at 44249 – a 35% decline from its peak of 67568

Source : http://www.bloomberg.com/markets/

Table 1
US Poverty Rate (%), Sorted by Race, 1959 – 2009

Year	All	White	Black	Asian	Hispanic
2009	14.3	12.3	25.8	12.4	25.3
2008	13.2	11.2	24.7	11.6	23.2
2007	12.5	10.5	24.5	10.2	21.5
2006	12.3	10.3	24.3	10.1	20.6
2005	12.6	10.6	24.9	10.9	21.8
2004	12.7	10.8	24.7	9.7	21.9
2003	12.5	10.5	24.4	11.8	22.5
2002	12.1	10.2	24.1	10	21.8
2001	11.7	9.9	22.7	10.2	21.4
2000	11.3	9.5	22.5	9.9	21.5
1999	11.9	9.8	23.6	10.7	22.7
1998	12.7	10.5	26.1	12.5	25.6
1997	13.3	11	26.5	14	27.1
1996	13.7	11.2	28.4	14.5	29.4
1995	13.8	11.2	29.3	14.6	30.3
1994	14.5	11.7	30.6	14.6	30.7
1993	15.1	12.2	33.1	15.3	30.6
1992	14.8	11.9	33.4	12.7	29.6
1991	14.2	11.3	32.7	13.8	28.7
1990	13.5	10.7	31.9	12.2	28.1
1989	12.8	10	30.7	14.1	26.2
1988	13	10.1	31.3	17.3	26.7
1987	13.4	10.4	32.4	16.1	28
1986	13.6	11	31.1		27.3
1985	14	11.4	31.3		29

Source: DeNavas-Walt Carmen, Bernadette D. Proctor and Jessica C. Smith (2010) U.S. Census Bureau, September 2010; Table B1

Table 1
US Poverty Rate (%), Sorted by Race, 1959 – 2009 (continued)

Year	All	White	Black	Asian	Hispanic
1984	14.4	11.5	33.8		28.4
1983	15.2	12.1	35.7		28
1982	15	12	35.6		29.9
1981	14	11.1	34.2		26.5
1980	13	10.2	32.5		25.7
1979	11.7	9	31		21.8
1978	11.4	8.7	30.6		21.6
1977	11.6	8.9	31.3		22.4
1976	11.8	9.1	31.1		24.7
1975	12.3	9.7	31.3		26.9
1974	11.2	9.6	30.3		23
1973	11.1	8.4	31.4		21.9
1972	11.9	9	33.3		22.8
1971	12.5	9.9	32.5		
1970	12.6	9.9	33.5		
1969	12.1	9.5			
1968	12.8	10	32.2		
1967	14.2	11	34.7		
1966	14.7	11.3	39.3		
1965	17.3	13.3	41.8		
1964	19	14.9			
1963	19.5	15.3			
1962	21	16.4			
1961	21.9	17.4			
1960	22.2	17.8			
1959	22.4	18.1			
			55.1		

Source: DeNavas-Walt Carmen, Bernadette D. Proctor and Jessica C. Smith (2010) U.S. Census Bureau, September 2010; Table B1

Table 2
Federal Spending as a % of GDP, 1795 – Present ; *Present – 2050 (projected)*

Year	Total Outlays	Year	Total Outlays
1795	1.90%	1935	9.21%
1800	2.07%	1940	9.78%
1805	1.50%	1945	41.88%
1810	1.01%	1950	15.59%
1815	2.84%	1955	17.35%
1820	2.12%	1960	17.80%
1825	1.52%	1965	17.21%
1830	1.39%	1970	19.33%
1835	1.10%	1975	21.29%
1840	1.41%	1980	21.67%
1845	1.21%	1985	22.85%
1850	1.56%	1990	21.85%
1855	1.54%	1995	20.60%
1860	1.41%	2000	18.20%
1865	14.20%	2005	19.90%
1870	3.65%	2008	20.66%
1875	3.04%	2009	24.71%
1880	2.40%	*2010*	*24.30%*
1885	2.21%	*2015*	*23.30%*
1890	2.35%	*2020*	*25.90%*
1895	2.43%	*2025*	*29.10%*
1900	2.79%	*2030*	*32.20%*
1905	2.16%	*2035*	*35.20%*
1910	2.22%	*2040*	*38.50%*
1915	2.06%	*2045*	*41.70%*
1920	7.33%	*2050*	*45.30%*
1925	3.22%	*2055*	*49.30%*
1930	3.41%	*2060*	*53.70%*

Source: Wall Street Journal- Shultz, Taylor, Cogan, Meltzer, 2010

Table 3
Selected Emerging European Economies : Real GDP, Consumer Prices, and Current

(Annual percent change, unless noted otherwise)

	Real GDP				Consumer Prices[1]				Current Account Balance[2]			
	2007	2008	2009	2010	2007	2008	2009	2010	2007	2008	2009	2010
Emerging Europe	**5.4**	**2.9**	**–3.7**	**0.8**	**6.2**	**8.0**	**4.7**	**4.2**	**–7.7**	**–7.6**	**–3.9**	**–3.4**
Turkey	4.7	1.1	–5.1	1.5	8.8	10.4	6.9	6.8	–5.8	–5.7	–1.2	–1.6
Excluding Turkey	5.9	4.1	–2.9	0.3	4.5	6.5	3.3	2.5	–9.0	–8.8	–5.6	–4.4
Baltics	**8.7**	**–0.7**	**–10.6**	**–2.3**	**7.3**	**12.2**	**3.6**	**–1.0**	**–18.0**	**–11.6**	**–5.4**	**–5.4**
Estonia	6.3	–3.6	–10.0	–1.0	6.6	10.4	0.8	–1.3	–18.1	–9.2	–6.5	–5.4
Latvia	10.0	–4.6	–12.0	–2.0	10.1	15.3	3.3	–3.5	–22.6	–13.2	–6.7	–5.5
Lithuania	8.9	3.0	–10.0	–3.0	5.8	11.1	5.1	0.6	–14.6	–11.6	–4.0	–5.3
Central Europe	**5.4**	**3.8**	**–1.3**	**0.9**	**3.7**	**4.6**	**2.4**	**2.6**	**–5.2**	**–6.1**	**–4.3**	**–3.8**
Hungary	1.1	0.6	–3.3	–0.4	7.9	6.1	3.8	2.8	–6.4	–7.8	–3.9	–3.4
Poland	6.7	4.8	–0.7	1.3	2.5	4.2	2.1	2.6	–4.7	–5.5	–4.5	–3.9
Southern and south-eastern Europe	**6.1**	**6.1**	**–3.6**	**–0.2**	**5.1**	**8.4**	**4.9**	**3.2**	**–14.2**	**–13.8**	**–8.2**	**–5.5**
Bulgaria	6.2	6.0	–2.0	–1.0	7.6	12.0	3.7	1.3	–25.1	–24.4	–12.3	–3.6
Croatia	5.5	2.4	–3.5	0.3	2.9	6.1	2.5	2.8	–7.6	–9.4	–6.5	–4.1
Romania	6.2	7.1	–4.1	0.0	4.8	7.8	5.9	3.9	–13.9	–12.6	–7.5	–6.5
Memorandum												
Slovak Republic	10.4	6.4	–2.1	1.9	1.9	3.9	1.7	2.3	–5.4	–6.3	–5.7	–5.0
Czech Republic	6.0	3.2	–3.5	0.1	2.9	6.3	1.0	1.6	–3.2	–3.1	–2.7	–3.0

[1]Movements in consumer prices are shown as annual averages. December/December changes can be found in Table A7 in the Statistical Appendix.
[2]Percent of GDP.

Source: World Economic Outlook, 2009.

REFERENCES

Blair,Tony (2010). *A Journey*. Hutchinson, Random House, London, UK.

Blinder and Zandi (2010). "How the Great Recession Was Brought to an End". http://www.economy.com/mark-zandi/documents/End-of-Great-Recession.pdf. Accessed on October 11, 2010.

Boyd, John H., Gianni de Nicolò, and Elena Loukoianova (2009). "Banking Crises and Crisis Dating: Theory and Evidence." *IMF Working Papers,* pp. 1-51, July 10.

Bloomberg. http://www.bloomberg.com/apps/quote?ticker=WIG:IND. Accessed on October 11, 2010.

Braunstein, Yale (2009). "The Role of Information Failures in the Financial Meltdown", School of Information, UC Berkeley, US.

Bruner, Robert F. and Sean D. Carr (2007). *The Panic of 1907: Lessons Learned from the Market's Perfect Storm*. John Wiley & Sons, Inc., Hoboken, New Jersey, US.

Bureau of Labor Statistics (BLS) (2010), Data Series: Employment, Hours, and Earnings from the Current Employment Statistics Survey (National).

Bordo, Michael D. (2003). "Historical Perspective on Booms, Busts and Recessions: When Bubbles Burst" *IMF World Economic Outlook*, pp. 64-66. Washington, DC, US.

Calomiris, Charles W. and Schweikart, Larry (1991). "The Panic of 1857 : Origins, Transmission, and Containment." *The Journal of Economic History*, Vol. 51, No. 4, 807-834.

Caporale, Tony and Barbara McKiernan (1998) "The Fischer Black Hypothesis: Some Time-Series Evidence," Southern Economic Journal, Vol. 64, No. 3 (January), pp. 765-771.

Cogan, John, Tobias Cwik, John B. Taylor, and Volker Wieland (2009). "New Keynesian versus Old Keynesian Government Spending Multipliers." *European Central Bank Working Paper Series,* No. 1090, pp. 1-27.Frankfurt am Main, DE.

Demirgüç-Kunt, Asli and Detragiache, Enrica, (1997). "The Determinants of Banking Crises in Developing and Developed Countries." *IMF Staff papers* 45, pp. 81-109. Washington, DC.

DeNavas-Walt Carmen, Bernadette D. Proctor and Jessica C. Smith (2010), "Income, Poverty, and Health Insurance Coverage in the United States: 2009 Current Population Reports, Consumer Income", U.S. Department of Commerce Economics and Statistics Administration, U.S. Census Bureau.

Conant, Charles (1915). A History of Modern Banks of Issue: With an Account of the Economic Crises of the Nineteenth Century, and the Crisis of 1907, Fourth Edition. Revised and Enlarged. G.P. Putnam & Sons, New York City, NY, US.

Federal Reserve Bank of Minneapolis Quarterly Review Vol. 24, No. 2. 2000.

Furceri, Davide and Aleksandra Zdzienicka (2010). "The Real Effect of Financial Crises in the European Transition Economies," *Economics of Transition*, August, pp. 1-25.

García-Herrero, Alicia and Pedro Del Rio-Lopez (2003). "Financial Stability and the Design of Monetary Policy," *Banco de España Working Paper,* No. 0315. Madrid, ES.

Hoffmann, Charles (1970). *The Depression of the Nineties: An Economic History*. Greenwood Publishing, Westport, CT, U.S.

Holton, James (1990). *The Reading Railroad: History of a Coal Age Empire, Vol. I: The Nineteenth Century*, Garrigues House Publishing, Laury's Station, PA, US.

International Monetary Fund (IMF): World Economic Outlook (WEO), April 2009. http://www.imf.org/external/pubs/ft/weo/2009/01/pdf/text.pdf. Accessed on October 11, 2010.

Izzo, Phil (2010), "August Unemployment Rates by State: Stagnant Labor Market", *The Wall Street Journal*. 21 September.

Kaminsky, Graciela L. and Carmen Reinhart (1999). "The Twin Crises: the Causes of Banking and Balance-of-Payments Problems." *The American Economic Review*, Volume 89, Issue 3 (June), pp. 473-500.

Klomp, Jeroen, 2010, “Causes of Banking Crises Revisited,” *The North American Journal of Economics and Finance*, Volume 21, Issue 1 (March), pp. 72-87.

Klein, Lawrence R. (1947). *The Keynesian Revolution*. Macmillan, New York, NY, US.

Krugman, Paul (2009). *The Return of Depression Economics and the Crisis of 2008, with a New Epilogue*. W. W. Norton & Company, New York, NY, US.

Labaton, Stephen (2008). “Agency's ’04 Rule Let Banks Pile Up New Debt”. *The New York Times*, 2 October.

Mason, David L. (2001). “From Building and Loans to Bail-Outs: A History of the American Savings and Loan Industry, 1831-1989”. Ph.D dissertation, Ohio State University, US.

McQueen, M.P. (2010). “Real Estate Investing : the Best and Worst Markets”, *The Wall Street Journal*. 21 August.

Oberholtzer, Ellis P. (1926). *A History of the United States Since the Civil War*. Macmillan, New York, NY, US.

Olsen, Henry (2010). “Unemployment: What Would Reagan Do?” *The Wall Street Journal*. 10 August.

Reinhart, Carmen M. and Kenneth S. Rogoff (2009). *This Time is Different: Eight Centuries of Financial Folly*. Princeton University Press, Princeton, NJ, US.

Reinhart, Carmen M. and Kenneth S. Rogoff (2009B). “The Aftermath of Financial Crises,” *American Economic Review*, Volume 99 (May), pp. 466-472.

Rolnick, Arthur J., Bruce D. Smith and Warren E. Weber (1998). “The Suffolk Bank and the Panic of 1837: How a Private Bank Acted as a Lender-of-Last-Resort.” Federal Reserve Bank of Minneapolis Research Department, US.

Rosenof, Theodore (1997). *Economics in the Long Run: New Deal Theorists and Their Legacies, 1933-1993*. University of North Carolina Press, Chapel Hill, NC, US.

Rothbard, Murray N. (1962). *The Panic of 1819: Reactions and Policies*. Columbia University Press, New York, NY, US.

Rousseau, Peter L. (2002). “Jacksonian Monetary Policy, Specie Flows, and the Panic of 1837”. *The Journal of Economic History*, Vol. 62, No. 2 (June), pp. 457-488.

Stiglitz, Joseph (2010) “Needed: A New Economic Paradigm,” *Financial Times.* 19 August.

Strunk, Norman and Fred Case (1988). *Where Deregulation Went Wrong: A Look at the Causes Behind Savings and Loan Failures in the 1980s*. United States League of Savings Institutions, Chicago, IL, US.

Shultz, George P., Michael J. Boskin, John F. Cogan, Allan Meltzer, and John B. Taylor (2010). “Principles for Economic Revival,” *The Wall Street Journal*. 16 September.

Tang, Helena, Edda Zoli and Irina Klytchnikova (2000). “Banking Crises in Transition Countries: Fiscal Costs and Related Issues”, *World Bank Policy Research Working Paper Series,* 30 November. Washington, DC, US.

Timberlake, Richard (1997). "Panic of 1893" in *Business Cycles and Depressions: An Encyclopedia*, ed. David Glasner. Garland, New York, NY, US.

"Timeline: A Selected Wall Street Chronology." Public Broadcasting Service. Retrieved 30 September, 2008.

United States Census Bureau, Data Series: Houses for Sale by Region and Months' Supply at Current Sales Rate. http://www.census.gov/const/fsalmon.pdf. Accessed on October 11, 2010.

United States Census Bureau, Data Series: Houses Sold by Region. http://www.census.gov/const/soldann.pdf. Accessed on October 11, 2010.

United States Census Bureau, Data Series: Median and Average Sales Prices of New Homes Sold in United States. http://www.census.gov/const/uspricemon.pdf. Accessed on October 11, 2010.

Wheeler, Keith (1973). *The Railroaders (Old West)*. Time Life Education, New York, NY, US.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XI, No. 1, 2010, pp. 24-36

MINIMIZING CARBON FOOTPRINT OF BIOMASS ENERGY SUPPLY CHAIN IN THE PROVINCE OF FLORENCE

Iacopo Bernetti, Christian Ciampi, Sandro Sacchelli
Department of Agricultural and Forest Economics, Engineering, Sciences and Technologies – University of Florence, Italy
e-mails: iacopo.bernetti@unifi.it, christian.ciampi@unifi.it, sandro.sacchelli@unifi.it

Abstract: The paper presents an approach for optimal planning of biomass energy system based on carbon footprint minimization. A geographical spatial demand driven approach is applied to assess the feasible ways for transferring energy from renewable sources to district heating plants in the Province of Florence (Italy). The proposed approach has been developed on three levels. In the first one, the Province of Florence is partitioned into a number of Regional Energy Cluster (REC) using a multidimensional algorithm of regionalization called SKATER. The variables used in SKATER model are related in order to realize sustainable policy for forest and agriculture biomass productions. In the second step a geographical fuzzy multiple attribute decision making model was applied to the selection of biomass district heating localization. Finally, in the third step a geo-referenced Mixed Integer Linear Programming model based on resource-supply-demand structure for carbon-minimization energy planning has been applied.

Keywords: carbon footprint, biomass, MILP, fuzzy MADM, regionalization, spatial analysis, GIS.

INTRODUCTION

Biomass is one of the key renewable energy sources. Using biomass to generate heat reduces emissions of greenhouse gases compared to the emissions of fossil fuel use. However, the dispersed nature of biomass resource involves complex transportation (and also environmental) problems within the supply chain. As a result, the environmental efficiency of a regional energy supply chain has the characteristics of spatial and geographical planning problem.

Several indicators for evaluating environmental impact of biomass supply process have been proposed. Carbon FootPrint (CFP) is defined as the total amount of CO_2 emitted over the full life cycle of biomass production process (Lam et al. 2010a). The CFP of a biomass supply chain is the total CO_2 amount emitted throughout the supply chain life cycle (Perry et al., 2008). Energy supplied from biomass cannot be considered truly carbon-neutral even though the direct carbon emissions from combustion have been offset by carbon fixation during feedstock photosynthesis (Anderson and Fergusson, 2006). The net CFP is mainly caused by the indirect carbon emissions generated along the supply chain, especially by harvesting, transportation and burning which release emissions. Especially transportation activities could contribute to the greater part of the CFP in the supply chain (Forsberg, 2000). The typical locations of biomass sources (farms, forest, etc.), the relatively low energy density and the distributed nature of the sources require extensive infrastructures and huge transport capacities for implementing the biomass supply networks. A solution to this problem is the utilization of biomass for heating houses, because this demand source is also dispersed in territory. District heating (less commonly called tele-heating) is an efficient system for distributing heat generated in a centralized location for residential and commercial heating requirements.

The paper presents an approach for optimal planning of biomass energy system based on CFP minimization in district-heating planning. A geographical spatial demand driven approach is applied to assess the feasible ways for transferring energy from renewable sources to district heating plants in the Province of Florence (Italy).

Section 2 presents the case study, Section 3 describes the methodological approach concerning how to minimize CFP in district heating planning at regional level. In Section 4 results are discussed.

THE CASE STUDY: DESCRIPTION OF THE PROVINCE OF FLORENCE

The territory of the Province of Florence covers an area of 3,514 km^2, with a population of 985,273 inhabitants (ISTAT, 2008) and a density of population of about 280 inhabitants/km^2, which sensitively increases to 1,644 inhabitants/km^2 in the urban area. The total surface covered by urban area is 183 km^2. The settlement morphology varies from the conurbation formed by Florence and by the others towns of the plain, with a higher urbanization, to the system of the small historical villages, farms, villas and parishes which are the evidence of the strong relationships which, in the past, described the rural organization of the countryside of Florence and, nowadays, "outline" the Chianti landscape.

The agro-forest environment of the Province is characterized by 1,640 km^2 of forest area, mainly covered by deciduous broad-leaf. The presence of agricultural land cultivated with permanent crops is relevant mainly in the hills surrounding the town of Florence, where the olive growing areas reach 278 km^2, while in the Florentine Chianti the vineyards reach 166 km^2.

METHODOLOGY

Biomass to energy projects are highly geographical dependent and the supply chain efficiency can be strongly influenced by location of district-heating, particularly in terms of CFP minimizations. The key element is to obtain sufficient biomass quantities in order to satisfy the energy plant demand, at least the carbon emission cost paid for transportation.

The problem of choosing the best locations for energy facilities is commonly assessed using a specific application based on the geographic informative system (GIS). Noon and Daly (1996), Bernetti et al. (2004) and Noon et al. (2002) used GIS, in order to identify the sites for only one facility, but without assessing site competition. However, when resources in the region are scarce, the district heating plants have to compete in order to meet their own demand. This leads to a location-allocation problem. Nord-Larsen and Talbot (2004) and Ranta (2005) applied linear programming in order to minimize transportation cost to detected units. Finally Panichelli and Gnansounou (2008) used a GIS based linear programming approach for selecting least cost bio-energy location.

Minimizing CFP in district heating issues requires the optimization both of demand location and of biomass network transfer links by the coupling of linear programming model and Geographic Information System. The main problem of this methodology regards the size of the model that, for a global optimization, can rise also to about several trillion of decision variables.

To face up to this problem Lam et al. (2010b) proposed a Regional Energy Clustering approach. The region under consideration is modeled as a collection of zones. The zones are smaller areas within the region, accounting for administrative or economic boundaries, which are considered atomic (i.e. non-divisible). For each zone corresponding rates of energy demands and biomass resource availability are specified.

In this paper, in order to obtain an environmentally efficient local optimization, the proposed approach has been developed on three levels.

In the first level, the region is partitioned into a number of Regional Energy Cluster (REC) using a multidimensional algorithm of regionalization called SKATER. The variables used in SKATER model are related in order to realize sustainable policy for forest and agriculture biomass productions. In the second step, a geographical fuzzy multiple attribute decision making model (FMADM) is applied to the selection of efficient biomass district heating localization. Finally, in the third step a geo-referenced Mixed Integer Linear Programming (MILP) model based on resource-supply-demand structure for carbon-minimization energy planning has been applied for each Regional Energy Cluster (Fig. 1).

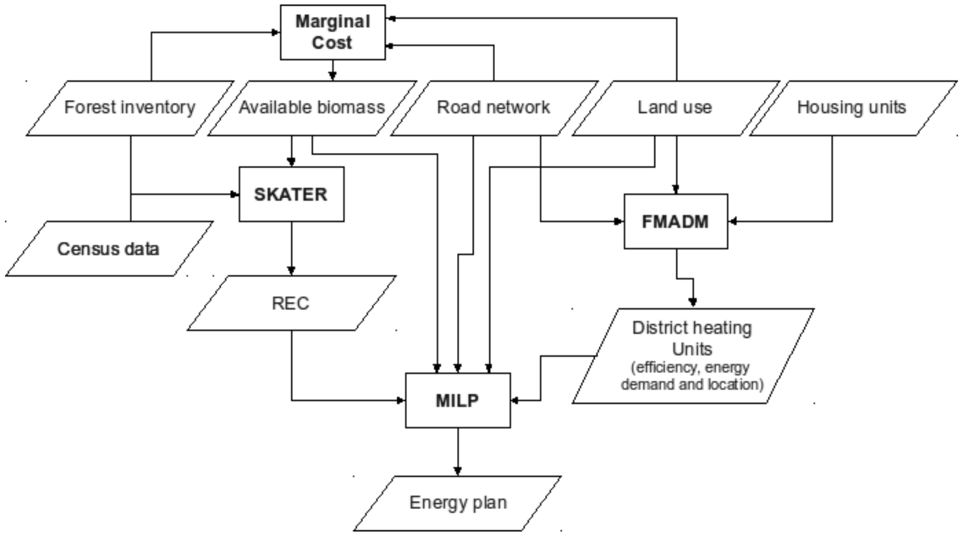


Figure 1. Methodology overview

As explained in figure 1, the GIS realized for the development of the model is composed by the following data bases: *a)* Forest Inventory of the Province of Florence, extracted by Tuscany's Inventory (vectorial format); *b)* Digital Terrain Model (raster format); *c)* Road system (vectorial format); *d)* Administrative boundaries (vectorial format); *e)* Housing units (vectorial format); *f)* Corine Land Cover land use map (raster format); *g)* Census data on Industry and Services (numerical format).

The REC step

In our approach the cluster combines smaller zone to assure efficient biomass energy policy. The zones are the smallest administrative units within the Province of Florence. Regionalization is to divide this large set of zones into a number of spatially contiguous regions while optimizing an objective function, which is a homogeneity intra-zones/heterogeneity inter-zones measure of the derived regions (Gou, 2008). In our model homogeneity/heterogeneity is related to indicators of sustainability of biomass harvest and supply chain logistics that are the characteristics of the agro-forest land and the potential supply of agro-energies and traditional timber assortments for each of the minimum administrative units.

The used sustainability indicators are the following ones:

- *land agro-forest characteristics,* defined through the quantification of the percentages of arable land areas, of permanent crops and forest areas;

- *biomass products,* including both the bio-fuel deriving from the agricultural crops and forest cultivations and from the traditional timber assortments (calculated with the methodology proposed by Bernetti et al., 2004);

- agricultural specialization index: it quantifies the importance of the agricultural sector, at municipal level, through the percentage of the workers involved in the same sector in relation to the total number of workers and in relation to the Region (Tuscany).

Given a set of spatial objects (e.g. administrative units) with a set of multivariate information a regionalization method aggregates the spatial objects into a number of spatially contiguous regions while optimizing an objective function, which is normally a measure of the attribute similarity in each region. With SKATER (Spatial 'K'luster Analysis by Tree Edge Removal) method a connectivity graph to capture the adjacency relations between objects is used. Figure 2 shows the connectivity graph for Province of Florence.

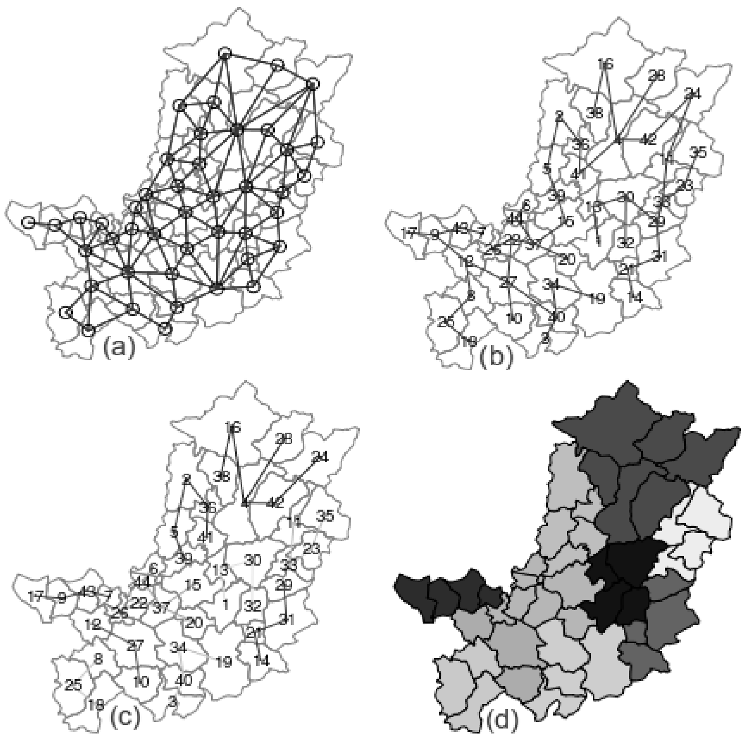


Figure 2. SKATER process

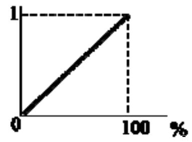
The cost of each edge is proportional to the dissimilarity between the objects it joins, where the dissimilarity is measured using the values of the attributes of the neighboring pair. It is possible to limit the complexity of the graph pruning the edge with high dissimilarity. An efficient method for pruning the graph is the so called "minimum spanning tree". A minimum spanning tree is a spanning tree with minimum cost, where the cost is measured as the sum of the dissimilarities over all the edges of the tree (Figure 2.b). By cutting the graph at suitable places, connected clusters (Figure 2.c and 2.d) are obtained (for further details see Assunçao et al. 2006).

The fuzzy FMADM step

District heating efficient location has been evaluated by the application of a model of decisional analysis based on a multi-attribute fuzzy approach (FMADM; Munda, 1995). The evaluation was done on the basis of the suitability to installing wood biomass plants for three different settlement typologies (ISTAT, 2001), which are: *i)* towns: aggregation of contiguous or nearby houses separated by roads, squares or similar; *ii)* residential complexes: aggregation of contiguous houses with the absence of squares or similar; *iii)* scattered houses: buildings scattered in the municipal territory having a distance such as neither creating a residential complex.

The indexes implemented in the model FMADM were elaborated on a GIS platform and then they were normalized with fuzzy functions (Cox, 1993; Zimmermann, 1987) on the basis of the information given by experts of the agro-energetic sector (table 1).

Table 1. Utilized Indicators in the model FMADM

Indicator	Description	Elaboration	Membership functions
Rurality Index (R)	It defines the potential easiness of wood bio-fuel supplying at a local scale.	It is calculated as the percentage of the agro-forest land with a *Density index* having a radius of 175 m around the settlement element.	
House density (Hd)	It defines the economic-logistic suitability for the creation of a district-heating plant.	It is calculated with a *Density index* having a radius from the buildings of 75 m, around the settlement element.	
Road distance (Rd)	It defines the access suitability to the plant, in terms of supplying costs of the raw material and of the CFP of the productive process.	It is calculated with a *distance* from the major and minor roads.	

The aggregation of the fuzzy criteria was done on the basis of two different techniques: one for the towns and the residential complexes and one for the scattered houses.

The following aggregation was applied for the towns and the residential complexes:

$$I^j_{tr} = \frac{1 + R^j}{2} \qquad (1)$$

with I_{tr} being the suitability for the localization of the district heating plant j within the town or the residential complex and R being the rural index.

The following formula was applied for the scattered houses:

$$I^j_{sh} = \frac{min\left(Hd^j, Rd^j\right) + R^j}{2} \qquad (2)$$

with I_{sh} being the suitability for the localization of the district heating plant j within the house cluster, Hd is the house density, Rd being road distance and R is the rural index.

The MILP step

The objective of MILP model is to minimize CFP within the boundary of each REC. The model structure is the following:

$$MIN \sum_{j=1}^{m} \sum_{i=1}^{n} t_{i,j} X_{i,j} \tag{3}$$

s.t

$$\begin{aligned} &\sum_{j=1}^{m}\sum_{i=1}^{n} p_i X_{i,j} \geq d_i Y_i \quad \forall j \\ &\sum_{j=1}^{m} X_{i,j} \leq 1 \quad \forall i,j \\ &Y_j = [0,1] \quad \forall j \end{aligned} \tag{4}$$

with: Y_j, being the district heating location; $X_{i,j}$, being the raster location of biomass source obtained from the regional forest inventory as data point representing 1 Km^2 of land and allocated to district heating j; d_j, being demand of biomass from district heating j; p_i being the supply of biomass from source i; $t_{i,j}$, being the carbon emission in full life cycle of biomass process (harvest + transportation) from source i to district heating j, calculated through a GIS cost surface procedure.

RESULTS AND DISCUSSION

The SKATER step

Figure 3 shows the results of the SKATER procedure in terms of the average value of the utilized indicators. From the analysis of the figure, it is possible to highlight that, among the ten obtained regions (R1, R2, …R10), the allocation of the agro-forest land shows greater values than the average value for the Province of Florence for: *a)* the arable land area in the R1, R5, R6, R8 and R10; *b)* the permanent crops in the R1, R3, R4, R7 and R10 and *c)* the woods in the R2, R5, R8 and R9. The availability of traditional timber assortments (other biomass) is directly linked with the wood area, while the wood splinter productive potentiality shows a tight correlation with the permanent crops areas. Thus, thanks to the clustering deriving from the SKATER analysis, it is possible to distinguish the territorial peculiarities of each region; for instance, the areas characterized by a high forest relevance (R9) and the ones in which the permanent crops are important, both in terms of cultivated area (R3 e R4) and of employment linked to the sector (R4).

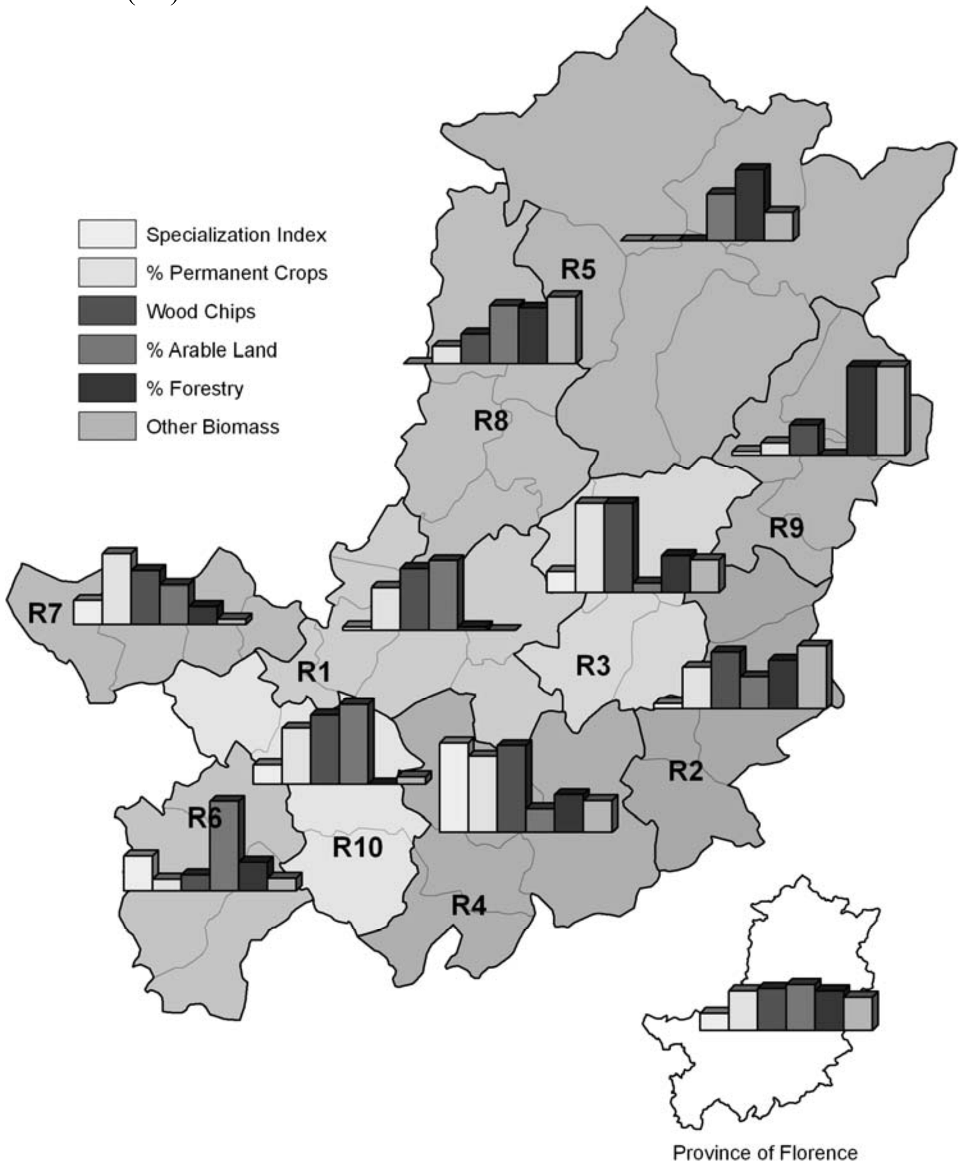


Figure 3. SKATER results

The FMADM step

The application of the FMADM model allowed the evaluation of the efficiency of the possible localizations of new plants. From the analysis of table 2, it is possible to highlight how the greater number of plants belongs to the medium-low, zero and low classes; these turn to be mainly the scattered houses, which is the least suitable settlement typology for the installation of district heating plants. For this reason, in order to optimize the agro-energetic planning process linked to the minimization of the CFP, the selected plants for the development of the MILP models are the ones with a highest possible efficiency.

Table 2. FMADM results

Efficiency class	Total plants	% scattered houses	% Towns and residential complexes	Total energetic requirement (MWh/year)	Potential installing power (MW)
0	5,887	90.8	9.2	7,876,097	4,653
0.01 - 0.25	2,441	97.5	2.5	554,842	267
0.26 - 0.5	7,454	99.5	0.5	1,245,963	619
0.51 - 0.75	33	3	97	98,656	28
0.76 - 1	187	0	100	302,612	156

In order to identify the localizations of new district-heating plants for the planning of the supply chain in every single REC, the district-heating localization in each REC has been ranked. Then, the localizations have been selected in order of decreasing efficiency until the biomass requirement (demand) came out to be equal to the annual yield (supply), considering also the sustainability of a wood ecological conservation. The results of the selection of each REC are showed in figure 4.

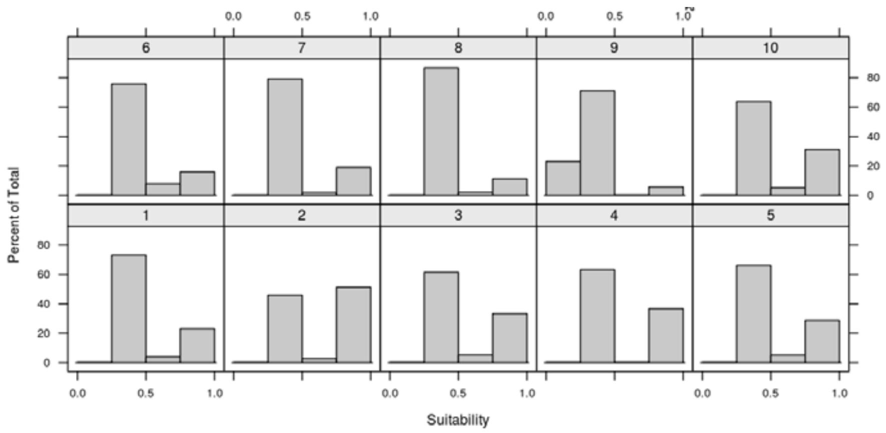
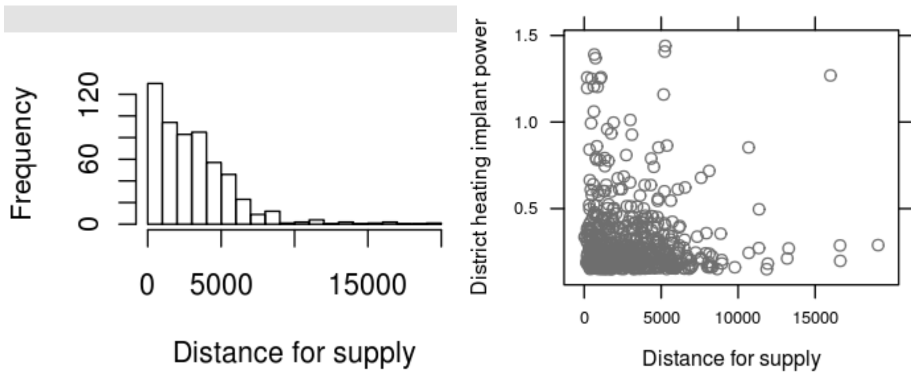


Figure 4. FMADM selection results

The MILP step

The most direct and immediate result of the adopted procedure is the possibility of planning with an high territorial degree an energetic supplying of the district heating plants in the REC, thus minimizing the environmental impact in terms of carbon dioxide emission.

By aggregating the results obtained from the 10 MILP models, it is possible to evaluate the global impact of the district-heating energy plan in the Province of Florence. As it is showed in figure 5, the optimization of the supply chain with the use of operational research methods allows the achievement of moderate CFP. In fact (see table in figure 5), the median of the transport distance is about 3 kilometers, with an interval between the first and the fourth quartile included between 2.7 and 4.5. The distribution of the frequency shows how values grater than 10 kilometers are rare. In addition, by examining the scatter plot between carbon footprint and power of district heating plant, it is possible to notice how the adopted procedure allows the achievement of limited distance of transportation values, also in the case of high power installations.



Mean Distance for supply (m)				
1st Qu.	Median	Mean	3rd Qu.	Max.
1053	2661	3124	4500	19090

Figure 5. CFP data by district-heating implant

By examining figure 6, which shows the previously explained elaboration for each REC, it is possible to notice how the lower CFP is obtained in REC 1 (Florence), 2 (Fiesole and the hills surrounding Florence), 6, 7 (hills of Vinci) and 10, corresponding to localizations characterized by a supplying based on the vineyards and olive growing pruning, which are generally close to the most suitable localizations for the rural district-heating plants. Worse results are evident for REC 5 and 9, characterized by a supplying based on forest resources often far from the towns.

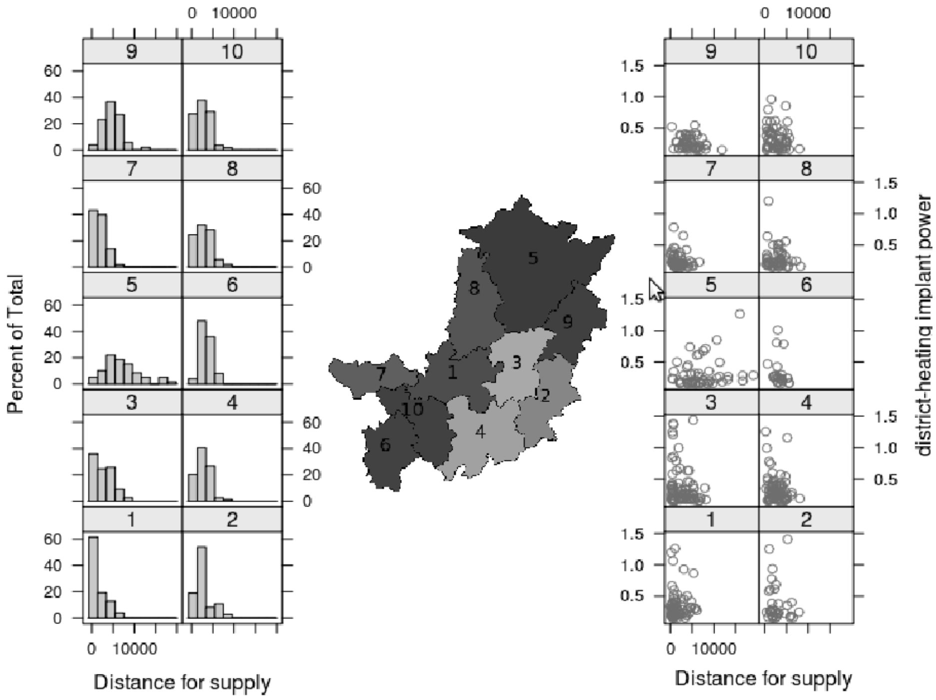


Figure 6. Carbon footprint by district-heating plant and by REC

CONCLUSIONS

When assessing simultaneous potential locations for energy facilities the location–allocation problem has to be solved in order to tackle resources competition among facilities. A multi-step procedure combined with GIS-based approach seems to be effective for selecting suitable energy facilities location. Addressing the problem through a convenient approach is fundamental to define facilities location as the optimal sites can vary in function of resources competition and environmental impact, evaluated through the carbon footprint concept.

The quantity and heterogeneity of the input data and the need for a structured analysis of the information make essential the need of integrated tools to assess the problem.

Further efforts have to be done to integrate software tools and data processing. The developed approach allows the planning of district-heating supply chain and select best energy facilities locations based on minimization of carbon footprint. Nevertheless, other models belonging to the logistic structure are being investigated, considering not only carbon footprint minimization but also environmental and social constraints.

REFERENCES

Anderson, G.Q.A., Fergusson, M.J. (2006). Energy from biomass in the UK: sources, processes and biodiversity implications. Ibis, 148, 180–183.

Assunçao, R.M., Neves, M.C., Camara, G.., Da Costa Freitas, C., (2006). Efficient regionalization techniques for socio-economic geographical units using minimum spanning trees. International Journal of Geographical Information Science, 20, 797–811.

Bernetti, I., Fagarazzi, C., Fratini, R., (2004). A methodology to analyze the potential development of biomass energy sector: an application in Tuscany. Forest Policy and Economic, 6, 415–432.

Cox, E., (1993). The fuzzy system handbook. Academic Press, London.

Forsberg, G. (2000). Biomass energy transport. Analysis of bioenergy transport chains using life cycle inventory method. Biomass and Bioenergy, 19, 17–30.

Guo, D., (2008). Regionalization with dynamically constrained agglomerative clustering and partitioning (REDCAP). International Journal of Geographical Information Science, 22, 801–823.

ISTAT (2001). VIII° Censimento sull'Industria e i Servizi. Available at: www.istat.it

ISTAT (2008). Bilancio demografico 2008. Available at: http://demo.istat.it/

Lam, H.L., Varbanov, P., Klemeš J., (2010a). Optimization of regional energy supply chains utilising renewables: P-graph approach. Computers and Chemical Engineering, 34, 782–792.

Lam, H.L., Varbanov, P., Klemeš, J., (2010b). Minimising carbon footprint of regional biomass supply chains. Resources, Conservation & Recycling, 54, 303–309.

Munda, G.., (1995). Multicriteria evaluation in a fuzzy environment. Theory and application in ecological economics. Springer–Verlag. Heidelberg.

Noon, C.E., Daly, J.M., (1996). GIS-based biomass resource assessment with BRAVO. Biomass and Bioenergy 10, 101–9.

Noon, C.E., Zhan, F.B., Graham, R.L., (2002). GIS-based analysis of marginal price variation with an application in the identification of candidate ethanol conversion plant locations. Networks and Spatial Economics, 2, 79–93.

Nord-Larsen, T., Talbot, B., (2004). Assessment of forest-fuel resources in Denmark: technical and economic availability. Biomass and Bioenergy, 27, 97–109.

Panichelli, L., Gnansounou, E., (2008). GIS-based approach for defining bioenergy facilities location: a case study in northern Spain based on marginal delivery costs and resources competition between facilities. Biomass and Bioenergy, 32, 289–300.

Perry, S., Klemeš, J., Bulatov, I., (2008). Integrating waste and renewable energy to reduce the carbon footprint of locally integrated energy sectors. Energy, 33, 1489–1497.

Ranta, T. (2005) Logging residues from regeneration fellings for biofuel production – a GIS-based availability analysis in Finland. Biomass and Bioenergy, 28, 171–182.

Zimmermann, H.J., (1987). Fuzzy sets, decision making and expert systems. Kluwer A.P., Boston.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XI, No. 1, 2010, pp. 37-44

PRICE VOLATILITY ON THE USD/JPY MARKET AS A MEASURE OF INVESTORS' ATTITUDE TOWARDS RISK

Katarzyna Banasiak
Department of Economics of Agriculture and International Economic Relations
Warsaw University of Life Sciences
e-mail: katarzyna_banasiak@sggw.pl

Abstract: The aim of the paper is to show the relationship between the value of Japanese yen and the investors' risk aversion. The correlation results from the application of carry trade strategies by investors. An increase in carry trade positions is associated with the decrease in risk aversion. The Japanese yen is one of the most popular carry trade funding currency and therefore the change in the value of this currency reflects the change in the investors' mood. This paper shows that there is a negative relationship between the USD/JPY and the risk aversion measured by volatility index (VIX).

Keywords: risk aversion, USD/JPY market, volatility index VIX

INTRODUCTION

During last years the carry trade phenomenon has started to be one of the main feature of the present financial markets. This strategy is based on borrowing in low interest yielding currency and using the funds to invest in high interest rate currencies [Fong, 2010]. In this way the investors obtain the money in the country where interest rates are low and then invest the capital in bonds, shares or commodities markets in the country with higher interest rates. This trading strategy can exhibit a favourable payoff but the risk involved in it is high. A carry traders are associated with the investors of low aversion towards risk. It is believed that the carry traders' activity on the market is going up when the level of investors' risk aversion is decreasing. Carry trade is applied by hedge funds, pension funds, investment banks, other financial institution and individual investors [Gagnon at al., 2007]. As a result the change in investors' involvement in this strategy has a huge impact on the price movements on the currency market. Japanese yen is one of the most popular carry trade funding currency and consequently the change in the value of this currency is likely to reflect the fluctuations in the investors' attitude towards risk. The growth in investors' risk aversion brings about the unwinding of carry trades [Brunnermeier at al., 2008]. It means that the investors start to withdraw their capital from the country of high interest rate and then buy Japanese yen to discharge a debt. Following, an increase in demand on the Japanese currency leads to the yen appreciation.

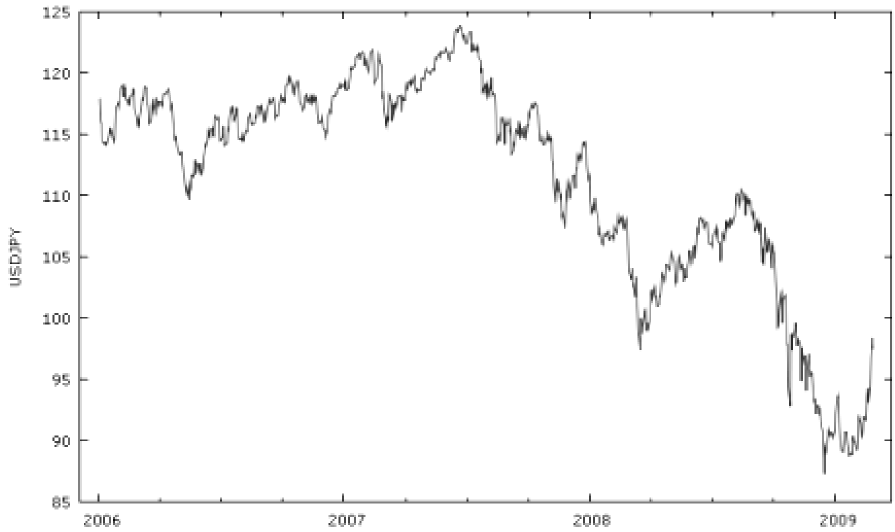
The aim of this paper is to investigate the relationship between the Japanese Yen exchange rate (USD/JPY) and the investor's risk aversion measured by the S&P500 option implied volatility index (VIX). It is shown that during the time in which the VIX decreases the Japanese yen depreciates. However, when there is unease on the market and when the risk aversion measured by VIX goes up then the Japanese currency appreciates. This articles studies the link between USD/JPY and volatility index VIX. By using the Autoregressive Conditional Heteroscedasticity (ARCH) model it provides evidence of negative relationship between the analyzed data. The data set used in research covers the period from January 2006 to February 2009.

THE IMPACT OF CARRY TRADE ON THE VALUE OF JAPANESE YEN

There is several factors which have a crucial impact on the price movement on the currency market. Some of the most important are inflation rates, interest rates, change in a country's price competitiveness, balance of payments and the economic growth of the country. The article shows that the currency exchange rate can be also significantly influenced by a change in the level of investor's risk aversion. The driving force behind it is the implementation of carry trade strategies by the investors. This paper is focused on the Japanese yen which is a prominent funding currency in carry trade. Japanese Yen is a funding currency for overseas investments because of prolonged low-interest rate policy of the Bank of Japan. In order to describe the relationship between value of the yen and the level of investors' risk aversion the author scrutinizes the USD/JPY market volatility and the change in investors' risk aversion measured by volatility index VIX.

Before the financial crisis of 21st century Japan enjoyed rapid economic growth. However, the Japanese currency was constantly depreciating, which means that the rate USD/JPY was going up. Between May 2006 and July 2007 the level of USD/JPY went up from 109,67 to 123,86 (Graph 1). It means that the yen depreciated 12,5 per cent against the U.S. dollar. There was several factors which had a crucial impact on the yen depreciation. For instance, the yen decrease in value was caused by the reduction in the share of yen-denominated assets held by the central banks. However, undoubtedly the depreciation of Japanese yen was exacerbated by the carry trade. The high investors' involvement in the carry trade led to an outflow of speculative investment from Japan [Winters, 2008]. Moreover, the carry traders activities contributed to the significant increase in supply of yen which also had a profound impact on the yen depreciation. The situation has changed when the sub-prime mortgage crisis began.

Graph 1. USD/JPY market between January 2006 and February 2009



Source: data - Reuters Information Agency

The collapse of U.S. sub-prime mortgage market had a ripple effect around the world. The global financial crisis of 21st century had a negative impact on the stock market. Many financial institution collapsed or were bought up. Because of the financial crisis of 21st century the market participants became more averse to risk. The carry traders simultaneously started to withdraw their funds and then buy Japanese yen to pay off a debt. It brought about the impressive increase of demand on the yen. Between July 2007 and February 2009 the level of USD/JPY went down from 123,86 to 87,32 (Graph 1). The Japanese currency became stronger although there were no sufficient fundamental reasons for it. It indicates that the carry trade accounts for the important factor which influences the value of Japanese yen.

THE RELATIONSHIP BETWEEN THE LEVEL OF INVESTORS' RISK AVERSION AND THE USD/JPY EXCHANGE RATE

The investors' risk aversion can be measured by the Chicago Board Options Exchange (CBOE) S&P 500 options implied volatility index (VIX) [Coudert at al., 2008]. It reflects the investors' expectations on future market volatility. The VIX value greater than 30 is associated with a high risk aversion among the market participants. The table below provides descriptive statistics of VIX before and during the financial crisis. The calculation is based on daily data and covers the period from January 2006 to February 2009.

Table 1. Descriptive statistics of VIX before and during the financial crisis of 21st century

	01.2006-07.2007	08.2007-02.2009
mean	13,13	31,24
standard deviation	2,44	14,57
maximum	24,17	80,86
minimum	9,89	16,12
kurtosis	2,49	0,93
skewness	1,41	1,39
coefficient of variation	0,19	0,47

Source: data - Reuters Information Agency

Both mean value and the standard deviation of the VIX have increased during the second period (08.2007-02.2009). It implies that during the financial crisis of 21st century the volatility on the market have increased significantly. Moreover, one may presume that between 08.2007 and 02.2009 the investors' aversion to risk swelled considerably. The maximum value of VIX was 80,86 in comparison to the first period when it was just 24,27.

Based on the VIX descriptive statistics and the graphs of USDJPY one may assume that there is a positive relationship between the value of the Japanese currency and the level of investor's risk aversion. Thereby, there is negative association between USD/JPY and the VIX. When the investor's aversion towards risk is going up, the yen is appreciating which means the USD/JPY exchange rate is decreasing. Moreover, it is shown that this relationship became even stronger during the financial crisis of 21st century. The table below presents the Spearman's rank correlation coefficient between USD/JPY and the Volatility Index VIX. The calculation is based on the daily data and covers the period from January 2006 to February 2009. Spearman rank correlation coefficient is a non-parametric statistic and thus can be used when the data have violated parametric assumptions such as non-normally distributed data [Field A., 2005]. Moreover, it is a better indicator than Pearson's correlation coefficient when the relationship between two variables is non-linear. Taking into account the features of the analyzed data, the Spearman's rank correlation can be used for an initial analysis of the relationship between USD/JPY and VIX.

Table 2. The Spearman's rank correlation coefficient between USD/JPY and the VIX

	01.2006-07.2007	08.2007-02.2009
Spearman rank correlation coefficient	-0,12	-0,64
p-value	0,0118	0,0000

Source: data - Reuters Information Agency

The figures in the table suggest that there is the significant relationship between these two variables. The correlation itself is negative. When the USD/JPY exchange rate goes up, the Volatility Index (VIX) decreases. Additionally, the correlation between USD/JPY and VIX is substantially stronger from August 2007 to February 2009 (-0,64 versus -0,12).

THE APPLICATION OF ARCH MODEL

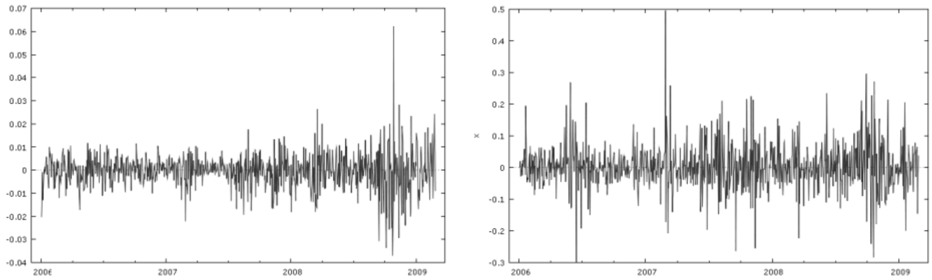
This paper adopts Autoregressive Conditional Heteroscedasticity (ARCH) model to explore the relationship between USD/JPY and VIX. The ARCH model was introduced by Engle (1982). This model is chosen mainly because it provides a way to solve the problem of heteroscedasticity. The volatility of USD/JPY is affected by change in investors' risk aversion. Therefore, the model additionally consists of the independent variable which expresses the investors' mood. The data series embrace daily closing values of USD/JPY and daily values of Volatility Index VIX. The daily series are generated from the following equation.

$$y = \ln(\frac{USD/JPY_t}{USD/JPY_{t-1}}) \tag{1}$$

$$x = \ln(\frac{VIX_t}{VIX_{t-1}}) \tag{2}$$

Where *ln* is the natural logarithm operator, *t* the time period, *y* is the outcome variable and *x* the independent variable. The data cover the period from January 2006 to February 2009. Both series are found to be stationary which was checked by the Augmented Dickey-Fuller unit root test (ADF). The graph below present fluctuations of the dependent variable y (left side) and independent variable x (right side) in the analysed time.

Graph 2. The volatility of dependent variable y and independent variable x



Source: data - Reuters Information Agency

The general form of ARCH is [Hughes at al., 2004 and Trzpiot, 2010]:

$$y_t = x_t'\beta + \varepsilon_t \tag{3}$$

$$\varepsilon_t = \xi_t \sigma_t \tag{4}$$

$$\sigma_t^2 = \alpha_0 + \sum_{t=1}^{q} \alpha_i \varepsilon_{t-i}^2 \tag{5}$$

Where y_t is the dependent variable, x_t is a $k \times 1$ vector of independent variables, ε_t is the disturbance term, ξ_t is the white noise process (with $E(\xi_t) = 0$ and $E(\xi_t^2) = 1$), σ_t^2 is the conditional variance, $\theta = (\beta', \alpha_0, \alpha_1, ..., \alpha_q)'$ is the vector of unknown parameters and q is the order of the ARCH model. By the time the unknown parameters are estimated the test for ARCH effects is carried out. The Lagrange Multiplier (LM) Test is applied to check the existence of ARCH effects. The null hypothesis is $H_0 : \alpha_0 = \alpha_1 = \alpha_2 = ... = \alpha_q$. The null hypothesis means that the ARCH effect does not exist. To verify this hypothesis the test statistics (LM) and the critical value ($\chi^2(q)$) are estimated for the q (order of ARCH) for 1, 5, 10, 15 lags. The results are presented in the table below.

Table 3. The Lagrange Multiplier Test's results

	LM	p-value
q = 1	16,1548	0,0000
q = 5	54,2893	0,0000
q = 10	60,7871	0,0000
q = 15	110,837	0,0000

Source: data - Reuters Information Agency

The LM test for ARCH(1), ARCH(5), ARCH(10) and ARCH(15) errors confirm the presence of ARCH effects in the analyzed data. The p-value

($P(\chi^2(q) > LM)$) is less than the required significance level which means that the null hypothesis is rejected.

Further, the investigation if the VIX has any explanatory power for USD/JPY exchange rate is carried out. The model is estimated and evaluated using daily data (variable x and variable y computed like in equations 1 and 2). The sample covers the period 02.01.2006 – 27.02.2009 which corresponds to 825 daily observations. The table below presents the results for ARCH(1) and ARCH(2). The ARCH(q) with q larger than 2 do not fulfil all requirements (e. g. the significance of all parameters) that is why they are not included in the Table 4.

Table 4. ARCH(1) and ARCH(2) models

	ARCH(1)	ARCH(2)
β_1	-0,04324**	-0,04064**
α_0	0,00003**	0,00003**
α_1	0,27066*	0,16423**
α_2		0,17506*
Akaike's Information Criterion (AIC)	-5934,06	-5964,08
Schwarz Criterion (SC)	-5915,17	-5940,51

*significant at the 0,05 level **at the 0,01 level

Source: data - Reuters Information Agency

In order to select the most appropriate model the Akaike's Information Criterion and the Schwarz Criterion are used. The lower the value of AIC and SC the better the model is. On the basis of Akaike's Information Criterion (AIC) and Schwarz Criterion (SC), the model ARCH(2) is chosen (Table 4). On the basis of equations 1 through 5, the form of ARCH(2) is written as:

$$\ln\left(\frac{USDJPY_t}{USDJPY_{t-1}}\right) = -0{,}04064\ \ln\left(\frac{VIX_t}{VIX_{t-1}}\right) + \varepsilon_t \tag{6}$$

$$\varepsilon_t = \xi_t \sigma_t \tag{7}$$

$$\sigma_t^2 = 0{,}0003 + 0{,}16423\varepsilon_{t-1}^2 + 0{,}17506\varepsilon_{t-2}^2 \tag{8}$$

The coefficient β_1 is negative and significant. Therefore, the results show that there is a statistically significant relationship between USD/JPY exchange rate and the Volatility Index VIX. At the same time, the outcomes indicate that the association between the value of Japanese yen and the level of investors' risk aversion exists. The negative coefficient β_1 suggests that when there is an increase in risk aversion among investors then the USD/JPY exchange rate decreases. Consequently, the growth of risk aversion brings about the appreciation of the Japanese currency. As a result, the change in investor's mood is observed in the USD/JPY market.

The volatility in the Japanese yen market reflects the change in investors' risk aversion. When the market expands, share prices increase, the investors have positive attitude towards risk then one can expect the depreciation of Japanese yen. However, when stock market crashes, the financial market is hit by crisis of confidence, investor's risk aversion is rising a drop of USD/JPY exchange rate follows.

CONCLUSIONS

1. There is statistically significant relationship between USD/JPY exchange rate and the Volatility Index VIX. The negative coefficient β_1 suggests that when there is growth in investors' risk aversion then the USD/JPY exchange rate is decreasing.
2. During the financial crisis of 21^{st} century the volatility in the market increased significantly. Between August 2007 and February 2009 investors' aversion to risk increased considerably. The maximum value of VIX was 80,86 in comparison to the time before the financial crisis (01.2006-07.2007) when the maximum value of the VIX was just 24,27. Moreover, the relationship between USD/JPY and VIX is substantially stronger during the financial crisis.
3. The yen market reflects the change in investor's attitude towards risk. The USD/JPY exchange rate is decreasing when there is an increase in investors' risk aversion. On the other hand, the Japanese currency depreciates when investors' attitude towards risk is positive.

REFERENCES

Brunnermeier M. K., Nagel S., Pedersen L., Carry trades and currency crashes, NBER Working Paper Series, November 2008, p. 24.

Coudert V., Gex M., Does risk aversion drive financial crisis? Testing the predictive power of empirical indicators, Journal of Empirical Finance, no 15, 2008, p. 168.

Engle R. F., (1982), Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflations, Econometrica, v50(4), p. 987-1008.

Field A., Discovering Statistics Using SPSS, SAGE Publications 2005, p. 129.

Fong W. M., A stochastic dominance analysis of yen carry trades, Journal of Banking & Finance, no 34, 2010, p. 1237.

Gagnon J. E., Chaboud A. P., What can the data tell us about carry trades in Japanese Yen?, International Finance Discussion Papers, no 899, 2007, p.2.

Hughes A., King M., Kwek K., Selecting the order of an ARCH model, Economics Letters, no 83, 2004, p. 269.

Trzpiot G., Wielowymiarowe metody statystyczne w analizie ryzyka inwestycyjnego, Polskie Wydawnictwo Ekonomiczne, Warszawa 2010, p. 165.

Winters C., The Carry Trade, Portfolio Diversification, and the Adjustment of the Japanese Yen, Discussion Paper, Bank of Canada 2008, p. 7.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XI, No. 1, 2010, pp. 45-59

RADAR MEASURES OF STRUCTURES' CONFORMABILITY

Zbigniew Binderman, Bolesław Borkowski
Department of Econometrics and Statistics, Warsaw University of Live Sciences
Wiesław Szczesny
Department of Informatics, Warsaw University of Live Sciences
e-mails: zbigniew_binderman@sggw.pl;boleslaw_borkowski@sggw.pl,
wieslaw_szczesny@sggw.pl

Abstract: In the following work a new method was proposed to study similarity of objects' structures. This method is an adaptation of radar methods of objects' ordering and cluster analysis, which are being developed by the authors. The value added by the authors is the construction of measures for conformability of structures of two objects. Those measures may also be used to define similarities between given objects. Proposed measures are independent of the order of features.

Key words: radar method, radar measure of conformability, measure of similarity, synthetic measures, classification, cluster analysis.

INTRODUCTION

Authors have for many years been researching into the problem of regional measurement of differentiation of agriculture in both dynamic and static aspects [Binderman, Borkowski, Szczesny 2008, 2009a, 2009b, 2010; Borkowski, Szczesny 2002]. In economic-agricultural research based on empirical data almost invariably there is a need of ordering, classification and clustering of homesteads, objects (units) of a multidimensional space of variables. Study of regional differentiation of agriculture is crucial now because of EU's politic of regionalization of allocation of funds. Presently, there are many methods used for classification and clustering of objects [Gatnar, Walesiak 2009, Hellwig 1968, Kukuła 2000, Malina 2004, Młodak 2006, Pociecha 2009, Strahl 1990, Zeliaś 2000].

Dynamic analysis of regional differentiation of agriculture based on a single feature was the common ground of those studies. Key differences in evaluation of similarity or spatial differentiation of agriculture formed by different authors using different measurement methods were apparent. Methods of measuring conformability (differentiation) of structures in a dynamic aspect were seldom used in economic-agricultural research. The basis of comparative analysis of structures is a set of *m* spatial units (in our scenario voivodeships) characterized by *n* features. The problem of examining structures' conformability is present in numerous scientific publications, e.g. [Binderman, Borkowski, Szczesny 2008, Binderman, Szczesny 2009, Ciok, Kowalczyk, Pleszczyńska, Szczesny 1995, Kukuła 2000, 2010, Ostasiewicz 1999]. In order to compare structures different methods are used, depending on the goals of research, possibility of evaluation, interpretation of analysis results and desired algebraic and statistical properties. Many methods are constructed intuitively, based on graphical analysis. Radar methods, which are used to display objects defined by a number of features, are an example of such methods. Synthetic index is constructed based on the area of a polygon which is used to illustrate objects in question. This method is simple and intuitive but has a serious flaw because the field value is dependent on the order of features. Our research is aimed at eliminating this flaw. Several proposed indices without this flaw are presented in [Binderman, Borkowski, Szczesny 2008] and [Binderman, Szczesny 2009]. In this work we present a manner in which the idea of measures based on the area of a polygon may be used to measure the conformability of two structures. This manner creates opportunities of using those measures to compare agricultural regions, which are characterized by many features. The use of methods given in this work is included in the article [Binderman, Borkowski, Szczesny 2010]. For entire collection of the conformability of two structures see [Grabiński, Wydymus, Zeliaś 1989, Kukuła 1989, 2010, Malina 2004, Strahl 1985, 1996, Walesiak 1983, 1984].

CONSTRUCTION OF RADAR MEASURES OF CONFORMABILITY

In their previous works authors used radar methods to order and classify objects [Binderman, Borkowski, Szczesny 2008, 2009, 2009a, 2010, Binderman, Szczesny 2009, Binderman 2009, 2009a]. Those methods are independent of the manner of ordering of features that describe a given object. In this work authors attempted to adapt radar methods to compare structures of given objects. Methods presented below may seem to be complicated in terms of calculations. However, with the beginning of the digital age that became inconsequential. Moreover, software to perform calculations for those methods is being developed.
Let Q and R be two objects described by sets of values of *n* ($n > 2$) features. We assume that objects Q,R are described by two vectors $\mathbf{x}, \mathbf{y} \in \Re_{+}^{n}$, where

$$\mathbf{x} = (x_1, x_2, ..., x_n), \mathbf{y} = (y_1, y_2, ..., y_n); x_i, y_i \geq 0; i = 1, 2, ..., n \text{ and } \sum_{i=1}^{n} x_i = 1, \sum_{i=1}^{n} y_i = 1.$$

For a geometric representation of the method we inscribe a regular n-gon into a unit circle (with a radius of one) centered at the origin in the polar coordinate system and connect the vertices of the polygon with the origin. Obtained line-segments with a length of one will be named, in sequence, O1, O2, …, On, for definiteness, beginning with the line-segment covering the w axis. Let's assume that at least two coordinates of each of the vectors x, y are nonzero. As features of objects x and y take on a value between 0 and 1, meaning

$$\mathbf{0}\le\mathbf{x}\le\mathbf{1}\equiv 0\le x_i\le 1,\ \mathbf{0}\le\mathbf{y}\le\mathbf{1}\equiv 0\le y_i\le 1,\ i=1,2,...n,\ \text{where } \mathbf{0}=(0,0,...,0),\ \mathbf{1}=(1,1,...,1),$$

it is possible to represent those values on a radar chart. To do this let x_i (y_i) denote points of intersections of axes Oi with circles centered at the origin of the coordinate system with a radius of x_i (y_i), i=1,2,...,n. By connecting points x_1 with x_2, x_2 with x_3, ..., x_n with x_1 (y_1 with y_2, y_2 with y_3, ..., y_n with y_1) we obtain n-gons S_Q and S_R, which areas $|S_Q|$, $|S_R|$ are given by

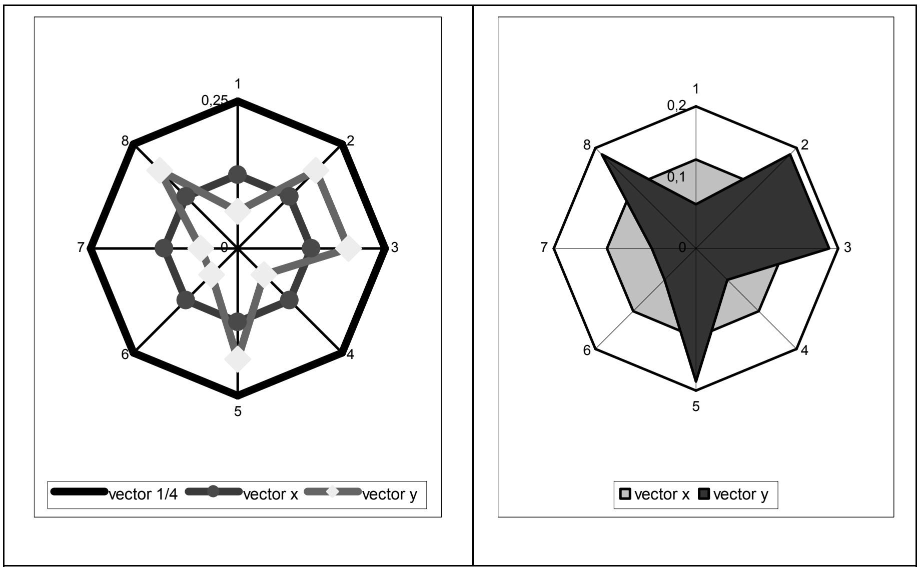
$$|S_Q| = |S_{\mathbf{x}}| = \sum_{i=1}^{n}\frac{1}{2}x_i x_{i+1}\sin\frac{2\pi}{n} = \frac{1}{2}\sin\frac{2\pi}{n}\sum_{i=1}^{n}x_i x_{i+1},\ \text{where } x_{n+1} := x_1\ ,$$

$$|S_R| = |S_{\mathbf{y}}| = \sum_{i=1}^{n}\frac{1}{2}y_i y_{i+1}\sin\frac{2\pi}{n} = \frac{1}{2}\sin\frac{2\pi}{n}\sum_{i=1}^{n}y_i y_{i+1},\ \text{where } y_{n+1} := y_1\ .$$

The following graph gives an illustration for vectors:

$$\mathbf{x} = \frac{\mathbf{1}}{\mathbf{8}} = (\frac{1}{8},\frac{1}{8},...,\frac{1}{8}),\quad \mathbf{y} = (\frac{1}{16},\frac{3}{16},\frac{3}{16},\frac{1}{16},\frac{3}{16},\frac{1}{16},\frac{1}{16},\frac{3}{16}),\ n = 8.$$

Illustration 1. Radar charts for vectors x and y



Source: our investigation

Given such a graphical illustration, each of the objects Q and R is defined by a polygon of vertices $Q_1, Q_2, \ldots, Q_n$ and $R_1, R_2, \ldots, R_n$, respectively. In a Cartesian coordinate system those points take on coordinates $Q_i(s_i, t_i)$, $R_i(w_i, z_i)$, $i=1,2,\ldots,n$; where

$$s_i = x_i \cos\varphi_i, \quad t_i = x_i \sin\varphi_i, \quad w_i = y_i \cos\varphi_i, \quad z_i = y_i \sin\varphi_i,$$

$$\varphi_i = (i-1)\frac{2\pi}{n}, \; i = 1, 2, \ldots, n.$$

Let us denote the areas set by vectors **x** and **y** (describing objects Q and R) by S_x and S_y, respectively, and their intersection by:

$$S_{\mathbf{x}\cap\mathbf{y}} := S_{\mathbf{x}} \cap S_{\mathbf{y}}$$

Let us consider one segment of the area $S_{\mathbf{x}\cap\mathbf{y}} - \Pi_i$, contained within an angle

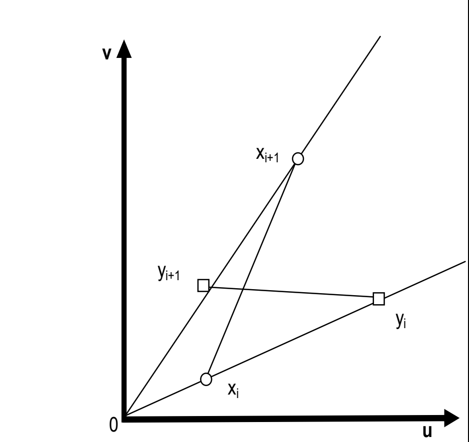
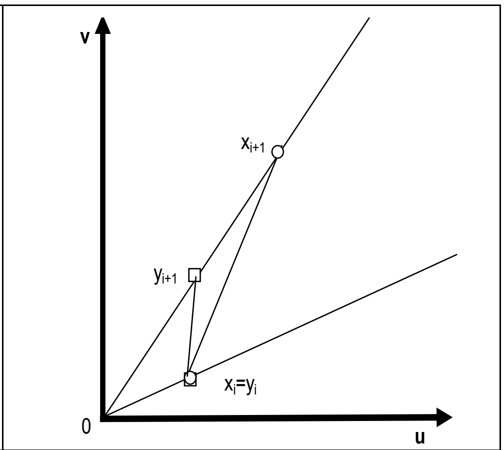
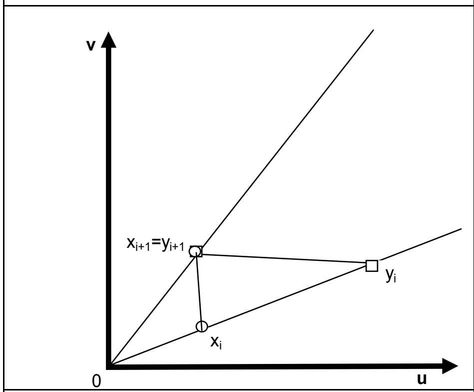
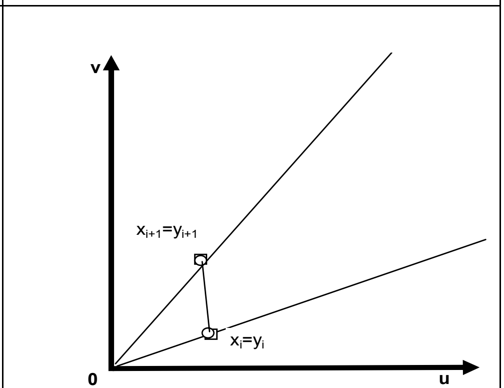
$[\frac{2\pi i}{n}, \frac{2\pi(i+1)}{n}]$. The following, mutually exclusive cases are possible:

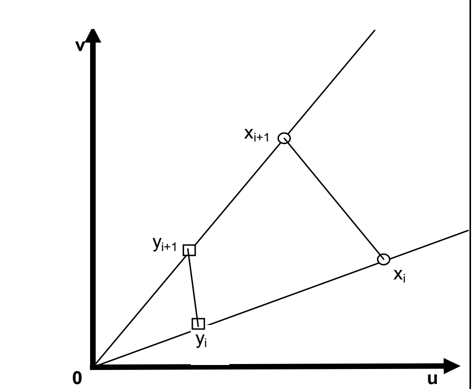
1. $y_i > x_i > 0 \wedge x_{i+1} > y_{i+1} > 0 \quad \vee \quad x_i > y_i > 0 \wedge y_{i+1} > x_{i+1} > 0,$
2. $y_i = x_i > 0 \wedge x_{i+1} > y_{i+1} > 0 \quad \vee \quad x_i = y_i > 0 \wedge y_{i+1} > x_{i+1} > 0,$
3. $x_{i+1} = y_{i+1} > 0 \wedge x_i > y_i > 0 \quad \vee \quad x_{i+1} = y_{i+1} > 0 \wedge y_i > x_i > 0,$
4. $y_i = x_i > 0 \wedge x_{i+1} = y_{i+1} > 0,$
5. $x_i > y_i > 0 \wedge x_{i+1} > y_{i+1} > 0 \quad \vee \quad y_i > x_i > 0 \wedge y_{i+1} > x_{i+1} > 0,$

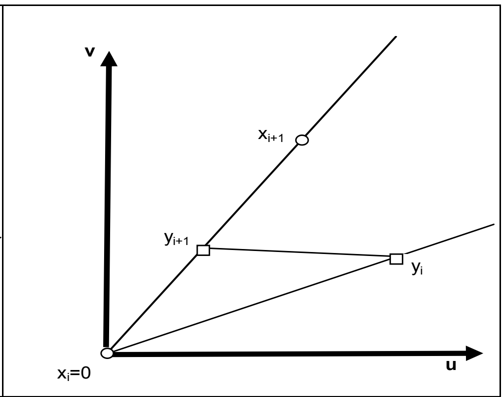
6. Product of coordinates $x_i y_i x_{i+1} y_{i+1} = 0$.

Below we provided representative cases of the above possible situations, linked with the manner of application of the equations.

Illustration 2. Graphical representation of possible cases.

	
1. $y_i > x_i > 0 \wedge x_{i+1} > y_{i+1} > 0$	**2**. $y_i = x_i > 0 \wedge x_{i+1} > y_{i+1} > 0$
	
3. $x_{i+1} = y_{i+1} > 0 \wedge x_i > y_i > 0$	**4.** $y_i = x_i > 0 \wedge x_{i+1} = y_{i+1} > 0$,

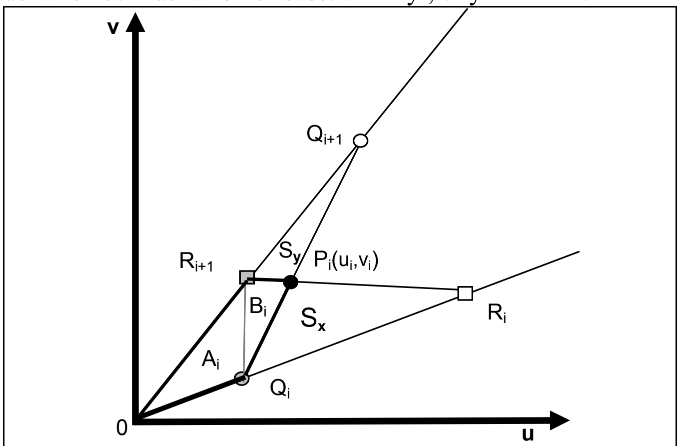




5. $x_i > y_i > 0 \wedge x_{i+1} > y_{i+1} > 0$	**6.** $x_i = 0$
Source: own investigation	

Let us consider the first case when one of the segments of the area $S_{\mathbf{x}\cap\mathbf{y}} - P_i$, is a quadrilateral, given by the origin of the coordinate system O(0, 0) and points Q_i, R_{i+1}, which satisfies our assumption that $0<x_i<y_i$, $0<y_{i+1}<x_{i+1}$ (see illus. 3) $i\in\{1,2,\ldots,n\}$.

Illustration 3. Illustration of case: 0<xi<yi, 0<yi+1<xi+1.



Source: own investigation

Area P consists of two triangles A_i and B_i. Value of the area of triangle A_i is defined as $\left|A_i\right| = \frac{1}{2} x_i y_{i+1} \sin\frac{2\pi}{n}$. Thus, to calculate the area of the quadrilateral S_i:=$OQ_iP_iR_{i+1}$ we must find the value of the area of triangle B_i. In order to do this we notice that the line containing points $R_i(w_i,z_i)$ and $R_{i+1}(w_{i+1},z_{i+1})$ is described by the equation:

$$\left|\begin{pmatrix} u-w_i & v-z_i \\ w_{i+1}-w_i & z_{i+1}-z_i \end{pmatrix}\right| = 0$$

Whereas the line containing points $Q_i(s_i,t_i)$ i $Q_{i+1}(s_{i+1},t_{i+1})$ is described by the equality:

$$\left|\begin{pmatrix} u-s_i & v-t_i \\ s_{i+1}-s_i & t_{i+1}-t_i \end{pmatrix}\right| = 0.$$

Coordinates of point $P_i(u_i,v_i)$, which is the point of intersection between the above lines, are the solution to the following system of equations:

$$(z_{i+1}-z_i)u-(w_{i+1}-w_i)v = w_i z_{i+1} - z_i w_{i+1},$$
$$(t_{i+1}-t_i)u-(s_{i+1}-s_i)v = s_i t_{i+1} - s_{i+1} t_i.$$

The solutions can be described using Cramer's rule as follows:

$$u_i = \frac{W_u}{W},\ v_i = \frac{W_v}{W},\quad \text{where } W = \left|\begin{pmatrix} z_{i+1}-z_i & w_i-w_{i+1} \\ t_{i+1}-t_i & s_i-s_{i+1} \end{pmatrix}\right|,$$

$$W_u = \left|\begin{pmatrix} w_i z_{i+1}-z_i w_{i+1} & w_i-w_{i+1} \\ s_i t_{i+1}-s_{i+1} t_i & s_i-s_{i+1} \end{pmatrix}\right|,\ W_v = \left|\begin{pmatrix} z_{i+1}-z_i & w_i z_{i+1}-z_i w_{i+1} \\ t_{i+1}-t_i & s_i t_{i+1}-s_{i+1} t_i \end{pmatrix}\right|.$$

Let us notice that a line containing two points $Q_i(s_i,t_i)$ i $R_{i+1}(w_{i+1},z_{i+1})$ can be described by the equation:

$$\left|\begin{pmatrix} u-s_i & v-t_i \\ w_{i+1}-s_i & z_{i+1}-t_i \end{pmatrix}\right| = 0,$$

which is identical to $(z_{i+1}-t_i)u-(w_{i+1}-s_i)v - z_{i+1}s_i + w_{i+1}t_i = 0$.

The distance h between point $P_i(u_i,v_i)$ and line containing points $Q_i(s_i,t_i)$ and $R_{i+1}(w_{i+1},z_{i+1})$ is determined by the equality:

$$h_i = \frac{\left|(z_{i+1}-t_i)u_i-(w_{i+1}-s_i)v_i - z_{i+1}s_i + w_{i+1}t_i\right|}{\sqrt{(z_{i+1}-t_i)^2+(w_{i+1}-s_i)^2}} = \frac{\left|(z_{i+1}-t_i)u_i-(w_{i+1}-s_i)v_i - z_{i+1}s_i + w_{i+1}t_i\right|}{\sqrt{y_{i+1}^2+x_i^2-2x_i y_{i+1}\cos\frac{2\pi}{n}}}$$

We utilize the cosine rule to calculate the length of a_i, a line-segment between points $Q_i(s_i,t_i)$ and $R_{i+1}(w_{i+1},z_{i+1})$

$$a_i = \sqrt{y_{i+1}^2+x_i^2-2x_i y_{i+1}\cos\frac{2\pi}{n}}$$

Thus, we obtain that the area of the triangle B_i is equal to:

$$|B_i| = \frac{1}{2}a_i h_i = \frac{1}{2}\left|(z_{i+1} - t_i)u_i - (w_{i+1} - s_i)v_i - z_{i+1}s_i + w_{i+1}t_i\right|$$

and so the area of quadrilateral S_i is described by the equality:

$$|S_i| = |A_i| + |B_i| = \frac{1}{2}x_i y_{i+1}\sin\frac{2\pi}{n} + \frac{1}{2}\left|(z_{i+1} - t_i)u_i - (w_{i+1} - s_i)v_i - z_{i+1}s_i + w_{i+1}t_i\right|.$$

In a similar manner one can obtain the area of segment S_i which is a quadrilateral given by the origin of the coordinate system O(0, 0) and points Q_{i+1}, R_i, what corresponds with the assumption that $0<y_i<x_i$, $0<x_{i+1}<y_{i+1}$ $i\in\{1,2,\ldots,n\}$.

In the case of $x_i = y_i > 0$ oraz $x_{i+1}y_{i+1} > 0$ $\{x_{i+1} = y_{i+1} > 0$ oraz $x_i y_i > 0\}$ segment S_i is a triangle with an area described by the equality:

$$|S_i| = \frac{1}{2}\sin\frac{2\pi}{n}x_i\min(x_{i+1}, y_{i+1})\ \left\{|S_i| = \frac{1}{2}\sin\frac{2\pi}{n}x_{i+1}\min(x_i, y_i)\right\}.$$

In the case of $x_i > y_i > 0$ oraz $x_{i+1} > y_{i+1} > 0$ $\{y_i > x_i > 0$ oraz $y_{i+1} > x_{i+1} > 0\}$ Segment S_i is a triangle with an area described by the equality:

$$|S_i| = \frac{1}{2}\sin\frac{2\pi}{n}y_i y_{i+1}\ \left\{|S_i| = \frac{1}{2}\sin\frac{2\pi}{n}x_{i+1}x_i\right\}.$$

In the case of $x_i y_i x_{i+1} y_{i+1} = 0$ segment S_i is a line-segment or a point and its area is equal to 0: $|S_i| = 0$.

The area of the intersection of polygons $S_\mathbf{x}$ and $S_\mathbf{y}$ is described by the equality:

$$|S_\mathbf{x} \cap S_\mathbf{y}| = \sum_{i=1}^{n}|S_i|.$$

Let us assume $\boldsymbol{\mu}_\mathbf{xy}$ as a measure of conformability of structures of two objects Q and R induced by vectors **x** and **y**. thus:

$$\mu_\mathbf{xy} = \begin{cases} \sqrt{\dfrac{|S_\mathbf{x} \cap S_\mathbf{y}|}{\sigma_\mathbf{xy}}} & \text{for } n=3 \\ \sqrt{\dfrac{|S_\mathbf{x} \cap S_\mathbf{y}|}{\omega_\mathbf{xy}}} & \text{for } n\geq 4 \end{cases} \tag{1}$$

$$\text{where } \sigma_\mathbf{xy} := \begin{cases} \min(|S_\mathbf{x}|, |S_\mathbf{y}|) & \text{gdy } |S_\mathbf{x}||S_\mathbf{y}|>0 \\ 1 & \text{gdy } |S_\mathbf{x}||S_\mathbf{y}|=0 \end{cases}, \quad \omega_\mathbf{xy} := \begin{cases} \max(|S_\mathbf{x}|, |S_\mathbf{y}|) & \text{gdy } |S_\mathbf{x}||S_\mathbf{y}|>0 \\ 1 & \text{gdy } |S_\mathbf{x}||S_\mathbf{y}|=0 \end{cases}.$$

Let us notice that the above measure of conformability satisfies $0\leq\mu_{x,y}\leq 1$ and is dependent on the order of features [Binderman, Borkowski, Szczesny 2008].

In order to define a measure of conformability which is independent of the order of features let us denote by p_j a j-th permutation of numbers 1, 2, …, n. There are *n!* such permutations. Each permutation corresponds to a permutation of coordinates of vectors **x** and **y**. Let $\mathbf{x}_j$, $\mathbf{y}_j$ denote the j-th permutation of coordinates of vectors **x** and **y** accordingly, where $\mathbf{x}_1=\mathbf{x}$ and $\mathbf{y}_1=\mathbf{y}$. E.g. if n=3, $\boldsymbol{x}=(x_1,x_2,x_3)$, $\boldsymbol{y}=(y_1,y_2,y_3)$ and $p_1=(1,2,3)$, $p_2=(1,3,2)$, $p_3=(2,1,3)$, $p_4=(2,3,1)$, $p_5=(3,1,2)$, $p_6=(3,2,1)$ then: $\boldsymbol{x}_1=(x_1,x_2,x_3)$, $\boldsymbol{y}_1=(y_1,y_2,y_3)$, $\boldsymbol{x}_2=(x_1,x_3,x_2)$, $\boldsymbol{y}_2=(y_1,y_3,y_2)$ $\boldsymbol{x}_3=(x_2,x_1,x_3)$, $\boldsymbol{y}_3=(y_2,y_1,y_3)$, $\boldsymbol{x}_4=(x_2,x_3,x_1)$, $\boldsymbol{y}_4=(y_2,y_3,y_1)$, $\boldsymbol{x}_5=(x_3,x_1,x_2)$, $\boldsymbol{y}_5=(y_3,y_1,y_2)$, $\boldsymbol{x}_6=(x_3,x_2,x_1)$, $\boldsymbol{y}_6=(y_3,y_2,y_1)$.
Based on our previous pondering we conclude that each j-th permutation $\mathbf{x}_j$, $\mathbf{y}_j$ of coordinates of vectors **x** and **y** corresponds to a measure of conformability of structures:

$$\mu^{j}_{Q,R} = \mu_{\mathbf{x}_j \mathbf{y}_j} \quad , \tag{2}$$

where, naturally, $\mu^1_{Q,R} = \mu_{\mathbf{xy}}$.

In accordance with the above let us define three different measures of conformability of considered objects Q and R:

$$\begin{aligned} \mathfrak{M}_{Q,R} &= \max_{1\le j\le n!} \mu^{j}_{Q,R}, \\ m_{Q,R} &= \min_{1\le j\le n!} \mu^{j}_{Q,R}, \\ S_{Q,R} &= \frac{1}{n!}\sum_{j=1}^{n!} \mu^{j}_{Q,R}. \end{aligned} \tag{3}$$

To compare structures of two objects

$$\mathbf{x} = (x_1, x_2, \ldots, x_n), \mathbf{y} = (y_1, y_2, \ldots, y_n): \; x_i, y_i \ge 0; \; i = 1, 2, \ldots, n \; ; \; \sum_{i=1}^{n} x_i = 1, \; \sum_{i=1}^{n} y_i = 1,$$

utilizing a popular and simple in use coefficient [Chomątowski, Sokołowski 1978]

$$W_{\mathbf{xy}} := \sum_{i=1}^{n} \min(x_i, y_i) . \tag{4}$$

In order to present the described above method of comparing structures we will consider three simple examples.

Example 1.

Let $Q = \mathbf{x} = \left(\frac{1}{2}, \frac{1}{2}, 0\right)$, $R = \mathbf{y} = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$. Let the following take on values:

$$\mathbf{x}_1 := \mathbf{x}_4 := \mathbf{x},\ \mathbf{x}_2 := \mathbf{x}_5 := \left(\frac{1}{2}, 0, \frac{1}{2}\right),\ \mathbf{x}_3 := \mathbf{x}_6 := \left(0, \frac{1}{2}, \frac{1}{2}\right),\ \mathbf{y}_1 := \mathbf{y}_2 := \mathbf{y}_3 := \mathbf{y}_4 := \mathbf{y}_5 := \mathbf{y}_6 = \mathbf{y}$$

Thus, we receive:

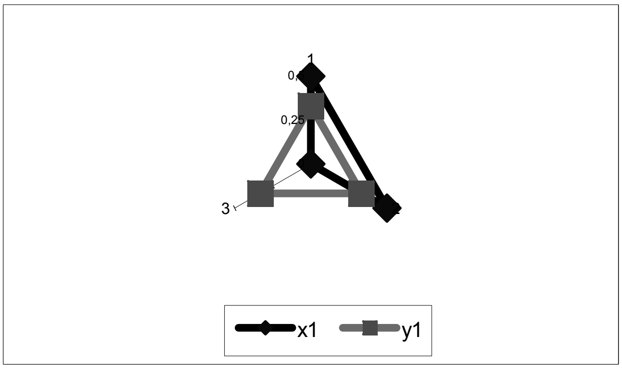
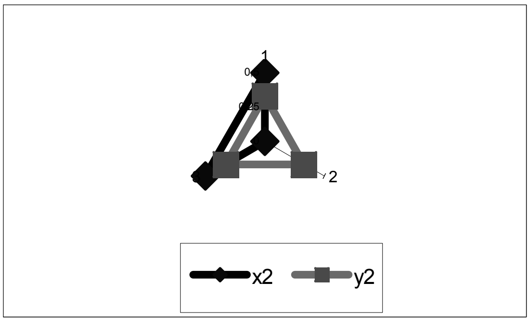
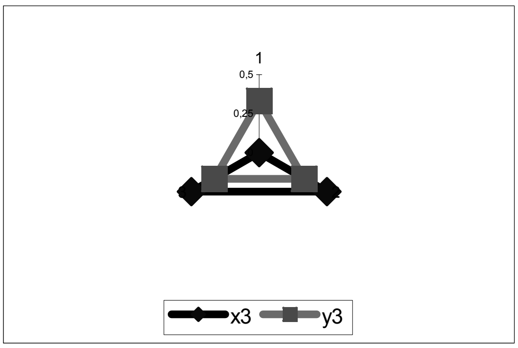
$$\left|S_{\mathbf{x}_i}\right| = \frac{1}{2}\sin\frac{2\pi}{3}\frac{1}{2}\frac{1}{2},\ \left|S_{\mathbf{y}_i}\right| = 3\frac{1}{2}\sin\frac{2\pi}{3}\frac{1}{3}\frac{1}{3},\ \left|S_{\mathbf{x}_i} \cap S_{\mathbf{y}_i}\right| = \frac{1}{2}\sin\frac{2\pi}{3}\frac{1}{3}\frac{1}{3},$$

$$\mu_{\mathbf{x}_i\mathbf{y}_i} = \sqrt{\frac{\frac{1}{2}\sin\frac{2\pi}{3}\frac{1}{3}\frac{1}{3}}{\frac{1}{2}\sin\frac{2\pi}{3}\frac{1}{2}\frac{1}{2}}} = \frac{2}{3},\ \text{for } i = 1, 2, ..., 6.$$

And so $\mathfrak{M}_{Q,R} = m_{Q,R} = S_{Q,R} = \frac{2}{3}$, where coefficients $\mathfrak{M}_{Q,R}, m_{Q,R}, S_{Q,R}$ are described by equations (3). It is worth noting that when a coefficient is described by the equation (4) then $W_{\mathbf{xy}} = \frac{1}{3} + \frac{1}{3} + 0 = \frac{2}{3}$.

The following illustrations present the considered example.

Illustration 4. Graphical presentation of the method for vectors
$\mathbf{x}=\left(\frac{1}{2},\frac{1}{2},0\right)$, $\mathbf{y}=\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)$.

$\mathbf{x}_1=\mathbf{x}_4=\left(\frac{1}{2},\frac{1}{2},0\right)$, $\mathbf{y}_1=\mathbf{y}_4=\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)$	
$\mathbf{x}_2:=\mathbf{x}_5:=\left(\frac{1}{2},0,\frac{1}{2}\right)$ $\mathbf{y}_2=\mathbf{y}_5=\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)$	
$\mathbf{x}_3:=\mathbf{x}_6:=\left(0,\frac{1}{2},\frac{1}{2}\right)$ $\mathbf{y}_3=\mathbf{y}_6=\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)$	
Source: own investigation	

Example 2.

Let $Q = \mathbf{x} = \left(\frac{1}{2}, \frac{1}{2}, 0, 0\right)$, $R = \mathbf{y} = \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$. Let the following take on values: $\mathbf{x}_1 := \mathbf{x}_2 := \mathbf{x}_3 := \mathbf{x}_4 := \mathbf{x}$, $\mathbf{x}_5 := \mathbf{x}_6 := \mathbf{x}_7 := \mathbf{x}_8 := \left(\frac{1}{2}, 0, \frac{1}{2}, 0\right)$,

$\mathbf{x}_9 := \mathbf{x}_{10} := \mathbf{x}_{11} := \mathbf{x}_{12} := \left(\frac{1}{2}, 0, 0, \frac{1}{2}\right), \mathbf{x}_{13} := \mathbf{x}_{14} := \mathbf{x}_{15} := \mathbf{x}_{16} := \left(0, \frac{1}{2}, \frac{1}{2}, 0\right), \mathbf{x}_{17} := \mathbf{x}_{18} := \mathbf{x}_{19} :=$ T

$:= \mathbf{x}_{20} := \left(0, 0, \frac{1}{2}, \frac{1}{2}\right), \mathbf{x}_{21} := \mathbf{x}_{22} := \mathbf{x}_{23} := \mathbf{x}_{24} := \left(0, \frac{1}{2}, 0, \frac{1}{2}\right)$, $\mathbf{y}_i := \mathbf{y}$ for $i = 1, 2, \ldots, 24$.

hus, we receive:

$$\left|S_{\mathbf{x}_j}\right| = \frac{1}{2}\sin\frac{\pi}{2}\cdot\frac{1}{2}\frac{1}{2},\quad \left|S_{\mathbf{y}_i}\right| = 4\cdot\frac{1}{2}\sin\frac{\pi}{2}\cdot\frac{1}{4}\frac{1}{4},\quad \left|S_{\mathbf{x}_j} \cap S_{\mathbf{y}_j}\right| = \frac{1}{2}\sin\frac{\pi}{2}\cdot\frac{1}{4}\frac{1}{4}, j = 1,\ldots,4; i = 1,2,\ldots,24;$$

$$\mu_{\mathbf{x}_j\mathbf{y}_j} = \sqrt{\frac{\frac{1}{2}\sin\frac{\pi}{2}\cdot\frac{1}{4}\frac{1}{4}}{4\cdot\frac{1}{2}\sin\frac{\pi}{2}\frac{1}{4}\frac{1}{4}}} = \frac{1}{2},$$ for $j = 1, 2, 3, 4$; where $\mu_{\mathbf{x}_j\mathbf{y}_j}$ is defined by formula (2).

It can be easily verified that

$\mu_{\mathbf{x}_j\mathbf{y}_j} = \frac{1}{2}$, for $j = 8, 9, \ldots, 20$ and $\mu_{\mathbf{x}_j\mathbf{y}_j} = 0$ for $j = 5, 6, 7, 8, 21, 22, 23, 24$.

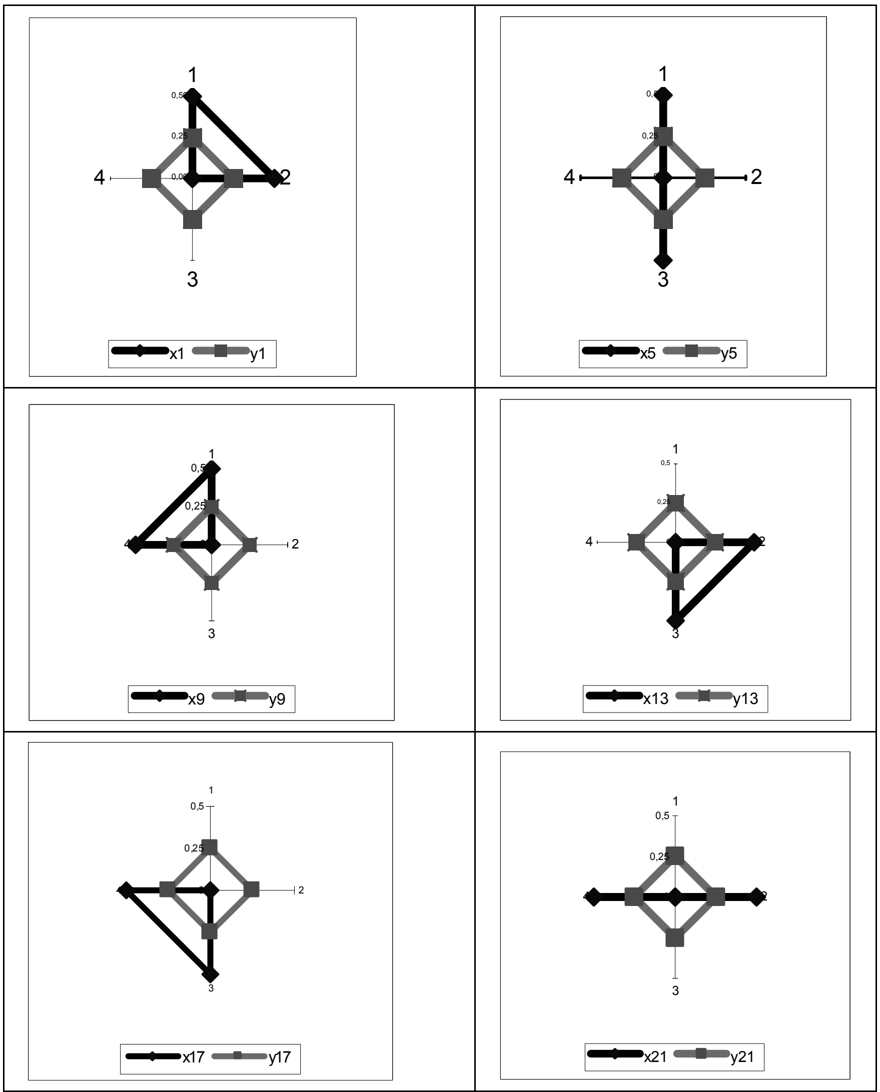
Hence $\mathfrak{M}_{Q,R} = \frac{1}{2}$, $m_{Q,R} = 0$, $S_{Q,R} = \frac{1}{3}$, where the coefficients $\mathfrak{M}_{Q,R}, m_{Q,R}, S_{Q,R}$ are described by equations (3). In this example the coefficient of structures conformability is equal to:

$$W_{\mathbf{xy}} = \frac{1}{4} + \frac{1}{4} + 0 + 0 = \frac{1}{2},$$

where $W_{\mathbf{xy}}$ is described by the formula (4).

The following illustrations present the considered example.

Illustration 5. Graphical presentation of the method for vectors $\mathbf{x} = \left(\frac{1}{2}, \frac{1}{2}, 0, 0\right)$, $\mathbf{y} = \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right)$.



Source: own investigation

Example 3.

Let $Q = \mathbf{x} = \left(\frac{1}{2}, \frac{1}{2}, 0, 0, ..., 0\right)$, $R = \mathbf{y} = \left(\frac{1}{n}, \frac{1}{n}, ..., \frac{1}{4}\right) \in \Re^n, n \geq 4$. It can be shown that $\mathfrak{M}_{Q,R} = \frac{2}{n}$, $m_{Q,R} = 0$, $S_{Q,R} = \frac{4}{n(n-1)}$, where coefficients $\mathfrak{M}_{Q,R}, m_{Q,R}, S_{Q,R}$

are described by equations (3). In this example the coefficient of structures conformability (described by the equation (4)) takes on the value:

$$W_{\mathbf{xy}} = \frac{1}{n} + \frac{1}{n} + 0 + ... + 0 = \frac{2}{n}.$$

SUMMARY

This work presents a mean to study conformability of structures but it is easily seen that introduced norms can also be used to analyze similarity of studied objects. Presented radar methods create a possibility of utilizing them in decision analysis on a local level. Radar methods are commonly used due to the ease of visualization of multidimensional data. However, some analyses incorrectly employ indices based solely on those illustrations, meaning they do not ensure the basic requirement of stability of the employed method – independence of the order of features [Jackson 1970]. The method presented by the authors does not have that flaw. As complicated as the presented methods may seem, in the digital age it remains largely inconsequential. Even more so with software for the presented methods is being developed.

REFERENCES

Binderman Z., Borkowski B., Szczesny W. (2008): O pewnej metodzie porządkowania obiektów na przykładzie regionalnego zróżnicowania rolnictwa, Metody ilościowe w badaniach ekonomicznych, IX, 39-48, wyd. SGGW 2008.

Binderman Z., Borkowski B., Szczesny W (2009), O pewnych metodach porządkowych w analizie polskiego rolnictwa wykorzystujących funkcje użyteczności, RNR PAN, Seria G, Ekonomika Rolnictwa , T. 96, z. 2, s. 77-90.

Binderman Z., Borkowski B., Szczesny W. (2009a) Tendencies in changes of regional differentiation of farms structure and area Quantitative methods in regional and sectored analysis/sc., U.S., Szczecin: - s. 33-50.

Binderman Z., Borkowski B., Szczesny W. (2010): The tendencies in regional differentiation changes of agricultural production structure in Poland, Quantitative methods in regional and sectored analysis, U.S., Szczecin, s. 67-103.

Binderman Z., Szczesny W., (2009), Arrange methods of tradesmen of software with a help of graphic representations Computer algebra systems in teaching and research, Siedlce : Wyd. WSFiZ, 117-131.

Binderman, Z. (2009): Ocena regionalnego zróżnicowania kultury i turystyki w Polsce w 2007 roku R. Wydz. Nauk Humanistycznych SGGW, T XII, s. 335-351.

Binderman Z. (2009a), Syntetyczne mierniki elastyczności przedsiębiorstw, Prace i Materiały Wydziału Zarządzania Uniwersytetu Gdańskiego 4/2, s. 257-267.

Borkowski B., Kasiewicz S. (2010), Konstrukcja metryk elastyczności przedsiębiorstw, Prace i Materiały Wydz. Zarządzania UG 4/2, s. 268-279.

Borkowski B., Szczesny W. (2002): Metody taksonomiczne w badaniach przestrzennego zróżnicowania rolnictwa. RNR PAN, Seria G, T 89, z. 2. s. 42. Chomątowski S., Sokołowski A. (1978), Taksonomia struktur, Przegląd Statystyczny, nr 2, s. 14-21.

Ciok A., Kowalczyk T., Pleszczyńska E., Szczesny W. (1995): Algorithms of grade correspondence-cluster analysis. The Coll. Papers on Theoretical and Aplied Computer Science, 7, 5-22.
Gatnar E., Walesiak M. (2009): Statystyczna analiza danych z wykorzystaniem programu R, PWN, Warszawa.
Grabiński T., Wydymus S., Zeliaś A. (1989), Metody taksonomii numerycznej w modelowaniu zjawisk społeczno-gospodarczych, PWN, Warszawa.
Hellwig Z. [1968]; Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju oraz zasoby i strukturę kwalifikowanych kadr, „Przegląd Statystyczny”, z. 4.
Jackson D. M. (1970): The stability of classifications of binary attribute data, Technical Report 70-65, Cornell University 1-13.
Kukuła K. (1989): Statystyczna analiza strukturalna i jej zastosowanie w sferze usług produkcyjnych dla rolnictwa, Zeszyty Naukowe, Seria specjalna Monografie nr 89, AE w Krakowie, Kraków.
Kukuła K. (2000): Metoda unitaryzacji zerowanej, PWN, Warszawa.
Kukuła K. (red.) (2010): Statystyczne studium struktury agrarnej w Polsce, PWN, Warszawa.
Malina A. (2004): Wielowymiarowa analiza przestrzennego zróżnicowania struktury gospodarski Polski według województw. AE, S. M. nr 162, Kraków.
Młodak A., (2006): Analiza taksonomiczna w statystyce regionalnej, DIFIN, Warszawa.
Ostasiewicz W. (red) (1999),: Statystyczne Metody Analizy Danych, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław.
Pociecha J. (2008), Rozwój metod taksonomicznych i ich zastosowań w badaniach społeczno-ekonomicznych, 90-lecie GUS, www.stat.gov.pl, 1-13.
Szczesny W. (2002),: Grade correspondence analysis applied to contingency tables and questionnaire data, Intelligent Data Analysis, vol. 6 , 17-51.
Strahl D. (1985), Podobieństwo struktur ekonomicznych, PN AE, nr 281, Wrocław.
Strahl D. (1996), Równowaga strukturalna obiektu gospodarczego [w:] Przestrzenno-czasowe modelowanie i prognozowanie zjawisk gospodarczych, red. A. Zeliaś, AE w Krakowie, Kraków.
Strahl D. (red.) (1998), Taksonomia struktur w badaniach regionalnych, Prace Naukowe AE we Wrocławiu, Wrocław.
Walesiak M. (1983), Propozycja rodziny miar odległości struktur udziałowych, „Wiadomości Statystyczne”, nr 10.
Walesiak M. (1984), Pojęcie, klasyfikacja i wskaźniki podobieństwa struktur gospodarczych, Prace Naukowe AE we Wrocławiu, nr 285, Wrocław.
Zeliaś A., red. (2000): Taksonomiczna analiza przestrzennego zróżnicowania poziomu życia w Polsce w ujęciu dynamicznym, AE w Krakowie, Kraków.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XI, No. 1, 2010, pp. 60-69

THE IMPORTANCE OF DEMOGRAPHIC VARIABLES IN THE MODELING OF FOOD DEMAND

Hanna Dudek
Department of Econometrics and Statistics, Warsaw University of Live Sciences
e-mail: hanna_dudek@sggw.pl

Abstract: The general objective of this research is to assess the impact of demographic variables on food demand in Poland. The empirical analysis of this paper is based on the household data, collected by GUS (Central Statistical Office) in the years 2001-2004 (household budget data).

Keywords: complete demand system, demographic variables, elasticity of demand

INTRODUCTION

During the last three decades, consumer demand analysis has moved toward system-wide approaches. Increasing attention has been given to the estimation of complete demand systems that consistently account for the interdependence in the choices made by consumers between a large number of commodities. Many algebraic specifications of demand systems have been developed, including the linear and quadratic expenditure systems, the Rotterdam model, Translog models and the Almost Ideal Demand System (AIDS) or its quadratic extension (QUAIDS) [Barnett, Serletis 2008].

The objective of this study is to estimate the impact of economic factors, such as the prices and the expenditures, and noneconomic factors, i.e. the demographic variables, on household demand for eight aggregated food items in Poland.

The issue analysed in this paper is particularly important in the case of the situation in Poland, where the **single-equation** approach to demand analysis dominates. The use of single-equation models suffers from some shortcomings, first of all it ignores the cross-equation restrictions implied by the neoclassical theory of consumer behavior. Following paper presents a struggling attempt to fill the gap in the literature in regard to system-wide approach to food expenditures in

Poland. Moreover, it emphasizes the incorporation of demographic variables in the analysis of food demand, which is very important due to changes demographic profile of consumers, i.e. the problem of aging society. As far as the author is concerned such research has not been conducted for the Polish household data.

ECONOMETRIC DEMAND ANALYSIS

The articles written by H. Working in the 1940s and C. Leser in the 1960s were one of the first papers on econometric demand analysis. Nonetheless, their significant contribution to the demand analysis was not consistent with the utility maximization theory. According to this theory, a sample of households behaves as a representative consumer who maximizes his utility function $u(\mathbf{q})$ subject to the budget constraint $\mathbf{p'q} = x$ where:

$\mathbf{q}$ is the vector of food demanded, $\mathbf{q}=[q_1, q_2,..., q_n]$,

$\mathbf{p}$ is the corresponding vector of prices, $\mathbf{p}=[p_1, p_2,..., p_n]$,

x is the total expenditures to consume $\mathbf{q}$.

By solving this maximization problem, we obtain a system of $n+1$ demand equations specified as follows (see e.g. [Varian 1992], [Barnett, Serletis 2008]):

$$\mathbf{q} = \mathbf{q}(\mathbf{p}, x) \tag{1}$$

The solution to the utility maximization problem yields a set of ordinary demand curves conditional on given prices and income. The system (1) is assumed to satisfy the theoretical plausibility conditions, especially adding up, homogeneity, symmetry and negativity[1] [Edgerton et al. 1996].

Only coherent demand systems allow to model various consumption patterns and behavior sufficiently, while simultaneously satisfying restrictions given by the economic theory. Such examples of coherent demand system present the Linear Expenditure System (LES) [Stone 1954], Transcendental Logarithmic System [Jorgenson, Lau Stoker 1982], Almost Ideal Demand System and its quadratic extension [Banks et al. 1997].

In order to illustrate the incorporation of the demographic variables the AIDS model (Almost Ideal Demand System) is used in this paper. It would also be reasonable to consider other demand systems such as QUAIDS (Quadratic Almost Ideal Demand System) and QES (Quadratic Expenditure System), however,

---

[1] These conditions represent the basic restrictions imposed on all demand functions $q_i = q_i(\mathbf{p}, x)$, $i = 1, 2, .., n$:

- the **adding up restriction** implies that the budget x is totally used;
- the homogeneity condition requires the demand functions to be homogeneous of degree zero in both prices and total expenditures;
- the symmetry and negativity **restrictions** imply that the substitution matrix should be symmetric and negative semidefinite.

estimation difficulties relating to the implementation of nonlinear numerical procedures, which are characteristic to these models, cause that no such attempt has been undertaken[2]. Though, it appears that in order to achieve the objective of this paper, it is sufficient to use a simpler model such as AIDS.

Apart from its flexibility, the main advantages of the AIDS model are as follows: first, it allows an exact aggregation among consumers; second, there is a possibility to estimate a non-linear model; third, it is a popular model because of its empirical validation. The general specification of the AIDS model is given by [Deaton, Muellbauer 1980]:

$$w_i = \alpha_i + \sum_{j=1}^{n} \gamma_{ij} \log p_j + \beta_i \log(x/P) \tag{2}$$

where:

w_i is the expenditure share associated with the ith good, $i=1, 2, ..., n$,

α_i is the constant coefficient in the ith share equation,

γ_{ij} is the slope coefficient associated with the jth good in the ith share equation,

p_j is the price of the jth good,

x is the total expenditure on the system of goods given by the following equation: $x = \sum_{i=1}^{n} p_i q_i$, where q_i is the quantity demanded for the ith good,

P is the general price index defined by:

$$\log P = \alpha_0 + \sum_{i=1}^{n} \alpha_i \log p_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij} \log p_i \log p_j \tag{3}$$

In empirical studies, in order to avoid the non-linearity and reduce the multi-colinearity effects in the model, the equation (3) is often approximated by a Stone index defined as [Deaton, Muellbauer 1980]:

$$\log P^* = \sum_{i=1}^{n} w_i \log p_i \tag{4}$$

Model AIDS (2) with the price index (4) instead (3) is called LA/AIDS model[3] (Linear Approximation of the Almost Ideal Demand System).

[2] Considering that eight expenditure groups were taken into account in the paper, the estimation of nonlinear models using an ordinary personal compute would be extremely time-consuming.

[3] The Stone index is one of a numerous of indices that could be used to define a LA/AIDS specification. The discussion relating to disadvantages of Stone index can be found in [Moschini 1995], [Asche, Wessells 1997] and [Barnett, Seck 2008]. Despite various shortcomings, this index is still applied in empirical researches (see e.g. [Suchecki 2006], [Regorsek, Erjavec 2007].

In our empirical estimation the usual theoretical restrictions derived from the utility maximization and the demand theory are directly imposed into LA/AIDS parameters. These restrictions are:

$$\sum_{i=1}^{n}\alpha_i = 1, \sum_{i=1}^{n}\beta_i = 0, \sum_{i=1}^{n}\gamma_{ij} = 0 \text{ for adding up,} \quad (5)$$

$$\sum_{j=1}^{n}\gamma_{ij} = 0 \text{ for homogeneity and,} \quad (6)$$

$$\gamma_{ij} = \gamma_{ji} \text{ for symmetry.} \quad (7)$$

According to Green and Alston (1990), elasticities in LA/AIDS can be expressed as: $e_i = 1 + \frac{\beta_i}{w_i}$ for income elasticity and $e_{ij}^{*} = -\delta_{ij} + w_j + \gamma_{ij} / w_i$ for compensated[4] (Hicksian) elasticity. The uncompensated (Marshallian) elasticity of relative expenditures on the commodity *i* relative to the price of commodity *j* is given as: $e_{ij} = -\delta_{ij} + \frac{\gamma_{ij}}{w_i} - \beta_i \frac{w_j}{w_i}$, where δ_{ij} is the Kronecker delta taking the value 1 for $i = j$ and 0 for $i \neq j$ [Dudek 2008].

DEMOGRAPHIC VARIABLES

There are general procedures for taking into account the demographic variables into classes of demand system [Pollack, Wales 1981; Ray 1983]. If we denote the originally demand system by $w_i = f(\mathbf{p}, logx)$, which depends on the vector of prices $\mathbf{p}$ and on the total expenditures *x*, then:

1) the method known as "demographic scaling" transforms it into $w_i = m_i \cdot f(\mathbf{p}, logx)$, where the *m*'s are scaling parameters, which depend on the demographic variables [Barten 1964];
2) the procedure named "demographic translating" replaces and translates the original demand system $w_i = f(\mathbf{p}, logx)$ into $w_i = d_i + f(\mathbf{p}, logx)$, where *d* is a parameter depending on the vector of demographic variables [Pollak, Wales 1979];

[4] The economic theory distinguishes two types of demand and thus the elasticities. *Compensated (or Hicksian) elasticity* derived from Hicks (compensated) demand, measures only the substitution effect, i.e. the change in demand due to the change in relative prices if the effect on real income due to the change in prices is compensated. *Marshallian demand and elasticity* considers, however, not only the substitution effect caused by the relative prices change, but also the income effect arising from the change in real income due to the change in prices.

3) the Gorman's specification proposed a method that replaces the original demand system in the following: $w_i = d_i + m_i \cdot f(\mathbf{p}, logx)$, where the d's and the m's depend on the demographic variables [Gorman 1976];
4) the "reverse Gorman" specification can be obtained by firstly by translating, then scaling to yield the following demand system: $w_i = m_i \cdot (d_i + f(\mathbf{p}, logx)$ [Pollak, Wales 1981];
5) the "price scaling" technique replaces the original demand system by $w_i = f\left(\mathbf{p}, \frac{\log x}{m(\mathbf{p},\mathbf{z})}\right) + \frac{\partial \log(m(\mathbf{p},\mathbf{z}))}{\partial \log p_i}$, where $m(\mathbf{p},\mathbf{z})$ - the scaling factor, $\mathbf{z}$ - the vector of demographic variables [Ray 1983].

The procedures mentioned above are general in the sense that they do not assume a particular form of the original demand system, but can be used in conjunction with any complete demand system. Arranging these procedures in an unambiguous ranking is not possible [Pollak, Wales 1981]. First of all, some of them are not nested. Furthermore, their assessment depends also on the functional form used to estimate the demand system. Note that the estimation of some procedures is computationally complex, thus only one relatively simple procedure, "demographic translating", was taken into account in this paper.

LA/AIDS models, as originally proposed by Deaton and Muellbauer (1980), did not consider the demographic variables. However, such variables appear to be crucial in household survey data, in which economic responses to the price changes can be considerably influenced by all the sorts of personal or household effects. The study takes into consideration the "demographic translating" in the LA/AIDS model by allowing the intercept term of each equation to be a function of the demographic variables:

$$w_i = d_i + \sum_{j=1}^{n} \gamma_{ij} \log p_j + \beta_i \log(x / P^*) \tag{8}$$

where $d_i = \alpha_{0i} + \sum_{k=1}^{K} \alpha_{ki} z_k$, K – a number of demographic variables.

The term demographic variables here denotes:

$$z_j = \frac{m_j}{m}, j=1,..,K-1, \qquad z_K = \ln(m) \tag{9}$$

where m_j – a number of household members in age group j,
m – a household size (number of household members).

DATA AND ESTIMATION

The estimation is based on a sample of the household data, derived from the Polish Household Budget Survey5 conducted by the Central Statistical Office of Poland annually conducted by the Central Statistical Office of Poland. The data mostly come from the survey on household monthly expenditures for the years 2001-2004, latest data is not available. Only the set of monthly prices indices was taken from publication *Prices in the National Economy in 2001-2004*.

Generally, data derived from National Household Budget Surveys often causes so-called problem of "zero expenditures" due to the fact that not every single household adhering to the sample buy at least one commodity from each of the aggregated groups. The reasons for this phenomenon are following: the infrequency of the purchase, the seasonality of some products, the self-production of some commodities, etc. Accordingly, data derived from the household of employees[6], in which the problem of "zero expenditures" was not so significant as in remaining groups, are only taken into account.

The specific food commodities within the food groups used in the empirical analysis are: 1) bread and cereals, 2) meat and fish, 3) milk, cheese and eggs, 4) oils and fats, 5) fruit, 6) vegetables, 7) sugar, jam, honey, chocolate and confectionery, 8) other food products. For such groups the ratio of households with "zero expenditures" did not exceed 3%, thus the problem of the censored data was passed over in this study.

The LA/AIDS model was estimated by the seemingly unrelated regression (SUR) technique in the STATA version 10 statistical package. The homogeneity and symmetry restrictions were imposed on the estimated model. To avoid singularity derived from adding-up constraint in the variance-covariance matrix one equation was deleted from direct estimation in the demand system. The parameters' estimates of this equation were recovered using homogeneity, symmetry and adding-up conditions.

RESULTS

We considered demand systems with various set of the demographic variables. In accordance with the statistical criteria, **i.e.** statistical significance of parameters, **Akaike** and **Schwarz** information criteria, one model with two

[5] The Household Budget Surveys data are considered superior compared to the available time-series data for the research because they include detailed demographic characteristics that allow heterogeneity in preferences across households. Additionally, the large sample size included in the NBS household survey data allows estimating a relatively large demand system.

[6] The number of households of employees participating in the survey in each year was about 12500.

demographic variables, **i.e** ratio number of household members aged 14 years and over[7] and household size, was chosen. Most of the parameter estimates were significant at the 0,05 significance level. For reasons of space, all detailed results could not be presented. The complete results are available from the author upon request. The obtained results[8] concerning the parameter estimates, standard errors and the goodness of fit measures are presented in Table 1.

Table 1. Estimation result of LA/AIDS model

Variable	Estimates of parameters in equation:							
	1	2	3	4	5	6	7	8
z_1	-0,0575*	0,0760*	-0,0371*	0,0066*	0,0165*	0,0223*	-0,0303*	0,0033*
	(0,0014)	(0,0021)	(0,0013)	(0,0006)	(0,0009)	(0,0011)	(0,0009)	(0,0006)
z_2	0,0018*	-0,0001	-0,0003	-0,0003	-0,0005*	-0,0019*	0,0010*	0,0003
	(0,0006)	(0,0001)	(0,0003)	(0,0002)	(0,0002)	(0,0005)	(0,0004)	(0,0002)
log(x/P)	-0,0310*	0,0416*	-0,0135*	-0,0059*	0,0046*	0,0051*	0,0008*	0,0684*
	(0,0006)	(0,0009)	(0,0006)	(0,0002)	(0,0004)	(0,0005)	(0,0004)	(0,0005)
$\log p_1$	0,1083*	-0,0653*	0,2459*	-0,2173*	0,0083*	-0,0359*	0,0130*	-0,0569*
	(0,0351)	(0,0094)	(0,0179)	(0,0100)	(0,0036)	(0,0021)	(0,0086)	(0,0161)
$\log p_2$	-0,0653*	0,5870*	-0,7930*	0,4655*	-0,0193*	0,0541*	-0,2102*	-0,0188*
	(0,0094)	(0,0114)	(0,0083)	(0,0043)	(0,0036)	(0,0025)	(0,0050)	(0,0042))
$\log p_3$	0,2459*	-0,7930*	0,8180*	-0,5008*	-0,0686*	-0,0462*	0,3024*	0,0423*
	(0,0179)	(0,0083)	(0,0160)	(0,0071)	(0,0031)	(0,0019)	(0,0066)	(0,0082)
$\log p_4$	-0,2173*	0,4655*	-0,5008*	0,4775*	0,0578*	0,0121*	-0,3176*	0,0227*
	(0,0054)	(0,0043)	(0,0071)	(0,0060)	(0,0015)	(0,0008)	(0,0042)	(0,0049)
$\log p_5$	0,0083*	-0,0193*	-0,0686*	0,0578*	[0,0115*	0,0370*	0,0057*	-0,0093*
	(0,0036)	(0,0036)	(0,0031)	(0,0015)	(0,0022)	(0,0011)	(0,0020)	(0,0015)
$\log p_6$	-0,0359*	0,0541*	-0,0462*	0,0121	0,0370*	0,0053*	-0,0263*	0,0001
	(0,0021)	(0,0025)	(0,0019)	(0,0008)	(0,0011)	(0,0014)	(0,0012)	(0,0002)
$\log p_7$	0,0130	-0,2102*	0,3024*	-0,3176*	0,0057*	-0,0263*	0,2414*	-0,0083*
	(0,0086)	(0,0050)	(0,0066)	(0,0042)	(0,0020)	(0,0012)	(0,0048)	(0,0043)
$\log p_8$	-0,0569*	-0,0188*	0,0423*	0,0227*	-0,0093*	0,0001	-0,0083*	0,0283*
	(0,0161)	(0,0042)	(0,0082)	(0,0049)	(0,0015)	(0,0002)	(0,0043)	(0,0088)
const	0,4288*	0,0100	0,2804*	0,0861*	0,0133*	0,0519*	0,0842*	0,0453*
	(0,0040)	(0,0061)	(0,0038)	(0,0017)	(0,0027)	(0,0034)	(0,0027)	(0,0017)
	Goodness of fit							
R^2	0,0680	0,0736	0,1636	0,1865	0,0038	0,0269	0,0139	0,0829

Source: the author's own computations in the STATA statistical package; z_1 denotes ratio number of adults, z_2 – logarithm of household size; standard error in parentheses; asterisk indicates significance at 0,05; food products are following: 1) bread and cereals, 2) meat and fish, 3) milk, cheese and eggs, 4) oils and fats, 5) fruit, 6) vegetables, 7) sugar, jam, honey, chocolate and confectionery, 8) other food products

[7] This is consistent with the OECD equivalence scale in which members of household aged less than 14 are considered children and members aged 14 and over - adults.
[8] STATA's *sureg* command was used for the LA/AIDS estimation.

The log likelihood ratio test[9] and the t-test show that the inclusion of a demographic variables was justified. The implication is that various types of a household, the different composition and the age structure, have an impact on food demand. Table 2 presents results of total food expenditure elasticities[10].

Table 2. Estimated food expenditure elasticities for chosen types of households[11]

	One adult without children	One adult + one member aged 14 and over Number of children (aged below 14) 0	1	2
bread and cereals	0,8508 [0,8442;0,8574]	0,8346 [0,8273;0,8419]	0,8669 [0,8610;0,8728]	0,8586 [0,8529;0,8648]
meat and fish	1,1483 [1,1416;1,1549]	1,1283 [1,1226;1,1341]	1,1583 [1,1512;1,1654]	1,1406 [1,1343;1,1469]
milk, cheese and eggs	0,9261 [0,9190;0,9331]	0,9189 [0,9111;0,9266]	0,9313 [0,9247;0,9378]	0,9279 [0,9211;0,9348]
oils and fats	0,8997 [0,8902;0,9093]	0,9018 [0,8924;0,9111]	0,8884 [0,8778;0,8991]	0,8934 [0,8837;0,9041]
fruit	1,0508 [1,0390;1,0625]	1,0567 [1,0436;1,0699]	1,0636 [1,0489;1,0783]	1,0690 [1,0530;1,0849]
vegetables	1,0317 [1,0217;1,0417]	1,0306 [1,0210;1,0402]	1,0323 [1,0222;1,0425]	1,0357 [1,0245;1,0470]
sugar, jam, honey, chocolate and confectionery	1,0011 [0,9889;1,0134]	1,0013 [0,9874;1,0402]	1,0010 [0,9904;1,0116]	1,0011 [0,9898;1,0123]
other food products	0,8976 [0,8841;0,9111]	0,8952 [0,8814;0,9090]	0,8784 [0,8623;0,8944]	0,8885 [0,8738;0,9032]

Source: the author's own computations in the STATA statistical package. The values included **in parentheses** are the **confidential intervals** (95%) of the food expenditure elasticities

The results presented in Table 2 reveal that all the food groups are fairly sensitive to the food expenditure changes. Moreover, the elasticities can vary according to the different demographic profiles that exist in the population[12].

The study has found some differences in the elasticity estimates for various demographic types of household. For example, the elasticity of meat and fish for households with two adults significantly differs from other considered types –

[9] The null hypothesis of overall absence of demographic effects is strongly rejected – LR=4433 and critical value $\chi^2(0,05;14) = 24$.

[10] For reasons of space, results of the price elasticities are not presented.

[11] All elasticities are evaluated at group's means.

[12] It should be noted that the demand for food in Poland, apart from demographic attributes of households, probably depends on other numerous relevant attributes of households such as, for example place of living, educational attainments and occupation of his members. The impact of such features on food demand in Poland can form a subject of research in the future.

households consisted of: a single person and two adults with one child or two children.

CONCLUDING REMARKS

The demographic variables played a major role in the analysis of the household budget data. Instead of assuming that all the households in the sample have identical tastes, only those with the same demographic profiles are assumed to have the same demand functions. The household size and its composition have been used as the demographic variables in demand studies, although seldom in the context of complete system of demand equations. In this study, we used the "demographic translation" procedure to incorporate demographic variables into the demand system. This procedure allows to explain the heterogeneous nature in the household consumption patterns.

Eventually, the results provide more insight for the understanding of the household consumption habits in Poland. It can be valuable for marketing strategies because a strong segmentation among households provides a more comprehensive picture of the food expenditures. In conclusion, this study indicates that the changing demographic profile of consumers in Poland has had a significant impact on food demand.

REFERENCES

Asche F., Wessells C. R. (1997) On Price Indices in the Almost Ideal Demand System, American Journal of Agricultural Economics, vol. 74, pp. 1182-1184.

Banks J., Blundell R., Lewbel A. (1997) Quadratic Engel Curves and Consumer Demand, The Review of Economics and Statistics, Vol. 74, pp. 527-539.

Barnett W. A., Seck O. (2008) Rotterdam Model versus Almost Ideal Demand System: Will the Best Specification Please Stand up?, Journal of Applied Econometrics, Vol. 23, pp. 699-728.

Barnett W. A., Serletis A. (2008) Consumer Preferences and Demand Systems, Journal of Econometrics, Vol. 147, No. 2, pp. 210-224.

Barten A. P. (1964) Family Composition, Prices and Expenditure Patterns, in: P.E. Hart, F.Mills, J.K. Whitaker (eds.), Econometric Analysis for National Economic Planning, Butterworths, London.

Deaton A. Muellbauer J. (1980) An Almost Ideal Demand System, American Economic Review, Vol. 70, No 3, pp. 312-316.

Dudek H. (2008) Price Elasticities of Food Demand - Analysis of LA/AIDS Model, Annals of the Polish Association of Agricultural and Agribusiness Economists, Vol. 10, No. 4, pp. 62-67 (in Polish).

Edgerton D. L., Assarsson B., Hummelmose A., Laurila I P., Rickertsen K., Vale P. H. (1996) The Econometrics of Demand Systems - With Applications to Food Demand in the Nordic Countries, Kluwer Academic Publisher, Dordrecht-Boston-London.

Gorman W. (1976) Tricks with Utility Functions, in: M. Artis, A. Nobay (eds.), Essays in Economic Analysis, Cambridge University Press, Cambridge.

Green R. Alston J. M. (1990) Elasticities in AIDS Models, American Journal of Agricultural Economics, Vol. 72, pp. 442-445.
Jorgenson D. W., Lau L. J., Stoker T. M. (1982) The Transcedental Logarithmic Model of Aggregate Consumer Behavior, in: R. Baumann, G. Rhodes, (eds.), Advances in Econometrics, JAI Press, Vol. 1, pp. 97-238.
Leser C. E. (1963) Forms of Engel Functions, Econometrica, Vol. 31, No. 4, pp. 694-703.
Moschini G. (1995) Units of Measurement and the Stone Index in Demand System, American Journal of Agricultural Economics, Vol. 77, No. 1, 63-68.
Pollak R. A., Wales T. J. (1979) Welfare Comparison and Equivalence Scales, American Economic Review, Vol. 69, pp. 216-221.
Pollak R. A., Wales T. J. (1981) Demographic Variables in Demand Analysis, Econometrica, Vol. 49, No. 6, pp. 1533-1551.
Ray R. (1983) Measuring the Cost of Children: an Alternative Approach, Journal of Public Economics, Vol. 22, pp. 89-102.
Regorsek D., Erjavec E. (2007) Food Demand in Slovenia, Acta Agriculturae Slovenica, Vol. 89, No. 1, pp. 221-232.
Stone J. R. N. (1954) Linear Expenditure Systems and Demand Analysis: an Application to the Pattern of British Demand, Economic Journal, Vol. LXIV, No. 255, pp. 511-527.
Suchecki B. (2006) Complete Demand Systems, Polskie Wydawnictwo Ekonomiczne, Warszawa (in Polish).
Varian H. R. (1992) Microeconomic Analysis, W.W. Norton&Company, New York.
Working H. (1943) Statistical Laws of Family Expenditure, Journal of the American Statistical Association, Vol. 38, pp. 43-56.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XI, No. 1, 2010, pp. 70- 80

AN APPLICATION OF BRANCHING PROCESSES IN STOCHASTIC MODELING OF ECONOMIC DEVELOPMENT

Marcin Dudziński, Konrad Furmańczyk, Marek Kociński, Krystyna Twardowska
Department of Applied Mathematics, Warsaw University of Life Sciences
e-mails: marcin_dudzinski@sggw.pl; konrad_furmanczyk@sggw.pl; marek_kocinski@sggw.pl; krystyna_twardowska@sggw.pl

Abstract. In our paper, a stochastic model of forecasting of the number of firms of a given type, acting on the market in a given year, is proposed. The model uses the probabilistic tools of the theory of branching processes. Our approach is an alternative method to the forecasting methods proposed so far, including those based on time series. The theoretical results presented in the paper may be applied in the forecasting of the market position of the firms of a given sector.

Key words: branching processes, moment generating function, forecasting of financial positions of firms.

INTRODUCTION

Forecasting of the number of firms - preliminaries

The forecasting of economic events belongs to the most important tasks of contemporary econometrics. Accurate econometric predictions help the company management and governmental authorities in taking good finacial decisions and solutions. A lot of methods of financial forecasting have been worked out so far (see [Cieślak 2001], [Gajek and Kałuszka 1999] and [Harvey 1989] among others). In our paper, we propose a new approach to this topic. It is based on the probabilistic model, while the starting point for the earlier proposed methods was an appropriate econometric function or some time series model. A probabilistic treatment of the problem of forecasting of economic phenomena is possible, if, except for the realizations of the random process, something is also known about the elements, which generate this process. In some stages of the construction of our forecasting model we also use some econometric models, but we avoid the situation, when some equalities appear a priori, without any explanation. The important features of our approach are the following: 1) it gives the possibility of economic interpretation of the obtained parameters, 2) it enables to look at the problems of economic forecasting from a different point of view than the methods proposed so far.

Analysis of the dynamics of the number of firms may be useful in evaluation of the current situation on the labour market, as well as in forecasting of its development in the future. The decrease in the dynamics of the number of firms may be caused by: too high taxes, strong market competition, bureaucracy, unclear law or financial regulations, lack of development plan or financial liquidity of firms, unsettled political situation. On the other hand, an increase in the dynamics of the number of firms may be caused by: the growing share of the private sector in the economy structure, an increase of export, an increase of activity of local communities. Such a variety of factors, which slow down or stimulate the process of creation of a new firm makes the forecasting of the number of firms fairly difficult.

There are not many publications concerning the issue of forecasting of the number of firms. One of a few exceptions is the paper of [Chybalski], which is devoted to the forecasting of the number of small and medium-sized enterprises in Poland.

In our work, we propose the model based on the generalization of the so-called branching processes. The definitions of the branching process and their certain generalization are given in section 2 of our paper. This section contains also the definition of generating function, as some properties of generating functions will be applied in constructing of our model. In section 3, we present the main goals of our investigations, as well as the proposed model of forecasting of the number of firms. In section 4, we describe the estimation procedure applied in the estimation of the parameters of our model and present the results of the forecasts obtained with the use of this model. In section 5, we compare these results with the results of predictions received by applying of some other time series models; our final conclusions are also included here.

BRANCHING PROCESSES AND GENERATING FUNCTIONS

Let us consider the following population. Suppose that at the beginning (time 0) it has c elements (individuals) and each element changes into the new element (or elements), so at time n we have the n th generation of elements. We assume that, for every n, the individuals from the n th generation change independently into the new individuals - called descendants - and these new individuals form the generation numbered by $n+1$. In addition, we assume that the individual disappears, if it has no offspring in the subsequent generation. Let us define the random variables (r.v.'s) Y_n and ξ_{in} as follows: Y_n - the number of individuals in the n th generation, $Y_0 = c$, and ξ_{in} - the number of descendants of the i th individual from the n th generation, $i = 1,2,...,Y_n$. We assume that $\{\xi_{in}$: $i = 1,2,...,Y_n$; $n = 1,2,...\}$ are independent, identically distributed (i.i.d.) r.v.'s, and $\forall i,n\ P(\xi_{in} = j) = b_j, j = 0,1,...,$ where b_j - the probability that the individual will change into j descendants (in particular, b_0 denotes the probability that the individual will have no offspring and disappear). Then, the number of individuals in the generation $n+1$ is given by $Y_{n+1} = \begin{cases} \sum_{i=1}^{Y_n} \xi_{in}, & \text{if } Y_n > 0, \\ 0, & \text{if } Y_n = 0. \end{cases}$

We call $(Y_n)_{n=0}^{\infty}$ the standard Bienayme-Galton-Watson process. It describes the development of descendents of c ancestors. We assume here that the individuals change into the new ones independently of each other and, for any n, i, the r.v.'s ξ_{in} are independent r.v.'s from a certain common distribution.

It seems that branching processes have not been widely used in economic studies so far. We partly try to fill this gap and show that a certain generalization of the Bienayme-Galton-Watson branching process may be useful in economic forecasting.

We now introduce some generalization of the metioned branching process. Let us allow the situation, when the individual may exist longer than one unit of time and that it can have descendants at the different moments in time. Then, Y_n - the number of individuals at the moment n (in the n th generation) - is given by $Y_n = \sum_{k=0}^{n} f_n^k$, where f_n^k - the number of individuals in the n th year existing from k years. Furthermore, the number of new individuals in the generation numbered by $n+1$ is described by $f_{n+1}^0 = \sum_{i=1}^{Y_n} \xi_{in}$, and the number of all individuals in this generation is expressed by $Y_{n+1} = \sum_{k=1}^{n+1} f_{n+1}^k + \sum_{i=1}^{Y_n} \xi_{in}$. In order to find the distribution of the random variable Y_n, the so-called generating function may be used. It is defined as follows: Suppose that X is a discrete random variable with the values in the set of natural numbers. Then, the funcion T, defined on the interval $[-1;1]$, by $T(s) = \sum_{r=0}^{\infty} P(X = r)s^r$ for $|s| \le 1$ is called the generating function of X. Clearly,

$$T(1) = \sum_{r=0}^{\infty} P(X = r) = 1 \text{ and } T'(1) = \sum_{r=0}^{\infty} rP(X = r) = EX. \qquad (1)$$

The properties in (1) will be used in further parts of our paper. For some more informations on the issue of branching processes and their applications, we refer to [Dawidowicz et al. 1995], [Epps 1996] and [Haccou et al. 2005].

FORMULATION OF THE PROBLEM AND THE PROPOSED MODEL

Our goals and empirical data

Our main purpose was to propose the model, which enables to forecast the number of firms. In our considerations we restricted ourselves to the firms of the building sector from the area of the Masovia Province (the Masovian Voivodeship) in Poland. We had the following two reasons for making such a choice: 1) the dynamics of development of the firms from the building sector is sensitive to the changes of both economical and political situation, 2) the Masovia Province belongs to the regions in Poland with the largest number of building companies. We carried out our forecasts for the years 2008 and 2009.

Our database consisted of the informations concerning the number of building companies from the Masovia Province, registered in the National Court Register (*Krajowy Rejestr Sądowy* - KRS) in the period 2001-2009. These informations included the date of registration of the firm in the KRS and the date of declaring bankruptcy. The empirical data are collected in the following table:

Table 1. The numbers of firms and their bankruptcies in the period 2001-2009 (the data marked by * include the number of firms established before 2001)

The year of registration of the firm	The number of firms at the end of 2009	The number of firms, which declared bankruptcy before the end of 2009	The total number of firms
2001	205*	24	229
2002	252*	28	280
2003	129	14	143
2004	144	9	153
2005	37	1	38
2006	24	0	24
2007	45	1	46
2008	44	0	44
2009	36	0	36

More formally, the purposes of our paper may by described as follows:
Let Y_n - the number of building firms from the area of the Masovia Province in the n th year. *Our main goal was to estimate* $E(Y_n)$ *for* $n > t$, *by the use of the historical data from the years* $0,1,...,t$, *where* $E(Y_n)$ *- the expected number of building companies in the Masovia Province in the* n th *year.* We carried out the

forecasts for the years 2008 and 2009 by applying of the historical data from the period 2001-2007. Thus, in our considerations: $n=0$ denoted the year 2001, $n=1$ the year 2002, ..., $n=t=6$ the year 2007, and $n>t$ (i.e., $n=7,8$) denoted the forecasting years 2008, 2009. After calculating (by means of our model) the forecasts of the number of firms for the years 2008 and 2009, we compared these forecasts with the real numbers of firms in those years and with the forecasts obtained by means of some other time series models. We also calculated the relative errors of our forecasts and the forecasts obtained for the other models.

The proposed model of the number of firms

Let: Y_n - the number of firms on the market in the n th year, f_n^k - the number of firms in the n th year existing from k years, $k \geq 0$ (in particular, f_n^0 - the number of new firms in the n th year). Obviously, we have $Y_n = \sum_{k=0}^{n} f_n^k$ and

$$E(Y_n) = \sum_{k=0}^{n} E(f_n^k) = E(f_n^0) + \sum_{k=1}^{n} E(f_n^k). \tag{2}$$

We made the following two complementary assumptions in our investigations: 1) establishing of a new firm may be connected with the existence of some other stable firms in the past (by stable firms we mean the firms existing from at least two years), 2) establishing of a new firm may be caused by some other reasons than the existence of the other firms in the previous years (the reasons for the creation of a new firm may be connected, for example, with the growth of demand for services of a certain type, with the development of high technologies, with the environmental changes, with the changes in law, etc.).

The assumptions above may be described by introducing the following notations: g_n^k ($k=2,...,n$) - the number of new firms in the n th year, the creation of which was connected with the existence of firms existing from $k-1$ years in the year $n-1$, φ_n - the number of new firms in the n th year, the creation of which was not connected with the existence of firms in the year $n-1$. Clearly, we have $f_n^0 = \sum_{k=2}^{n} g_n^k + \varphi_n$, and

$$E(f_n^0) = \sum_{k=2}^{n} E(g_n^k) + E(\varphi_n). \tag{3}$$

Our goal now is to derive the formula for $E(f_n^0)$ in terms of f_{n-1}^{k-1}, $k=1,...,n$. Due to (3), in order to do it, we need to obtain the formulas for $E(g_n^k)$, $E(\varphi_n)$.

Denote by ξ_{ikn} the number of new firms in the n th year, the creation of which was connected with the existence of the i th firm among the firms existing from $k-1$ years in the year $n-1$. We assume that $\{\xi_{ikn}\}$ are i.i.d. and $\forall i,k,n\ P(\xi_{ikn}=j)=b_j, j=0,1,...,$ where b_j - the probability that the number of new firms in the n th year, the creation of which was connected with the existence of the i th firm among the firms existing from $k-1$ years in the year $n-1$ is equal to j. It is clear that $g_n^k=\sum_{i=1}^{f_{n-1}^{k-1}}\xi_{ikn}$. Let: $\hat{B}$ - the generating function of the r.v.'s ξ_{ikn}, G_n^k - the generating function of the r.v. g_n^k. Then, the conditional generating function of the r.v. g_n^k given the event $f_{n-1}^{k-1}=r$ is given by $G_n^k\left(x \mid f_{n-1}^{k-1}=r\right)=\sum_{i=0}^{\infty}P\left(g_n^k=i \mid f_{n-1}^{k-1}=r\right)x^i=\left[\hat{B}(x)\right]^r$, since, under such conditioning, g_n^k is the sum of r i.i.d. r.v.'s ξ_{ikn} with a common generating function $\hat{B}$. The derived formula for $G_n^k\left(x \mid f_{n-1}^{k-1}=r\right)$ implies that

$$G_n^k(x)=\sum_{r=0}^{\infty}G_n^k\left(x \mid f_{n-1}^{k-1}=r\right)P\left(f_{n-1}^{k-1}=r\right)=\sum_{r=0}^{\infty}\left[\hat{B}(x)\right]^r P\left(f_{n-1}^{k-1}=r\right)=F_{n-1}^{k-1}\left(\hat{B}(x)\right), \quad (4)$$

where F_{n-1}^{k-1} - the generating function of f_{n-1}^{k-1}.

In view of (4) and the properties of generating functions in (1), we obtain

$$E(g_n^k)=\left(G_n^k(x)\right)'_{x=1}=\left(F_{n-1}^{k-1}\left(\hat{B}(x)\right)\right)'_{x=1}=\left(\hat{B}'(1)\right)\left(\left(F_{n-1}^{k-1}\right)'\left(\hat{B}(1)\right)\right)=\left(\hat{B}'(1)\right)\left(\left(F_{n-1}^{k-1}\right)'(1)\right)$$

$$=\hat{B}'(1)E(f_{n-1}^{k-1})=bE(f_{n-1}^{k-1}),\ \text{ where } b=E\left(\xi_{ikn}\right)=\sum_{j=0}^{\infty}jb_j. \quad (5),$$

Our next task is to derive the formula for $E(\varphi_n)$. Observe that assuming

$$\forall n\ \ P\left(\varphi_n=r \mid f_{n-1}^0=q\right)=e^{-(\alpha q)}\frac{(\alpha q)^r}{r!}, \text{where } \alpha \text{ is a certain parameter,} \quad (6)$$

we have $E\left(\varphi_n \mid f_{n-1}^0=q\right)=\alpha q$, which implies that

$$E(\varphi_n)=\sum_{q=0}^{\infty}E\left(\varphi_n \mid f_{n-1}^0=q\right)P\left(f_{n-1}^0=q\right)=\alpha\sum_{q=0}^{\infty}qP\left(f_{n-1}^0=q\right)=\alpha E(f_{n-1}^0). \quad (7)$$

The relations in (3), (5) and (7) yield

$$E(f_n^0)=\sum_{k=2}^{n}bE\left(f_{n-1}^{k-1}\right)+\alpha E(f_{n-1}^0). \quad (8)$$

Thus, in view of (2), in order to find the recursive formula for $E(Y_n)$, we need to derive the recursive formula for $E\left(f_n^k\right)$, where $k \geq 1$ Let us introduce the following notations:

h_n^k - the number of firms that declared bankruptcy in the n th year, for which it was the k th year of activity, p_{k-1} - the probability that the firm existing from $k-1$ years will declare bankruptcy in the k th year of its activity.

Assuming that the probabilities of bankruptcy of the firms in the k th year of activity are identical for all these firms and the events of bankruptcy of the firms are independent, we may write that

$$\forall n \;\; P(h_n^k = r \mid f_{n-1}^{k-1} = q) = \binom{q}{r} p_{k-1}^r (1 - p_{k-1})^{q-r}. \tag{9}$$

Therefore, $E\left(h_n^k \mid f_{n-1}^{k-1} = q\right) = q p_{k-1}$ and

$$E\left(h_n^k\right) = \sum_{q=0}^{\infty} E(h_n^k \mid f_{n-1}^{k-1} = q) P\left(f_{n-1}^{k-1} = q\right) = p_{k-1} \sum_{q=0}^{\infty} q P\left(f_{n-1}^{k-1} = q\right) = p_{k-1} E\left(f_{n-1}^{k-1}\right). \tag{10}$$

Due to the identity $f_n^k = f_{n-1}^{k-1} - h_n^k$, for $k \geq 1$, and the relation in (10), we have

$$E(f_n^k) = E(f_{n-1}^{k-1}) - E(h_n^k) = E(f_{n-1}^{k-1}) - p_{k-1} E\left(f_{n-1}^{k-1}\right) = (1 - p_{k-1}) E\left(f_{n-1}^{k-1}\right), \text{ if } k \geq 1. \tag{11}$$

By (2), (8) and (11), we conclude that

$$E(Y_n) = \sum_{k=2}^{n} (b + 1 - p_{k-1}) E\left(f_{n-1}^{k-1}\right) + (\alpha + 1 - p_0) E(f_{n-1}^0), \tag{12}$$

where $b = E(\xi_{ikn}) = \sum_{j=0}^{\infty} j b_j$ and α is such as in (6).

The recursive formula in (12), obtained with the use of (6), (9), establishes our model, which enables to estimate the expected number of firms in the n th year.

In the next part of our paper, we will estimate (by the use of the data from Table 1) the parameters of the model in (12), as well as we will present the results of the predictions of the number of building firms in the Masovia Province for the years of 2008 and 2009.

ESTIMATION OF THE MODEL PARAMETERS AND THE RESULTS OF FORECASTS

The table below presents the numbers of firms registered in the KRS, which declared bankruptcy in the subsequent years:

Table 2. The numbers of firms, which declared bankruptcy

The year of registration	'01	'02	'03	'04	'05	'06	'07	'08	'09
The year of bankruptcy									
'10	0	1	0	0	0	0	0	0	0
'09	0	3	0	1	0	0	1	0	0
'08	1	7	3	0	1	0	0	0	0
'07	4	2	2	3	0	0	0	0	0
'06	1	3	1	2	0	0	0	0	0
'05	4	3	2	0	0	0	0	0	0
'04	2	1	2	3	0	0	0	0	0
'03	5	6	4	0	0	0	0	0	0
'02	7	2	0	0	0	0	0	0	0

As we have already mentioned, in order to estimate our model (12), we used as the historical data ($n = 0,1,..., t$, t is the current time) the data from the period 2001-2007 (thus, $n = 0,..., t, t = 6$). After we had estimated the parameters of our model, we calculated the forecasts of the number of firms for the years 2008 and 2009 ($n = 7,8$). At the beginning, we put f_n^k for $E(f_n^k)$ for our historical data. The values of f_n^k for our historical data ($n = 0,...,6$) were as follows:

Table 3. The values of f_n^k for the historical data from the period 2001-2007

$f_0^0 = 229$	-	-	-	-	-	-
$f_1^0 = 278$	$f_1^1 = 222$	-	-	-	-	-
$f_2^0 = 139$	$f_2^1 = 272$	$f_2^2 = 217$	-	-	-	-
$f_3^0 = 150$	$f_3^1 = 137$	$f_3^2 = 271$	$f_3^3 = 215$	-	-	-
$f_4^0 = 38$	$f_4^1 = 150$	$f_4^2 = 135$	$f_4^3 = 268$	$f_4^4 = 211$	-	-
$f_5^0 = 24$	$f_5^1 = 38$	$f_5^2 = 148$	$f_5^3 = 134$	$f_5^4 = 265$	$f_5^5 = 210$	-
$f_6^0 = 46$	$f_6^1 = 24$	$f_6^2 = 38$	$f_6^3 = 145$	$f_6^4 = 132$	$f_6^5 = 263$	$f_6^6 = 206$

We used the data in Table 3 to estimate p_{k-1} - the probability that the firm existing from $k-1$ years will declare bankruptcy in the kth year of its activity. Since, for any $i = k,..., t$; $t = 6$, each quantity h_i^k / f_{i-1}^{k-1} is (on the condition that the kth year of the firm activity is the ith year among the considered years) a natural estimate of p_{k-1}, we put for $\hat{p}_{k-1}$ the average value of these quantities, i.e.,

$$\hat{p}_{k-1} = \frac{1}{6-k+1}\sum_{i=k}^{6} h_i^k / f_{i-1}^{k-1} = \frac{1}{6-k+1}\sum_{i=k}^{6}\left(1 - f_i^k / f_{i-1}^{k-1}\right) \tag{13}$$

The values of $\hat{p}_{k-1}$, obtained from (13) for $k = 1,2,\ldots,6$, were as follows: 0,01109; 0,01083; 0,01199; 0,01491; 0,00614; 0,01905, respectively. It is obvious that, if we estimate p_{k-1} according to the scheme given above, we can estimate p_{k-1} only for $k \leq t$, where $t = 6$ is the current time. There is no method, which enables to estimate p_{k-1} for $k \geq t+1$ only with the use of the historical data. For this reason, we assumed that the position of the firm existsting for $t = 6$ years is so stable that it seemed reasonable to make the additional condition $\hat{p}_{k-1} = \hat{p}_t$ for $k - 1 \geq t = 6$. Thus, in our considerations, we made the assumption $\hat{p}_{k-1} = \hat{p}_5$ for $k - 1 \geq 6$. After we had estimated p_{k-1}, we estimated the parameters b, α of the model (12). Observe that b, α are also the parameters in the equation (8): $E(f_n^0) = b\sum_{k=2}^{n} E\left(f_{n-1}^{k-1}\right) + \alpha E(f_{n-1}^0)$. Therefore, in order to estimate b and α, we applied the following model of multiple regression

$$f_n^0 = bx_{1,n} + \alpha x_{2,n} + \varepsilon_n \text{, where : } x_{1,n} = \sum_{k=2}^{n} f_{n-1}^{k-1},\ x_{2,n} = f_{n-1}^0. \tag{14}$$

By using of the historical data, we obtained the following input data for the estimation of the parameters b and α of the regression function (14)

Table 4. The empirical data for estimation of the regression function (14)

Rok	f_n^0	$x_{1,n} = \sum_{k=2}^{n} f_{n-1}^{k-1}$	$x_{2,n} = f_{n-1}^0$
'03 ($n = 2$)	$139 = f_2^0$	$222 = f_1^1$	$278 = f_1^0$
'04 ($n = 3$)	$150 = f_3^0$	$489 = f_2^1 + f_2^2$	$139 = f_2^0$
'05 ($n = 4$)	$38 = f_4^0$	$623 = f_3^1 + f_3^2 + f_3^3$	$150 = f_3^0$
'06 ($n = 5$)	$24 = f_5^0$	$764 = f_4^1 + f_4^2 + f_4^3 + f_4^4$	$38 = f_4^0$
'07 ($n = 6$)	$46 = f_6^0$	$795 = f_5^1 + f_5^2 + f_5^3 + f_5^4 + f_5^5$	$24 = f_5^0$

By applying of the data from Table 4, we obtained (by the method of least squares) the following estimates for the parameters b and α of the regression function in (14): $\hat{b} = 0{,}025$, $\hat{\alpha} = 0{,}4986$, where the coefficient of determination and the adjusted coefficient of determination were equal to: $R^2 = 0{,}83$, $R^2 adj = 0{,}72$.

Thus, since $\hat{p}_{k-1} = \hat{p}_5$ for $k-1 \geq 6$, we obtained - by the model in (12) - the following predicted value of the number of firms in 2008 ($n = 7$)

$$E(Y_7) = \sum_{k=2}^{7} \left(\hat{b} + 1 - \hat{p}_{k-1}\right) f_6^{k-1} + \left(\hat{\alpha} + 1 - \hat{p}_0\right) f_6^0 = 884{,}11.$$

In order to estimate the number of firms in 2009 ($n = 8$), we needed to obtain the estimates for f_7^0. By (14), we calculated that $\hat{f}_7^0 = \hat{b}\sum_{k=1}^{6} f_6^k + \hat{\alpha} f_6^0 = 43{,}2406$. Next, we calculated $\hat{f}_7^1$ - $\hat{f}_7^7$, from (11). The results were as follows: The values of $\hat{f}_7^k$ for k = 1,2,…,7 were: 45,49; 23,74;

37,54; 142,84; 131,18; 257,99; 202,08. By applying of the estimates for f_7^k , and the assumption $\hat{p}_{k-1} = \hat{p}_5$, for $k-1 \geq 6$, we received the following predicted values of the number of firms in the year 2009 ($n = 8$)

$$E(Y_8) = \sum_{k=2}^{8} \left(\hat{b} + 1 - \hat{p}_{k-1}\right)\hat{f}_7^{k-1} + \left(\hat{\alpha} + 1 - \hat{p}_0\right)\hat{f}_7^0 = 912{,}84.$$

MODEL ASSESSMENT AND FINAL CONCLUSIONS

Below, we present the predicted numbers of building firms from the Masovia Province, obtained with the use of our model and the five chosen time series models (*the real numbers of firms are given in parantheses*), together with the relative errors of all the forecasts (for the details concerning the chosen models see [Gajek and Kałuszka 1999] and [Harvey 1989]):

Table 5. The predicted numbers of firms in 2008 and 2009

Model	Forecasts'08 and '09	Relative errors of forecasts
The proposed model	884,11 (868), 912,84 (916)	1,86%, -0,34%
Linear trend	1022,3 (868), 1129,4 (916)	17,78%, 23,3%
Quadrating trend	793,89 (868), 729,19 (916)	-8,54%, -20,39%
Moving average of rank 2	804,5 (868), 814,75 (916)	-7,32%, -11,05%
Holt's model	920,63 (868), 1015,48 (916)	6,06%, 10,86%
Structural TS model	866,88 (868), 908,76 (916)	-0,13%, -0,79%

We conclude that: 1) in the case of forecasts for the year 2008, the only model, for which the relative error of prediction was smaller than the relative error of forecast obtained with the use of the proposed model was the Structural TS model, 2) the results of forecasts for the year 2009 show that in this case, the relative error of forecast obtained with the use of our model was the smallest of all the calculated errors. The established model may be applied not only to forecast the number of firms. For example, if we add to our data concerning the firms, the informations on the numbers of employees or the paid tax amounts, we may construct models, which enable to forecast the rate of employment or the potential tax revenues to the budget. The results of our predictions seem to be promising and show that the models based on a certain generalization of branching processes may become an interesing alternative for the models so far applied in economic forecasting.

REFERENCES

Cieślak M. (2001) "Prognozowanie gospodarcze. Metody i zastosowanie", PWN.

Chybalski F. "Tendencje rozwojowe sektora MSP w Polsce", available on Internet.

Dawidowicz A. L., Kulczycki P., Tumidajowicz D. (1995) "A stochastic model of the development of alpine rhododendron", Univ. Iagellonicae Acta Math., XXXII, pp. 37-55.

Epps T. W. (1996) "Stock proces as branchimg processes", Commun. Statist. - Stochastic models, 12 (6), pp. 529-558.
Gajek L., Kałuszka M. (1999) "Wnioskowanie statystyczne. Modele i metody", WNT.
Haccou P., Jagers P, Vatutin V. A. (2005) "Branching Processes", Cambridge Studies in Adaptive Dynamics, 5, Cambridge.
Harvey A. C. (1989) "Forecasting, Structural Time Series and the Kalman Filter", Cambridge University Press.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XI, No. 1, 2010, pp. 81-89

EXPECTED SHORTFALL AND HARELL-DAVIS ESTIMATORS OF VALUE-AT-RISK

Leszek Gadomski [1], Vasile Glavan [1, 2]
[1] Collegium Mazovia in Siedlce;
[2] University of Natural Sciences and Humanities in Siedlce, Institute of Mathematics and Physics
e-mails: Leszek.Gadomski@mazovia.edu.pl, vglavan@uph.edu.pl

Abstract: The most widely used estimator for the *Value-at-Risk* is the corresponding order statistic. It relies on a single historic observation date, therefore it can exhibit high variability and provides little information about the distribution of losses around the tail. In this paper we purpose to replace this estimator of *VaR* by an appropriately chosen estimator of the Expected Shortfall. We also consider the Harrel-Davis estimator of *VaR* and give some comparative analysis among these estimators.

Key words: Risk management, tail loss, VaR, Expected Shortfall, Harrel-Davis estimator

INTRODUCTION

Risk measures appeared as a response to the necessity of quantifying the risk of potential losses on some asset, or a portfolio of assets. Among them Value-at-Risk (*VaR*) has become a standard risk measure for financial risk management due to its conceptual simplicity, easy of computation, and ready applications.

Most banks calculate daily 99% confidence interval *VaR* figures. To do this they look at a discrete distribution of simulated revenues. *VaR* at the 99% confidence level is estimated, for example, by the 14th worst loss across 1305 daily observations from 5yr historical data. It relies on a single historic observation date and therefore can exhibit high variability. This both reduces its efficiency and provides little information about the distribution of losses around the tail.

The process of risk management requires not only estimating the *VaR* but also examining the sensitivity of its positions comprising the portfolio. Taking a single order statistic such as the 14th worse loss may be inadequate for this purpose. Computing a weighted average of the dates in the tail will produce more robust risk analysis. The use of quantile estimators ensures a more stable and accurate measure of tail losses and regulatory capital requirement.

This paper offers a coherent estimate for *VaR* and makes no distributional assumptions whatsoever in doing so. One of the major limitations of *VaR*, and it has been severely criticized through the crisis for not being additive. Using *HD* as an estimator for *Var* solves this problem, while still managing to keep *VaR* as a risk measure, which is a hard requirement by regulations and Basel rules for capital calculations. In our opinion, using the *HD* estimator solves the problem of *VaR* not being coherent, while at the same time adheres to Basel rules by keeping *VaR* for capital calculation.

VaR AS RISK MEASURE

Current regulations from finance (Basle II) or insurance (Solvency II) business formulates risk and capital requirements in terms of quantile based measures (see, e.g., Dowd and Blake 2006). The upper quantile of the loss distribution is called Value-at-Risk (*VaR*). In other words, *VaR* is defined as the maximum potential loss in value of a portfolio over a given holding period within a fixed confidence level: the riskier the portfolio, the larger are the minimal losses during the holding period and for a certain probability level.

More formally, given a random variable Y and a probability level $\alpha \in (0,1)$ denote by $Q(Y,\alpha)$ the α-quantile , i.e.,

$Q(Y,\alpha) = \inf\{y \in R, P(Y \le y) > \alpha\} = F_Y^{-1}(\alpha)$, where "minus one" denotes the right-continuous generalized inverse of the cumulative distribution function *F*.

Recall that in the Gaussian model $Y \approx N(\mu,\sigma^2), Q(Y,\alpha) = \mu + u_\alpha \sigma$.

For a confidence level $\alpha \in (0,1)$ the Value-at-Risk at level α for log-returns *X* is defined as

$VaR(X,\alpha) = Q(-X,\alpha)$.

There are some conceptual problems with *VaR* , an important one is that the *Value-at-Risk* disregards any loss beyond the *VaR* level, also called the problem "of the tail risk". As mentioned in [Artzner et al. 1999], *VaR* has major drawback by not being coherent. By a coherent risk measure [Artzner et al. 1999] mean any real-valued function ρ of real-valued random variables *X*, which models the losses, and with the following characteristics:

- $X \ge Y \Rightarrow \rho(X) \le \rho(Y)$ a.s. (*Monotonicity*)
- $\rho(X+Y) \le \rho(X) + \rho(Y)$ (*Subadditivity*)
- $\rho(\lambda X) = \lambda \rho(X)$ (*Positive homogeneity*)
- $\rho(X+\lambda) = \rho(X) - \lambda$ (*Translation equivariance*).

The *Value-at-Risk* does not meet subadditivity in some cases. For a counter-example see, e.g., (Dowd and Blake 2006).

In response to a coherent equivalent to *VaR*, a series of *VaR*-related risk measures were proposed. Among them the Expected Shortfall as an alternative to *VaR* is mentioned [Rockafeller and Uryasev, 2000].

The Expected Shortfall [Acherbi and Tasche, 2002] at level α is defined as

$$ES(X,\alpha) = E(-X \mid -X > VaR(X,\alpha)).$$

It equals the conditional expected loss given that it exceeds $VaR(X,\alpha)$, and is also called *Tail Value-at-Risk* by [Artzner et al.1999], *Conditional Tail Expectation* [Wirch and Hardy, 1999], or *Conditional Value-at-Risk* [Rockafeller and Uryasev, 2000]. An alternative definition of *ES* is the mean of the tail distribution of the *VaR* losses.

EMPIRICAL ESTIMATORS FOR *VaR* AND *ES*

Let $X_{(n)}, X_{(n-1)}, ..., X_{([n\alpha+1])}, X_{([n\alpha])}, ... X_{(2)}, X_{(1)}$ denote the log-returns of a portfolio in the sample period arranged in increasing order. Then, for a sample large enough, the estimator of $VaR(X,\alpha)$, at a given level of confidence will be the statistics $X_{([n\alpha+1])}$. Similarly,

$$EstimateES(X,\alpha) = \frac{X_{(1)} + X_{(2)} + ... + X_{[n\alpha+1]}}{[n\alpha+1]}.$$

The estimator of $VaR(X,\alpha)$, as the corresponding sample statistic, has the advantage of simplicity and no specific distributional assumption. It is an unbiased estimator, but neither efficient, nor consistent. The above mentioned estimator for the Expected Shortfall, is an unbiased, efficient and consistent estimator.

ESTIMATING *VAR* USING *ES*

A different way to estimate the 99% percentile of the distribution is to use an Expected Shortfall approach. As we are not looking at 99% *ES*, but estimating the *VaR* percentile using *ES*, we need to determine the confidence level for which *ES* is equivalent to a 99% *VaR*. This approach has no closed solution and the equivalent confidence interval is dependent on the distribution assumptions of the underlying losses.

Assuming the losses are Normally Distributed

Let: *X ~ N(0,1)* and recall the corresponding cumulative distribution function, and density, respectively:

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{y^2}{2}} dy$$

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \qquad VaR_p(X) = -F^{-1}(p)$$

In this case the Expected Shortfall has the following form:

$$ES(X,p) = E[X | X \le VaR_p] = E[X \le VaR_p] \cdot \frac{1}{P(X \le VaR_p)} =$$

$$\int_{-\infty}^{VaR_p} y \cdot f(y) dy . \frac{1}{F(VaR_p)} = f(VaR_p) \cdot \frac{1}{(1-p)}$$

The problem is to find a probability p for which $ES_p = VaR_{99}$. Since this problem has no closed solution, we have to find a numerical one. The numerical solution p for a $N(0,1)$ distribution is then applied to our discrete distribution, in order to find the number of the worst observations we need to use for the ES_p calculation. The numerical solution is $p = 94.72\%$. 13

If the empirical distribution of our losses has fatter tails than that of the normal distribution, then the actual confidence interval for using *ES* as an estimate for *VaR* will be lower than the one used for the normal distribution. Hence the *ES* using the 'actual' confidence interval is lower than that calculated by using the Gaussian distribution, and hence empirical *VaR*(99%) is lower than *ES*(94.72%). Most studies show that financial time series exhibit fat tail, hence using the normal distribution confidence interval is conservative whenever we have a fat tail.

Examples of other distributions: t-student distribution

Under the assumption that the losses follow a t-student distribution, we have that the equivalent confidence interval *x*, (*ES*(*x*%) = *VaR*(99%)) is lower than 94.72%, the confidence interval for the normal distribution. The confidence level *x* converges towards 94.72% as the degrees of freedom increase, as expected, because the *t*-student distribution converges in distribution to the Gaussian one as the degrees of freedom approach infinity.

[13] Notation: Let N be the total number of observations in our historical window.
Number of worst historical observations to be used for *ES* calculation equals $\text{int}(N(1-p))+1$.

Even though, the family of *t*-student distributions has fatter tails than normal distribution does, it would be unrealistic to assume that the loss distribution follows a t-student distribution. It is widely accepted that empirical loss distribution varies significantly depending on the positions in the portfolio and hence one cannot make reasonable assumptions that it follows a certain t-student distribution with a specific degree of freedom.

Table 1. Confidence intervals vs. degrees of freedom for t-student distributions

Degrees of freedom	Confidence interval *X* such that ES(*X*%) = VaR(99%)
1.1	86.1%
1.5	94.2%
2	96.0%
3	96.7%
4	96.9%
5	97.0%
8	97.2%
15	97.3%
50	97.4%
200	97.4%
1000	97.4%

Source: own calculations

ESTIMATING VAR USING THE HARELL-DAVIS ESTIMATOR

The *Harrel-Davis* quantile estimator was proposed by [Harrell and Davis, 1982]. It makes no assumptions about the underlying loss distribution (just that the observations are i.i.d). It is in general close to an *ES* measure, just that the weights are not a step function, but given by a beta function. The *Harrel-Davis* estimator is in essence the bootstrap estimator of the expected value of the *(n+1)p-th* order statistic, with *p* - the quantile and *n* the sample size. It is based on the fact that as the sample size increases, the expected value of the (*n*+1)*p*-th order statistic converges to the *p* quantile. Another advantage of using the *HD* estimator is that it gives confidence intervals regarding how good the *VaR* estimator is. The *Harrel-Davis* estimators are defined as follows:

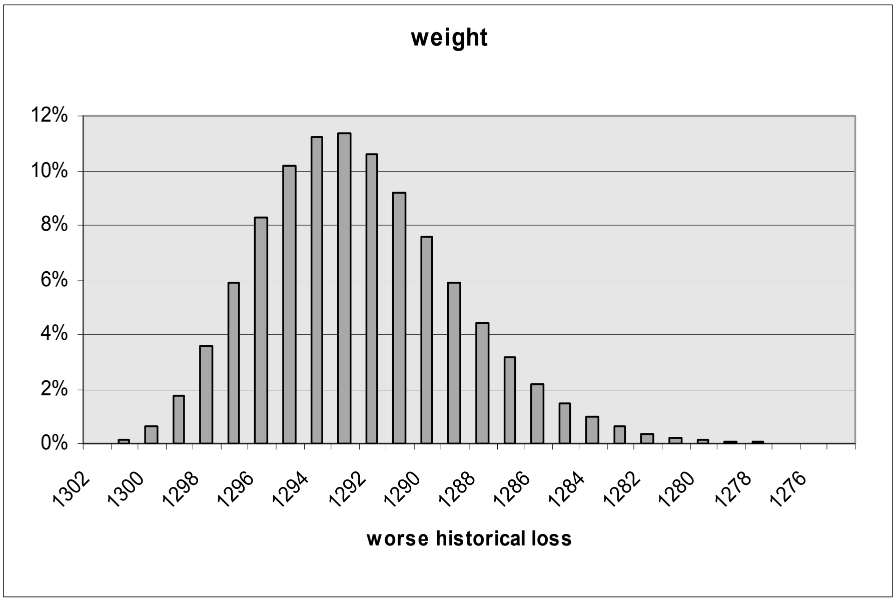
$$HD_y = \sum_{k=1}^{n} w_k x_k \quad \text{with}$$

$$w_k = \frac{1}{\beta\{(n+1)p,(n+1)(1-p)\}} \int_{(k-1)/n}^{k/n} y^{(n+1)p-1}(1-y)^{(n+1)(1-p)-1} dy =$$
$$= I_{k/n}\{(n+1)p,(n+1)(1-p)\} - I_{(k-1)/n}\{(n+1)p,(n+1)(1-p)\}$$

where $I_{k/n}\{(n+1)p,(n+1)(1-p)\}$ is the incomplete beta function.

Figure 1 plots the *HD* weights for estimating the 0.99 quantile from a sample of 1305 observations. Note that unlike the 14th worse loss estimator, which places the total weight on the 1292th order statistic in this case, the *HD* estimator distributes the weights among a range of order statistics. It is worth noting that the weights depend only on the sample size and on the quantile.

Figure1.The weights of the *HD* estimator for p=0.99 and N=1305



Source: own calculations

In what follows we calculate the *VaR*(99%) estimates for some major stock indices, using 1305 daily observations from 5yr historical data taken from Bloomberg data services. The estimates in Table 2 represent *VaR* numbers as a percentage loss of the entire portfolio if we are to hold it entirely in the respective stock / index. The estimates *ES*(94.7%) are the numbers using the *ES* as an estimate, the 14th worse loss represents the actual *VaR* number when we have 5 years history of observations which is 1305 data points as relative returns, and in this case *VaR*(99%) is the 14th worse loss of the empirical distribution, while *HD* is the *VaR* estimate using the Harell-Davis estimator.

Table 2. *VaR*(99%) estimates by ES(94.7%), 14th worse loss and *HD*

ES (94.7%)	-5,17%	-5,03%	-5,91%	-4,42%	-5,47%	-2,50%	-4,57%	-4,36%	-5,34%	-5,93%	-4,75%	-5,00%	-3,20%
14th worse loss	-5,12%	-5,37%	-5,35%	-4,31%	-5,19%	-2,36%	-4,72%	-3,73%	-4,89%	-5,61%	-4,58%	-4,75%	-3,32%
HD Estimate	-5,14%	-5,34%	-5,75%	-4,45%	-5,39%	-2,43%	-4,74%	-3,96%	-5,16%	-6,06%	-4,76%	-4,93%	-3,35%
Stock Indices	AEX	ATG	ATX	BFX	BGLI	BHSE	BMV	BSI	BUX	BVSP	CAC	CCSI	CFG25

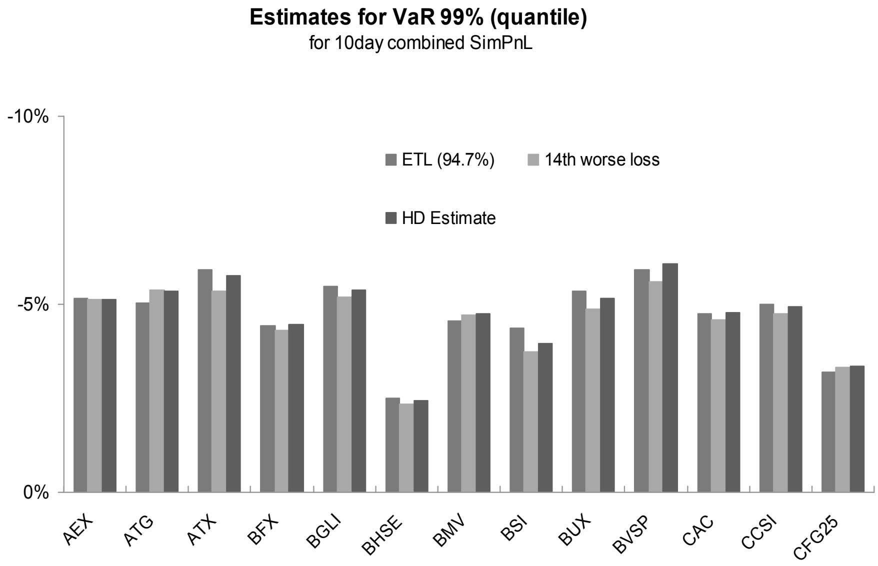
Source: own calculations

The stock indices in Table 2 have the following description:

AEX	Amsterdam Exchanges Index
ATG	Athens General Composite
ATX	Austrian Traded Index
BFX	BEL 20 Index
BGLI	Bulgarian Index (WDR Sofia 30)
BHSE	Bahrain All Share Index
BMV	IPC General Index
BSI	Beirut Stock Index
BUX	Budapest S.E. Index
BVSP	Bovespa Index
CAC	CAC 40 Index
CCSI	Egyptian Stock Index
CFG25	Casablanca 25

The following figure represents graphically table 2. Here *ETL* stands for Expected Tail Loss, another term for *ES*. From this diagram one can observe that, as was mentioned above, the 14th worse loss underestimates the *Value-at-Risk* , while *ES*(94.7%) and *HD* agree quite well as estimators, mainly due to sample size of observed data and to the asymptotic normality of the underlying statistics for calculating Expected Shortfall.

Figure 2. Graphical representation of Table 2



Source: own calculations

CONCLUSIONS

VaR as risk measure together with the corresponding sample statistic as empirical quantile estimator are widely used in the financial risk management, due to their conceptual simplicity, easy of computation and using no specific distributional assumption. At the same time, *VaR* suffers major drawback not taking into account the losses beyond the *VaR* level, on one hand, and not being coherent, on another hand. Also, taking a single order statistic as estimator, it can exhibit high variability. Expected Shortfall avoids these shortcomings and the (uniform) average of data in the tail gives a more stable and accurate estimate of the tail losses. In this paper, for given α, we solve the equation $ES_p = VaR_\alpha$ to find the confidence level *p* with the aim to replace the *[nα+1]* sample statistic by the uniformly averaged data from the tail as estimator for *VaR* at the α confidence level. We also discuss the *Harell-Davis* estimators as *beta*-averaged data from the tail and give some comparative analysis of these three estimators. Using *HD* as an estimator for *Var* solves the problem of *VaR* not being coherent, while at the same time adheres to Basel rules by keeping *VaR* for capital calculation.

REFERENCES

Acherbi C., Tasche D., (2002). On the coherence of expected shortfall. J. Bank. Finance, 26:1487-1503.

Artzner P., Delbaen F., Eber J.M., Heath D. (1999). Coherent measures of risk. Math. Finance, 9:203-228.

Dowd K., Blake D., (2006). After VaR: The theory, estimation, and insurance applications of quantile-based risk measure. J. Risk Insur., vol.73, No.2, 193-229.

Harrell, F.E., Davis, C.E. (1982). A new distribution-free quantile estimator, Biometrika, 69(3): 635-640.

Mausserr H., (2001). Calculating quantile-based risk analytics with L-estimators. Algo Research Quarterly, vol.4, No.4, 33-47.

Rockafeller R., Uryasev S., (2000). Optimization of conditional value-at-risk. J. Risk, 2: 21-41.

Wirch J., Hardy M., (1999). A synthesis of risk measures for capital adequacy. Insur. Math. Econ. 25:337-348.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XI, No. 1, 2010, pp. 90-99

STRUCTURAL CHANGES IN EGGS PRICES IN EUROPEAN UNION MEMBER STATES

Stanisław Jaworski
Department of Econometrics and Statistics
Warsaw University of Life Sciences
e-mail: stanislaw_jaworski@sggw.pl

Abstract: The work relates to changes of the eggs prices in European Union member states since 2004 to 2010. The analysis is based on annually, monthly and weekly average eggs prices. Correspondence analysis is applied to analyze the direction and structure of the changes with reference to all considered states. The unexpected and violent price changes are captured with respect to particular states. Moreover in the reference to chosen states, the model of structural time series analysis is applied to show the price changes in a more detail.

Keywords: Correspondence analysis, Correlation, Dendrogram, Structural time series models.

INTRODUCTION

The work relates to changes of eggs prices in European Union Member States. The prices are collected since 2004 to 2010. The source of the data is the website of Ministry of Agriculture and Rural Development. The principal of the work is to reflect the structure of the price changes in the states.

The problem is the observed price series is multidimensional, as we have 27 states. If any stochastic statistical time series model had been used, the variance-covariance matrix with over 300 various elements would have to be estimated. The length of the collected time series isn't greater than 65 and the task is not feasible. Thus it seems reasonable to use one of the explanatory multidimensional technique and the correspondence analysis is applied in the work.

The bottom line of the work is to compare relative prices, because the mean level of each price series strongly depends on the state. An example of annually averaged eggs prices in two states (in Belgium and Denmark) is given in table 1. The prices are given in EURO/100kg.

Table 1. Annually averaged prices in Denmark and in Belgium

Year	2004	2005	2006	2007	2008	2009
Denmark	136	133	137	142	158	173
Belgium	61	65	77	93	93	103

The prices in Belgium are considerably less than in Denmark in the given consecutive years. The difference is influenced by many factors, which are difficult to capturing and thus the idea of relative prices raised.

DYNAMICS OF ANNUAL CHANGES

General insight in the dynamics of the eggs prices is given in the chapter. The clue is the form of the data given in table 2. The rows of the table consist of the distributions of eggs prices over a period of six years in European Union Member States. Such rows are called row profiles in correspondence analysis.

Table 2. Relative eggs prices over a period of six years.

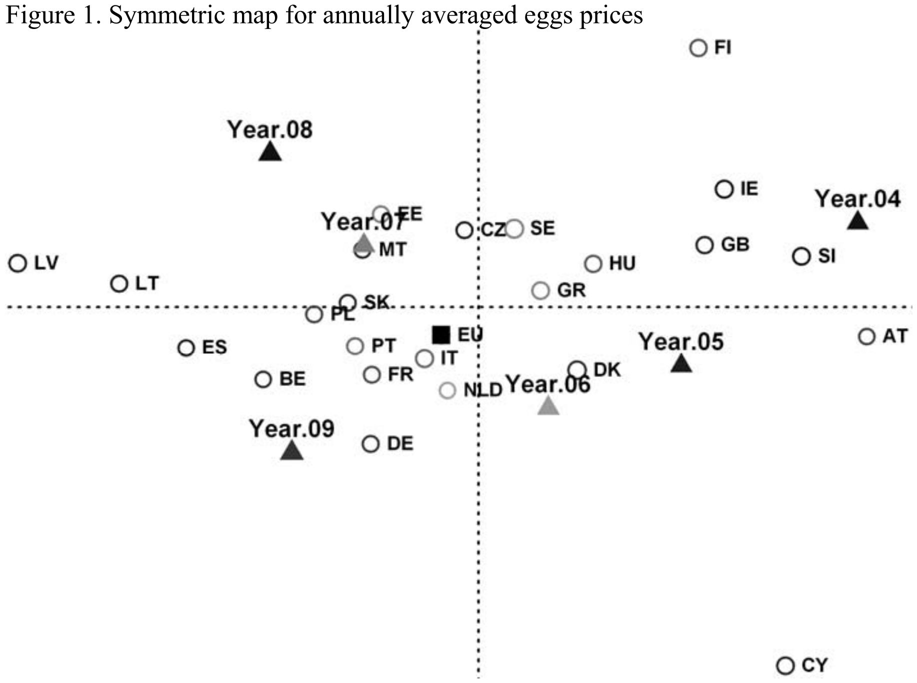
Year	2004	2005	2006	2007	2008	2009
Belgium	12%	13%	16%	19%	19%	21%
Czech Republic	15%	14%	15%	18%	20%	19%
Denmark	15%	15%	16%	16%	18%	20%
- - -	- - -	- - -	- - -	- - -	- - -	- - -
EU	14%	14%	16%	18%	19%	20%
Mean	**15%**	**14%**	**15%**	**17%**	**19%**	**19%**

The mean profile consists of column averages. It's close to European Union profile. Thus it represents the average eggs prices in Europe. The distance between two profiles is measured relatively to the mean profile. An example of the Belgium-Denmark distance is calculated as follows

$$\sqrt{\frac{(0.12-0.15)^2}{0.15}+\frac{(0.13-0.15)^2}{0.14}+\cdots+\frac{(0.21-0.20)^2}{0.19}}$$

The distance is a weighted Euclidean distance. It measures difference between two profiles relative to mean eggs prices in European Union. Large distance between the state profile and the mean profile represents dynamically changing eggs prices in the state. Relatively similar changes of eggs prices in two states imply small distance. The distances between all states can be represented graphically in so called symmetric map (figure 1).

Figure 1. Symmetric map for annually averaged eggs prices



Source: own calculations

The point which represents the European Union profile is close to the origin. The origin represents the mean profile. Arrangement of the points for Italy, Netherlands, Portugal and France suggests that its profiles, that is the distributions of annually averaged eggs prices are close to each other. But it's not the case of Latvia. It's point is far from the origin.

Table 3. Relative eggs prices in Latvia

	R2004	R2005	R2006	R2007	R2008	R2009
Latvia	11%	12%	13%	18%	23%	23%
Mean	**15%**	**14%**	**15%**	**17%**	**19%**	**19%**

The values of Latvian profile are less then values in mean profile in 2004, 2005 and 2006 an then they are greater (table 3). It means that it was a dynamic period of changes of eggs prices in Latvia in comparison with Italy, Netherlands, Portugal or France. The correspondence between the prices of the states can be easily found in figure 1. There are in the figure two clouds of points: the first one for states and the second one for years. The joint display indicates the correspondence between the clouds. Geometrically a particular row profile tends to a position which corresponds to the years which are prominent in that row profile. The purpose is to give the management all the essentials facts in the most concise form, to show the increase or decrease in eggs prices in particular states and their relation to average price in Europe.

Note that Latvia is close to the position of 2008 and 2009. The values in its profile are greater in the years than corresponding values in the mean profile.

In conclusion let us briefly interpret the display: States which are positioned to the right: Austria, Slovenia, Great Britain, Ireland had relatively large prices in 2004 and 2005 and small in 2008 and 2009. States which are to the left: Latvia, Lithuania, Spain and Belgium had relatively large prices in 2008 and 2009. Cyprus had relatively the largest prices in 2006 and 2005. Note that the Germane profile is rather close to the origin.

DYNAMICS OF MONTH CHANGES

Monthly averaged eggs prices provide additional reference to Germany, Austria, Netherlands and Estonia. Conclusion relates to changed arrangement of prices given in table 4. The row profile of the table denotes now price distribution across states in a particular time. Thus we can describe prices' proportion dynamics in a more detail.

Table 4.Price distribution across states

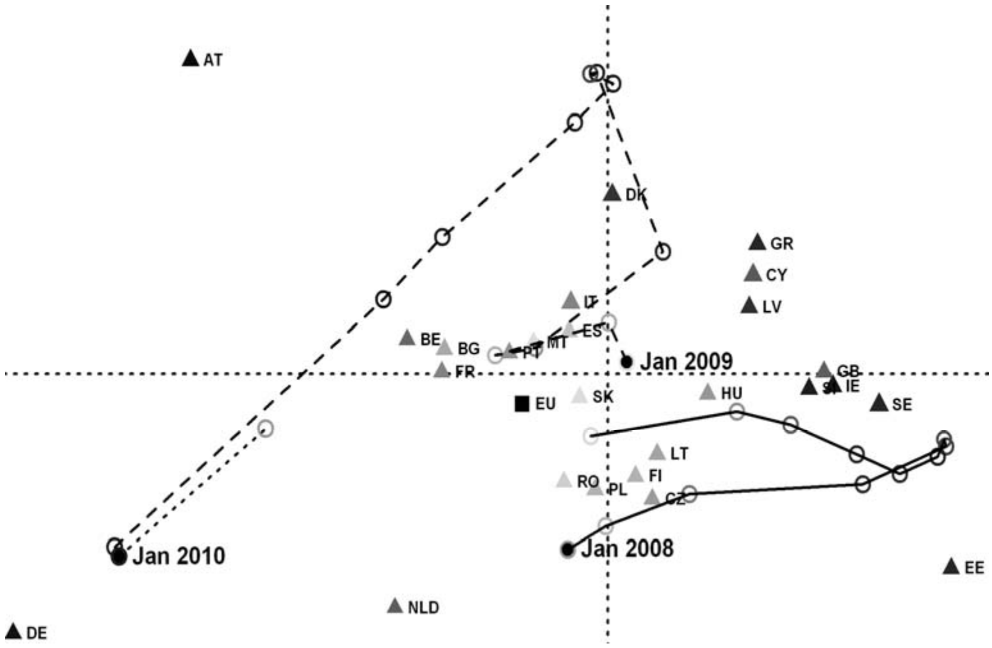
	Belgium	Bulgaria	Czech Republic	- - -	Denmark
Jan 2008	3.2%	3.3%	3.7%	- - -	4.5%
Feb 2008	3.2%	3.3%	3.7%	- - -	4.5%
- - -	- - -	- - -	- - -	- - -	- - -
Feb 2010	3.2%	3.7%	3.2%	- - -	5.0%
Mean	**3.0%**	**3.4%**	**3.2%**	- - -	**5.0%**

Source: own calculations

The profiles of the table 4 are represented by joined with a line points in figure 2. Solid line represents the year 2008, the dashed one the year 2009 and the dotted one the beginning of the year 2010. It reflects the very purpose of the graph. We can see clearly and instantly any upward or downward movement. In particular it allows of several price changes be demonstrated simultaneously in time. Thus we see the difference between the year 2008 and 2009. In the middle of 2008 we have relatively large eggs prices in Estonia, Sweden, Slovenia, Ireland and in Great Britain. Then the prices decreased in 2009.

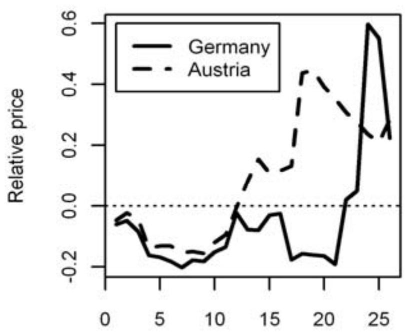
The most dramatic changes in eggs prices occurred in Germany, Austria, Netherlands and Estonia. It is seen that in Germany and Austria the prices increased at the end of the year 2009 and in Estonia in the middle of 2008. The lowest prices in Netherlands occurred in the middle of 2008 and 2009.

Figure 2. Symmetric map for monthly averaged prices



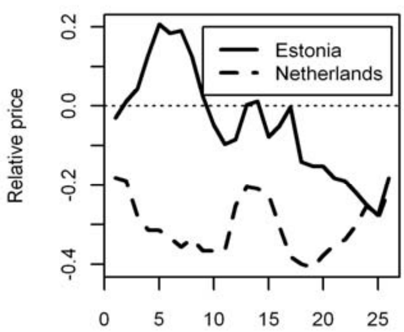
The four States may be examined in a more detail. The figures 3 and 4 reflect the changes. The time factor is shown in the bottom line of each figure and is given in weeks. The price of eggs in Germany and Austria increased over 40% with respect to the mean profile.

Figure 3. Relative prices in Austria and Germany



Source: own calculations

Figure 4. Relative prices in Estonia and Netherlands

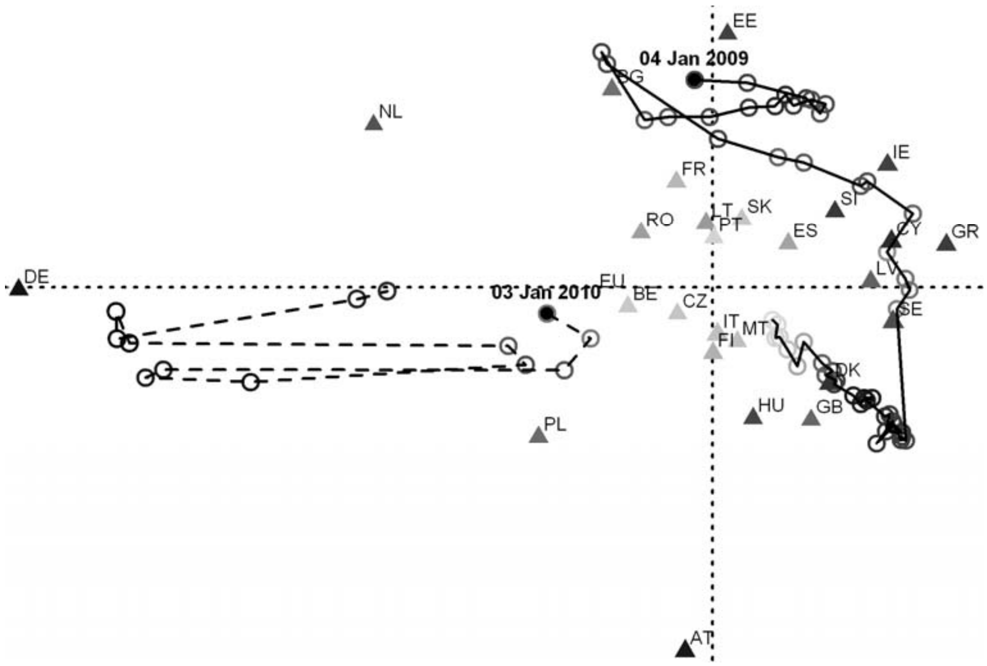


Source: own calculations

DYNAMICS OF WEEK CHANGES

Weekly averaged prices are represented in figure 5. The solid line in figure 5 represents changes in 2009. Relatively large prices occurred in Bulgaria and Estonia at the beginning of the year, in Cyprus, Greece, Lithuania and Sweden in the middle of the year and in Denmark and Great Britain in the end of the year.

Figure 5. Symmetric map for weekly averaged prices

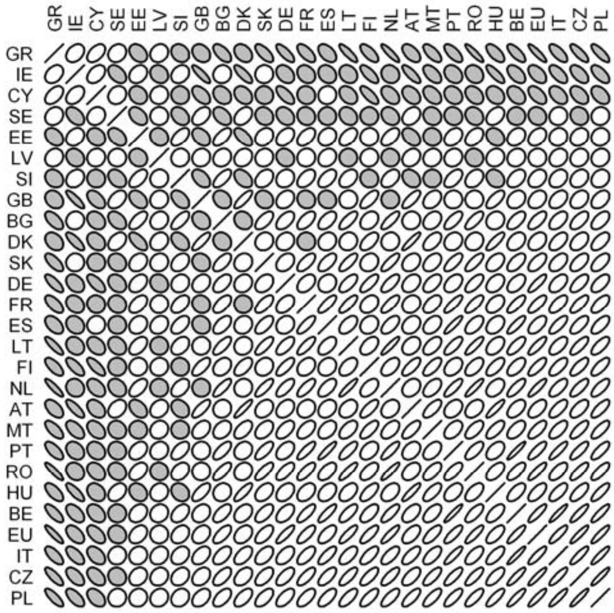


Source: own calculations

The dashed line represents changing prices in 2010. Considerably large changes in eggs prices occurred in Germany at the beginning of the year 2010. The prices were really unstable. The prices have been changing week by week.

The length of weekly averaged price series is equal to 65. It's over two times longer than in the previous case of annually and weekly averaged prices. Thus it gives possibility to construct a reasonable stochastic time series model for some of the series. The idea was to separate simultaneous changes in the series and simple correlations between all pairs of the series was calculated. The appropriate correlation matrix is graphically represented in figure 6.

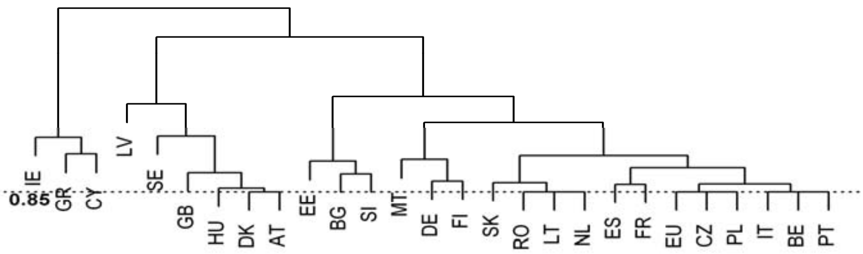
Figure 6. Graphical correlation matrix representation



Source: own calculations

Thin ellipsis represents highly correlated price series. Negative correlation is represented by gray ellipsis and positive by white one. Let's note that the majority of eggs prices in European states are positively correlated. Highly positively correlated price series are represented at the bottom of the dendrogram in figure 7.

Figure 7. Dendrogram

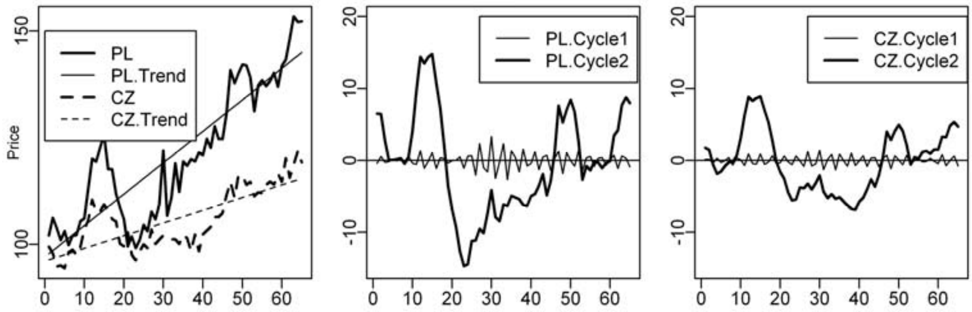


Source: own calculations

Four groups, with simple correlation exceeding 0.85 within each group, were sepqareted. The first group consists of Denmark and Austria, the second: Romania, Latvia and Netherlands, the third: Czech Republic and Poland, and the forth: Italy, Belgium and Portugal.

The three groups of four can be analyzed with the use of structural time series model. The best fit of the model was obtained in the case of Poland and Czech Republic. Note that the eggs prices are highly correlated with the mean prices in Europe. The price series can be separated into deterministic simple trend and two stochastic cycles (figure 8).

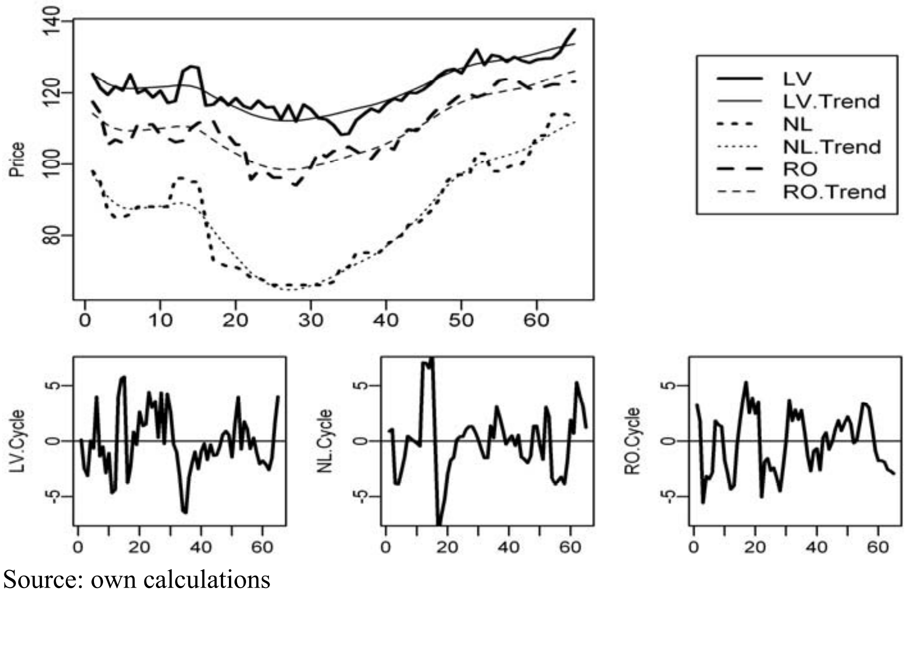
Figure 8. Average prices in Poland and Czech Republic



Source: own calculations

Slightly worse fit was obtained in the rest of the groups. In the case of Latvia, Netherlands an Romania the model consists of stochastic smooth trend and stochastic cycle (figure 9). The trends are similar.

Figure 9. Average prices in Latvia, Netherlands and Romania

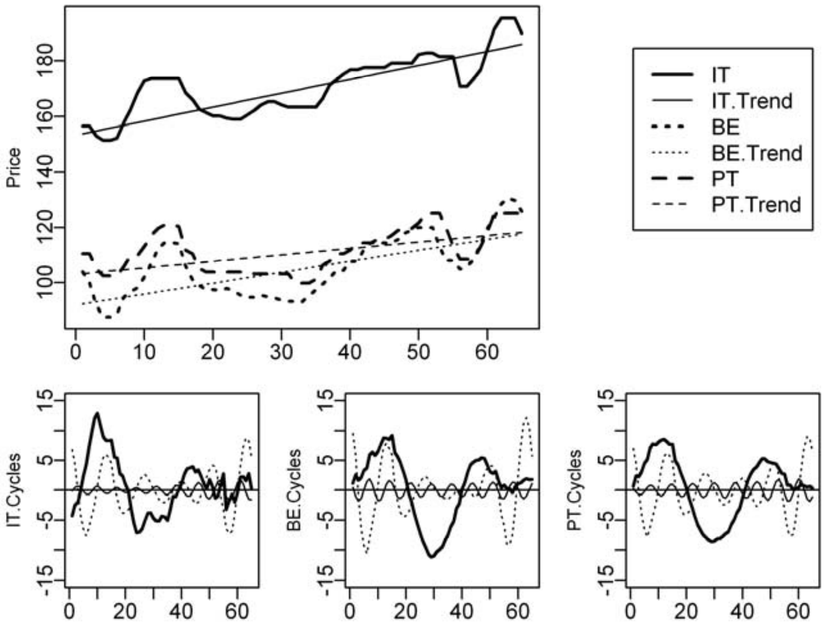


Source: own calculations

The fitted model is not adequately capturing the dynamic structure of the series and the normality assumption is disturbed.

The same problem was found in the group of Italy, Belgium and Portugal. For the group it managed to separate simple trend and three cycles for each series (figure 10). Two of them have relatively high amplitude. Note that the cycles are similar with respect to considered states.

Figure 10. Average prices in Italy, Belgium and Portugal



Source: own calculations

For the last group no model was done, because the series are almost parallel and steady with a one single joint change. Probably the prices are strictly controlled.

THEORETICAL BACKGROUND AND TECHNICAL NOTE

In the work theory of correspondence analysis is used, an approach that has become more used and appreciated over years. The formalization of the analysis can be found in Greenacre (1984) and in the framework of abstract linear algebra in Le Rouan and Rouanet (2004).

Theory of structural time series is used (linear Gaussian form) to separate trends and cycles in the investigated price series. The theory from the standpoint of statistics an econometrics can be found in Harvey(1989) and in Durbin and Koopman (2005).

Necessary calculation are made with R (http://www.r-project.org/) and STAMP (http://stamp-software.com)

REFERENCES

Durbin J., Koopman S.,J. (2005). Time Series Analysis by State Space Methods, Oxford: Oxford University Press.

Greenacre M.J. (1984) Theory and Application of Correspondence Analysis, Academic Press, London.

Harvey A.,C. (1989), Forecasting, Structural Time Series Models and the Kalman Filter, Cambridge: Cambridge University Press.

Le Roux B., Rouanet H. (2004). Geometric Data Analysis: From Correspondence Analysis to Structural Data Analysis, Kluwer Academic Publishers.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XI, No. 1, 2010, pp. 100-109

ASSESSING THE IMPACT OF TRAINING ON UNEMPLOYMENT DURATION USING HAZARD MODELS WITH INSTRUMENTAL VARIABLES[14]

Joanna Landmesser
Department of Econometrics and Statistics
Warsaw University of Life Sciences
e-mail: Joanna_landmesser@sggw.pl

Abstract: The goal of the study is to prove the effectiveness of training programs directed to the unemployed on the local labor market in Poland. We estimate a semiparametric hazard model to assess the impact of training on the individual's unemployment duration. To resolve the potential sample selection problem, the participation in a training program is instrumented using a probit model. The main question of this paper is whether the training significantly raises the transition rate from the unemployment into the employment state.

Key words: program's evaluation, instrumental variable method, hazard models

INTRODUCTION

In 2002, the government of Poland implemented a special program, whose main objective was the vocational activation of people belonging to risk groups in local job markets. The program is administered by the Polish Ministry of Labor and Social Policy and funded by the Labor Fund. The expenditure on the active labor market policy has increased, especially on trainings, apprenticeships and vocational training at the workplace. Program beneficiaries are selected from the unemployed workers who register in the state labor offices.

As far as the impact of the active labor market policy in Poland is concerned, literature is modest. The effectiveness of the policy has been studied by the World

[14] This work was supported by grant No. N N111 209436 from Polish Ministry of Science and Higher Education.

Bank in 1997, as part of a project covered by Czech Republic, Poland, Hungary and Turkey, and by Kluve, Lehmann and Schmidt [Kluve et al. 2000] and also by Puhani [Puhani 1998]. The recent program has been evaluated in 2008 by the Polish Ministry of Labor and Social Policy. The research indicated that only trainings and business incentives increased the chances of finding jobs, whereas programs such as intervention and public works were ineffective [Bukowski 2008]. The part of this research was a microeconometric analysis based on the logit model and propensity score matching.

The selection bias problem, which is crucial for good evaluations of the training's program, was considered by Landmesser [Landmesser 2010]. This evaluation employed matching methods to find a control group for the group of trainees, and it assessed the impact of the vocational training on the unemployment duration using a hazard model. A positive effect of training on reemployment probabilities was found. Although the study was carefully implemented, the method used to control for endogeneity was in this case rough and the matching was imprecise. For every treated one, only one untreated one, that resembled it as much as possible in terms of observable pre-training characteristics, was selected. As a result, too many individuals were excluded from the pooled sample and the method was not enough sufficient for evaluating the impact of the program. The matching methods are not robust against "hidden bias" arising from unobserved variables that simultaneously affect assignment to treatment.

In this article we would like to use an alternative method for evaluating the impact of the vocational training on the duration of unemployment. The goal of the study is to prove the effectiveness of training programs directed to the unemployed on the local labor market in Poland. We identify the effectiveness with the impact of training on chances of finding jobs. To resolve the potential sample selection problem, the participation in a training program is instrumented using a probit model. Then, a semiparametric hazard model is estimated to assess the impact of training on the length of the employment search. We investigate whether the training significantly raises the transition rate from the unemployment into the employment state in the short- and the long-run.

In our research study we try to analyze the situation on the local labor market. Therefore, the study is based on the data obtained from the District Labor Office in Słupsk in Poland from 2000 to 2007.

METHODOLOGICAL CONSIDERATIONS

The aim of the evaluation of training program effects is to assess the difference between the level of the outcome variable (i.e. duration of the unemployment period) at time t for a given participant having received training and the level of that variable at time t for the same individual without participation in the training program [Hujer et al. 1999]. Let Y_i^1 be the unemployment duration after training, and Y_i^0 - the unemployment duration without training. The effect of training for individual *i* is then defined as $Y_i^1 - Y_i^0$. But it is impossible to observe individual treatment effect since we do not know the outcomes for untreated observations when it is under treatment, and for treated when it is not under treatment. We can only observe either Y_i^1 or Y_i^0, never both (compare with [Lalonde, 1986], [Dehejia, Wahba, 1999]).

If the group of treated and the group of untreated are random samples from the population, the outcomes are independent of treatment: $Y^0, Y^1 \perp P$. In such a case, the average treatment effect could be obtained by comparing the expected level of the outcome for the two groups.

For the non-experimental data sets like ours, the independence assumption is not valid and we have to cope with sample selection problem [Heckman, et al. 1998]. The comparison between the outcomes of the two groups requires some assumptions. The conditional independence assumption states that conditional on the relevant covariates X, the outcomes are independent of treatment variable P:

$$Y^0, Y^1 \perp P | X \tag{1}$$

Consider the simple linear model

$$Y = \beta X_Y + \alpha P + u \tag{2}$$

The error term *u* embodies all omitted (observed and unobserved) factors that determine *Y*. If the assumption (1) is not fulfilled, there may be a correlation between the treatment variable *P* and *u*. The variable *P* is then endogenous and OLS gives biased estimates of parameters (selection bias). The solution to the problem, for instance, is to estimate simultaneously the equation for treatment and then the outcome equation. It is also possible, to use instrumental variable (IV) methods to handle endogenous treatment variable. In the IV approach, the participation is substituted with a variable *IP* (an instrument) that is correlated with participation *P* but not with error term *u*. If we denote $\widetilde{Z} = (X_Y, IP)$ and $\widetilde{X} = (X_Y, P)$ the IV estimator for linear model equals $(\widetilde{Z}^T \widetilde{X})^{-1} \widetilde{Z}^T Y$ ([Bowden, Turkington 1984], [Bijwaard 2008]).

To tackle the problem of sample selection in this study, we are substituting participation in the training program *P* with a variable that is correlated with participation but not with error term *u*. We implement probit regressions for men and women separately to analyze the determinants of participation in the training course. For individual *i* the past participation in the training program P_i is defined as:

$$P_i = \begin{cases} 1 & \text{if } P_i^* > 0 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

The latent variable P_i^* is defined as a function of a vector of individual-level variables X_{Yi}, a vector of exogenous variables X_{Pi} and an error component v_i:

$$P_i^* = \gamma X_{Yi} + \delta X_{Pi} + v_i \tag{4}$$

We propose the propensity to participate in a training course as a suitable instrumental variable.

The outcome variable, we are interested in, is the duration of time an individual spends in the state of being unemployed. Therefore, in the next step we analyze the impact of training program on the length of the employment search. An appropriate approach, which considers right censoring of unemployment spells, and which controls characteristics of individuals that influence the unemployment duration, is the use of hazard models (see, e.g. [Kalbfleisch, Prentice, 1980], [Hosmer, Lemeshow, 1999], [Cameron, Trivedi, 2005]).

In the terminology of survival analysis, the survivor function *S*(*t*) is the probability that the length of the unemployment after training exceeds a time point *t* and is defined by

$$S(t) = \Pr[T > t] = 1 - F(t) \tag{5}$$

where *T* is a random variable, which represents the duration in the unemployment state with a density function *f*(*t*).

Given *S*(*t*), the hazard function *h*(*t*) denoting the chance of leaving the unemployment state at time *t* among the individuals who were not yet employed at that time is

$$h(t) = \frac{f(t)}{S(t)} = -\frac{d(\log S(t))}{dt} = \lim_{dt \to 0} \frac{\Pr[t \leq T < t + dt \mid T \geq t]}{dt} \tag{6}$$

In other words, the hazard function *h*(*t*) is the limit of the probability that the unemployment episode is completed during the interval [*t*, *t+dt*], given that it has not been completed before time *t*, for *dt*→0. The hazard rate – the value of hazard function – describes the intensity of transition from one state to another.

The survivor curve can be specified as a function of individual characteristics for unemployed people and the program participation, so that $h = h(t; X_Y, P)$. The widely applied semiparametric method of analyzing the effect of covariates on the hazard rate is the Cox's proportional hazard model [Cox 1972]. In Cox model we have:

$$h(t; X_{Yi}, P_i) = h_0(t)\exp(\beta X_{Yi} + \alpha P_i) \tag{7}$$

Cox proposed a partial maximum likelihood estimation of this model. The model is estimated non-parametrically and there is no need to make assumptions about the baseline hazard $h_0(t)$. It can be stated that in the Cox model the hazard functions for two individuals *i* and *j* are multiplicatively related, that is their ratio is constant. One subject's hazard is a multiplicative replica of another one. If parameter α is positive, the individual receiving a training is likely to find employment before the individual who received no training.

CHARACTERISTICS OF THE DATA SET

The data used in our analysis concern the unemployed registered in the District Labor Office in Słupsk in Poland in the period from January 2000 to August 2007. The selected sample consists of 3513 persons, who were registered as unemployed at least for one day. On the basis of the history of events for each person registered in the labor office we can state the period of time a person was looking for a job or the period of time during which an unemployed is actually looking for a job (in days). The time spent in the unemployment state is called a spell. The spell is completed when the event occurs (finding a job). Otherwise, unemployment spells are treated as right censored. While our data basis contains multiple spells for 3513 persons we have got 6198 episodes. Descriptive statistics for the resulting spell data set can be found in Table 1.

Table 1. Descriptive statistics for the data set

number of:		mean duration in days:	
individuals	3513		
spells	6198	all spells	349,45
censored spells	870	censored spells	714,54
spells of trainees	625	spells of trainees	404,19
spells of non-trainees	5573	spells of non-trainees	343,31

Source: own computations

The participation in a vocational training seems to increase the unemployment duration. However, a simple comparison between the averages has to be done carefully since it is subject to potential selection effects.

To model the participation in a vocational training the set of covariates for the hazard models includes dummy variables capturing the whole, the short-run and the long-run effect of the participation in the training:

tr – participation in a vocational training during the last 3 years prior to the unemployment beginning,

trs – participation in a vocational training during the last 12 months prior to the unemployment beginning, *trl* – participation in a vocational training between 13 and 36 months prior to the unemployment beginning.

EMPIRICAL RESULTS

To tackle the problem of sample selection, we estimate probit regressions for the whole sample and for men and women separately to analyze the determinants of the participation in the training program. The empirical studies on training participation suggest that important determinants of training are: age, sex, caring for children, belonging to minority groups, educational degrees or occupational status (e.g. [Blundell et al. 1994], [Hujer et al. 1999]). Our empirical findings show that the only significant variables for the participation in training are:
period – dummy variable: 1 for the time period 2004-2007 and 0 for the time period 2000-2003 (with this variable we prove the training course availability for the unemployed in the time span),
edu1 – dummy variable: 1 if individual has incomplete primary, primary, lower secondary or basic vocational education level,
edu2 – dummy variable: 1 if individual has general secondary, vocational secondary or post-secondary education level,
edu3 – dummy variable: 1 if individual has tertiary education level,
language – dummy variable: 1 if individual declares any foreign language skills.
The hypothesis that age, sex or the marital status has influence on training participation could not be confirmed. The results of probit models estimation for men and women separately are given in Table 2.

Individual participation in the vocational training is influenced by the variable *period*, which confirms that the variable could be a valid instrumental variable. Individuals who have primary or secondary education levels, in comparison with the tertiary education level, or individuals who declare any foreign language skills tend to participate more.

Table 2. Results of probit models estimation for participation in training during the last 3 years (variable *tr*), last 12 months (*trs*) and between 13 and 36 last months (*trl*)

Covariates	Whole sample		Men		Women	
	Coef.	P>|z|	Coef.	P>|z|	Coef.	P>|z|
Probit regression for variable *tr*						
period	0.329	0.000	0.333	0.000	0.339	0.000
edu1	0.555	0.000	0.794	0.004	0.314	0.085
edu2	0.470	0.001	0.519	0.066	0.462	0.006
language	0.232	0.001	0.230	0.010	0.204	0.058
cons	-2.563	0.000	-2.727	0.000	-2.479	0.000
	No. of obs. = 6,198 Pseudo R2 = 0.029		No. of obs. = 3,247 Pseudo R2 = 0.032		No. of obs. = 2,951 Pseudo R2 = 0.034	
Probit regression for variable *trs*						
period	0.272	0.002	0.233	0.056	0.330	0.011
edu1	0.546	0.006	0.699	0.055	0.397	0.115
edu2	0.421	0.033	0.205	0.598	0.504	0.030
language	0.400	0.000	0.404	0.001	0.371	0.012
cons	-3.005	0.000	-3.092	0.000	-2.973	0.000
	No. of obs. = 6,198 Pseudo R2 = 0.037		No. of obs. = 3,247 Pseudo R2 = 0.043		No. of obs. = 2,951 Pseudo R2 = 0.049	
Probit regression for variable *trl*						
period	0.314	0.000	0.344	0.000	0.284	0.014
edu1	0.460	0.009	0.729	0.037	0.188	0.397
edu2	0.420	0.017	0.616	0.084	0.335	0.103
language	0.073	0.374	0.079	0.453	0.030	0.821
cons	-2.625	0.000	-2.830	0.000	-2.501	0.000
	No. of obs. = 6198 Pseudo R2 = 0.021		No. of obs. = 3247 Pseudo R2 = 0.026		No. of obs. = 2951 Pseudo R2 = 0.019	

Source: own computations using Stata Statistical Software

Now we consider the impact of training on the length of the unemployment duration. The estimated hazard models as a determinant for the probability of leaving the unemployment state will comprise usual socio-demographic characteristics of individuals and variables capturing the effect of participation in training. The additional new covariates are dummies:
age 25 – with 1 if individual is 25 or younger,
age 2640 – with 1 if individual is 26 or older, but younger than 41,
age41 – with 1 if individual is 41 or older,
marr – with 1 if individual is married,
town – with 1 if the place of residence is town,
disabled – with 1 if individual is disabled,
benefit – with 1 if individual receives unemployment benefit.

The results of Cox regressions are given in Table 3. We estimated two types of models: models A with the covariate *tr* for investigation of effects of any training in the past, and models B with covariates *trs* and *trl* for investigation of effects of training in the short- and the long-run.

Table 3. Results of Cox models estimation for participation in training

Models A								
Covariates	Naive				Control			
	Men		Women		Men		Women	
	HR	P>|z|	HR	P>|z|	HR	P>|z|	HR	P>|z|
age25	1,530	0,000	0,882	0,033	1,556	0,000	0,885	0,037
age2640	1,190	0,000	0,984	0,769	1,224	0,000	0,985	0,772
marr	1,316	0,000	0,910	0,035	1,352	0,000	0,914	0,048
edu1	0,699	0,000	0,540	0,000	0,479	0,000	0,526	0,000
edu2	0,844	0,046	0,704	0,000	0,646	0,000	0,662	0,000
town	1,109	0,007	1,053	0,232	1,110	0,007	1,050	0,258
disabled	0,548	0,000	0,788	0,039	0,552	0,000	0,786	0,037
benefit	0,629	0,000	0,630	0,000	0,636	0,000	0,635	0,000
*tr, tr**	1,447	0,000	1,216	0,097	1,827	0,000	1,169	0,157
	No. of obs. = 3,247 ln L = -20,872.4		No. of obs. = 2,951 ln L = -17,076.1		No. of obs. = 3,247 ln L = -20,859.6		No. of obs. = 2,951 ln L = -17,076.4	
Models B								
Covariates	Naive				Control			
	Men		Women		Men		Women	
	HR	P>|z|	HR	P>|z|	HR	P>|z|	HR	P>|z|
age25	1,529	0,000	0,881	0,031	1,541	0,000	0,848	0,006
age2640	1,190	0,000	0,982	0,734	1,219	0,000	0,980	0,713
marr	1,315	0,000	0,910	0,036	1,351	0,000	0,919	0,063
edu1	0,698	0,000	0,539	0,000	0,503	0,000	0,556	0,000
edu2	0,844	0,045	0,702	0,000	0,711	0,005	0,710	0,000
town	1,109	0,007	1,052	0,239	1,105	0,010	1,038	0,388
disabled	0,546	0,000	0,786	0,038	0,553	0,000	0,801	0,056
benefit	0,629	0,000	0,630	0,000	0,637	0,000	0,633	0,000
*trs, trs**	1,367	0,045	1,544	0,014	1,499	0,002	1,846	0,000
*trl, trl**	1,492	0,000	1,042	0,791	1,262	0,152	0,436	0,001
	No. of obs. = 3,247 ln L = -20,872.2		No. of obs. = 2,951 ln L = -17,074.7		No. of obs. = 3,247 ln L = -20,859.0		No. of obs. = 2,951 ln L = -17,069.4	

Source: own computations using Stata Statistical Software; HR – hazard rates

The first columns under the "Naive" heading were obtained by using the hazard function $h(t;X_{Yi},P_i)=h_0(t)\exp(\beta X_{Yi}+\alpha P_i)$, where P_i denotes the participation in the training program, and contains estimates for hazard rates. The columns under the "Control" heading were obtained by using instead $h(t;X_{Yi},P_i)=h_0(t)\exp(\beta X_{Yi}+\alpha IP_i)$, where instrument $IP_i=P_i^*$ denotes the index value obtained from the estimation (4) of the probit model. The index values are the right-hand sides of probit equations less the residuals (not the expected probabilities computed using the normal distribution) [Wodon, Minowa 2001].

The age coefficients imply that men older than 41 are at a disadvantage to find a job. Married women are at the disadvantage at the job market, but for the married men the effect is the opposite. Primary or secondary education level has a significant negative effect on the opportunity to break unemployment. The disabled have a significant lower reemployment chance. There is a greater tendency to leave the unemployment state if the registered person receives no unemployment benefit.

The naive estimates indicate that every training during the last three years reduces the length of unemployment. Training in the past has a positive effect on reemployment probabilities; the transition rate for men increases by about 45%, and for women increases by about 22%. There are positive impacts of training in the short-run for both men and women and in the long-run for men only.

When we use the index values from probit models instead of variable *tr* (see estimates in models A under the "Control" heading), we can detect still positive impacts in the case of both men and woman. These effects are greater for men and are smaller for women (although for women not statistically significant). In the short-run we still observe strong positive impacts on employment for both men and women (see models B under the "Control" heading). The recent training seems to provide the unemployed with modern knowledge which positively distinguishes them from the other unemployed when searching for a job. Surprisingly, in the long-run this effect is statistically insignificant for men; for women the impact is significantly negative.

CONCLUSIONS

The goal of our study was to prove the effectiveness of training programs directed to the unemployed on the local labor market. We estimated hazard models to assess the impact of vocational training on the duration of unemployment spells. To resolve the potential sample selection problem, the participation in a training program was instrumented using a probit model. The IV procedure provides an answer to the question to what extent the effect of the training program was a result of the effectiveness of the program and to what extent it was due to the fact that program participants had different characteristics than other unemployed. The empirical results obtained confirm that training courses proved not fully effective.

In particular, our results indicate that, mostly, there is a positive effect of the training in the past on reemployment probabilities. In the short-run this positive effect is much bigger and statistically significant for both men and women.

Surprisingly, we detect a negative impact of training for women in the long run. Such an effect can be called the stigmatization effect. The long-run effect for men is statistically insignificant. Other results obtained for the whole sample - not presented in the article - also show the lack of training impact in the long-run. This is called a deadweight loss effect and it occurs when a training participant would have reached the same result without participating in the program (it will be the case when company hires a subsidized employee but would also do so, if there is no subsidy) [Bukowski 2008].

REFERENCES

Bijwaard G. (2008) Instrumental Variable Estimator for Duration Data, Tinbergen Institute Discussion Paper, No. 032/4.

Blundell R., Dearden L., Meghir C. (1994) The Determinants and Effects of Work Related Training in Britain, Working Paper, Institute for Fiscal Studies, London.

Bowden R.J., Turkington D.A. (1984) Instrumental Variables, Cambridge University Press, Cambridge.

Bukowski M. (ed.) (2008) Employment in Poland 2007: Security on Flexible Labour Market, Ministry of Labour and Social Policy, Warsaw.

Cameron A.C., Trivedi P.K. (2005) Microeconometrics: Methods and Applications, Cambridge University Press, New York.

Cox D.R. (1972) Regression Models and Life Tables (with Discussion), Journal of the Royal Statistical Society, Series B 34.

Dehejia R.H., Wahba S. (1999) Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs, Journal of the American Statistical Association, 94, pp. 1053-1062.

Heckman J.J., Ichimura, H., Smith, J.A. and Todd, P. (1998) Characterising Selection Bias Using Experimental Data, Econometrica, 66(5), pp. 1017-98.

Hosmer D., Lemeshow S. (1999) Applied Survival Analysis: Regression Modeling of Time to Event Data, John Wiley and Sons, New York.

Hujer R., Maurer K.-O., Wellner M. (1999) Estimating the Effect of Vocational Training on Unemployment Duration in West Germany – A Discrete Hazard-Rate Model with Instrumental Variables, Jahrbücher für Nationalölkonomie und Statistik, Vol. 218, No. 1.

Kalbfleisch J., Prentice R. (1980) The Statistical Analysis of Failure Time Data, John Wiley and Sons, New York.

Kluve J., Lehmann H., Schmidt Ch.M. (2000) Disentangling Treatment Effects of Polish Active Labor Market Policies: Evidence from Matched Samples, Centre for Economic Reform and Transformation, Discussion Paper No. 2000/07, Edinburgh.

Lalonde R.J. (1986) Evaluating the Econometric Evaluations of Training Programs, American Economic Review, 76(4), pp. 604–620.

Landmesser J. (2010) Ocena skuteczności szkoleń dla bezrobotnych z obszarów wiejskich, in: Roczniki Naukowe Stowarzyszenia Ekonomistów Rolnictwa i Agrobiznesu, Zeszyt 5, Tom XII, pp. 109-114.

Puhani A.P. (1998) Advantage through Training? A Microeconometric Evaluation of the Employment Effects of Active Labour Market Programmes in Poland, ZEW Discussion Paper No. 98-25.

Wodon Q., Minowa M. (2001) Training for the Urban Unemployed: A Reevaluation of Mexico’s Training Program, Probecat, World Bank Economists' Forum, Washington D.C., pp. 197-215.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XI, No. 2010, pp. 110-117

ON UPPER GAIN BOUND FOR TRADING STRATEGY BASED ON COINTEGRATION

Rafał Łochowski
Department of Mathematical Economics, Warsaw School of Economics
e-mail: rlocho@sgh.waw.pl

Abstract: A long-run trading strategy based on cointegration relationship between prices of two commodities is considered. A linear combination of the prices is assumed to be a stationary AR(1) process. In some range of parameters, AR(1) process is obtained by discrete sampling of Ornstein-Uhlenbeck process. This allows to calculate approximate number of transactions in long run trade horizon and obtain approximate upper bound for possible gain.

Key words and phrases: cointegration, AR(1) process, Ornstein-Uhlenbeck process, trading strategy.

INTRODUCTION

The Engle-Granger [Engle, Granger 1987] idea of cointegration deepened understanding of two central properties of many economic time series – nonstationarity and time-varying volatility. Two nonstationary series may be related so that the values of one of them can not go (after appropriate scaling) too far from the values of the second. This relationship may often be observed for prices of two commodities (e.g. crude oil and heating oil). When we consider series of differences between appropriately scaled prices of such commodities, we observe that it reverts to its mean. In this paper we investigate trading strategy based on this phenomenon. We assume AR(1) structure of the series of differences. However, it often appears that it is much easier to investigate properties of discrete time series through their continuous counterparts - continuous-time stochastic processes. There already exists large literature concerning continuous time ARMA and GARCH processes, also driven by Levy processes and fractionally integrated (see for example [Brockwell, Marquardt 2005]). We will use this approach and find continuous counterpart for AR(1) process - an Ornstein-Uhlenbeck process. Since statistical properties of Ornstein-Uhlenbeck process are subject of interest of many authors. we will be able to calculate all necessary quantities using their results.

We will use results of Thomas [Thomas 1975] and Ricciardi & Sato [Riccardi, Sato 1988] concerning first hitting time of Ornstein-Uhlenbeck process (for more recent survey see [Alili et al 2005]).

Author presumes that the used approach – investigation of properties of discrete time series through their continuous counterparts - may be very useful, since it is much easier to handle with continuous-time stochastic processes than with discrete-time ones. However, this approach shall be used with caution, since we always shall prove that the properties of continuous-time processes are good for approximation of the properties of discrete time series. As far as author knows, first step in this direction for the problem of approximating stopping times of crossing barriers by AR(1) discrete process with stopping times of hitting barriers by Ornstein-Uhlenbeck process was author's paper [Łochowski 2007].

INTRODUCTION OF APPROXIMATE GROSS GAIN

Let $(P_n, n \geq 1)$ and $(Q_n, n \geq 1)$ be two non-stationary time series representing evolution of the prices of futures contracts for two commodities P and Q. We will assume that $(P_n, n \geq 1)$ and $(Q_n, n \geq 1)$ are cointegrated i.e. for some positive α, β the process $R_n = \alpha P_n - \beta Q_n$ is stationary. Moreover, we will assume that it is mean zero AR(1) process, i. e.

$$R_{n+1} = \gamma R_n + Z_n, \tag{1}$$

where $(Z_n, n \geq 1)$ is i.i.d. sequence, independent from R_1, with $Z_1 \sim N(0, \sigma^2)$ ($N(\mu, \sigma^2)$ denotes here normal distribution with mean μ and variance σ^2). An equivalent form of (1) is

$$\Delta R_{n+1} = -(1-\gamma) R_n + \sigma \varepsilon_n, \tag{2}$$

where $\varepsilon_n = Z_n / \sigma, n \geq 1$.

It is easy to see that the stationarity of R_n holds iff $\gamma \in (-1;1)$. Stationarity implies that $R_n \sim N(0, \sigma^2/(1-\gamma^2))$ and $\mathrm{Cov}(R_n, R_{n+h}) = \gamma^h \sigma^2/(1-\gamma^2)$.

From stationarity of R_n one may derive long-run trading strategy based on selling α commodity P contracts and buying β commodity Q contracts when R_n exceeds certain threshold value a and doing opposite, when R_n goes below $-a$. If we enter the market with α contracts of commodity P and are interested in leaving it with the same volume of commodity P contracts after long time horizon

T then the gross gain obtained from the strategy equals $2a \cdot N(a)$, where $N(a)$ denotes the number of pairs of transactions:

- when $R_n = \alpha P_n - \beta Q_n \geq a$ sell α commodity P contracts and simultaneously buy β commodity Q contracts,
- when $R_n = \alpha P_n - \beta Q_n \leq -a$ buy α commodity P contracts and simultaneously sell β commodity Q contracts.

The problems which one faces deciding for the described strategy is the estimation of $N(a)$ and then the choice of an optimal threshold value a.

Firstly we will try to estimate $N(a)$ for large T and positive, fixed (but not too small) a. Let $\tilde{T}_1, \tilde{T}_2$ be the following stopping times

$$\tilde{T}_1 = \inf\{n \,:\, R_n \geq a\},\ \tilde{T}_2 = \inf\{n \geq \tilde{T}_1 \,:\, R_n \leq -a\}.$$

Let us assume that the process R_n has small jumps (much smaller then a) and it is obtained by discrete sampling of a continuous-time process with continuous trajectories, $(U_t, t \geq 0)$, i. e. $R_n = U_n$ for $n = 1, 2, ?$. Let T_1, T_2 be continuous counterparts to $\tilde{T}_1, \tilde{T}_2$, i. e.

$$T_1 = \inf\{t \,:\, U_t \geq a\},\ T_2 = \inf\{t \geq T_1 \,:\, U_t \leq -a\}.$$

In fact, from continuity of U_t we have

$$T_1 = \inf\{t \,:\, U_t = a\},\ T_2 = \inf\{t \geq T_1 \,:\, U_t = -a\}.$$

Let us denote $T(a) = E(T_2 - T_1)$. From the assumption about R_n we get $\tilde{T}_1 \approx T_1, \tilde{T}_2 \approx T_2$ (cf. [Łochowski 2007]). Now from theory of renewal processes (cf. [Rolski et al 1998]) and symmetry of the process R_n we may approximate number of transactions $N(a)$ in long time horizon by $T/\left(2T\left(a\right)\right)$. Thus the gross gain in long time horizon equals $2a \cdot N\left(a\right) \approx a \cdot T/T\left(a\right)$. Based on this reasoning let us define.

Definition. Approximate gross gain, *$AGG(a,T)$, for a threshold value $a > 0$ and time horizon $T > 0$ is defined by the formula*

$$AGG(a,T)=\frac{T\cdot a}{T(a)}.$$

ORNSTEIN-UHLENBECK PROCESS AS A CONTINUOUS VERSION OF AR(1) PROCESS

The natural candidate for process U_t is Ornstein-Uhlenbeck process being a solution of the following stochastic differential equation - the continuous counterpart of (2)

$$dU_t=-(1-\gamma)U_t dt+\sigma dW_t, \tag{3}$$

where W_t denotes a standard Wiener process.

Equation (3) has the following solution

$$U_t=e^{-(1-\gamma)t}U_0+\sigma\int_0^t e^{-(1-\gamma)(t-s)}dW_s.$$

From bilinearity of covariance, independence of increments of Wiener process and then from isometry formula for stochastic integrals we have

$$\begin{aligned}&Cov\left(\int_0^t e^{-(1-\gamma)(t-s)}dW_s,\int_0^u e^{-(1-\gamma)(u-s)}dW_s\right)\\&=e^{-(1-\gamma)(t+u)}Cov\left(\int_0^t e^{(1-\gamma)s}dW_s,\int_0^u e^{(1-\gamma)s}dW_s\right)\\&=e^{-(1-\gamma)(t+u)}\int_0^{\min(t,u)}e^{2(1-\gamma)s}ds=e^{-(1-\gamma)(t+u)}\frac{e^{2(1-\gamma)\min(t,u)}-1}{2(1-\gamma)}.\end{aligned}$$

Assuming that U_0 is independent from $(W_t, t\geq 0)$, we get

$$\mathrm{Cov}(U_t,U_{t+h})=e^{-(1-\gamma)(2t+h)}\mathrm{Var}(U_0)+\sigma^2 e^{-(1-\gamma)(2t+h)}\frac{e^{2(1-\gamma)t}-1}{2(1-\gamma)}.$$

If $U_0\sim N(0,\sigma^2/(2(1-\gamma)))$ then $U_t\sim N(0,\sigma^2/(2(1-\gamma)))$ and $\mathrm{Cov}(U_t,U_{t+h})=\sigma^2 e^{-(1-\gamma)h}/(2(1-\gamma))$.

Comparing the distribution of U_n with the distribution of R_n we see that they are different. But when $\gamma\in(0;1)$, taking process $(V_t, t\geq 0)$ defined by the equation

$$dV_t = -\ln(1/\gamma)V_t dt + \sigma\sqrt{\frac{2\ln(1/\gamma)}{1-\gamma^2}}dW_t$$

with $V_0 \sim N(0,\sigma^2/(1-\gamma^2))$ independent from $(W_t, t \geq 0)$, we obtain such a process that for positive integers $n_1, n_2, ?, n_k$ vector $(V_{n_1}, V_{n_2}, ?, V_{n_k})$ has the same distribution as vector $(R_{n_1}, R_{n_2}, ?, R_{n_k})$.

Remark. Ornstein-Uhlenbeck process may also be introduced as a time-space scaled Wiener process. Defining

$$\tilde{V}_t = \frac{\sigma}{\sqrt{1-\gamma^2}} e^{-\ln(1/\gamma)t} W(e^{2\ln(1/\gamma)t})$$

we get a process with the same finite distributions as the process $(V_t, t \geq 0)$.

UPPER BOUND FOR APPROXIMATE GROSS GAIN

Now we are ready to calculate $T(a)$. Let us denote

$$T_{b,c} := \inf\{t \geq 0 \, : \, V_t = c \,|\, V_0 = b\}.$$

From results of Thomas [Thomas 1975] and Ricciardi & Sato [Riccardi, Sato 1988] as well as from scaling properties of V_t we have that for $a > 0$

$$E(T_{a,0}) = \frac{\sqrt{\pi}}{2\ln(1/\gamma)} \int_0^{a\sqrt{1-\gamma^2}/\sqrt{2\sigma^2}} \left(1 + \operatorname{erf}(t)\right) e^{t^2} dt$$

and

$$E(T_{0,a}) = \frac{\sqrt{\pi}}{2\ln(1/\gamma)} \int_0^{a\sqrt{1-\gamma^2}/\sqrt{2\sigma^2}} \left(1 + \operatorname{erf}(-t)\right) e^{t^2} dt,$$

where $\operatorname{erf}(t)$ stands for error function defined as $\operatorname{erf}(t) = \frac{2}{\sqrt{\pi}} \int_0^t e^{-s^2} ds$. From strong Markov property of V_t and since $\operatorname{erf}(t) + \operatorname{erf}(-t) = 0$

we see that

$$T(a) = E(T_{a,0}) + E(T_{0,-a}) = E(T_{a,0}) + E(T_{0,a})$$
$$= \frac{\sqrt{\pi}}{\ln(1/\gamma)} \int_0^{a\sqrt{1-\gamma^2}/\sqrt{2\sigma^2}} e^{t^2} dt.$$

Since $\int_0^u e^{t^2} dt = \int_0^u \sum_{k=0}^{\infty} \frac{t^{2k}}{k!} dt = \sum_{k=0}^{\infty} \frac{u^{2k+1}}{(2k+1)\cdot k!}$, we have

$$T(a) = \sum_{k=0}^{\infty} d_{2k+1} a^{2k+1},$$

where

$$d_{2k+1} = \frac{\sqrt{\pi}}{\ln(1/\gamma)} \frac{1}{(2k+1)\cdot k!} \left(\frac{1-\gamma^2}{2\sigma^2} \right)^{k+1/2} > 0, \text{ for } k = 0,1,?.$$

Now we have

$$\frac{a}{T(a)} = \frac{a}{d_1 a + d_3 a^3 + \cdots} = \frac{1}{d_1 + d_3 a^2 + \cdots}$$

and we see that approximate gross gain, $AGG(a,T) = \frac{T\cdot a}{T(a)}$, is a decreasing function of a. We also have

$$\sup_{a>0} \frac{a}{T(a)} = \lim_{a?0} \frac{a}{T(a)} = \frac{1}{d_1} = \sqrt{\frac{2}{\pi}} \frac{\ln(1/\gamma)\sigma}{\sqrt{1-\gamma^2}}.$$

The above calculations imply

Theorem. *If* $\gamma \in (0;1)$ *the approximate gross gain,* $AGG(a,T)$, *for any positive* a *and* T *is bounded from above by* T/d_1, *i. e.*

$$AGG(a,T) \le \sqrt{\frac{2}{\pi}} \frac{\ln(1/\gamma)\sigma}{\sqrt{1-\gamma^2}} \cdot T. \qquad (4)$$

Moreover, there is no positive a *for which the above value is attained.*

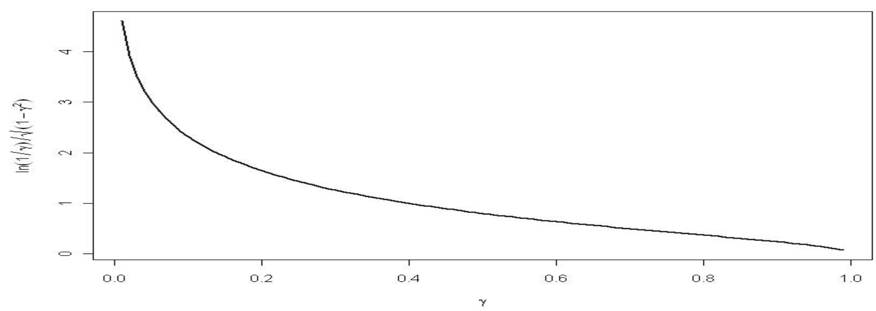
FINAL REMARKS

The results of the previous section are in some sense negative, since there is no optimal postive threshold value a maximizing the approximate gross gain. Moreover, the smaller is $a > 0$, the greater is $AGG(a,T)$. But for very small threshold values the relationships $\tilde{T}_1 \approx T_1, \tilde{T}_2 \approx T_2$ on which our reasoning was based, may fail. Moreover, when a is small we have to change our positions very often which may be not technically possible and if we take into account transaction costs then it would appear that for small a our trading startegy leads to loss.

However, it seems that the formula for upper gain bound in our simple model has natural interpretation in terms of quantities appearing on the right side of (4). The maximal posible gain is proportional to the duration T of the investment and to standard deviation σ of the random term Z_n. It is known phenomenon (used sometimes in so called volatility trading), that the bigger volatility, represented here by σ, the bigger profits in short term are possible.

The more sophisticated seems to be the dependence of $AGG(a,T)$ on parameter γ. This parameter determines the speed of reverting AR(1) process to its mean value. The bigger γ the longer time is needed for AR(1) process to revert to its mean. The dependence between γ and $AGG(a,T)$ is represented by the function $\gamma \mapsto \frac{\ln(1/\gamma)}{\sqrt{1-\gamma^2}}$. This is decreasing function on the interval $(0;1)$ and its graph is presented below.

Figure 1. Graph of the function $\gamma \mapsto \frac{\ln(1/\gamma)}{\sqrt{1-\gamma^2}}$.



REFERENCES

Alili, L., Patie, P. and Pedersen, J. L. (2005). Representation of the first hitting time density of an Ornstein-Uhlenbeck process. Stoch. Models 21 967--980.

Brockwell, P. J. and Marquardt, T. (2005). Lévy driven and fractionally integrated ARMA processes with continuous time parameter Statistica Sinica 15 474--494.

Engle, R. F. and Granger, C. W. J. (1987). Cointegration and Error-Correction: Representation, Estimation and Testing. Econometrica 55 251--276.

Łochowski, R. (2007) Stopping time for Ornstein-Uhlenbeck process vs. stopping time for discrete AR(1) process, reprint available from web page http://akson.sgh.waw.pl/~rlocho/stoppingtime2.pdf

Rolski T., Schmidli H., Schmidt V. and Teugels J. (1998) Stochastic Processes for Insurance and Finance. Wiley, Chichester New York Weinheim Brisbane Singapore Toronto.

Ricciardi, L. M. and Sato, S. (1988). First-passage time density and moments of the Ornstein-Uhlenbeck process. J. Appl. Probab. 25 43--57.

Thomas, M. U. (1975). Some mean first-passage time approximations for the Ornstein-Uhlenbeck process. J. Appl. Probab. 12 600--604.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XI, No. 2010, pp. 118-127

DOES SIMULTANEOUS INVESTING ON DIFFERENT STOCK MARKETS ALLOW TO DIVERSIFY RISK? THE COINTEGRATION ANALYSIS WITH MAIN FOCUS ON WARSAW STOCK EXCHANGE

Anna Barbara Misiuk, Olga Zajkowska
Department of Econometrics and Statistics
Warsaw University of Life Sciences
e-mails: annamisiuk@gmail.com, o.zajkowska@gmail.com

Abstract: This paper aims at examining the bilateral linkage between daily stock market indices, in which the leading index of WSE (WIG20) is the reference. Thus, the study is limited to pairs including WIG20 and indices which are listed on the financial centers of WSE's main foreign investors. The relationship between the markets is investigated throughout the cointegration theory. Further, the Granger causality is carried out in order to distinguish the directions of influence across the stock market environments. The obtained results shall explain the investor's tendencies in portfolio diversification.

Keywords: market stock exchange, stock exchange indices, WIG20, cointegration theory, Granger causality, portfolio diversification.

INTRODUCTION

The international portfolio diversification, as one of general techniques for reducing investment risk, has been among the most celebrated concepts in finance for more than half a century. Starting with the pioneering work in Modern Portfolio Theory presented by Markowitz [Markowitz 1952, 1959] together with later findings of Grubel [Grubel 1968], the concept of modern portfolio analysis was irrevocably extended from domestic to international capital markets. Since then there have been a numerous empirical studies which showed substantial advantages of international diversification.

In spite of different methods applied, the conclusions of early empirical studies, namely [Levy et al., 1970], [Grubel et al., 1971], [Lessard, 1973] and

[Solnik, 1974], were consistent with both Markowitz's and Grubel's predictions [Markowitz, 1952, 1959], [Grubel, 1968]. In fact, each of these studies indicated that movements of stock prices within different countries are characterized by low correlation, hence are almost unrelated to each other. This leads to the conclusions that simultaneous capital investing across these countries can bring benefits.

Nevertheless, it is worth mentioning that the studies cited above were written in times when the movement of capital between market exchanges was relatively limited by political and legal restrictions. Indeed, the relationship between the stock markets started to be visible when these restrictions were lifted with time.

Recently, numerous studies focus on cointegration techniques to investigate the existence of long-run benefits from international diversification, for instance: [Gilmore et al. 2005] and [Voronkova, 2004][15]. These researches examine both bilateral and multilateral cointegration properties. According to these studies, the equity markets move in the same direction in the long-run and therefore there is no advantage of international diversification.

The U.S. market is found to be the dominant market in the financial world. What is interesting, it did not influence the market's behaviour of "new EU members"[16] before its EU enrolment [Gilmore et al. 2002][17], [Kanas, 1998][18]. On the other hand, the long-run linkages between Central European markets[19] and the developed markets in Western Europe and the U.S can be found after the EU accession [Rousova, 2009].

The aim of this paper is to investigate long-run benefits from international portfolio diversification for foreign investors who invest within Warsaw Stock Exchange (WSE).

The question is whether the international investors present on WSE tend to exploit all benefits from portfolio diversification. Specifically, examined is the case of foreign investors who diversify portfolio by holding assets on both, WSE and on his/her home financial centre. The gains of portfolio diversification is investigated by bilateral relationship between Polish and other market. To achieve the goals, the framework of cointegration theory is adopted.

---

[15] [Gilmore et al. 2005] and [Voronkova, 2004] show that process of integration of the Central and Eastern European countries into the EU resulted in their equity markets' comovements with other major EU countries and even with USA stock market (in case of Varonkova).

[16] "new EU members" refer to Central and Eastern European countries which joined the EU on 1st May 2004, that is: Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovakia, Slovenia and the Mediterranean islands of Malta and Cyprus.

[17] Paper shows that Czech Republic, Hungary and Poland were not cointegrated with the U.S equity market during the period from 1995 to 2001.

[18] The results for the period from 3rd January 1983 to 29th November 1996 show that the US market was not pairwise cointegrated with any of the European markets, namely the UK, Germany, France, Switzerland, Italy, and the Netherlands.

[19] that is specifically Czech Republic, Poland and Hungary.

In fact, integration among stock markets is not constant over time mainly due to unexpected events, such as financial crises, for instance. For this reason, the study distinguishes additional period after the financial crisis that started at the end of 2008. This division helps to find out the eventual effects of the financial crisis on cointegration relationship. In other words, the study tries to answer the question whether the crisis caused the contagion or contrary weaken the linkage between the stock markets.

Additionally, the Granger causality is carried out in order to distinguish the directions of influence across the stock market environments.

UNIT ROOT, COINTEGRATION AND GRANGER CAUSALITY

The research on the relationship between the markets is performed with use of cointegration theory. Since the stationarity of time series is the precondition for cointegration analysis, the first step is to test the stationarity of each time series. The Augmented Dickey-Fuller test (ADF), as an extension of the Dickey and Fuller method, is used [Dickey and Fuller, 1979; 1981]. It tests for the presence of unit root, which alternative of stationarity of the series investigated. Additionally, for the cointegration analysis purposes, all the series must be integrated of the same order.

The next step is application of cointegration analysis to test the presence of long-run equilibrium relationships. This paper uses Banerjee, Dolado and Mestre cointegration test [Banerjee et al., 1998]. Null hypothesis is that there is no cointegrating relationship among the variables.

Additionally, the simple Granger causality test is performed in order to distinguish directions of influence across the stock market environments [Granger, 1969]. Granger causality test allows to determine whether one index is useful in forecasting another. If the past values of the index A can be used to predict another index B more accurately than using just the past values of B index, it can be argued that A Granger-cause B. This means that if past values of A statistically improve the prediction of the B, then we can conclude that A Granger-causes B. The null hypothesis assumes no Granger causality and is verified by using F-tests.

In this study, the Granger causality test and cointegration tests are proceeded pairwise with up to 10 lags tested.

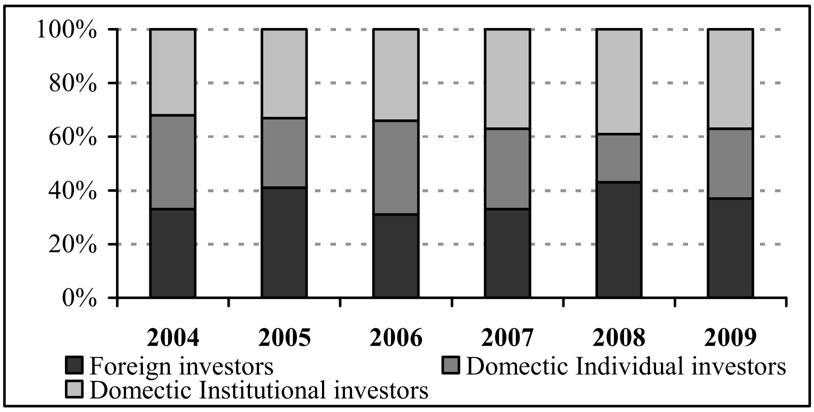
Confirming existence of Granger causality and cointegration can be considered as evidence against portfolio diversification opportunities.

FOREIGN INVESTORS ON WSE

In view of the fact that WSE is of particular interest for this paper analysis, the stock market indices of countries of its main foreign investors are investigated. Thus, the decision on countries chosen as investors' origin was driven primarily by data availability to the researches conducted by WSE. In fact, these regular studies provide information about the structure of its investors[20]. It is worth noticing that foreign investors are identified here with the brokerage houses. Although we realize that nationality of brokerage house's customers does not necessarily coincide with placement of itself, it is the only source of data available.

Over the past years, according to the data published by WSE, foreign investors had the highest share in equity trading, followed by domestic financial institutions and individual investors [see Figure 1].

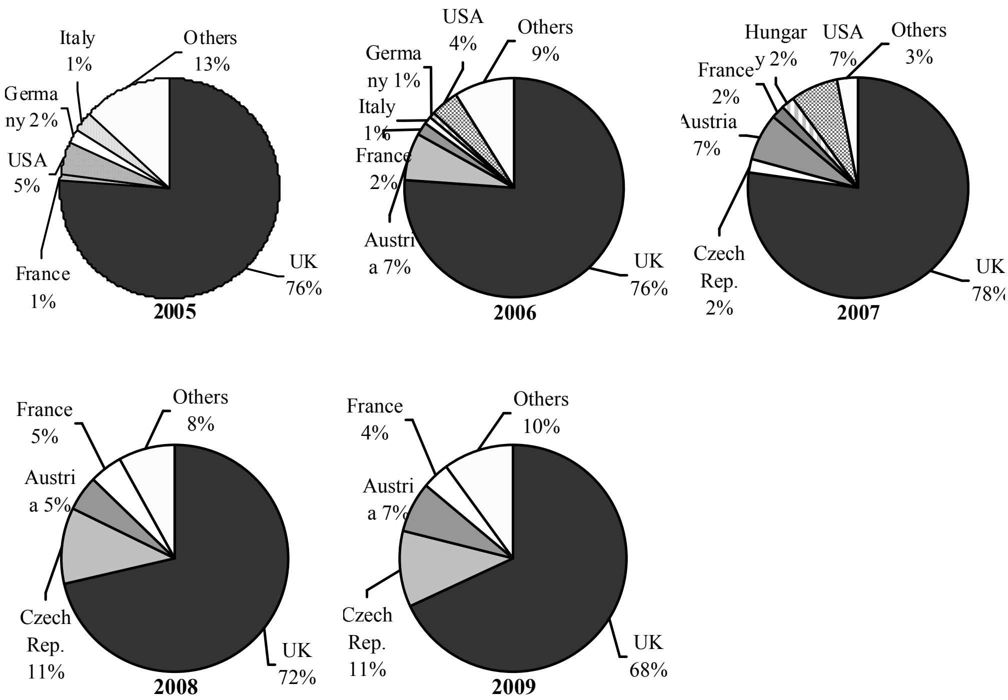
Figure 1. Investor structure on WSE (% shares in equity trading) for period between 2004 and 2009 (as of the end of the year)



Source: own investigation based on WSE data

[20] The analyses are based on Warsaw Stock Exchange Fact Book 2009. Official publication of WSE can be found at *www.wse.com.pl.*

Figure 2. Foreign investors structure on WSE (% shares in equity trading) for period between 2005 and 2009 (as of the end of the year)



Source: own investigation based on WSE data

What is more, there have been no changes in national origin of foreign brokers trading in shares [see Figure 2]. Indeed, brokers from Great Britain have been prevailing among foreign investors since the beginning of the period under investigation. The rest of shares in equity trading has been almost equally distributed among other foreign participants and when compared to Great Britain, gives a negligible contribution of these investors (less than 11% each).

To sum up, during the period under consideration the main investors that were present on Polish market were from: United Kingdom, Czech Republic, Austria, France, Italy, Germany, Hungary and the USA, and are identified in this study with main foreign investors groups present on WSE.

Therefore, in order to investigate the stock market linkage between WSE and home financial centres of its main foreign investors, daily closing quotes of WIG20 and leading indices from seven corresponding market stock exchange is used [see Table 1].

Table 1. Stock Market Indices

Country	Stock Market Exchange	Webpage	Index
Austria	Vienna Stock Exchange	www.wienerborse.at/	ATX
Czech Republic	Prague Stock Exchange	www.pse.cz/	PX
France	Euronext Paris	www.euronext.com/	CAC40
Germany	Frankfurt Stock Exchange	www.boerse-frankfurt.de/	DAX
Hungary	Budapest Stock Exchange	www.bse.hu/	BUX
Italy	Borsa Italiana	www.borsaitaliana.it	FTSE MIB
Poland	Warsaw Stock Exchange	www.wse.com.pl/	WIG 20
United Kingdom	London Stock Exchange	www.londonstockexchange.com/	FTSE 100
United States of America	New York Stock Exchange	www.nyse.com/	S&P 500

Source: own Investigation based on official Stock Market Exchange webpages

DATA DESCRIPTIONS

All data used in the study was obtained from the same source, namely from the webpage: *www.finance.yahoo.com*. Study covers the period between 1st January 2004 and 12th March 2010 (1575 observations for each stock market index). The additional subperiod covers the series between 12th March 2009 and 12th March 2010 which represent the data after the financial turmoil that started by the end of 2008.

The basic statistical characteristics of the sampled indices for the whole period under consideration and subperiod are presented respectively in Table 2 and Table 3.

Table 2. Main descriptive statistics (1st January 2004 – 12th March 2010)

Country	AUSTRIA	CZECH	FRANCE	GERMANY	HUNGARY	ITALY	POLAND	UK	USA
Index	ATX	PX	CAC 40	DAX	BUX	FTSE MIB	WIG20	FTSE 100	SP500
Mean	3 147.00	1 271.80	4 390.50	5 543.10	19 265.00	31 125.00	2 494.40	5 321.40	1 211.80
Median	3 179.60	1 274.00	4 321.60	5 519.80	20 781.00	31 878.00	2 384.80	5 330.50	1 213.90
Minimum	1 412.00	628.50	2 519.30	3 647.00	9 380.00	12 621.00	1 327.60	3 512.10	676.53
Maximum	4 981.90	1 936.90	6 168.10	8 105.70	30 118.00	44 364.00	3 917.90	6 732.40	1 565.20
Standard Deviation	1 027.30	349.81	874.74	1 233.50	5 251.10	7 447.80	680.30	771.98	185.95
Skewness	0.08	-0.03	0.18	0.36	-0.28	-0.27	0.36	-0.05	-0.36
Kurtosis	1.62	1.84	1.98	2.06	1.99	2.13	1.90	1.89	2.74
Jarque-Bera test	126.06	87.87	78.38	91.71	87.14	68.95	112.97	82.26	38.14

Source: own calculations

For the whole period measures for skewness and excess kurtosis show that PX, BUX, FTSE MIB and FTSE are negatively skewed, whereas ATX, CAC 40, DAX, WIG20 are positively skewed. All series are leptokurtic and are not normally distributed.

For the subperiod all indices have characteristics typical for financial series that is negative skewness, positive kurtosis and time series are not normally distributed.

Table 3. Main descriptive statistics (12th March 2009 – 12th March 2010)

Country	AUSTRIA	CZECH	FRANCE	GERMANY	HUNGARY	ITALY	POLAND	UK	USA
Index	ATX	PX	CAC 40	DAX	BUX	FTSE MIB	WIG20	FTSE 100	SP500
Mean	2 310.20	1 044.10	3 513.40	5 301.20	18 074.00	21 026.00	2 105.10	4 807.00	1 005.00
Median	2 453.10	1 117.60	3 640.10	5 481.30	19 457.00	21 627.00	2 187.30	4 987.70	1 032.20
Minimum	1 467.30	677.30	2 694.30	3 953.60	9 461.30	13 804.00	1 452.80	3 712.10	750.74
Maximum	2 752.40	1 220.30	4 045.10	6 048.30	23 210.00	24 426.00	2 489.40	5 599.80	1 150.20
Standard Deviation	312.07	140.00	349.71	535.11	3 597.60	2 340.10	269.96	518.81	105.35
Skewness	-0.75	-0.80	-0.52	-0.73	-0.63	-0.97	-0.58	-0.36	-0.52
Kurtosis	2.47	2.31	2.09	2.56	2.15	3.45	2.25	1.78	2.08
Jarque-Bera test	26.74	32.33	20.39	24.88	24.42	42.13	20.29	21.02	20.63

Source: own calculations

RESULTS

As first step, the ADF test was applied both to the levels and first differences of each series [see Table 4]. Appropriate number of lags for the ADF test was selected according to the Schwarz information criterion. For the levels, the results show that the null hypothesis of a unit root cannot be rejected at 5% significance level. The first-differenced series rejects the null hypothesis, indicating that time series are stationary. Consequently, all series are integrated I(1).

Given that the first differences are integrated of the same order (1), hence stationary, the Granger causality tests are performed [see Table 5]. WIG20 Granger-causes DAX (null hypothesis rejection for all 10 lags), PX (null hypothesis rejection for lags from 2 to 10), BUX (null hypothesis rejected for lags from 4 to 10), ATX (null hypothesis rejected for lags 7 and 8) and S&P 500 (null hypothesis rejection for all 10 lags).

WIG20 is Granger-caused by CAC (null hypothesis rejected for lags 3 and from 7 to 10), PX (for all 10 lags), BUX (for all 10 lags), ATX (for all 10 lags), FTSE (for lags from 8 to 10) and FTSE MIB (for lags 3, 4 and from 7 to 10).

Results indicate that WIG20 is influenced by strong and developed markets in Germany, UK and Italy. Moreover, WIG20 influences and is simultaneously influenced by stock markets of Central and Eastern Europe (Austria, Hungary and Czech Republic). In other words changes in the CEE indexes cause changes in WIG20 index with up to 10 days lags. The reverse situation is also true with less WIG20's lags. Granger-causing S&P500 by WIG20 doesn't necessarily indicative.

In chosen subperiod some important changes occurred and significantly less dependency can be observed [see Table 5]. There are no causalities between WIG20 and CAC, FTSE, FTSE MIB. There is also a decrease in number of lags in WIG20's influence on DAX or ATX and an increase in WIG20's influence on PX. A decrease in influence on WIG20 is noted for BUX and ATX.

Table 4. ADF Test results (1st January 2004 – 12th March 2010)

Country	Index	levels[21]	First difference	Conclusion
Austria	ATX	-1,316	-14,1099	I(1)
Czech Republic	PX	-1,551	-13,9403	I(1)
France	CAC 40	-1,327	-14,3275	I(1)
Germany	DAX	-1,302	-14,6469	I(1)
Hungary	BUX	-1,719	-13,4408	I(1)
Italy	FTSE MIB	-1,269	-14,0153	I(1)
Poland	WIG20	-1,301	-14,6806	I(1)
United Kingdom	FTSE 100	-1,551	-14,2512	I(1)
United States of America	SP500	-1,372	-14,5581	I(1)

Source: own calculations

Table 5. Granger Test results

Index	1st January 2004 – 12th March 2010		12th March 2009 – 12th March 2010	
	WIG20 Granger causes…	**…Granger causes WIG20**	**WIG20 Granger causes…**	**…Granger causes WIG20**
CAC	x	3, from 7 lags	x	x
DAX	to 10 lags	x	to 2	x
PX	from 2 lags	to 10 lags	3,from 5 to 10	to 10
BUX	from 4 lags	to 10 lags	x	from 1 to 4, 6, 9
ATX	from 7 to 8 lags	to 10 lags	7	from 1 to 4
SP500	to 10 lags	x	to 10	x
FTSE	x	from 8 lags	x	x
FTSE MIB	3	3-4, from 7 lags	x	x

Source: own calculations

Cointegration tests results are shown in Table 6. In period 2004-2010 WIG20 is cointegrated with PX (null hypothesis rejected for lags 1 and 2), ATX

[21] The critical values of the ADF t-statistic as reported by STATA, the econometric software used for performing the unit root test, are -2.329, -1.646, -1.282 at the 1%, 5% and 10% levels of significance, respectively.

(for lags 1 and 2) and FTSE (null hypothesis rejection for all 10 lags). In subperiod 2009-2010 WIG20 is cointegrated with CAC (for lags from 1 to 8), DAX (for all 10 lags), BUX (for lags from 1 to 5 and for 7th lag), S&P500 (for all 10 lags) and FTSE (for lags from 1 to 5 and for 7th lag). No cointegration with FTSE MIB was observed. Therefore stronger cointegration WIG20 with most of the indices considered in our study can be observed.

In some cases, the results for the cointegration and Granger causality test seem to be in conflict with each other. The explanation is based on the consideration that the Granger causality test explores the short-term relationships among variables whereas cointegration tests are used to examine long-term relationship.

Table 6 Cointegration Relationships

	1st January 2004 – 12th March 2010	12th March 2009 – 12th March 2010
CAC	1	from 1 to 8
DAX	x	from 1 to 10
PX	from 1 to 2	x
BUX	x	from 1 to 5 and 7
ATX	from 1 to 2	x
SP500	x	from 1 to 10
FTSE	from 1 to 10	from 1 to 5 and 7
FTSE MIB	x	x

Source: own calculations

CONCLUSION

The aim of this paper was to investigate possible interactions between Warsaw Stock Exchange index - WIG20 - and indices of home country financial centres of WSE's main foreign investors.

The results suggest that since the year of Polish accession to the European Union, Polish market has shown particularly strong relationships with stock markets from the CEE region. The influence is two-sided and is visible in short and long horizon, which was proved by the results of the Granger and cointegration test, respectively. These findings seem to be in line with the previous studies conducted by [Gilmore et al. 2005] and [Voronkova, 2004]

Furthermore, the findings reveal that every stock market were sensitive to financial turmoil which started at the end of 2008. In fact, in the short term strength of relationship between WIG20 and other markets seem to decrease. WIG20 is Granger-caused by indices of neighbouring stock markets of the CEE region. However in the long-run the Polish market reacts strongly at signals from the most matured and developed markets, that is the USA, United Kingdom, Germany and France.

Finally, taking into account that comovement can be found for Polish market and other European markets (especially LSE and all CEE stock markets), increasing integration among these financial markets gradually reduces benefits derived from international diversification in the long term perspective. However, the US investors can still benefit from investing in emerging markets, like WSE, both in short and long time horizon.

In conclusion incentives for investing on WSE are not connected with portfolio diversification due to strong dependencies between analysed stock markets.

REFERENCES

Banerjee A., Dolado J., Mestre R. (1998) Error correction mechanism tests for cointegration in a single equation framework, Journal of Time Series Analysis, 19(3), 267-283.

Dickey D., Fuller W. (1979) Distribution of the Estimators for Autoregressive Time Series with a Unit Root, Journal of the American Statistical Association, 74, 427-431.

Dickey D., Fuller W. (1981) Likelihood Ratio Tests for Autoregressive Time Series with a Unit Root, Econometrica, 49, 1057-1072.

Gilmore C. G. McManus G. M. (2002) International Portfolio Diversification: US and Central European Equity Markets, Emerging Markets Review, 3, 69–83.

Gilmore G. C, Lucey M. B, MacManus M. G. (2005) The dynamics of Central European Equity Market Integration, IIIS Discussion Paper, 69.

Granger C.W.J. (1969) Investigating causal relations by econometric model, Econometrica, 37, 424-438.

Grubel H.G. (1968) Internationally Diversified Portfolios: Welfare Gains and Capital Flows, American Economic Review 58, 1299-1314.

Grubel H.G. Kenneth F. (1971) The Interdependence of International Equity Markets, Journal of Finance 26 (1), 89-94.

Lessard D.R. (1973)International Portfolio Diversification:A Multivariate Analysis for a Group of Latin American Countries, Journal of Finance 28 (3), 619-633.

Levy H. Sarnat M. (1970) International Diversification of Investment Portfolios, The American Economic Review 60 (4), 668-675.

Kanas A. (1998) Linkages between the US and European equity markets: further evidence from cointegration tests, Applied Financial Economics 8 (6),607 – 614.

Markowitz H.M. (1952) Portfolio Selection The Journal of Finance 7 (1), 77–91.

Markowitz H.M. (1959) Portfolio Selection: Efficient Diversification of Investments, John Wiley & Sons New York.

Solnik B.H. (1974) Why Not Diversify Internationally Rather Than Domestically?, Financial Analysts Journal 30, 48-54.

Rousova L. (2009) Are the Central European Stock Markets Still Different? A Cointegration Analysis, Economics 15, University of Munich, Dep. of Economics.

Voronkova S. (2004) Equity Market Integration in Central European Emerging Markets: A Cointegration Analysis with Shifting Regimes, International Review of Financial Analysis, Vol.13, 633– 647.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XI, No. 1, 2010, pp. 128-138

ORTHOGONALIZED FACTORS IN MARKET-TIMING MODELS OF POLISH EQUITY FUNDS[22]

Joanna Olbryś
Department of Informatics, Bialystok University of Technology
e-mail: j.olbrys@pb.edu.pl

Abstract: The main goal of this paper is to examine the influence of factor orthogonalization in modified versions of classic market-timing models with the Fama and French spread variables *SMB* and *HML*, which have been introduced in [Olbryś 2010]. We construct the orthogonal market factors using the Busse procedure [Busse 1999]. The market-timing and selectivity abilities of 15 equity open-end mutual funds have been evaluated for the period January 2003 – December 2009 based on the panel data estimation using the SUR method. We compare the regression results of the models with common and orthogonal market factors and investigate their statistical properties.

Keywords: mutual fund, multifactor market-timing model, orthogonalized factor, SUR method

THREE-FACTOR MARKET-TIMING MODELS WITH FAMA AND FRENCH SPREAD VARIABLES

E. Fama was the first to propose a formalized theoretical methodology for the decomposition of total return into the components of timing and selectivity [Fama 1972]. Treynor and Mazuy develop a procedure for detecting timing ability that is based on a regression analysis of the managed portfolio's realized returns, which includes a quadratic term [Treynor & Mazuy 1966]. Henriksson and Merton propose a theoretical structure that allows for the formal distinction of managers' forecasting skills into timing and selectivity [Henriksson & Merton 1981]. By assuming that the market timer's forecasts take two possible predictions: either

[22] Financial support in 2009 – 2011 from the Polish Committee for Scientific Research within the grant No. N N113 173237 is gratefully acknowledged

stocks will outperform bonds or bonds will outperform stocks, Merton derives an equilibrium theory that shows that the return patterns resulting from a market-timing strategy are similar to the return pattern of an option strategy (of the put-protective type) [Merton 1981]. Based on this model, Henriksson and Merton develop statistical procedures to investigate market-timing abilities of portfolios' managers. Fama and French find that two variables, the market value (MV) and the ratio of book value to market value (BV/MV) capture much of the cross – section of average stock returns [Fama & French 1993]. They form portfolios meant to mimic the underlying risk factors in returns related to size and book-to-market equity. These mimicking portfolios (*SMB* and *HML*) have been introduced as explanatory variables into regressions of Polish equity mutual funds' portfolios excess returns in [Olbryś 2010]. The size (*SMB*) and book-to-market (*HML*) mimicking portfolios on the Polish market have been constructed using the Fama and French procedure. The market-timing and selectivity abilities of the funds' managers have been evaluated for the period January 2003 – December 2009, based on the modified three-factor market-timing models, using Newey-West robust HAC estimators or the SUR method, respectively.

In [Olbryś 2010] the modified three-factor Treynor - Mazuy model with Fama and French spread variables (T-M-FF model) has been expressed as:

$$r_{P,t} = \alpha_P + \beta_P \cdot r_{M,t} + \delta_{1P} \cdot r_{SMB,t} + \delta_{2P} \cdot r_{HML,t} + \gamma_P \cdot (r_{M,t})^2 + \varepsilon_{P,t} \qquad (1)$$

where:

$r_{P,t} = R_{P,t} - R_{F,t}$ is the excess return of the portfolio P in the period t,

$r_{M,t} = R_{M,t} - R_{F,t}$ is the excess return of the portfolio M in the period t,

$R_{P,t}$ is the one-period return of the portfolio P,

$R_{M,t}$ is the one-period return of the market portfolio M,

$R_{F,t}$ is the one-period return of riskless securities,

Jensen's α_P measures selectivity skills of the portfolio's P manager [Jensen 1968],

β_P is the systematic risk measure of the portfolio P,

γ_P measures market-timing skills of the portfolio's P manager [Henriksson & Merton 1981],

$\varepsilon_{P,t}$ is a residual term, with the following standard CAPM conditions: $E(\varepsilon_{P,t}) = 0$, $E(\varepsilon_{P,t} | \varepsilon_{P,t-1}) = 0$.

$r_{SMB,t} = R_{SMB,t} - R_{F,t}$ is the excess return of the portfolio *SMB*,

$r_{HML,t} = R_{HML,t} - R_{F,t}$ is the excess return of the portfolio *HML*,

δ_{1P} is a sensitive measure of the portfolio P returns due to the changes in the *SMB* factor returns,

δ_{2P} is a sensitive measure of the portfolio P returns due to the changes in the *HML* factor returns.

In a way analogous to (2), Olbryś expressed the modified three-factor Henriksson - Merton model with Fama and French spread variables (H-M-FF model) as:

$$r_{P,t} = \alpha_P + \beta_P \cdot r_{M,t} + \delta_{1P} \cdot r_{SMB,t} + \delta_{2P} \cdot r_{HML,t} + \gamma_P \cdot y_{M,t} + \varepsilon_{P,t} \quad (2)$$

where:

$r_{P,t}$, $r_{M,t}$, $r_{SMB,t}$, $r_{HML,t}$, α_P, β_P, γ_P, δ_{1P}, δ_{2P}, $\varepsilon_{P,t}$ are as in the equation (1),

$y_{M,t} = \max\{0, R_{F,t} - R_{M,t}\} = \max\{0, -r_{M,t}\}$.

ORTHOGONALIZED FACTORS IN MARKET-TIMING MODELS

We orthogonalize the *SMB* and *HML* indices to maintain consistency with the theoretical and practical development that there is a correlation between the market factor *M* and mimicking portfolios *SMB* or *HML* ([Busse 1999], [Fama & French 1993]).

We take the orthogonal *SMB* factor (call it *SMBO*) to be the intercept plus the *SMB* factor regression residuals on the simple excess returns of the main index of Warsaw Stock Exchange companies, given as:

$$r_{SMB,t} = \alpha_{SMB} + \beta_{SMB} \cdot r_{M,t} + \varepsilon_t \quad (3)$$

The sum of the intercept and the residuals in (3):

$$r_{SMBO,t} = \alpha_{SMB} + \varepsilon_t \quad (4)$$

is uncorrelated with the explanatory market variable in (3).

Similarly, we take the orthogonal *HML* factor (call it *HMLO*) to be the intercept plus the *HML* factor regression residuals on the simple excess returns of the main index of Warsaw Stock Exchange companies and the orthogonal *SMBO* factor, which is expressed as:

$$r_{HML,t} = \alpha_{HML} + \beta_{HML} \cdot r_{M,t} + \gamma_{HML} \cdot r_{SMBO,t} + e_t \quad (5)$$

The sum of the intercept and the residuals in (5):

$$r_{HMLO,t} = \alpha_{HML} + e_t \quad (6)$$

is uncorrelated with the explanatory variables in (5).

Then the three-factor T-M-FF model (1), with the orthogonalized *SMBO* and *HMLO* factors can be expressed as:

$$r_{P,t} = \alpha_P + \beta_P \cdot r_{M,t} + \delta_{1P} \cdot r_{SMBO,t} + \delta_{2P} \cdot r_{HMLO,t} + \gamma_P \cdot (r_{M,t})^2 + \varepsilon_{P,t} \quad (7)$$

where the notations are as in the equation (1) but the explanatory variables $r_{SMBO,t}$ and $r_{HMLO,t}$ are given by the equations (4) or (6), respectively.

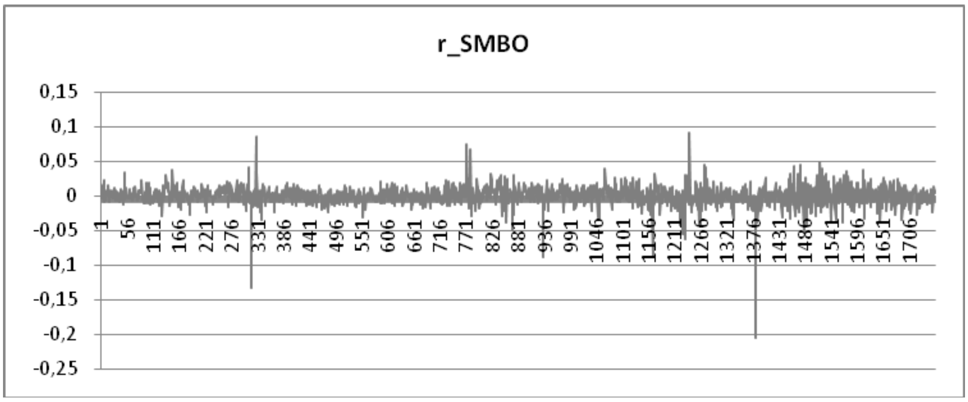
Similarly, the three-factor H-M-FF model (2) with the orthogonalized *SMBO* and *HMLO* factors can be given as:

$$r_{P,t} = \alpha_P + \beta_P \cdot r_{M,t} + \delta_{1P} \cdot r_{SMBO,t} + \delta_{2P} \cdot r_{HMLO,t} + \gamma_P \cdot y_{M,t} + \varepsilon_{P,t} \quad (8)$$

where the notations are as in the equation (2) but the explanatory variables $r_{SMBO,t}$ and $r_{HMLO,t}$ are given by the equations (4) or (6).

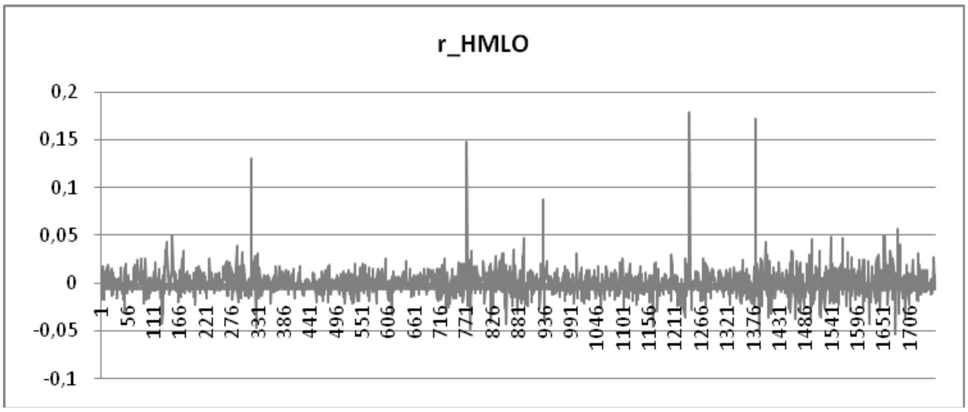
Fig. 1 and Fig. 2 present the exogenous variables $r_{SMBO,t}$ (4) and $r_{HMLO,t}$ (6) in the form of charts, respectively. We have detected (based on Dickey – Fuller test) that the analysed series are stationary.

Figure 1. The exogenous variable $r_{SMBO,t}$ from Jan 2003 to Dec 2009



Source: author's calculations

Figure 2. The exogenous variable $r_{HMLO,t}$ from Jan 2003 to Dec 2009



Source: author's calculations

ESTIMATION METHOD AND EMPIRICAL RESULTS

The SUR (seemingly unrelated regression) method was described by Zellner [Zellner 1962]. SUR is a way of estimating panel data models that are long (large T) but not wide (small N). The assumptions underlying the SUR model are the following [Marshall & Young 2003]:

1) All disturbances have a zero mean;
2) In a given cross-sectional unit, the disturbance variance is constant over time, but each cross-sectional unit can have a different variance;
3) Two disturbances in different cross-sectional units but corresponding to the same time period are correlated (contemporaneous correlation);
4) Disturbances in different time periods, whether they are in the same cross-sectional unit or not, are uncorrelated (autocorrelation does not exist).

In the basic SUR model, the errors are assumed to be homoskedastic and linearly independent within each equation. Each equation is correlated with the others in the same time period. This assumption is called contemporaneous correlation, and it is this property that sets SUR apart from other models [Adkins 2009]. Given that it is very likely that equity funds' portfolios from the same market are contemporaneously correlated, the SUR model seems to be appropriate for this case. If contemporaneous correlation does not exist, the LSR method applied separately to each equation (fund's portfolio) is quite efficient.

We use daily data following evidence that daily data provide better inferences than monthly data regarding timing ability [Bollen & Busse 2001]. This evidence has been examined in the case of Polish equity mutual funds in [Olbryś 2008b]. We examine the performance of 15 selected equity open-end mutual funds. We study daily simple excess returns from Jan 2003 to Dec 2009. Daily returns on the main index of Warsaw Stock Exchange companies are used as the returns on the market portfolio. The average daily returns on 52-week Treasury bills are used as the riskless asset returns. Daily return rates on spread factors *SMB* and *HML* are used as the values of the additional exogenous variables in the T-M-FF (1) and H-M-FF (2) models. In the data panel the number of funds is equal to N=15 and the number of time periods is T=1760.

Tables 1 and 2 provide details on the estimated T-M-FF (1) and H-M-FF (2) market-timing models, respectively. The SUR method has been used to consider the contemporaneous correlation effects. In all of the tables: * denotes coefficients that are significantly different from zero at the ten percent level; ** denote coefficients that are significantly different from zero at the five percent level and *** denote coefficients that are significantly different from zero at the one percent level.

Table 1. Three-factor T-M-FF model (1) (Jan 2, 2003 - Dec 31, 2009)

	Equity funds	$\hat{\alpha}_P$	$\hat{\beta}_P$	$\hat{\delta}_{1P}$	$\hat{\delta}_{2P}$	$\hat{\gamma}_P$	R^2
1	Arka BZ WBK Akcji FIO	0.0006***	0.730***	0.070***	0.046***	-2.03***	0.623
2	Aviva Investors FIO Polskich Akcji	0.0005***	0.751***	-0.002	0.021*	-1.92***	0.691
3	BPH FIO Akcji	0.0001	0.716***	0.011	0.031***	-0.96***	0.716
4	DWS Polska FIO Top 25 Małych Spółek	0.0002	0.445***	0.216***	0.070***	-1.54***	0.297
5	DWS Polska FIO Akcji	0.0002	0.630***	0.033	0.017	-1.30**	0.373
6	DWS Polska FIO Akcji Plus	0.0002	0.555***	0.092***	0.025	-1.27**	0.355
7	ING FIO Akcji	0.0001	0.746***	0.009	0.027**	-0.92**	0.695
8	Legg Mason Akcji FIO	0.0003*	0.698***	0.017	0.035***	-1.05***	0.700
9	Millennium FIO Akcji	0.0000	0.683***	0.033***	0.050***	-1.07***	0.673
10	Pioneer Akcji Polskich FIO	0.0000	0.814***	-0.001	0.035***	-1.52***	0.698
11	PKO/CREDIT SUISSE Akcji FIO	0.0003	0.559***	0.029*	0.028*	-2.20***	0.422
12	PZU FIO Akcji KRAKOWIAK	0.0001	0.702***	-0.003	0.035***	-1.39***	0.689
13	SEB 3 – Akcji FIO	0.0004	0.522***	0.085***	0.017	-1.64***	0.315
14	Skarbiec – Akcja FIO	0.0003	0.457***	0.068***	0.013	-0.52	0.279
15	UniKorona Akcja FIO	0.0004	0.519***	0.091***	0.019	-1.06*	0.309

Source: author's calculations (using *Gretl 1.8.5*)

Table 2. Three-factor H-M-FF model (2) (Jan 2, 2003 - Dec 31, 2009)

	Equity funds	$\hat{\alpha}_P$	$\hat{\beta}_P$	$\hat{\delta}_{1P}$	$\hat{\delta}_{2P}$	$\hat{\gamma}_P$	R^2
1	Arka BZ WBK Akcji FIO	0.0010***	0.650***	0.071***	0.046***	-0.17***	0.623
2	Aviva Investors FIO Polskich Akcji	0.0010***	0.672***	-0.002	0.021*	-0.17***	0.690
3	BPH FIO Akcji	0.0004*	0.673***	0.011	0.031***	-0.09***	0.716
4	DWS Polska FIO Top 25 Małych Spółek	0.0005	0.387***	0.217***	0.071***	-0.12**	0.297
5	DWS Polska FIO Akcji	0.0005	0.579***	0.034	0.017	-0.11*	0.373
6	DWS Polska FIO Akcji Plus	0.0005	0.501***	0.092***	0.025	-0.11**	0.355
7	ING FIO Akcji	0.0004	0.702***	0.008	0.026**	-0.09***	0.695
8	Legg Mason Akcji FIO	0.0006**	0.651***	0.016	0.035***	-0.10***	0.700
9	Millennium FIO Akcji	0.0004*	0.631***	0.032***	0.049***	-0.11***	0.674
10	Pioneer Akcji Polskich FIO	0.0005*	0.746***	-0.001	0.034***	-0.14***	0.698
11	PKO/CREDIT SUISSE Akcji FIO	0.0008**	0.467***	0.029*	0.028*	-0.19***	0.422

12	PZU FIO Akcji KRAKOWIAK	0.0005**	0.640***	-0.003	0.035***	-0.13***	0.689
13	SEB 3 – Akcji FIO	0.0007*	0.462***	0.085***	0.017	-0.13**	0.314
14	Skarbiec – Akcja FIO	0.0005	0.429***	0.067***	0.012	-0.06	0.279
15	UniKorona Akcja FIO	0.0007*	0.475***	0.091***	0.019	-0.09*	0.309

Source: author's calculations (using *Gretl 1.8.5*)

Tables 1-2 include the estimation results of the three-factor T-M-FF (1) and H-M-FF (2) models. Results of the T-M-FF tests (Table 1) show that the estimates of Jensen's measure of performance ($\hat{\alpha}_P$) are positive, but not significant in the case of twelve funds. We can observe that in the case of H-M-FF models (Table 2) ten out of fifteen funds present a significant positive estimate of selectivity. According to Jensen's interpretation of the $\hat{\alpha}_P$ value, this measure could be positive for two reasons: (1) the extra returns actually earned on the portfolio due to the manager's ability, and (2) the positive bias in the estimate of $\hat{\alpha}_P$ resulting from the negative bias in the $\hat{\beta}_P$ estimate [Jensen 1968, pp. 396]. The systematic risk levels ($\hat{\beta}_P$) are significantly positive (Tables 1-2). Unfortunately, the empirical results show no statistical evidence that Polish equity funds' managers have outguessed the market. Almost all of the funds (except Skarbiec – Akcja FIO in Tables 1-2) present significantly negative estimates of market-timing skills ($\hat{\gamma}_P < 0$). We find evidence of negative market-timing. Significant negative estimates of market-timing indicate that, contrary to what would be expected of rational investors, the managers increase the exposition of their portfolios to the market in down markets and act inversely in up markets [Romacho & Cortez 2006]. There is a statistically significant negative relationship between selectivity ($\hat{\alpha}_P$) and timing ($\hat{\gamma}_P$). As for the sensitive measure of the fund's portfolio P returns due to the changes in the *SMB* factor returns, only eight out of fifteen funds (in Tables 1-2) exhibit positive and statistically significant coefficients $\hat{\delta}_{1P}$. The spread variable *HML* is positive and statistically significant in the case of ten out of fifteen funds (coefficients $\hat{\delta}_{2P}$ in Tables 1-2).

Tables 3-4 include the estimation results of the three-factor T-M-FF (7) and H-M-FF (8) models with the orthogonalized *SMBO* (4) and *HMLO* (6) factors as the explanatory variables.

Table 3. Three-factor T-M-FF model (7) with the orthogonalized *SMBO* and *HMLO* factors (Jan 2, 2003 - Dec 31, 2009)

	Equity funds	$\hat{\alpha}_P$	$\hat{\beta}_P$	$\hat{\delta}_{1P}$	$\hat{\delta}_{2P}$	$\hat{\gamma}_P$	R^2
1	Arka BZ WBK Akcji FIO	0.0006***	0.712***	0.049***	0.046***	-2.03***	0.623
2	Aviva Investors FIO Polskich Akcji	0.0005***	0.748***	-0.011	0.021*	-1.92***	0.691
3	BPH FIO Akcji	0.0001	0.709***	-0.003	0.031***	-0.96***	0.716
4	DWS Polska FIO Top 25 Małych Spółek	0.0002	0.399***	0.184***	0.070***	-1.54***	0.297
5	DWS Polska FIO Akcji	0.0002	0.622***	0.026	0.017	-1.30**	0.373
6	DWS Polska FIO Akcji Plus	0.0002	0.535***	0.081***	0.025	-1.27**	0.355
7	ING FIO Akcji	0.0001	0.741***	-0.003	0.027**	-0.92**	0.695
8	Legg Mason Akcji FIO	0.0003*	0.690***	0.0009	0.035***	-1.05***	0.700
9	Millennium FIO Akcji	0.0000	0.671***	0.010	0.050***	-1.07***	0.673
10	Pioneer Akcji Polskich FIO	0.0000	0.809***	-0.016	0.035***	-1.52***	0.698
11	PKO/CREDIT SUISSE Akcji FIO	0.0003	0.550***	0.017	0.028*	-2.20***	0.422
12	PZU FIO Akcji KRAKOWIAK	0.0001	0.698***	-0.018*	0.035***	-1.39***	0.689
13	SEB 3 – Akcji FIO	0.0004	0.505***	0.077***	0.017	-1.64***	0.315
14	Skarbiec – Akcja FIO	0.0003	0.444***	0.062***	0.013	-0.52	0.279
15	UniKorona Akcja FIO	0.0004	0.500***	0.082***	0.019	-1.06*	0.309

Source: author's calculations (using *Gretl 1.8.5*)

Table 4. Three-factor H-M-FF model (8) with the orthogonalized *SMBO* and *HMLO* factors (Jan 2, 2003 - Dec 31, 2009)

	Equity funds	$\hat{\alpha}_P$	$\hat{\beta}_P$	$\hat{\delta}_{1P}$	$\hat{\delta}_{2P}$	$\hat{\gamma}_P$	R^2
1	Arka BZ WBK Akcji FIO	0.0010***	0.631***	0.050***	0.046***	-0.17***	0.623
2	Aviva Investors FIO Polskich Akcji	0.0010***	0.669***	-0.011	0.021*	-0.17***	0.690
3	BPH FIO Akcji	0.0004*	0.667***	-0.003	0.031***	-0.09***	0.716
4	DWS Polska FIO Top 25 Małych Spółek	0.0005	0.340***	0.185***	0.071***	-0.12**	0.297
5	DWS Polska FIO Akcji	0.0005	0.571***	0.026	0.017	-0.11*	0.373
6	DWS Polska FIO Akcji Plus	0.0005	0.482***	0.081***	0.025	-0.11**	0.355
7	ING FIO Akcji	0.0004	0.697***	-0.003	0.026**	-0.09***	0.695
8	Legg Mason Akcji FIO	0.0006**	0.643***	0.0006	0.035***	-0.10***	0.700
9	Millennium FIO Akcji	0.0004*	0.618***	0.098	0.049***	-0.11***	0.674
10	Pioneer Akcji Polskich FIO	0.0005*	0.742***	-0.017	0.034***	-0.14***	0.698

11	PKO/CREDIT SUISSE Akcji FIO	0.0008**	0.458***	0.017	0.028*	-0.19***	0.422
12	PZU FIO Akcji KRAKOWIAK	0.0005**	0.636***	-0.019*	0.035***	-0.13***	0.689
13	SEB 3 – Akcji FIO	0.0007*	0.445***	0.078***	0.017	-0.13**	0.314
14	Skarbiec – Akcja FIO	0.0005	0.416***	0.062***	0.012	-0.06	0.279
15	UniKorona Akcja FIO	0.0007*	0.457***	0.082***	0.019	-0.09*	0.309

Source: author's calculations (using *Gretl 1.8.5*)

It can be observed that our initial conclusions concerning $\hat{\alpha}_P$ and $\hat{\gamma}_P$ remain unaltered. The levels of the systematic risk $\hat{\beta}_P$ in Tables 3-4 down somewhat relative to the values in Tables 1-2, respectively. However, the two sets of regressions produce the same R-squared values. As for the sensitive measure of the fund's portfolio P returns due to the changes in the *SMBO* factor returns, only seven out of fifteen funds (in Tables 3-4) exhibit positive and statistically significant coefficients $\hat{\delta}_{1P}$. The evidence is that the values of $\hat{\delta}_{1P}$ coefficients in Tables 3-4 significantly differ from these in Tables 1-2. On the other hand, in the case of all models and all funds, we have received the same estimator values of $\hat{\delta}_{2P}$ coefficients in Tables 1, 3 and Tables 2, 4, respectively.

The three-factor T-M-FF (7) and H-M-FF (8) models with the orthogonalized *SMBO* (4) and *HMLO* (6) factors as the explanatory variables have also been estimated using logarithmic excess returns. A logarithmic excess return is given by the equation:

logarithmic rate=ln(1+simple rate)

Table 5 reports the estimation results of T-M-FF (7) market-timing models using logarithmic excess returns. It can be observed that in the case of all funds, we have received almost the same estimator values as when using simple excess returns (see Table 3). The R-squared values in Table 5 up somewhat relative to the values in Table 3. We have received similar regression effects of H-M-FF (8) market-timing models using logarithmic excess returns but due to the space restriction, we do not report full results.

Table 5. Three-factor T-M-FF model (7) with the orthogonalized *SMBO* and *HMLO* factors (logarithmic excess returns from Jan 2, 2003 to Dec 31, 2009)

	Equity funds	$\hat{\alpha}_P$	$\hat{\beta}_P$	$\hat{\delta}_{1P}$	$\hat{\delta}_{2P}$	$\hat{\gamma}_P$	R^2
1	Arka BZ WBK Akcji FIO	0.0006***	0.713***	0.049***	0.049***	-2.06***	0.626
2	Aviva Investors FIO Polskich Akcji	0.0005***	0.749***	-0.011	0.023**	-1.96***	0.693
3	BPH FIO Akcji	0.0001	0.710***	-0.002	0.031***	-0.95***	0.718
4	DWS Polska FIO Top 25 Małych Spółek	0.0002	0.400***	0.182***	0.077***	-1.45***	0.299
5	DWS Polska FIO Akcji	0.0002	0.623***	0.025	0.018	-1.34**	0.374
6	DWS Polska FIO Akcji Plus	0.0002	0.537***	0.080***	0.027	-1.22**	0.356
7	ING FIO Akcji	0.0001	0.741***	-0.002	0.028**	-0.89**	0.697
8	Legg Mason Akcji FIO	0.0003	0.690***	0.0009	0.036***	-1.02***	0.702
9	Millennium FIO Akcji	0.0000	0.671***	0.011	0.051***	-1.05***	0.675
10	Pioneer Akcji Polskich FIO	0.0000	0.810***	-0.015	0.035***	-1.61***	0.701
11	PKO/CREDIT SUISSE Akcji FIO	0.0003	0.552***	0.017	0.029*	-2.23***	0.424
12	PZU FIO Akcji KRAKOWIAK	0.0001	0.699***	-0.018	0.036***	-1.38***	0.690
13	SEB 3 – Akcji FIO	0.0003	0.506***	0.075***	0.019	-1.64***	0.316
14	Skarbiec – Akcja FIO	0.0003	0.444***	0.060***	0.014	-0.43	0.279
15	UniKorona Akcja FIO	0.0004	0.501***	0.080***	0.021	-1.01*	0.310

Source: author's calculations (using *Gretl 1.8.5*)

Table 5 reports the estimation results of T-M-FF (7) market-timing models using logarithmic excess returns. It can be observed that in the case of all funds, we have received almost the same estimator values as when using simple excess returns (see Table 3). The R-squared values in Table 5 up somewhat relative to the values in Table 3. We have received similar regression effects of H-M-FF (8) market-timing models using logarithmic excess returns but due to the space restriction, we do not report full results.

CONCLUSION

In this paper we have examined the usefulness of the orthogonalized *SMBO* and *HMLO* factors as explanatory variables in market-timing models for the investment managers' performance evaluation. We have confirmed that the quality increase of the models is rather small. To summarize, basing on the empirical analysis it can be concluded that there is no clear reason to prefer the three-factor T-M-FF (7) and H-M-FF (8) models over those given by the equations (1) and (2). This evidence is consistent with the literature, for example [Fama & French 1993, pp. 31].

REFERENCES

Adkins L.C. (2009) Using Gretl for Principles of Econometrics, Third edition, Version 1.31, July 20.

Bollen N. P. B., Busse J.A. (2001) On the timing ability of mutual fund managers, The Journal of Finance, 56 (3), pp. 1075-1094.

Busse J.A. (1999) Volatility timing in mutual funds: evidence from daily returns, The Review of Financial Studies, Vol. 12, No. 5, pp. 1009-1041.

Cambell J.Y., Lo A.W., MacKinlay A.O. (1997) The Econometric of Financial Markets, Princeton University Press, New Jersey.

Fama E.F. (1972) Components of investment performance, The Journal of Finance, 27, pp. 551-567.

Fama E.F., French K.R. (1993) Common risk factors in the returns on stocks and bonds, Journal of Financial Economics, 33, pp. 3-56.

Henriksson R., Merton R. (1981) On market timing and investment performance. II. Statistical procedures for evaluating forecasting skills, Journal of Business, 54, pp. 513-533.

Henriksson R. (1984) Market timing and mutual fund performance: an empirical investigation, Journal of Business, 57, pp. 73-96.

Jensen M. (1968) The performance of mutual funds in the period 1945-1964, Journal of Finance, 23, pp. 389-416.

Kufel T. (2009) Ekonometria. Rozwiązywanie problemów z wykorzystaniem programu Gretl, PWN, Warszawa.

Maddala G.S. (2008) Ekonometria, PWN, Warszawa.

Marshall B.R., Young M. (2003) Liquidity and stock returns in pure order-driven markets: evidence from the Australian stock market, International Review of Financial Analysis, 12, 173-188.

Merton R. (1981) On market timing and investment performance. I. An equilibrium theory of value for market forecasts, Journal of Business, 54, 363-406.

Olbryś J. (2010) Three-factor market-timing models with Fama and French spread variables, Badania Operacyjne i Decyzje, 2/2010, in progress.

Olbryś J. (2008a) Parametric tests for timing and selectivity in Polish mutual fund performance, Optimum. Studia Ekonomiczne, Wydawnictwo Uniwersytetu w Białymstoku, 3(39), pp. 107-118.

Olbryś J. (2008b) Ocena umiejętności stosowania strategii market-timing przez zarządzających portfelami funduszy inwestycyjnych a częstotliwość danych, Studia i Prace Wydziału Nauk Ekonomicznych i Zarządzania 10, Uniwersytet Szczeciński, pp. 96-105.

Romacho J. C., Cortez M. C. (2006) Timing and selectivity in Portuguese mutual fund performance, Research in International Business and Finance, 20, 348-368.

Treynor J., Mazuy K. (1966) Can mutual funds outguess the market?, Harvard Business Review, 44, pp. 131-136.

Zellner A. (1962) An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias, Journal of American Statistical Association, 57, pp. 348-368.

QUANTITATIVE METHODS IN ECONOMICS
VOL. XI, No. 1, 2010, pp. 139-144

INTRADAY VOLATILITY MODELING: THE EXAMPLE OF THE WARSAW STOCK EXCHANGE

Magdalena Sokalska
Queens College, CUNY
e-mail: msokalska@qc.cuny.edu

Abstract: We present an intraday volatility model for equally spaced data and apply it for the WIG Index- a broad market index of the Warsaw Stock Exchange. The current study is an application and extension of the model proposed by Engle and Sokalska [2010]. We decompose the conditional variance of intraday returns into components that have a natural interpretation and can be easily estimated.

Keywords: Volatility, ARCH, Intra-day Returns.

INTRODUCTION

As recent developments on world's stocks markets have shown, intraday asset price movements can be very dramatic. During the crisis of 2008-2009, it was not unusual to observe spectacular stock market rallies and sudden plunges of Dow Jones Industrial Index worth several hundred of points on a single day. Therefore there is great value in being able to forecast volatility on intraday basis. A paper by Andersen and Bollerslev [1997] documents that an application of standard ARCH-type volatility models [Engle, 1982] to intraday data gives unsatisfactory results. This is because there are pronounced periodic patterns in volatility throughout a day.

A number of papers have presented work on intraday returns related the current study. Andersen and Bollerslev [1997, 1998] propose models for 5-minute returns on Deutschemark-dollar exchange rate and the S&P500 index. They build a multiplicative model of daily and diurnal volatility. Andersen and Bollerslev [1998] add an additional component which takes account of the influence of macro-economic announcements on the foreign exchange volatility. For most of their models, the intra-daily volatility components are deterministic. The model applied in this paper contains stochastic intraday variance component.

THE MODEL

Our paper uses the volatility model for high frequency intraday financial returns proposed by Engle and Sokalska [2010]. We employ the following notation. Days in the sample are indexed by t (t =1,…, T). Each day is divided into 10 minute intervals referred to as bins and indexed by i (i =1,…, N). The current period is $\{t,i\}$. Price of an asset at day t and bin i is denoted by $P_{t,i}$. The continuously compounded return $r_{t,i}$ is modeled as: $r_{t,i} = \ln\left(P_{t,i} / P_{t,i-1} \right) \; for \;\; i \geq 2$, whereas the return for the first bin of the day i=1 is $r_{t,i} = \ln\left(P_{t,1} / P_{t-1,N} \right)$. In our model, the conditional variance is a multiplicative product of daily, diurnal and stochastic intraday variance components. Intraday equity returns are described by the following process:

$$r_{t,i} = \sqrt{h_t s_i q_{t,i}}\,\varepsilon_{t,i} \text{ and } \quad \varepsilon_{t,i} \sim N(0,1) \tag{1}$$

where:

h_t is the daily variance component,

s_i is the diurnal (periodic) variance pattern,

$q_{t,i}$ is the intraday volatility component, and

$\varepsilon_{t,i}$ is an error term.

The model is estimated in several steps. Similarly to Andersen and Bollerslev [1997, 1998], daily variance component will be estimated from a daily ARCH –type specification for a longer sample, going back a number of years.

Empirical analysis indicates that GARCH (1,1) process [Bollerslev, 1986] proved to be the most successful daily volatility model.

$$\begin{aligned} r_t &= \sqrt{h_t}\,\zeta_t \qquad \zeta_t \sim N(0,1) \\ h_t &= w_0 + \alpha_d r_{t-1}^2 + \beta_d h_{t-1}^2 \end{aligned} \tag{2}$$

Where ζ_t is an error term for daily returns r_t, whereas w_0, α_d and β_d are parameters of the variance equation.

The diurnal component is calculated as the variance of returns in each bin after deflating by the daily variance component.

$$E\left(\frac{r_{i,t}^2}{h_t} \right) = s_i E\left(q_{i,t} \right) = s_i \tag{3}$$

Then we model the residual volatility as a GARCH(1,1) process [23]:

$$q_{\{t,i\}} = \omega + \alpha (r_{\{t,i-1\}} / \sqrt{h_t s_{i-1}})^2 + \beta \ q_{\{t,i-1\}} \tag{4}$$

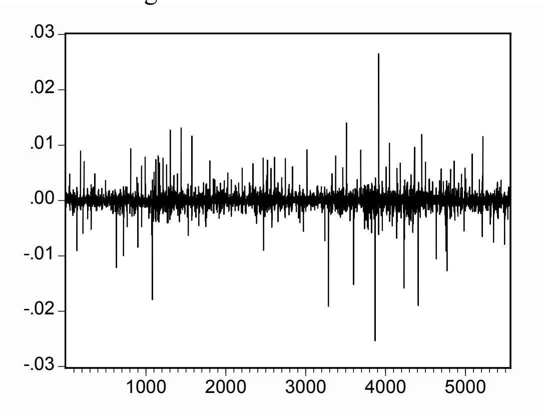
Engle and Sokalska [2010] show that this multistep estimator is consistent and asymptotically normal.

EMPIRICAL ANALYSIS

Our data consists of 10 minute logarithmic returns on the WIG index for the period 4 January 2010 – 31 May 2010. We estimated daily variance component using daily data on the WIG index between October 1994 and May 2010. Both intraday and daily series come from the Bloomberg database.

We exclude overnight returns from our analysis. Figure 1 and 2 depict intraday logarithmic returns of WIG including and excluding overnight returns, respectively. As can be seen from Figure 1 intraday returns are dominated by substantially negative and positive overnight changes; all of the returns that exceed +/- 1% are in fact overnight. Although the inclusion of overnight returns could yield a more complete analysis, it would require a far more complex model. Such a complete model would need to concentrate more on economic or global factors and this approach was not followed in this paper.

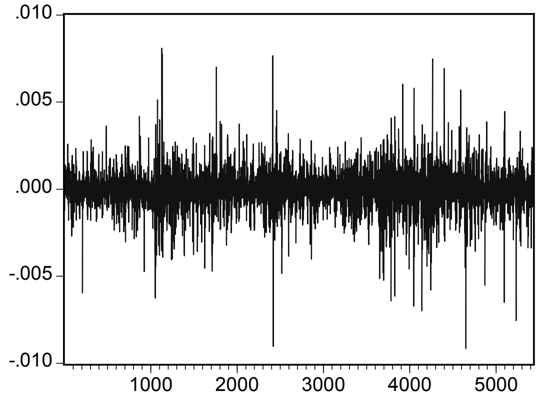
Figure 1. Intraday WIG 10-min logarithmic returns



Labels of the horizontal axis denote observation number.
Source: own calculations

[23] Higher order GARCH models were fitted as part of the specification search for scaled intraday returns but GARCH(1,1) performed best.

Figure 2. Intraday WIG 10-min logarithmic returns excluding the overnight period



Labels of the horizontal axis denote observation number.
Source: own calculations

Table 1 presents estimation results of the daily GARCH (1,1) model (2). Attempts to fit higher order models yielded statistically insignificant coefficients for the respective parameters. The sum of coefficients $\alpha_d+\beta_d$ (0.9896) is close to but smaller than one. This is indicates high persistence (high degree of volatility clustering).

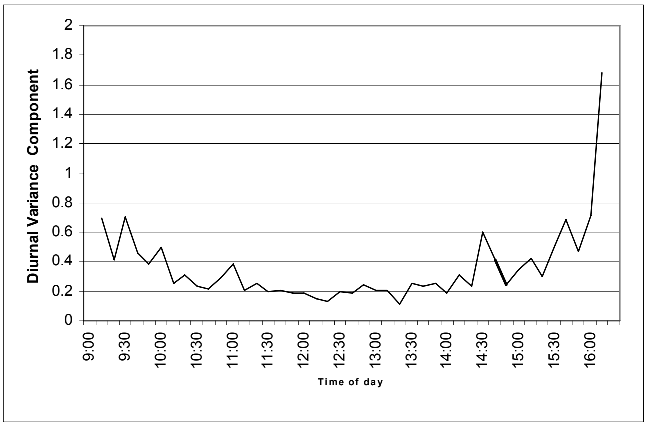
Table 1. GARCH Results for Daily Data

GARCH Results for Daily WIG Index			
	Coefficient	Std. Error	t-Statistic
ω_d	3.52E-06	5.09E-07	6.925
α_d	0.0845	0.00498	16.984
β_d	0.9051	0.00479	188.941

Notes: This table presents estimation results for GARCH(1,1) model for the WIG index. Sample period for daily data: October 1994 -May 2010. Symbols α_d , β_d and ω_d denote GARCH parameters from the variance equation (2).
Source: own calculation

Figure 3 depicts the estimates of diurnal variance component (3). We may observe a typical U-shaped pattern in volatility - an increased variability at the beginning of the day followed by a calm period in the middle and then a rise in variation at the close of the session. A small spike around 14.30 coincides with the time of macroeconomic announcements in the US. The volatility of the overnight period, which is excluded from our analysis, would actually much exceed the volatility at the close.

Figure 3. Diurnal Variance Component Throughout a Day (Excluding Overnight Period)



Source: own calculations

Finally Table 2 presents results of estimation of the parameters of the stochastic intraday variance component (4). The persistence measure, $\alpha+\beta$, equals to 0.8884 and is smaller than $\alpha_d+\beta_d$ (0.9896) obtained for the daily component. The degree of volatility clustering for the intraday component is smaller than for the daily component because we estimated the intraday GARCH on scaled returns (compare equation (4)). In other words we have already taken account of some of the volatility persistence present in the data by scaling squared returns by the daily and diurnal variance components.

Table 2. GARCH Results for Intraday Data

GARCH Results for Intraday Daily WIG Index			
	Coefficient	Std. Error	t-Statistic
ω	1.61E-07	9.16E-09	17.555
α	0.1984	0.0108	18.354
β	0.6900	0.0136	50.886

Notes: This table presents estimation results for GARCH(1,1) model for WIG returns that have been previously scaled by the squared root of the daily and diurnal variance components. Sample period January-May 2010. Symbols α , β and ω denote GARCH parameters from the variance equation (4).

Source: own calculation

CONCLUSION

We have estimated an intraday volatility model for equally spaced logarithmic returns on the WIG index. In our specification, the conditional variance is a multiplicative product of daily, diurnal and stochastic intraday variance components. As the next step of our analysis, we plan to conduct forecasts comparison between our model and a selection of alternative benchmarks.

REFERENCES

Andersen T.G., T. Bollerslev T. (1997) Intraday Periodicity and Volatility Persistence in Financial Markets, Journal of Empirical Finance, 4, 115-158.

Andersen T.G., and Bollerslev T. (1998) DM-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer-Run Dependencies, Journal of Finance, 53, 219-265.

Bollerslev, T (1986) Generalized Autoregressive Conditional Heteroscedasticity. Journal of Econometrics, 31, 307-327.

Engle R. F. (1982) Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," Econometrica 50(4), 987-1007.

Engle R.F. and M.E. Sokalska (2010), Forecasting Intraday Volatility in the US Equity Market. Multiplicative Component GARCH, forthcoming in Journal of Financial Econometrics.

QUANTITATIVE METHODS IN ECONOMICS
Vol. XI, No. 1, 2010, pp. 145-157

CHANGES OF EXCHANGE RATE BEHAVIOR DURING AND AFTER CRISIS

Ewa Marta Syczewska
Institute of Econometrics, Warsaw School of Economics
e-mail: ewa.syczewska@sgh.waw.pl

Abstract: This study extends earlier analysis, in which behavior of daily exchange rates during the global crisis was compared to that before crisis. We repeat similar comparison for data set extended until the end of April 2010, use ARMA/ARMAX and GARCH models with stock indices as additional regressors, for volatility and returns of EURPLN, EURUSD, USDPLN exchange rates. Marked increase in volatility during crisis, negatively affected quality of models. After crisis volatility and returns seem to stabilize, hence exchange rate risk seems to decline gradually. There is a slight improvement in quality of models after the crisis.

Key words: Exchange rates, stock indices, crisis, risk, autoregressive and conditional heteroskedasticity models, Granger causality.

INTRODUCTION

The aim of research presented is to study the effects of the current crisis on exchange rate behavior, and on quality of exchange rate models. We use daily data of exchange rates USDPLN, EURPLN and EURUSD, and stock indices S&P500 and WIG20, since 4th January 2000 until 30th April 2010. We note stabilization of exchange rate behavior since spring 2009 in comparison to previous period (2007-2008), and assume that this may ease exchange rate modeling and improve quality of models.

We study the behavior of variance and volatility of models, test Granger-causality from stock indices towards exchange rate variances and returns. Next we estimate ARMA and ARMAX models for exchange rate volatility, and ARMA and GARCH models for logarithmic returns. We use S&P500 and WIG20 volatility or returns as additional variables in models describing respectively volatility or returns of the exchange rates.[24]

In [Syczewska 2010] we compared behavior of rates and quality of modeling for two subperiods: before (up to September 2008) and during crisis (up to end of July 2009). We have shown that[25]

- Volatility of returns, hence errors of forecasts from the ARMA and GARCH models of returns, hugely increased during the crisis.
- Introduction of corresponding stock indices returns led to a slight improvement of models and forecast performance.

There are several symptoms showing improvement of economic performance in current period. The Polish economy in particular during the crisis did well in comparison to other European economies. Whether the crisis ended can of course be argued, but let us treat the year (Spring 2009 – Spring 2010) as period "after crisis" and check if behavior of series modeled had stabilized enough to improve the quality of models.

QUALITATIVE DESCRIPTION OF THE DATA

We use daily data from 4th January 2000 until 30th April 2010 for the indices and exchange rates.[26] The typical measure of returns is:

(1) $z_t = 100 * (\ln y_t - \ln y_{t-1})$

where y_t – closing values of an instrument; we use also logarithm of proportion of daily maximum and minimum

(2) $\sigma_t^2 = \ln(y_{\max,t} / y_{\min,t})$

as a measure of variance/volatility (see [Brooks 2008]). Fig. 1 shows a typical behavior of stock index returns, an increase of volatility during both 2001 dotcom crisis and even higher increase during the last crisis. Fig. 2 shows a similar picture for exchange rate returns. Fig. 3 shows volatility defined by equation (2) for the corresponding exchange rate.

We compare both the whole sample and two equal subsamples: 16.10.2007 – 23.01.2009 as "crisis period" and 27.01.2009 – 30.04.2010 as "post-crisis period" (each consists of 312 observations). Choice of "crisis" period is to some extent arbitrary, but we follow [Reinhart and Rogoff 2008] as to characteristic symptoms of crisis.

---

[24] This was suggested by specification of models for daily returns of Norwegian krona [Bauwens, Rime and Succarat 2008].

[25] [Syczewska 2010] paper was presented at the International Conference "Zagadnienia aktuarialne – teoria i praktyka" in Warsaw, 2nd-4th September 2009.

[26] http://stooq.pl database, opening, closing, minimum and maximum daily quotes. We use only the dates, for which all quotes (Polish and American) were available.

Table 1. Comparison of variance in the two samples

Variable	Quote	Mean		Variance		**Proportion of variance 2^{nd} to 1^{st}**
		1st sample	2nd sample	1st sample	2nd sample	
S&P500	open	1244.6	1000.8	45552.4	17524.5	**38.5%**
	max	1257.7	1145.5	43986.7	16848.0	**38.3%**
	min	1227.3	1117.7	47624.3	18068.7	**37.9%**
	close	1242.2	1131.6	45817.4	17540.4	**38.3%**
WIG20	open	2695.8	2413.1	396358.4	117889.2	**29.7%**
	max	2723.8	2440.6	394610.1	115613.6	**29.3%**
	min	2657.3	2378.3	395741.6	119639.9	**30.2%**
	close	2690.0	2409.0	393756.3	117703.9	**29.9%**
USDPLN	open	2.4547	2.7100	0.09736	0.07607	**78.1%**
	max	2.4786	2.7390	0.10827	0.08185	**75.6%**
	min	2.4370	2.6860	0.09174	0.06955	**75.8%**
	close	2.4582	2.7120	0.10051	0.07587	**75.5%**
EURUSD	open	1.4608	1.4270	0.00953	0.00435	**45.7%**
	max	1.4697	1.4360	0.00909	0.00421	**46.3%**
	min	1.4519	1.4180	0.00999	0.00444	**44.4%**
	close	1.2970	1.4270	0.00959	0.00435	**45.3%**

Source: stooq.pl

Table 1 shows *decreases in variances* of the indices (between 30–40 percent) and exchange rates (between 75–80 percent for USD, approximately 45% for EURO); in contrast to comparison of "crisis" and "pre-crisis period" [Syczewska 2010] when it increased 2.6 times for S&P500, 2 times for WIG20, 2.8 times for USDPLN exchange rate, 1.7 times for EURUSD.

Volatility defined as in (2), i.e., log difference between maximum and minimum daily quotes for both stock indices has decreased to 70-90% of previous value, for USD exchange rate increased by 14 percent, for EURUSD – decreased to 89% of "crisis" volatility. For USDPLN, EURUSD exchange rates and for both stock indices difference in means is significant.

Table 2. Comparison of volatility during and after the crisis

	USDPLN	EURUSD	EURPLN	SP500	WIG20
Mean during crisis	.01598	.01248	.02846	.02674	.0264597
Mean after crisis	.01815	.01114	.02929	.01799	.0235659
Proportion	1.14	0.89	1.03	0.67	0.89
t statistic for difference in the means	-2.63	2.74	-0.67	6.66	2.53
Median during crisis	.01086	.01016	.02101	.01955	.02265
Median after crisis	.01630	.01038	0.02716	.01514	.02051
Proportion	1.50	1.02	1.29	0.77	0.91

Source: computations based on stooq.pl data.

GRANGER CAUSALITY TESTS

Variable x is defined as Granger-cause for another variable y, if lagged values of x used as additional regressors in a model describing y can improve quality of modeling/forecasting. There are several tests of this property. The Granger test of Granger causality is performed in the following way: we estimate VAR – type equation and check joint significance of lagged x parameters:

(3) $$y_t = a_{11} y_{t-1} + ... + a_{1k} y_{t-k} + b_{11} x_{t-1} + b_{12} x_{t-2} + ... + b_{1k} x_{t-k} + \varepsilon_t$$

The null H_0: $b_{11} = b_{12} = ... = b_{1k} = 0$ means that the x *does not* Granger-cause the y variable. This test is reported to work well for stationary variables, for non-stationary series, it should be used with caution. The Augmented Dickey-Fuller (in short, ADF) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests show stationarity of logarithmic returns for the whole sample. Results for subsamples ("crisis", "after crisis") are slightly different (see Table 3).

Table 3. The ADF test results for returns and volatility of variables

ADF	S&P500	WIG20	USDPLN	EURUSD
Returns	-12.22 (.00)	-21.11 (.00)	-9.89 (.00)	-10.72 (.00)
Volatility				
whole sample	-4.48 (.0001)	-3.4889 (.008)	-4.9505 (.00)	-3.4384 (.0097)
- crisis	-3.1005 (.03)	**-2.6481 (.08)**	**-0.8928 (.79)**	**-1.3654 (.60)**
- after crisis	**-2.4920 (.12)**	**-2.5312 (.11)**	**-2.0222 (.28)**	-3.0360 (.03)

Source: own computations; bold denotes insignificant values.

To check whether stock indices volatility/returns Granger-cause respective measures for the exchange rates, we estimate VAR(5) models.[27] For bilateral USDPLN exchange rate we check whether stock indices of respective economies, i.e. American and Polish indices, Granger-cause the exchange rate. Joint significance test statistic for lagged values of S&P500 index volatility F(5,2508) = 3.8661 with p-value 0.0017. Hence we reject the null of lack of causality: the US stock index Granger-causes volatility in USDPLN exchange rate, as expected. There is no causality from the exchange rate towards volatility of index. Full results of the Granger test of Granger causality are given in Table 4.

[27] Number of lags chosen by reduction of insignificant lags.

Table 4. Granger-causality tests from stock indices towards exchange rates

Causal relationship	Whole sample	Crisis	After crisis
WIG20→USDPLN	2.3771 [.0367]	**2.1216 [.0628]**	**1.5200 [.1833]**
SP500→USDPLN	3.8661 [.0017]	**1.8947 [.0951]**	**1.8937 [.0952]**
WIG20→EURUSD	303.41 [.0000]	**2.0457 [.0722]**	**1.0351 [.3970]**
SP500→EURUSD	7.1983 [.0000]	40.887 [.0000]	4.4195 [.0007]

Source: own computations; p-values in brackets, greater than 0.05 marked in bold.

FRACTIONAL INTEGRATION, PERSISTENCY AND LONG MEMORY MEASURES

As shown by the ADF test for the whole sample, log returns of exchange rates are stationary. Fractional integration parameter is perhaps more accurate indicator of time series behavior, either stationary or nonstationary. It generalizes the Engle and Granger [1987] definition of integrated series (Hosking [1981], Granger and Joyeux [1980]), it can take any real values (not only integer, as in the ADF test), and is defined with use of binomial series expansion or the Gamma function. The fractional integration parameter is often estimated with periodogram regression (one of the variants was introduced by [Geweke and Porter-Hudak 1983]). Another, semiparametric, method is the Whittle local estimator, introduced by [Künsch 1987] and [Robinson 1995] (see e.g. [Phillips and Shimotsu 2000]).

Classification of d values shows whether series in question is stationary or not, but more accurately, if it is persistent or antipersistent, whether results of external shocks diminish in time etc.

- For $d = 1$ a series is nonstationary, with infinite variance.
- For $1 \leq d$ variance is infinite, moreover results of a shock increase with time;
- If $0.5 \leq d < 1$, the process is nonstationary, but in a long-term reverts to its mean [Hosking 1981];
- For $0 < d < 0.5$, the process is stationary, with finite variance, and is mean reverting;
- For $d = 0$ it is mean-reverting in the short term, has finite variance and effects of shocks diminishes quickly;
- For $-0.5 < d < 0$ it is stationary, but mean-averting (antipersistent).

Quite similar classification can be done with use of the Hurst exponent [Hurst 1951]: if H=0.5 we have a random walk, if H is in (0; 0.5) – a mean-reverting process; for H in (0.5; 1) – mean-averting process with a trend.

Table 5. Hurst exponents for logarithmic returns

	Hurst exponent				
Logarithmic returns of:	SP500	WIG20	USDPLN	EURUSD	EURPLN
Whole sample	0,553 [3,10]	0,578 [4,84]	0,574 [10,89]	0,567 [9,24]	0,547 [3,77]
First subsample	*0,513 [0,75]*	*0,547 [1,69]*	0,636 [2,93]	0,608 [4,69]	*0,544 [0,62]*
Second subsample	0,578 [5,15]	0,520 [1,99]	*0,511 [0,28]*	0,609 [3,06]	0,459 [-3,22]

Source: own computations; t statistics in brackets

The Hurst exponent value of 0.5 corresponds to a white noise process, values greater than 0.5 but less than 1 suggest persistency and stationarity of a series. Table 5 shows computed values of the Hurst exponent for logarithmic returns of stock indices and exchange rates, computed for the whole sample and for two subsamples – during crisis and after crisis. Critical value of the Student t statistics is 2.44 for the whole sample, 2.22 for both shorter subsamples. We test the null of H = 0.5. Computed values of the t statistics show that the null cannot be rejected for stock indices and the log returns of EURPLN exchange rate in the first sample, and for WIG20 and USDPLN in the second sample. In other cases H is slightly greater than 0.5. Hence the Hurst exponents suggest stationarity and persistency of all logarithmic returns.

Table 6. Estimates of fractional integration parameter for logarithmic returns[28]

Returns of:	Method:	Whole sample	Subsample	
			First	Second
SP500	GPH	0,0013 [0,98]	-0,0374 [0,76]	0,1222 [0,69]
	Whittle	-0,0118 [0,81]	-0,0603 [0,51]	0,0913 [0,94]
WIG20	**GPH**	**0,1208 [0,07]**	0,0976 [0,49]	0,1393 [0,48]
	Whittle	0,0499 [0,30]	-0,0871 [0,34]	**0,0913 [0,08]**
USDPLN	GPH	0,0636 [0,32]	0,0396 [0,80]	0,1557 [0,79]
	Whittle	**0,0861 [0,07]**	0,0014 [0,99]	0,0913 [0,40]
EURUSD	GPH	0,0493 [0,42]	0,2026 [0,16]	0,1390 [0,73]
	Whittle	0,0560 [0,24]	0,0886 [0,33]	0,0913 [0,51]
EURPLN	GPH	0,0754 [0,23]	**-0,2265 [0,06]**	0,1170 [0,25]
	Whittle	0,0611 [0,20]	-0,0398 [0,66]	**0,0913 [0,05]**

Source: own computations

28 In brackets there are p-values of t statistics in case of Geweke and Porter-Hudak method, z statistics in case of the Whittle estimator, both for a null of insignificance.

More accurate are results of the GPH and Whittle methods, allowing for tests of significance of the fractional integration parameter. Results presented in table 6, with p-values of statistics for null of d =0 in brackets, show that all returns series are stationary. Estimates of the fractional integration parameter are in most cases insignificant, as shown by the t or z-statistics p-values. Only in few cases the null of insignificance is rejected – note in particular difference between results for EURPLN, with positive insignificant values for the whole sample, negative significant value for the first subsample, suggesting stationary antipersistent behavior, and positive significant value for the second subsample, suggesting persistent stationary behavior. For stationary series and ARMA model with finite number of parameters can be estimated.

FRACTIONAL INTEGRATION PARAMETERS FOR VOLATILITY

We estimate in a similar way fractional integration parameters for volatility defined as in (2). The results are as follows (see table 7); estimates of the fractional integration parameter are in interval (0.5; 1), suggesting nonstationarity of a series. The Hurst exponents, H, are close to 1. These may pose problems with the choice of number of lags for the ARMA model. Fig. 4 and 5 show periodogram for log returns and for volatility (2) for the USDPLN exchange rate. The first is similar to spectrum of a stationary series, the second has relatively high values for lower frequencies, which is closer to a behavior for nonstationary series.

We test whether H =1 and whether the fractional integration parameter, d is equal first, to 0.5 and second, to 1.[29] The null hypothesis H = 1 is rejected for WIG20 and USDPLN. The null hypothesis d=0.5 cannot be rejected for WIG20 and EURUSD in case of the GPH estimation, and for WIG20, USDPLN and EURUSD in case of the Whittle method. The null hypothesis d=1 is rejected in all cases. Hence volatility is in all cases at least mean-reverting in long term, nonstationary if d >0.5, stationary if d = 0.5.

Table 7. Hurst exponents, the Geweke-Porter-Hudak and Whittle estimators of fractional integration parameter for volatility

	SP500	WIG20	USDPLN	EURUSD	EURPLN
Hurst exponent	.9702 (.020)	.8845 (.015)	.9405 (.025)	.9741 (.038)	.9702(.031)
GPH estimator	.7031 (.058)	.4904 (.066)	.6248 (.050)	.5809 (.059)	.6894 (.055)
Whittle estimator	.6548 (.048)	.5240 (.048)	.5867 (.048)	.5878 (.048)	.6333(.048)

Source: own computations; standard errors in parentheses

[29] The Hurst exponent was estimated with use of R/S regression with 9 degrees of freedom, hence critical value t* = 2.26. The GPH and Whittle estimators are based on regressions with 109 degrees of freedom, hence critical value t* = 1.98.

ARMA/ARMAX MODEL

We expect that stock indices can improve quality of ARMA models for volatility (2). We apply ARMA formulation, with 4 lags as starting point, then try and reduce the model on the basis of significance tests and Schwarz Bayesian Information Criterion (procedure of reduction if similar, e.g., to [Matuszewska, Witkowska 2007], albeit their starting point is an autoregressive distributed lags model).

Model estimated for pre-crisis period has been reduced to ARMA(1,1) [Syczewska 2010]. For USDPLN exchange rate volatility 3rd lags are significant. We use S&P500 and WIG20 volatility (2) as additional regressors; both are significant. Similar results are obtained for other exchange rates. The roots of AR and MA polynomials have absolute values greater than 1, hence all ARMA/ARMAX models estimated are stable. For all exchange rates models with the stock indices variables have lower values of Akaike, Schwartz and Hannan-Quinn information criteria.

GARCH MODELS FOR LOGARITHMIC RETURNS

We estimate next the ARMA and ARMAX models for the logarithmic returns of exchange rates, starting first with an ARMA(4,4,) model, and then adding logarithmic returns of both indices. The added variable parameters prove to be significant. For both models we could not reject the ARCH effect:
The Engle test of the ARCH effect is based on the regression

$$e_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \alpha_2 e_{t-2}^2 + ... + \alpha_k e_{t-k}^2 + u_t$$

where *e* are error terms of the model in question. We check whether lagged error squares are jointly significant: the null $H_0 : \alpha_1 = \alpha_2 = ... = \alpha_k = 0$ corresponds to lack of the ARCH effect. Under the null, the test statistic is asymptotically distributed as $\chi^2(k)$.

For USDPLN, computed values of the test statistics in case of ARMA and ARMAX models estimated for the whole sample, are equal respectively to TR^2 = 396.238 and 322.041, respectively. Hence the null hypothesis of no ARCH effect is rejected.

As a result we estimate GARCH model for logarithmic returns of exchange rates, with and without log returns of stock indices as additional variables: starting with the GARCH model with 10 lags in autoregressive equation of mean value of the USDPLN log returns, and reducing insignificant lags, we reduce this model to one with AR(6) equation for expected value and GARCH(1,1) for variance.[30]

---

[30] GARCH(1,1) is in most cases well suited for stock indices and exchange rates modeling, see e.g. [Brzeszczyński, Kelm 2002].

Starting with the GARCH model, 10 lags in autoregressive equation for the mean and log returns of the S&P500 and the WIG20 indices as additional variables, we reduce the model to only AR(1) with SP500 and WIG20 as additional variables, and GARCH(1,1) for conditional variance.

To compare results of forecasting exercise for both versions (with and without additional variables), we decide to use one lag for expected value equation. We reestimate the models for shorter sample, up to 2009/01/26, and compute forecasts for end of the whole sample (up to the end of April 2010). Forecasts quality is still not impressive, according to mean absolute error MAE, mean squared error MSE, mean absolute percentage error MAPE, and the Theil U indicator, given by

$$U = \frac{\sqrt{\frac{1}{m}\sum_{t=1}^{m}(y_t - y_t^p)^2}}{\sqrt{\frac{1}{m}\sum_{t=1}^{m} y_t^2} + \sqrt{\frac{1}{m}\sum_{t=1}^{m}(y_t^p)^2}}$$

(where y_t, y_t^p denote observation of the series and value of forecast in period *t*, and m denotes forecast horizon).[31] Models with additional regressors perform slightly better.

CONCLUSIONS

We perform analysis for logarithmic returns and for volatility (defined as log difference between maximum and minimum of daily quotes) for daily values of exchange rates. We compare their behavior in period 2007-2010 (during crisis) and since January 2009 until April 2010. We note that the series have slightly stabilized, although volatility is still quite high in comparison to the period before the crisis.

Long memory of the series and values of the fractional integration parameter indicating nonstationarity for the crisis period result in greater number of lags in the ARMA model in comparison to the earlier period (where the ARMA model specification has been chosen using significance tests and information criteria).

Granger-causality tests show that the corresponding measures of stock indices Granger-cause returns or volatility of exchange rates. The Engle test shows presence of ARCH effect.

Indeed, results of estimation and of in-sample forecasting exercise, show that S&P500 and WIG20 stock indices measures used as additional regressors in mean equation, improve slightly the quality of ARMAX and GARCH models for either

[31] See M. Gruszczyński and M. Podgórska (eds.), „Ekonometria", Warsaw School of Economics, Warsaw 2004, p. 117.

returns or volatility of the exchange rates. The volatility of exchange rates until the end of April 2010 is quite high, hence to improve quality of modeling we should wait for hopefully further stabilization of the series and work towards improvement of specification of the estimated econometric models.

REFERENCES

Bauwens, L., Pohlmeier, W., Veredas, D. (2008) *High Frequency Financial Econometrics. Recent Developments*, Physica-Verlag A Springer Company, Heidelberg.

Bauwens, L., Rime, D., Succarat, G. (2008) *Exchange Rate Volatility and the Mixture of Distribution Hypothesis*, [in:] Bauwens et al., 7–29.

Brooks, Ch. (2008) *Introductory Econometrics for Finance*, 2nd ed., Cambridge University Press, New York.

Brzeszczyński, J., R. Kelm (2002) *Ekonometryczne modele rynków finansowych. Modele kursów giełdowych i kursów walutowych*, WIG Press, Warsaw 2002.

DiMartino, D., J.V. Duca, *The Rise and Fall of Subprime Mortgages*, Federal Reserves Bank of Dallas, Economic Letter, 2(11), November 2007.

Engle, R.F., Granger, C.W.J. (1987), *Cointegration and error correction: representation, estimation and testing*, Econometrica, 55, 251–276.

Geweke, J., Porter-Hudak, S. (1983), *The estimation and application of long-memory time series models*, Journal of Time Series Analysis, 4, 221–228; reprinted in: Robinson (2003).

Granger, C.W.J., Joyeux R. (1980), *An introduction to long memory time series models and fractional differencing*, Journal of Time Series Analysis, 1, 15–29.

Gruszczyński, M., Podgórska M. (eds.) (2004) *Ekonometria*, VIIth ed., Oficyna Wydawnicza SGH, Warsaw 2004.

Hosking, J.R.M. (1981), *Fractional differencing*, Biometrika, 68(1), 165–176.

Hurst, H. (1951), *Long term storage capacity of reservoirs*, Transactions of the American Society of Civil Engineers, 116, 770–799.

Künsch, H. (1987), *Statistical aspects of self-similar processes*, In: *Proceedings of the First World Congress of the Bernoulli Society* (Yu. Prokhorov and V. V. Sazanov, eds.) 1 67-74. VNU Science Press, Utrecht.

Kuszewski, P., Podgórski J. (2008) *Statystyka. Wzory i tablice*, Warsaw School of Economics, Warsaw 2008.

Kwiatkowski, D., Phillips P.C.B., Schmidt P., Shin Y. (1992) *Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?*, Journal of Econometrics, 54, 159–178.

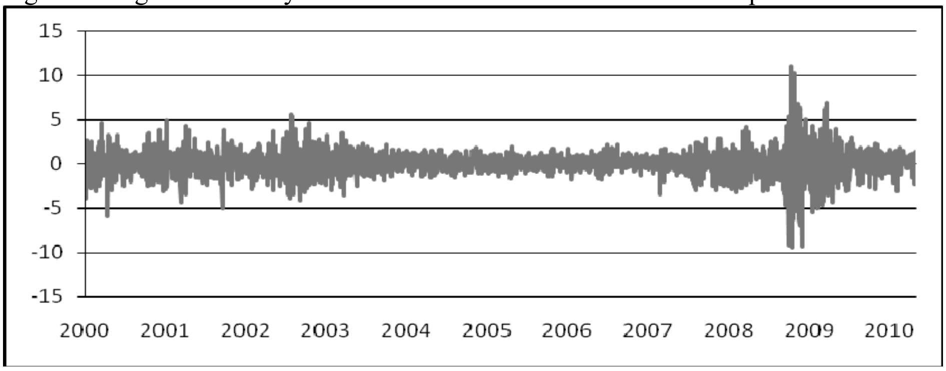
Matuszewska, A., Witkowska, D. (2007), *Wybrane aspekty analizy kursu euro/dolar: Modele autoregresyjne z rozkładami opóźnień i sztuczne sieci neuronowe*, Metody ilościowe w badaniach ekonomicznych VIII, Modele ekonometryczne, 203–212.

Phillips, P.C.B., K. Shimotsu (2000), *Local Whittle estimation in nonstationary and unit root cases*, Cowles Foundation Discussion Paper No. 1266, New Haven, http://cowles.econ.yale.edu/P/cd/d12b/d1266.pdf

Reinhart, C.M, Rogoff K.S.(2008) *Is the 2007 US Sub-prime Financial Crisis So Different? An International Historical Comparison*, American Economic Review, American Economic Association, 98(2), 339–44, May 2008.

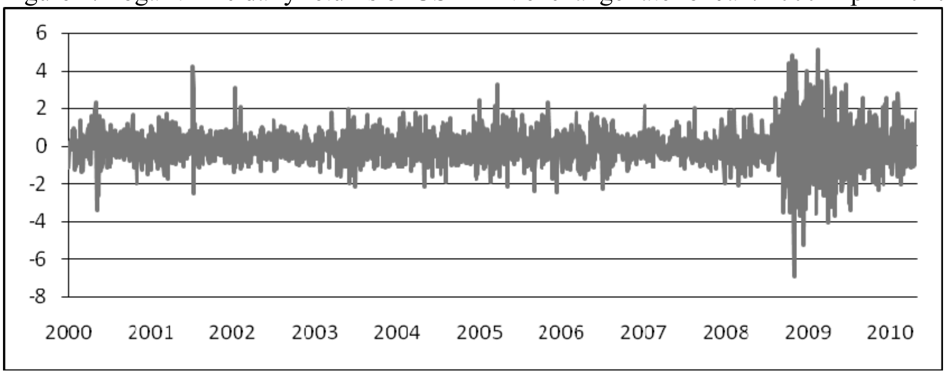
Reinhart, C.M., Rogoff, K.S. (2009) *The Aftermath of Financial Crises, American Economic Review*, American Economic Association, 99(2), 466-472, May 2009.
Robinson, P. M. (1995). *Gaussian semiparametric estimation of long range dependence*. Ann. Statist., 23, 1630-1661.
Robinson, Peter M. (ed.) (2003), *Time series with long memory*, Oxford University Press, Oxford.
Schwert, G.W.(1989) *Tests for unit roots: A Monte Carlo investigation*, Journal of Business and Economic Statistics, 2, 147–159.
Syczewska, E.M. (2010) *Increase of exchange rate risk during current crisis*, Roczniki Kolegium Analiz Ekonomicznych, Warsaw School of Economics, [in print].

Figure 1. Logarithmic daily returns of S&P500 index for Jan. 2000–April 2010



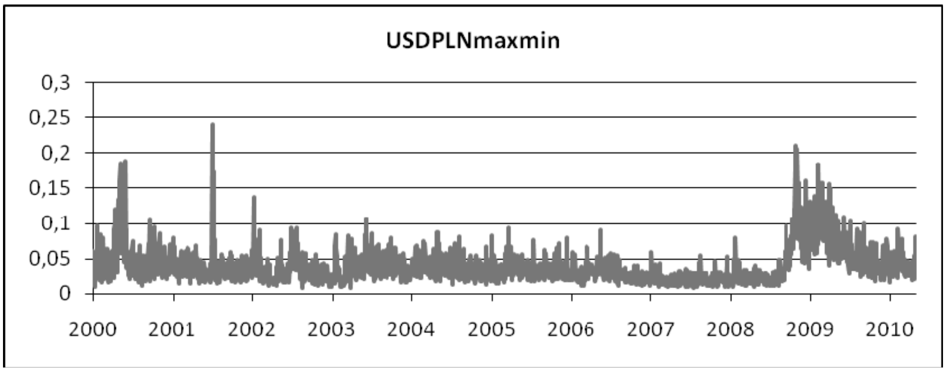
Source: http://stooq.pl

Figure 2. Logarithmic daily returns of USDPLN exchange rate for Jan. 2000–April 2010



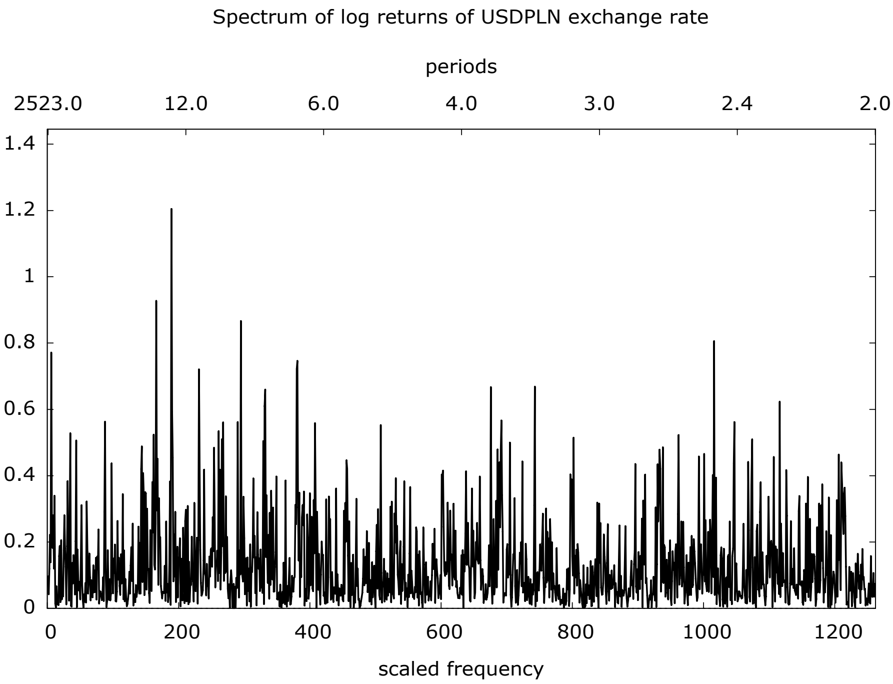
Source: http://stooq.pl

Figure 3. Same day max-min volatility of USDPLN exchange rate for Jan. 2000–April 2010



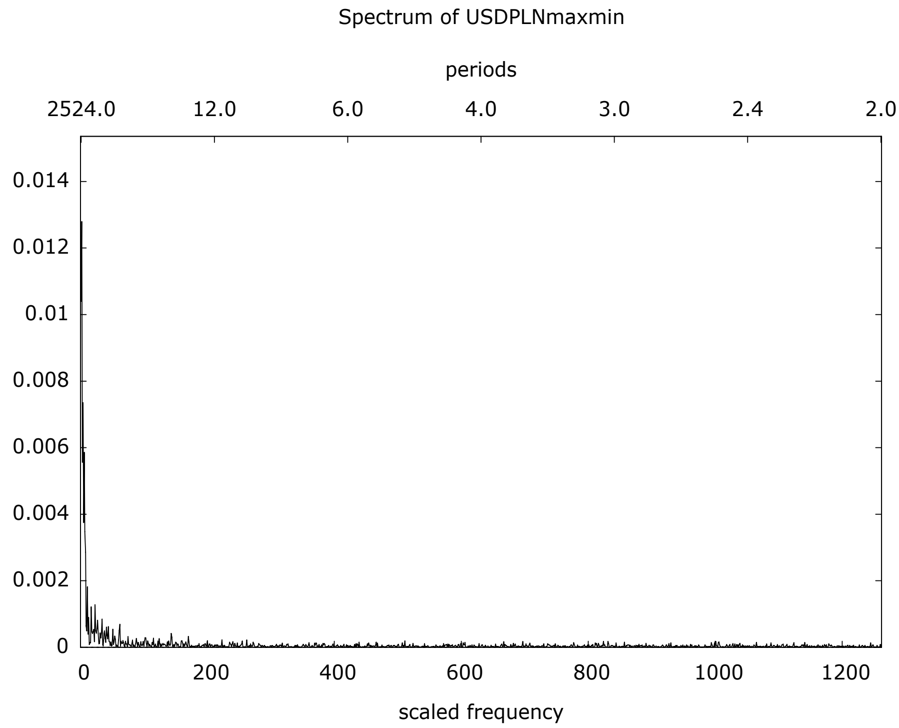
Source: own computations

Figure 4. Periodogram for logarithmic returns of USDPLN exchange rate for Jan. 2000–April 2010



Source: own computations

Figure 5. Periodogram for volatility of USDPLN exchange rate for Jan. 2000–April 2010



Source: own computations

QUANTITATIVE METHODS IN ECONOMICS
Vol. XI, No. 1, 2010, pp.158 -164

STATISTICAL PROPERTIES OF A CONTROL DESIGN OF CONTROLS PROVIDED BY SUPREME CHAMBER OF CONTROL

Wojciech Zieliński
Katedra Ekonometrii i Statystyki SGGW
Nowoursynowska 159, PL-02-767 Warszawa
e-mail: wojtek.zielinski@statystyka.info

Abstract. In statistical quality control objects are alternatively rated. It is of interest to estimate a fraction of negatively rated objects. One of such applications is a quality control provided by Supreme Chamber of Control (NIK) to find out a percentage of abnormalities in the work among others of tax offices. Mathematical details of experimental designs for alternatively rated phenomena are given in Karliński (2003). There are given requirements for sample sizes, numbers of negative rates, accuracy of estimation and error risks. In the paper, some statistical properties of given plans are investigated.

Key words: statistical quality control, alternative rating, experimental design

One of the problem of the statistical quality control is the problem of the estimation of the fraction of defective products. Generally speaking, the products are alternatively rating and one is interested in estimation of a fraction of negatively rated objects. In this approach, the binomial statistical model is applied, i.e. if ξ is a random variable counting negative rated in a sample of size n, then ξ is binomially distributed

$$P_\theta\{\xi = x\} = \binom{n}{x}\theta^x(1-\theta)^{n-x}, \quad x = 0,1,\ldots,n,$$

where $\theta \in (0,1)$ is a probability of drawing a defective product. The aim of the statistical quality control is to estimate θ.

In many norms and books devoted to different applications there are given exacts designs of experiments, i.e. requirements for ample sizes, number of negative rates in the sample, accuracy of estimation and error risks. One of such applications are quality controls provided by Supreme Chamber of Control, the goal of which is finding abnormalities in tax offices. Karliński (2003) gives mathematical details of such controls. There are given methods of providing experiments and rules of statistical inference. In what follows, statistical properties of given experimental designs are investigated.

The aim of a control is the interval estimation of a percentage of the defective objects. NIK guidelines are based on approximate solutions. There is given a method of calculating a minimal sample size:

$$n = \frac{Nu_\alpha^2\theta(1-\theta)}{\varepsilon^2(N-1)+u_\alpha^2\theta(1-\theta)},$$

where N is the size of the population, u_α is the critical value of the standard normal distribution for $1-\alpha$ confidence level, ε is the given accuracy of estimation and θ is the real (assumed in control) theoretical percentage of defective objects. Let the number of observed defective objects in a sample be m .Then

$$e = u_\alpha\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}\frac{N-n}{N-1}}$$

and an interval is obtained

$$\left(\hat{\theta}-e,\hat{\theta}+e\right).$$

The interval is considered as a confidence interval for the fraction of failures.

Let us investigate the statistical properties of the above method. In what follows it is assumed that the population is infinite. Under the assumption the given formulae take on the form

$$n = \frac{u_\alpha^2\theta(1-\theta)}{\varepsilon^2} \quad \text{and} \quad e = u_\alpha\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}.$$

The formulae are derived from asymptotic approximations of e Binomial distribution. Central Limit Theorem states that for large n we have

$$P_\theta\{\xi \le k\} \approx \Phi\left(\frac{\xi - n\theta}{\sqrt{n\theta(1-\theta)}}\right),$$

where $\Phi(\cdot)$ denote the cumulative distribution function of the standard normal distribution $N(0,1)$. Hence, it assumed that $(\xi - n\theta)/\sqrt{n\hat{\theta}(1-\hat{\theta})}$ is asymptotically normal $N(0,1)$, it means that

$$P_\theta\left\{\frac{|\hat{\theta}-\theta|}{\sqrt{\hat{\theta}(1-\hat{\theta})}}\sqrt{n}\le u_\alpha\right\}\approx 1-\alpha$$

From the above we obtain given earlier interval.

Taking $1-\alpha=0.95$, $\varepsilon=0.05$ and $\theta=0.05$ we obtain

$$u_{0.05}=1.96, n=73, e=0.0499967.$$

The number of failures in the sample is a random variable ξ distributed as Binomial $B(73,0.05)$. Numerical results which demonstrate statistical properties of the above confidence are given In the Table 1.

Table 1.

m	$\left(\hat{\theta}-e,\hat{\theta}+e\right)$		$P_{0.05}\{\xi=m\}$
0	(-0.049997,0.049997)	0	0.0236
1	(-0.036298,0.063695)	1	0.0909
2	(-0.022599,0.077394)	1	0.1722
3	(-0.008901,0.091093)	1	0.2144
4	(0.004798,0.104791)	1	0.1975
5	(0.018496,0.118490)	1	0.1435
6	(0.032195,0.132188)	1	0.0856
7	(0.045894,0.145887)	1	0.0431
8	(0.059592,0.159586)	0	0.0187
9	(0.073291,0.173284)	0	0.0071
10	(0.086990,0.186983)	0	0.0024
⋮	⋮	⋮	⋮

In the column before last one denotes that the obtained interval covers the estimated value 0.05. Note that, for small values of m the left end of the interval is a negative number. Moreover, multiplying the last columns the real confidence level is obtained: 0.9471. This is smaller value than nominal 0.95. In consequence, by the application of the above method more correct populations are considered as wrong ones. Note that the expected length is 0.1 and equals given accuracy ε of estimation.

Clopper and Pearson (1934) give the confidence interval for θ, based on the exact distribution of ξ. Because

$$P_\theta\{\xi \le x\} = \beta(n-x, x+1; 1-\theta) \quad \text{and} \quad P_\theta\{\xi \ge x\} = \beta(x, n-x+1; \theta),$$

where $\beta(a,b;\cdot)$ denotes the CDF of Beta distribution with parameters (a,b), hence the confidence interval has the form $\left(\theta_L(x), \theta_U(x)\right)$, where

$$\theta_L(x) = \beta^{-1}\left(x, n-x+1; \frac{\alpha}{2}\right), \qquad \theta_U(x) = \beta^{-1}\left(x+1, n-x; 1-\frac{\alpha}{2}\right).$$

For $x=0$ we take $\theta_L(0)=0$, and for $x=n$ is taken $\theta_U(n)=1$.

Here $\beta^{-1}(a,b;\cdot)$ denotes the quantile of the Beta distribution with parameters (a,b).

In the Table 2 there are given analogous results as in the Table 1, but for Clopper—Pearson confidence interval.

Table 2.

m	(left, right)		$P_{0.05}\{\xi = m\}$
0	(0.0000,0.0493)	0	0.0236
1	(0.0003,0.0740)	1	0.0909
2	(0.0033,0.0955)	1	0.1722
3	(0.0086,0.1154)	1	0.2144
4	(0.0151,0.1344)	1	0.1975
5	(0.0226,0.1526)	1	0.1435
6	(0.0308,0.1704)	1	0.0856
7	(0.0394,0.1876)	1	0.0431
8	(0.0485,0.2046)	1	0.0187
9	(0.0580,0.2212)	0	0.0071
10	(0.0677,0.2375)	0	0.0024
⋮	⋮	⋮	⋮

True confidence level of the classical confidence interval equals 0.9659, so it is higher than nominal 0.95. Unfortunately expected length of the confidence interval is 0.1119 and is bigger than postulated precision $2\varepsilon = 0.1$. To gain the accuracy, the sample size should be enlarged. It is easy to calculate that minimal sample size is $n = 90$. The results of calculations for that sample size are given in the Table 3.

Table 3.

m	(left, right)		$P_{0.05}\{\xi = m\}$
0	(0.0000,0.0402)	0	0.0099
1	(0.0003,0.0604)	1	0.0468
2	(0.0027,0.0780)	1	0.1097
3	(0.0069,0.0943)	1	0.1694
4	(0.0122,0.1099)	1	0.1939
5	(0.0183,0.1249)	1	0.1755
6	(0.0249,0.1395)	1	0.1309
7	(0.0318,0.1537)	1	0.0827
8	(0.0392,0.1677)	1	0.0451
9	(0.0468,0.1814)	1	0.0216
10	(0.0546,0.1949)	0	0.0092
11	(0.0626,0.2082)	0	0.0035
⋮	⋮	⋮	⋮

The true confidence level equals 0.9756, and its expected length is 0.0998. It is seen that expected length is smaller than required. For the sample of the size $n = 89$ expected length of the confidence interval equals 0.1004, which is a little bigger that assumed. To obtain precision exactly 2ε, the randomization should be applied in the following way

$$n = \begin{cases} 89, \text{ with probability } 0.3461 \\ 90, \text{ with probability } 0.6539 \end{cases}$$

Expected length equals now

$$0.0998 \cdot 0.3461 + 0.1004 \cdot 0.6539 = 0.1.$$

Of course, drawing sample size should be done before realization of the proper experiment. Any random number generator may be applied, for example the one in Excel.

In both cases, i.e. for sample size 89 and 90, the real confidence level is greater than nominal one. This inconvenience may be suppressed by second randomization. This randomization is applied in the construction of the confidence interval. If the number of observed failures is m two random numbers u_1, u_2 from uniform distribution $U(0,1)$ are drawn and ends of the confidence interval are obtained by solving two equations:

$$\textit{left} \quad u_1\beta(m-1, n-m+2; \theta) + (1-u_1)\beta(m, n-m+1; \theta) = 0.025, \quad (L)$$

$$\textit{right} \quad u_2\beta(m+1, n-m; \theta) + (1-u_2)\beta(m+2, n-m-1; \theta) = 0.975. \quad (R)$$

In that way, obtained confidence interval has the confidence level exactly 0.95.
In practice, the above method may be realized as follows. In the spreadsheet Excel three numbers are drawn according to the uniform $U(0,1)$ distribution:

$$u_0 = 0.3820, u_1 = 0.1006, u_2 = 0.5964.$$

Number u_0 is used to drawing a sample size: because u_0 is greater than 0.3461, hence the sample size is $n = 90$.

In the experiment $m = 10$ defective objects were observed. The two following equations are solved (Addin Solver in Excelu may be used):

$$\textit{left} \quad 0.1006\beta(9, 82; \theta) + 0.8994\beta(10, 81; \theta) = 0.025, \quad (L)$$

$$\textit{right} \quad 0.5964\beta(11, 80; \theta) + 0.4036\beta(12, 79; \theta) = 0.975. \quad (R)$$

The interval is obtained

$$(0.0535, 0.2012).$$

Drawn numbers u_0, u_1 and u_2 should be added to the report.
As it was mentioned, all calculations may be done in Excel. There are following useful functions.
BETADISTRIBUTION(x;alfa;beta;A;B): where alfa and beta are the parameters of the distribution. The function gives a values of CDF at point x. Numbers A and B defines a support of the distribution: default values are 0 and 1.

BETAINV(probability;alpha;beta;A;B): where alfa and beta are parameters of the distribution. The function gives the probability quantile of the Beta distribution. Numbers A and B defines a support of the distribution: default values are 0 and 1.

The confidence interval in the binomial model may be calculated in the following way.

	A	B
1	100	Sample size
2	10	Number of succeses
3	0.95	Confidence level
4	=IF(A2=0;0;BETAINV((1-A3)/2;A2;A1-A2+1))	Left end
5	=IF(A2=A1;1;BETAINV((1+A3)/2;A2+1;A1-A2))	Wright end

In cells A4 and A5 the values of the left and the right end of the interval are obtained.

The worksheet calculating a randomized confidence interval is a little bit complicated.

	A	B
1	100	Sample size
2	10	Number of succeses
3	0.95	Confidence level
4	0.1006	u1
5	0.5964	u2
6	0.1	Left end
7	=A4* BETAINV (A6;A2-1;A1-A2+2) +(1-A4)* BETAINV (A6;A2;A1-A2+1)-(1-A3)/2	equation (L)
8	0.1	Right end
9	=A5* BETAINV (A8;A2+1;A1-A2) +(1-A5)* BETAINV (A8;A2+2;A1-A2-1)-(1+A3)/2	equation (R)

To obtain the randomized confidence interval the Addin Solver should be used twice. Firstly, the goal is cell A7 by changing A6, secondly the goal is the cell A9 by changing A8. In cells A6 and A8 ends of randomized confidence interval are calculated. Numbers in A4 and A5 are obtained from a random number generator (Addis DataAnalysis in Excel).

More information on the confidence level may be fund in Zieliński (2009), and on randomized confidence intervals in Bartoszewicz (1996).

REFERENCES

Bartoszewicz, J. 1996: Wykłady ze statystyki matematycznej, wyd. II, PWN Warszawa.

Clopper C. J., Pearson E. S. 1934: The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial, Biometrika 26, 404-413.

Karliński W. 2003: http://nik.gov.pl/docs/kp/kontrolapanstwowa-200305.pdf

Zieliński R. 2009: Przedział ufności dla frakcji, Matematyka Stosowana 10(51), 51-67.

To the memory of Leonid Hurwicz (1917-2008)
Graduate of the University of Warsaw (1938)
Nobel Laureate in Economics (2007)

From Econometrics to Game Theory

edited by: *Aleksander Strasburger*

Quantitative Methods in Economics
Vol. XI, No. 1, 2010, pp. 167–188

ASYMPTOTIC NASH EQUILIBRIA IN DISCOUNTED STOCHASTIC GAMES OF RESOURCE EXTRACTION

Łukasz Balbus

Institute of Mathematics and Computer Sciences, Wroclaw University of Technology
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
e-mail: lukasz.balbus@pwr.wroc.pl

Abstract: A class of two person nonzero-sum nonsymmetric stochastic games of capital accumulation/resource extraction is considered. It is shown that the Nash equilibrium in the discounted games has a limit when the discount factor tends to 1. Moreover, this limit is an epsilon-equilibrium in the discounted game with sufficiently large discount factor.

Key words: Nonzero-sum stochastic games. capital accumulation problems. resource extraction games. Nash equilibrium.

Introduction

Nonzero-sum dynamic games of capital accumulation (or resource extraction) studied in this paper belong to a class of stochastic games with uncountable state space. Unlike in the games with a countable state space (see [9]), the question of the existence of stationary Nash equilibria in such games has a positive answer only in some special cases of interest. For a survey of the existing literature on this topic the reader is referred to [1, 2, 7, 12, 13, 14] and [16]. The special case of these games is a class of concave games of resource extraction or capital accumulation. The pioneering work on this field is [10]. Similar class is studied in [1, 2, 3, 4, 12, 14] and [16]. Our main assumption on the transition probability function in the game says that it is a combination of finitely many probability measures on the state space with the coefficients depending on the investment. Similar form of it we can find in [2] and in [16].

In our model, we restrict our assumptions to two person game. Asymptotic properties of Nash equilibria in the discounted stochastic game, with respect to discount factor tending to 1, are the main results of this paper. We accept the assumptions, from the model considering in [14]. There is proven existence of Nash equilibria in finite and infinite horizon game. Moreover, there is also proven the uniform convergence of equilibria payoffs with respect to horizon of the game. The limit is an equilibrium payoff in infinite horizon game. The convergence of Nash equilibria is proven as well, but there is only pointwise convergence. In this paper we can complete those results and prove uniform convergence. This also leads us to uniform convergence of Nash equilibria and adequate equilibria payoffs with respect the discount factor tending to 1 in finite and in infinite horizon game. These limits it is shown to be Nash equilibrium and adequate equilibrium payoff in the undiscounted model. Similar problems are often described in the literature. However, unlike in [8, 11] and [17] we also obtain further results. Uniform convergence yields us more properties. Taking an arbitary small ε, the Nash equilibrium in the undiscounted model is ε - equilibrium in the discounted games with sufficiently large discount factor. The results obtained in this paper would be useful in the discounted model, if we do not know exact value of the discount factor. Then we could approximate Nash equilibria in discounted models by the Nash equilibrium in the undiscounted model.

This paper is organized in a following way: in the next section is presented a model with main assumptions. In the third section the model of auxillary one shot game and some properties of it are described. This model is a special case of that from Section 3 in [3]. The main results are discreibed in the fourth and fifth section. Fourth section is about finite horizon game and fifth about infinite horizon game. At the end of the last section there is an example in which the results obtained in this paper would be useful.

The model and basic assumptions

Consider an two-person nonzero-sum nonsymmetric stochastic game in which:

(i) $S \subset R_+ := [0, \infty)$ is a compact interval containing zero called the *state space* or the set of all possible stocks of a resource. Denote $S := [0, \bar{s}]$, for some strictly positive constant $\bar{s}$.

(ii) $A_i(s) \subset S$, $A_i(s) := [0, a_i(s)]$ is the *space of actions* available to player i in the state $s \in S$, where $a_i(\cdot)$ is called *capacity function.* Assume that $a_i(\cdot)$ is nonnegative, nondecreasing and continous function such that $a_1(s) + a_2(s) \leq s$. For $\bar{x} := (x_1, x_2)$ let

$$X(s) := A_1(s) \times A_2(s) \quad \text{and} \quad D = \{(s, \bar{x}) : s \in S, \bar{x} \in X(s)\}.$$

(iii) $u_i : S \to R_+$ is a bounded *instantaneous utility function* for player i.

(iv) q is a Borel measurable transition probability from D to S, called the *law of motion* among states. If s is a state (resource stock) at some stage of the game and the players select an $\bar{x} = (x_1, x_2) \in X(s)$, then $q(\cdot|s, x_1, x_2)$ is the probability distribution of the next stock.

Our further assumptions are:

C1: The utility function for player i $u_i : S \to R_+$ is a strictly concave twice continuously differentiable and increasing function such that $u_i(0) = 0$.

C2: The transition probability is of the form

$$q(\cdot|s, \bar{x}) = \sum_{l=1}^{L} g_l\,(s - x_1 - x_2)\,\lambda_l(\cdot|s) + g_0(s - x_1 - x_2)\delta_0(\cdot),$$

where

(a) δ_0 is the Dirac measure concentrated at 0,

(b) λ_l is a Borel measurable transition probability from S to S. Moreover,

- assume additionally that there exists a probability measure μ such that, μ stochasticaly dominates $\lambda_l(\cdot|s)$ for all $s \in S$, $l = 1, ..., L$, and holds

  $$\int_S g_0(s)\mu(ds) > 0,$$

  or equivalentely

  $$\int_S \sum_{l=1}^{L} g_l(s)\mu(ds) < 1.$$

- for each $l = 1, ..., L$ and for each Borel measurable and bounded function $v(\cdot)$ the function $s \to \int_S v(s')\lambda_l(ds'|s)$ is continous. Clearly λ_l is a stochasticaly continous measure.

(c) $g_0(0) = 1$ and for $l = 1\ldots, L$ $g_l : S \to [0, 1]$ is strictly concave, increasing and twice continuously differentiable. Obviously $\sum_{l=0}^{L} g_l \equiv 1$.

Remark. By assumption **C2** it follows that 0 is an absorbing state since

$$q(\{0\}|0, 0, 0) = g_0(0)\delta_0(\{0\}) + \sum_{l=1}^{L} g_l(0)\lambda_l(\{0\}|0) = \delta_0(\{0\}) = 1.$$

Remark. Note that assumption **C2** it satisfied when

$$\sum_{l=1}^{L} g_l(\bar{s}) < 1.$$

Then as μ we can take a Dirac measure concentrated at $\bar{s}$.

Remark. Typical examples of the functions g_l are:

$$g_l(y) := \alpha(1 - e^{-y}) \quad \text{or} \quad g_l(y) := \alpha\sqrt{y}, \quad y = s - x_1 - x_2,$$

where $\alpha > 0$ are small enough, S is bounded from above.

Remark. Typical example of the collection of measures λ_l are

$$\lambda_l(A|s) := \int_A \rho_l(s, s')\nu_l(ds'),$$

where $\nu_l(\cdot)$ is some probability measure on S and $\rho_l(s, \cdot)$ is a density for each $s \in S$. Moreover, for each $s' \in S$ $\rho_l(\cdot, s')$ is a continous function.

The game is played in discrete time with past history as common knowledge for all the players. A *strategy* for a player is a Borel mapping which associates with each given history an action available to him. For $i = 1, 2$ let F_i be set of Borel measurable functions $f_i : S \to R_+$ such that $f_i(s) \in A_i(s)$ for each $s \in S$. A *Markov strategy* for player i is a sequence $\pi_i = (f_{i,1}, f_{i,2}, \dots,)$ where each $f_{i,n}$ belongs to F_i. A *stationary strategy* for player i is a constant sequence π_i where each $f_{i,n} = f_i$ for some $f_i \in F_i$. Let $F := F_1 \times F_2$ be a set of all *profiles*. In the sequence, a stationary strategy $(f_i, f_i, ...)$ of player i will be identified with f_i.

Let $H^\infty := D \times D \times \dots$ by the set of all possibles histories of the game endowed with the product σ-field. For any profile of strategies $\pi = (\pi_1, \pi_2)$, and every initial state $s \in S$, a probability measure P_s^π and a stochastic proces $\{s_n, \bar{x}_n\}$ are defined on H^∞ in the canonical way, where the random variables s_n and $\bar{x}_n = (x_{n1}, x_{n2}) \in X(s_n)$ describe the state and the actions chosen by the players, respectively, on the n-th stage of the game (see Chapter 7 in Bertsekas and Shreve (1978)). Thus, for each profile π of strategies and any initial state s, one can define the operator of the expected value E_s^π with respect to the probability measure P_s^π. In the n-stage β - *discounted* model the total expected utility for player i is

$$\gamma_{i,n}^\beta(\pi)(s) = E_s^\pi \left(\sum_{k=1}^n \beta^{k-1} u_i(x_{ki}) \right).$$

The value $\beta \in (0, 1]$ is said to be a *discount factor*. If $\beta = 1$ then we have *undiscounted* model. Let us denote $\gamma_{i,n}(\pi)(s) := \gamma_{i,n}^1(\pi)(s)$. The total expected utility for player i in the infinite horizon game is

$$\gamma_i^\beta(\pi)(s) = E_s^\pi \left(\sum_{k=1}^\infty \beta^{k-1} u_i(x_{ki}) \right) = \lim_{n\to\infty} \gamma_{i,n}^\beta(\pi)(s).$$

In the undiscounted model, the total expected utility for player i is

$$\gamma_i(\pi)(s) = E_s^\pi \left(\sum_{k=1}^\infty u_i(x_{ki}) \right).$$

Remark. This criterion above makes a sens, because as shows Corollary 12, the series under the expected value are convergents are limited by a common value.

Let $\pi = (\pi_1, \pi_2)$ and σ_i be a strategy for player i. Then, as usual, (π_{-i}, σ_i) is the strategy profile π with π_i replaced by σ_i.

Definition 1. A strategy profile $\pi^\star = (\pi_1^\star, \pi_2^\star)$ is called a *Nash equilibrium* in the β - discounted stochastic game if and only if no unilateral deviations from it are profitable, that is, for every player i, his strategy σ_i and $s \in S$,

$$\gamma_i^\beta(\pi^\star)(s) \geq \gamma_i^\beta(\pi_{-i}^\star, \sigma_i)(s).$$

Similarly, Nash equilibria are defined in the finite horizon case and in the undiscounted case. Obviously, a strategy for player i in a n-step game consists of n components only.

Definition 2. Fix $\varepsilon > 0$. A strategy profile $\pi^\varepsilon = (\pi_1^\varepsilon, \pi_2^\varepsilon)$ is called an *ε - equilibrium* in the β - discounted stochastic game if and only if for every player i, his strategy σ_i and $s \in S$,

$$\varepsilon + \gamma_i^\beta(\pi^\varepsilon)(s) \geq \gamma_i^\beta(\pi_{-i}^\varepsilon, \sigma_i)(s).$$

Similarly we define ε - equilibrium in the finite horizon game. It is easy to see that unilateral deviation from ε - equilibrium can take the profit but no greater then ε.
Notation: If ρ is arbitrary game, then by $NE\,\rho$ we denote the set of Nash equilibria in this game.

Auxilliary one shot game

In this section we introduce an auxilliary one shot model $G(s, \eta)$ of two person game. The payoff function for player i $(i = 1, 2)$ is

$$w_i(\eta_i, s, x_1, x_2) = u_i(x_i) + \sum_{l=1}^{L} \eta_{i,l}(s) g_l\left(s - x_1 - x_2\right), \tag{1}$$

for $s \in S$, $x_i \in A_i(s)$, where each $\eta_{i,l} : S \to R_+$ is a continous function, $\eta_i := (\eta_{i,1}, ..., \eta_{i,L})$ and η is a matrix with the rows η_1 and η_2. Let $x^\eta(s) := (x_1^\eta(s), x_2^\eta(s))$ be a Nash equilibrium in $G(s, \eta)$. By Proposition 1 in [3] follows that this definition is well.

Lemma 3. For each continous η, the function $x_i^\eta(\cdot)$ is continous.

Proof. Let $s_n \to s_0$. Then by definition of x^η we have

$$\begin{aligned} w_1(\eta_1, s_n, x_1^\eta(s_n), x_2^\eta(s_n)) &\geq w_1(\eta_1, s_n, x_1, x_2^\eta(s_n)), \\ w_2(\eta_2, s_n, x_1^\eta(s_n), x_2^\eta(s_n)) &\geq w_2(\eta_2, s_n, x_1^\eta(s_n), x_2). \end{aligned} \tag{2}$$

for arbitrary $(x_1, x_2) \in X(s)$. Let $x^* := (x_1^*, x_2^*)$ be an arbitrary cumulation point of $(x^\eta(s_n))$. Consider subsequence of s_n which leads $(x^\eta(s_n))$ to x^*. By Assumption **C1** and **C2** and definition of w_i, taking a limit in (2) we obtain

$$\begin{aligned} w_1(\eta_1, s_0, x_1^*, x_2^*) &\geq w_1(\eta_1, s_0, x_1, x_2^*), \\ w_2(\eta_2, s_0, x_1^*, x_2^*) &\geq w_2(\eta_2, s_0, x_1^*, x_2). \end{aligned}$$

This implies that x^* is a Nash equilibrium in $G(s_0, \eta)$. As we have mentioned before this game has unique Nash equilibrium $x^\eta(s_0)$, hence $x^* = x^\eta(s_0)$. ■

Lemma 4. For all $i = 1, 2$ and $l = 1, ..., L$, let $\eta_{i,l}^n(\cdot)$ be a sequence of continous function on S and

$$\lim_{n \to \infty} \eta_{i,l}^n(\cdot) = \eta_{i,l}(\cdot)$$

in sup - norm on S. Then

$$\lim_{n \to \infty} x_i^{\eta^n}(s) = x_i^\eta(s), \tag{3}$$

and

$$\lim_{n \to \infty} w_i \left(\eta_i^n, s, x_1^{\eta^n}(s), x_2^{\eta^n}(s) \right) = w_i \left(\eta_i, s, x_1^\eta(s), x_2^\eta(s) \right), \tag{4}$$

in sup - norm on S.

Proof. *Step 1* First we prove uniformly convergence in (3). We show that there is exactly one cumulation point of the set of couples $K := \left\{ (s_n, x^{\eta^n}(s_n)) : n \in N \right\}$, where $(s_1, s_2, ...)$ is a sequence such that

$$\sup_{s \in S} \left| x^{\eta^n}(s) - x^\eta(s) \right| = \left| x^{\eta^n}(s_n) - x^\eta(s_n) \right|.$$

This sequence exists by Lemma 3. Let (s^*, x^*) be an arbitrary cumulation point of the set K. Then there exisists a sequence containing in K such that

$$x^* := \lim_{k \to \infty} x^{\eta^k}(s_k) \quad \text{and} \quad s^* = \lim_{k \to \infty} s_k.$$

Note that in this equation above, we denote x^{η^k} instead $x^{\eta^{n_k}}$ and s_k instead s_{n_k} for simplify the notation. Note that

$$\begin{aligned} w_1 \left(\eta_1^k, s_k, x_1^{\eta^k}(s_k), x_2^{\eta^k}(s_k) \right) &\geq w_1 \left(\eta_1^k, s_k, x_1, x_2^{\eta^k}(s_k) \right), \\ w_2 \left(\eta_2^k, s_k, x_1^{\eta^k}(s_k), x_2^{\eta^k}(s_k) \right) &\geq w_2 \left(\eta_2^k, s_k, x_1^{\eta^k}(s_k), x_2 \right), \end{aligned} \tag{5}$$

when $x_i \in A_i(s_k)$ is arbitrary. Since the sequence η^k uniformely converges to η, and η is continous, hence $\eta^k(s_k) \to \eta(s^*)$. By continouity of w_i and by (5) we obtain

$$\begin{aligned} &w_1 \left(\eta_1, s^*, x_1^*, x_2^* \right) \geq w_1 \left(\eta_1, s^*, x_1, x_2^* \right), \\ &w_2 \left(\eta_2, s^*, x_1^*, x_2^* \right) \geq w_2 \left(\eta_2, s^*, x_1^*, x_2 \right), \end{aligned}$$

which means that x^* is Nash equilibrium in the game $G(s_0, \eta)$. By Proposition 1 in [3] there is exactly one Nash equilibrium in this game. Hence each cumulation point of the set K is in a form $(s^*, x^\eta(s^*))$ i.e. is on the graph of $x^\eta(\cdot)$. Hence and by Lemma 3, we immediately obtain that the unique cumulation point of the sequence

$$\left| x^{\eta^n}(s_n) - x^\eta(s_n) \right|$$

is 0. This implies that $x^{\eta_n}(\cdot)$ is uniformely convergent to $x^\eta(\cdot)$.

Step 2. The uniform convergence in (4) follows directly from uniform convergence of the sequences $x^{\eta^n}(\cdot)$ and $\eta^n_{i,l}(\cdot)$, and uniform continouity of the functions u_i and g_l. ■

In [14] it is proven following lemma.

Lemma 5. For $i = 1, 2$ $X_i = [0, d_i]$ be an action spaces. Let

- u_i satysfy condition **C1**,
- the functions $\xi_i : [0, d_1 + d_2] \to R_+$ and $\zeta_i : [0, d_1 + d_2] \to R_+$ are twice continously differentables, strictly concave and decreasing,
- for $t \in [0, d_1 + d_2]$ we have $\xi_i(t) \le \zeta_i(t)$ and $\xi_i'(t) \ge \zeta_i'(t)$.

Consider two games: ρ_1 in which the payoff function for player i is on the form

$$W_i^1(x_1, x_2) = u_i(x_i) + \xi_i(x_1 + x_2),$$

and ρ_2 with the payoff function

$$W_i^2(x_1, x_2) = u_i(x_i) + \zeta_i(x_1 + x_2)$$

for $(x_1, x_2) \in X_1 \times X_2$. Then there exists Nash equilibrium in the game ρ_1 say $x^* := (x_1^*, x_2^*)$ and Nash equilibrium in the game ρ_2 say $y^* := (y_1^*, y_2^*)$ such that

$$W_i^1(x_1^*, x_2^*) \le W_i^2(y_1^*, y_2^*).$$

By Lemma 5 we immediately obtain following lemma:

Lemma 6. If for each $l = 1, ..., L$, and $s \in S$ holds $\eta^1_{i,l}(s) \le \eta^2_{i,l}(s)$, then

$$w_i\left(\eta_i^1, s, x_1^{\eta^1}(s), x_2^{\eta^1}(s)\right) \le w_i\left(\eta_i^2, s, x_1^{\eta^2}(s), x_2^{\eta^2}(s)\right).$$

Proof. Fix $s \in S$. Note that the games $G(s, \eta^j)$ $j = 1, 2$ can be described as ρ_i where

$$W_i^1(x_1, x_2) := w_i(\eta_i^1, s, x_1, x_2) = u_i(x_i) + \xi_i(x_1 + x_2)$$

with

$$\xi_i(t) := \sum_{l=1}^{L} \eta^1_{i,l}(s) g_l(s - t),$$

and

$$W_i^2(x_1,x_2) := w_i(\eta_i^2, s, x_1, x_2) = u_i(x_i) + \zeta_i(x_1+x_2)$$

with

$$\zeta_i(t) := \sum_{l=1}^{L} \eta_{i,l}^2(s) g_l\,(s-t)\,, t := x_1 + x_2.$$

Clearly u_i, ξ_i, ζ_i are strictly concave, twice continously differentable and strictly monotone (u_i is increasing ξ_i and ζ_i are decreasing). Since $\eta_{i,l}^1(s) \le \eta_{i,l}^2(s)$ for all $l = 1, ..., L$, hence $\xi(t) \le \zeta(t)$ for all $t \in [0, d_1 + d_2]$. Note that

$$\xi_i'(t) = -\sum_{l=1}^{L} \eta_{i,l}^1(s) g_l'\,(s-t) \ge -\sum_{l=1}^{L} \eta_{i,l}^2(s) g_l\,(s-t) = \zeta_i'(t).$$

Therefore conditions of Lemma 5 are satisfied. By Proposition 3 in [3] Nash equilibria in both ρ_i are unique, hence the proof is complete. ■

Asymptotic Nash equilibria in the n - step model

In this section we consider finite horizon game. Define $B_0(S) := \{v : S \to R_+ : v(0) = 0\}$. For every $(v_1, v_2) \in B_0(S) \times B_0(S)$, $s \in S$, $\beta \in (0,1]$, we define auxiliary two person one shot game $\Gamma(\beta, v_1, v_2, s)$, in which the payoff function for each player i is

$$k_i(\beta, v_i, s, x) = u_i(x_i) + \beta \int_{S_+} v_i(s') q(ds'|s, x),$$

where $x = (x_1, x_2) \in X(s)$. Since $v_i \ge 0$, by the assumptions **C1** and **C2** the payoff function this game has the same form as in section 3 in [3]. Hence by Proposition 1 in [3] we can conclude that for every $s \in S$, this game have an unique proper Nash equilibrium $NE\Gamma(\beta, v_1, v_2, s)$.
Obviously $NE\Gamma(\beta, v_1, v_2, s) = (0,0)$ for $s = 0$.

Let $\beta \in (0,1]$ be a discount factor, and n be a horizon of the finite step game. For $i = 1,2$ i $s \in S$ let $f_{i,1}^\beta(s) := a_i(s)$, and

$$v_{i,1}^\beta(s) := \max_{a_i \in A_i(s)} u_i(a_i) = u_i(f_{i,1}^\beta(s)).$$

Clearly $v_{i,1}^\beta \in B_0(S)$. If $v_{i,0}^\beta(s) := 0$ for arbitrary $s \in S$, then

$$\bar{f}_1^\beta := (f_{1,1}^\beta(s), f_{2,1}^\beta(s)) = NE\Gamma(\beta, v_{1,0}^\beta, v_{2,0}^\beta, s).$$

Therefore $\bar{f}_1^\beta$ is a Nash equilibrium in the one-step game, $v_{i,1}^\beta$ is an equilibrium function for the player i and $v_{i,1}^\beta = k_i(\beta, v_{i,0}, s, (a_1(s), a_2(s)))$. Analogously as in [14] and in section 4 in [2], we can define $f_{i,2}^\beta, ..., f_{i,n}^\beta \in F_i$ and $v_{i,2}^\beta, ..., v_{i,n}^\beta \in B_0(S)$ in the following way

$$\begin{aligned} \bar{f}_k^\beta := (f_{1,k}^\beta, f_{2,k}^\beta) &:= NE\Gamma(\beta, v_{1,k-1}^\beta, v_{2,k-1}^\beta, s) \quad \text{and} \\ v_{i,k}^\beta(s) &:= k_i(\beta, v_{i,k-1}^\beta(s), s, \bar{f}_k^\beta(s)), \end{aligned}$$

where $s \in S$ and $k = 2, ..., n$. By Proposition 1 in [3], these definitions above are well. Let $\pi_i^{(n),\beta}$ be a n-step strategy for the player i which is defined as

$$\pi_i^{(n),\beta} = (f_{i,1}^{\beta *}, f_{i,2}^{\beta *}, ..., f_{i,n}^{\beta *}) := (f_{i,n}^{\beta}, f_{i,n-1}^{\beta}, ..., f_{i,1}^{\beta}).$$

(Clearly, $f_{i,k}^{\beta} = f_{i,n-k+1}^{\beta *}$.) Let $\pi^{(n),\beta} := \left(\pi_1^{(n),\beta}, \pi_2^{(n),\beta}\right)$. We denote $f_{i,n}^{\star} := f_{i,n}^{\beta}$, $v_{i,n}^{\star} := v_{i,n}^{\beta}$, and $\pi^{(n)} := \left(\pi_1^{(n)}, \pi_2^{(n)}\right) := \left(\pi_1^{(n),\beta}, \pi_2^{(n),\beta}\right)$ when $\beta = 1$.

By the construction above and Bellman equations in the dynamic programming in the finite horizon game (see [5, 6] or [15]), it follows that $\pi^{(n),\beta}$ is a Nash equilibrium in the n - step β-discounted game.

Main Theorem 7. *For each $n \in N$ and $i = 1, 2$ hold*

$$\lim_{\beta \to 1} f_{i,n}^{\beta}(\cdot) = f_{i,n}^{\star}(\cdot), \tag{6}$$

and

$$\lim_{\beta \to 1} v_{i,n}^{\beta}(\cdot) = v_{i,n}^{\star}(\cdot). \tag{7}$$

Both convergences are uniform on S (i.e. in sup - norm on S).

Proof. Clearly, the hypothesis is true for $n = 1$. Suppose that for some $n \in N$ the hipothesis is satisfied i.e. if $\beta \to 1$ then hold

$$f_{i,n}^{\beta}(\cdot) \to f_{i,n}^{\star}(\cdot)$$

and

$$v_{i,n}^{\beta}(\cdot) \to v_{i,n}^{\star}(\cdot)$$

uniformely on S. By Bellman equations (see [5, 15] or [15]) and conditions **C2**, we conclude that $n + 1$ step game is on the form $G(s, \eta^{\beta})$ with

$$\eta_{i,l}^{\beta}(s) := \beta \int_S v_{i,n}^{\beta} \lambda_l(ds'|s). \tag{8}$$

Let $\eta_{i,l}(\cdot) := \eta_{i,l}^{1}(\cdot)$. By induction hipothesis we obtain $\sup_{s \in S} \left| v_{i,n}^{\beta}(s) - v_{i,n}^{\star}(s) \right| \to 0$, as

$n \to \infty$. Hence

$$\begin{aligned}
\sup_{s\in S}\left|\eta^{\beta}_{i,l}(s)-\eta_{i,l}(s)\right| &\leq \sup_{s\in S}\left\{\int\limits_S \left|v^{\beta}_{i,n}(s')-v^{\star}_{i,n}(s')\right|\lambda_l(ds'|s)\right\}\\
&+ (1-\beta)\sup_{s\in S}\left\{\int\limits_S v^{\star}_{i,n}(s')\lambda_l(ds'|s)\right\} \qquad (9)\\
&\leq \sup_{s\in S}\left|v^{\beta}_{i,n}(s)-v^{\star}_{i,n}(s)\right| + (1-\beta)n\,||u_i||_{\infty} \to 0 \quad \text{as } \beta\to 1.
\end{aligned}$$

The thesis for $n+1$ follows directly from (10) and Lemma 4. Hence uniform convergences in (6) and (7) hold. ■

Remark. If we additionaly assumed that the capacity functions $a_i(\cdot)$ are Lipshitz - continous with a constant 1 and that no measures $\lambda_l(\cdot|s)$ depends on s (i.e. $\lambda_l(\cdot|s) = \lambda_l(\cdot)$ for each l), then the transition probability would be a special case of that from Amir (1996). Then we immediately would obtain that the Nash equilibria are Lipshitz continous with a constant 1, and uniform continouity in Main Theorem 7 and further in Main Theorem 15 would be satisfied immediately.

Lemma 8. For arbitrary $n \in N$ let

$$\psi_i^{(n),\beta} := (\phi_i^{(n),\beta}, \phi_i^{(n-1),\beta}, ..., \phi_i^{(1),\beta})$$

be a certain collection of Markov strategies for player i depending on $\beta \in (0,1]$. Moreover, assume that there exists a limit

$$\psi_i^{(n)} := \lim_{\beta\to 1}\psi_i^{(n),\beta}. \qquad (10)$$

If the convergence in (10) is uniform, then

$$\lim_{\beta\to 1}\left(\sup_{s\in S}\left|\gamma^{\beta}_{i,n}(\psi_1^{(n),\beta},\psi_2^{(n),\beta})(s)-\gamma^{\beta}_{i,n}(\psi_1^{(n)},\psi_2^{(n)})(s)\right|\right)=0. \qquad (11)$$

Remark. Since Markov strategy for player i in n - step game can be treaten as n - element vector from the space F_i^n, uniform convergence of Markov strategy means uniform convergence of each component.

Proof. Clearly for $n = 1$ the hypothesis is true. Suppose that (11) holds for some n, and this convergence is uniform. Note that by Bellman equations for finite horizon game (see [5, 6] or [15]) we have

$$\left|\gamma^{\beta}_{i,n+1}\left(\psi_1^{(n+1),\beta},\psi_2^{(n+1),\beta}\right)(s)-\gamma^{\beta}_{i,n+1}\left(\psi_1^{(n+1)},\psi_2^{(n+1)}\right)(s)\right| \leq$$

$$\left|u_i(\phi_i^{(n+1),\beta}(s))-u_i(\phi_i^{(n+1)}(s))\right| + \beta\Delta_{\beta}(s), \qquad (12)$$

where

$$\Delta_\beta(s) := \left|\omega_1^\beta(s) - \omega_2^\beta(s)\right|,$$

$$\omega_1^\beta(s) = \int\limits_S \gamma_{i,n}^\beta\left(\psi_1^{(n),\beta}, \psi_2^{(n),\beta}\right)(s')q(ds'|s, \phi_1^{(n+1),\beta}(s), \phi_2^{(n+1),\beta}(s))$$

and

$$\omega_2^\beta(s) = \int\limits_S \gamma_{i,n}^\beta\left(\psi_1^{(n)}, \psi_2^{(n)}\right)(s')q(ds'|s, \phi_1^{(n+1)}(s), \phi_2^{(n+1)}(s)).$$

By uniform convergence in (10) and uniform continuity of the function u_i we know that the first part of (12) uniformely converges to 0. It is sufficient to prove that $\left|\left|\Delta^\beta(\cdot)\right|\right|_\infty$ converge do 0, when $\beta \to 1$. Note that by condition **C2**, we obtain

$$\omega_1^\beta(s) = \sum_{l=1}^L \left(\int\limits_S \gamma_{i,n}^\beta\left(\psi_1^{(n),\beta}, \psi_2^{(n),\beta}\right)(s')\lambda_l(ds'|s)g_l\left(s - \phi_1^{(n+1),\beta}(s) - \phi_2^{(n+1),\beta}(s))\right)\right),$$

and

$$\omega_2^\beta(s) = \sum_{l=1}^L \left(\int\limits_S \gamma_{i,n}^\beta\left(\psi_1^{(n)}, \psi_2^{(n)}\right)(s')\lambda_l(ds'|s)g_l\left(s - \phi_1^{(n+1)}(s) - \phi_2^{(n+1)}(s))\right)\right)$$

Hence we have

$$\begin{aligned}
\sup_{s\in S}\left|\Delta^\beta(s)\right| &\leq \sup_{s\in S}\left|\gamma_{i,n}^\beta\left(\psi_1^{(n),\beta}, \psi_2^{(n),\beta}\right)(s) - \gamma_{i,n}^\beta\left(\psi_1^{(n)}, \psi_2^{(n)}\right)(s)\right| \\
&+ \sup_{s\in S}\left|\sum_{l=1}^L \int\limits_S \gamma_{i,n}^\beta\left(\psi_1^{(n)}, \psi_2^{(n)}\right)(s')\lambda_l(ds'|s)\left|\tilde{g}_l^\beta(s) - \tilde{g}_l(s)\right|\right| \\
&\leq \sup_{s\in S}\left|\gamma_{i,n}^\beta\left(\psi_1^{(n),\beta}, \psi_2^{(n),\beta}\right)(s) - \gamma_{i,n}^\beta\left(\psi_1^{(n)}, \psi_2^{(n)}\right)(s)\right| \\
&+ n\left|\left|u_i\right|\right|_\infty \sum_{l=1}^L \sup_{s\in S}\left|\tilde{g}_l^\beta(s) - \tilde{g}_l(s)\right|
\end{aligned}$$

with

$$\tilde{g}_l^\beta(s) := g_l\left(s - \phi_1^{(n+1),\beta}(s) - \phi_2^{(n+1),\beta}(s))\right)$$

and

$$\tilde{g}_l(s) := g_l\left(s - \phi_1^{(n+1)}(s) - \phi_2^{(n+1)}(s))\right).$$

Clearly, by induction hypothesis it is sufficient to show that $\tilde{g}_l^\beta(\cdot) \to \tilde{g}_l(\cdot)$ uniformly on S. But it is also clear, because of uniform continouity of g_l and uniform convergence of ϕ_i^β. ■

Denote $\mathcal{M}_i^n$ as a set of all Markow strategies for player i in n step game.

Lemma 9. Let

$$\mathcal{T}_{i,n}^{\beta} := \sup_{\psi^{(n)}\in\mathcal{M}_i^n} \sup_{s\in S} \left|\gamma_{i,n}^{\beta}(\pi_{-i}^{(n,\beta)}, \psi^{(n)})(s) - \gamma_{i,n}^{\beta}(\pi_{-i}^{(n)}, \psi^{(n)}))(s)\right|.$$

Then we have

$$\lim_{\beta\to 1} \mathcal{T}_{i,n}^{\beta} = 0.$$

Proof. Denote $\psi^{(n)} := (\phi^{(n+1)}, \phi^{(n)}, ..., \phi^{(1)})$. We prove this theorem by induction. Clearly this hipothesis is satisfied for $n = 1$. Suppose that the thesis of the theorem is satisfied for some $n \in N$. We have

$$\begin{aligned}
\mathcal{T}_{i,n+1}^{\beta} &= \sup_{\psi^{(n)}\in\mathcal{M}_i^n} \sup_{s\in S} \left|\gamma_{i,n+1}^{\beta}(\pi_{-i}^{(n+1,\beta)}, \psi^{(n+1)})(s) - \gamma_{i,n+1}^{\beta}(\pi_{-i}^{(n+1)}, \psi^{(n+1)}))(s)\right| \\
&= \beta \sup_{\psi^{(n)}\in\mathcal{M}_i^n} \sup_{s\in S} \Bigg|\sum_{l=1}^{L} \int_S \gamma_{i,n}^{\beta}(\pi_{-i}^{(n,\beta)}, \psi^{(n)})(s')\lambda_l(ds'|s) g_l\left(s - f_{i,n+1}^{\beta}(s) - \phi^{(n+1)}(s)\right) \\
&\quad - \sum_{l=1}^{L} \int_S \gamma_{i,n}^{\beta}(\pi_{-i}^{(n)}, \psi^{(n)})(s')\lambda_l(ds'|s) g_l\left(s - f_{i,n+1}^{\star}(s) - \phi^{(n+1)}(s)\right)\Bigg| \\
&\le \mathcal{T}_{i,n}^{\beta} \\
&+ n\,||u||_\infty \sum_{l=1}^{L} \sup_{\phi_i\in F_i} \sup_{s\in S} \left|g_l\left(s - f_{i,n+1}^{\beta}(s) - \phi_i(s)\right) - g_l\left(s - f_{i,n+1}^{\star}(s) - \phi_i(s)\right)\right|
\end{aligned}$$

By indyction hypothesis $\mathcal{T}_{i,n}^{\beta} \to 0$ when $\beta \to 1$. By Main Theorem 7 and uniform continouity of $g_l(\cdot)$, the second term of the right side of the inequality above tends to 0 when $\beta \to 1$.

■

Main Theorem 10. *For arbitrary ε there exist a constant β_0 such that if $\beta > \beta_0$, the profile $\left(\pi_1^{(n)}, \pi_2^{(n)}\right)$ is ε- equilibrium in the β - discounted n - step game.*

Proof. Fix $n \in N$. From Main Theorem 7 it follows that, for each $n \in N$ $\pi^{(n),\beta}$ is uniformly convergent to $\pi^{(n)}$ (when $\beta \to 1$). Let $j \neq i$ and $(i, j = 1, 2)$. Denote σ_i as an arbitrary Markov strategy for player i. Let ϵ be also arbitrary. From Main Theorem 7 and Lemma 8 we conclude the existence of β_1, such that for $\beta > \beta_1$ holds

$$\gamma_{i,n}^{\beta}\left(\pi^{(n)}\right)(s) \ge \gamma_{i,n}^{\beta}\left(\pi^{(n),\beta}\right)(s) - \frac{\epsilon}{2}. \tag{13}$$

Since $\pi^{(n),\beta}$ is the Nash equilibrium in n step game we obtain

$$\gamma_{i,n}^{\beta}\left(\pi^{(n),\beta}\right)(s) \ge \gamma_{i,n}^{\beta}\left(\pi_{-i}^{(n),\beta}, \sigma_i\right)(s). \tag{14}$$

By Lemma 9 we conclude the existing a constant $\beta_2 > \beta_1$ such that for each $\beta > \beta_2$ we have

$$\gamma_{i,n}^{\beta}\left(\pi_{-i}^{(n),\beta}, \sigma_i\right)(s) \ge \gamma_{i,n}^{\beta}\left(\pi_{-i}^{(n)}, \sigma_i\right)(s) - \frac{\epsilon}{2}. \tag{15}$$

Combining (13), (14) and (15) we obtain

$$\gamma_{i,n}^{\beta}\left(\pi^{(n)}\right)(s) \geq \gamma_{i,n}^{\beta}\left(\pi_{-i}^{(n)}, \sigma_i\right)(s) - \epsilon.$$

■

Asymptotic Nash equilibria in infinite step model

For player i let $\mathcal{S}_i$ be a set of all stationary strategies. Let $\mathcal{S} := \mathcal{S}_1 \times \mathcal{S}_2$. Note that there is one to one correspondence between $\mathcal{S}_i$ and the set of the functions F_i. Hence if the stationary multi - strategy ψ we can describe as $\psi = (\phi, \phi, ...)$ for some borel function $\phi \in F$, we denote $\gamma_i^{\beta}(\phi)(s) := \gamma_i^{\beta}(\psi)(s)$. Similarly we can define $\gamma_{i,n}^{\beta}(\phi)(s) := \gamma_{i,n}^{\beta}(\psi)(s)$, when the profile ψ is used in n step model.

Lemma 11. Let $\psi := (\psi_1, \psi_2) \in \mathcal{S}$ be arbitrary. Let $\psi = (\phi, \phi, ...)$. Then

$$\sup_{\phi\in\mathcal{S}} \sup_{\beta\in(0,1]} \sup_{s\in S} \left|\gamma_{i,n}^{\beta}(\phi)(s) - \gamma_i^{\beta}(\phi)(s)\right| \to 0, \quad \text{when} \quad n \to \infty. \tag{16}$$

Moreover, for each stationary strategy ϕ holds

$$\gamma_i^{\beta}(\phi)(s) \leq ||u||_{\infty} \frac{C}{1-C} \tag{17}$$

with

$$C = \int_S \left(\sum_{l=1}^{L} g_l(s')\right) \mu(ds').$$

Proof. Let $\phi \in F$. Define $(s_0, s_1, ..., s_t, ...)$ as a sequence of the states generated by the stationary strategy profile ϕ. Let $s_0 = s$. It is easy to see that

$$\gamma_i^{\beta}(\phi)(s) = \sum_{t=1}^{\infty} E_s^{\phi}\left(u_i(\phi_i(s_t))\beta^{t-1}\right). \tag{18}$$

From the assumption **C2** for $t > 1$ holds

$$\begin{aligned} z_t^{\beta,\phi}(s) &:= E_s^{\phi}\left(u_i(\phi_i(s_t))\beta^{t-1}\right) \\ &= \sum_{l=1}^{L} \beta^{t-1} E_s^{\phi}\left(\int_S u_i(\phi_i(s_{t-1}))\lambda_l(ds'|s_{t-1}) g_l\left(s_{t-1} - \phi_1(s_{t-1}) - \phi_2(s_{t-1})\right)\right) \\ &\leq ||u||_{\infty} \sum_{l=1}^{L} E_s^{\phi}\left(g_l(s_{t-1})\right) = ||u||_{\infty} h_{t-1}^{\beta}(s), \end{aligned} \tag{19}$$

where

$$h_t^{\beta}(s) := E_s^{\phi}\left(\sum_{l=1}^{L} g_l(s_t)\right)$$

for $t > 1$ and

$$h_0^\beta(s) := \sum_{l=1}^{L} g_l(s).$$

To prove this lemma we just need to show that the series $z_t^{\beta,\phi}(s)$ are uniformely convergent in (β, ϕ, s). Clearly $0 \leq h_0^\beta(\cdot) \leq 1$. Let $t > 0$ From assumption **C2** we have

$$\begin{aligned} h_t^\beta(s) &= \sum_{l=1}^{L} E_s^\phi \left(\int_S \sum_{l=1}^{L} g_l(s')\lambda_l(ds'|s_{t-1}) g_l\left(s_{t-1} - \phi_1(s_{t-1}) - \phi_2(s_{t-1})\right) \right) \\ &\leq \int_S \left(\sum_{l=1}^{L} g_l(s') \right) \mu(ds') E_s^\psi \left(\sum_{l=1}^{L} g_l(s_{t-1}) \right) \\ &= C h_{t-1}^\beta(s). \end{aligned}$$

Hence we have

$$h_t^\beta(s) \leq C * h_{t-1}^\beta(s) \leq ... \leq C^{t-1} h_1^\beta(s) \leq C^t, \tag{20}$$

By Assumption **C2** $0 < C < 1$. Obviously C is a constant independent on s, ϕ and β. Combining (19) and (20) we obtain

$$z_t^{\beta,\phi}(s) \leq ||u||_\infty h_{t-1}^\beta(s) \leq ||u||_\infty C^{t-1}. \tag{21}$$

Hence by Weierstrass criterion the series $z_t^{\beta,\phi}(s)$ are uniformly convergent in (s, β, ϕ), which complete the proof that the condition in (16) is satisfied. Condition (17) follows from (18), (19), (20) and (21). ■

Corollary 12. *By Lemma 11 is easy to see that $\gamma_i(\pi)(s) < \infty$ for arbitrary profile π. Moreover, repeating the reasoning in the proof of Lemma 11 we can obtain (17) for arbitrary profile π.*

Lemma 13. For arbitrary $\beta \in (0, 1]$ let $\left(\psi_1^\beta, \psi_2^\beta\right) \in \mathcal{S}$ be stationary multi - strategy. If for player i holds

$$\lim_{\beta \to 1} \sup_{s \in S} \left|\psi_i^\beta(s) - \psi_i(s)\right| = 0, \tag{22}$$

then

$$\lim_{\beta \to 1} \sup_{s \in S} \left|\gamma_i^\beta \left(\psi_1^\beta, \psi_2^\beta\right)(s) - \gamma_i^\beta\left(\psi_1, \psi_2\right)(s)\right| = 0.$$

Proof. Let $\epsilon > 0$ be arbitrary. By Lemma 11 there exists n_0 such that for $n > n_0$ we have

$$\left|\gamma_i^\beta \left(\psi_1^\beta, \psi_2^\beta\right)(s) - \gamma_i^\beta\left(\psi_1, \psi_2\right)(s)\right| \leq \left|\gamma_{i,n}^\beta \left(\psi_1^\beta, \psi_2^\beta\right)(s) - \gamma_{i,n}^\beta\left(\psi_1, \psi_2\right)(s)\right| + \epsilon. \tag{23}$$

Fix $n > n_0$. By (22) and Lemma 8 if we take a limit with $\beta \to 1$, we obtain

$$\lim_{\beta\to 1}\left(\sup_{s\in S}\left|\gamma_{i,n}^{\beta}\left(\psi_1^{\beta},\psi_2^{\beta}\right)(s)-\gamma_{i,n}^{\beta}\left(\psi_1,\psi_2\right)(s)\right|\right)=0.$$

Hence and by (23) we have

$$\limsup_{\beta\to 1}\left(\sup_{s\in S}\left|\gamma_{i}^{\beta}\left(\psi_1^{\beta},\psi_2^{\beta}\right)(s)-\gamma_{i}^{\beta}\left(\psi_1,\psi_2\right)(s)\right|\right)\leq \epsilon.$$

Since ϵ is arbitrary this proof is complete. ■

By the main results in [14] we can conclude uniform convergence of equilibria payoffs in the finite horizon β - discounted game to the stationary equilibrium payoff in the infinite horizon β - discounted game, when the horizon tends to infinity. Moreover, we also know pointwise convergence of Nash equilibria adequates to these equilibria payoffs. This theorem bellow shows that convergence of Nash euilibria is uniform as well.

Theorem 14. For arbitrary $i = 1, 2$, and $\beta \in (0, 1)$ there exist a limits

$$f_i^{\beta}(s) := \lim_{n\to\infty} f_{i,n}^{\beta}(s), \tag{24}$$

and

$$v_i^{\beta}(s) := \lim_{n\to\infty} v_{i,n}^{\beta}(s), \tag{25}$$

Moreover, these convergences above are uniform on S (i.e. in sup-norm).
Further, the stationary multi-strategy $\pi^{\beta} := (\pi_1^{\beta}, \pi_2^{\beta})$, $\pi_i^{\beta} := \left(f_i^{\beta}, f_i^{\beta}, \dots\right)$ is a Nash equilibrium in β - discounted infinie horizon game and v_i^{β} is equilibrium payoff adequate to $f^{\beta} := \left(f_1^{\beta}, f_2^{\beta}\right)$.

Proof. Fix β. In [14] it is prooven uniform covergence $\lim_{n\to\infty} v_{i,n}^{\beta}(\cdot) = v_i^{\beta}(\cdot)$. We also know that pointwise convergence in (24) holds, and $f^{\beta} := \left(f_1^{\beta}, f_2^{\beta}\right)$ is Nash equilibrium and $v_i^{\beta}(s) = \gamma_i\left(f^{\beta}\right)(s)$. We just need to show that the convergence in (24) is uniform.

From Bellman equations for finite horizon game ([5, 6] and [15]) we conclude that $n+1$ horizon β - discounted game can be described as $G(s, \eta^{\beta,n})$ with

$$\eta_{i,l}^{\beta,n}(s) := \beta \sum_{l=1}^{L} \int_S v_{i,n}^{\beta}(s')\lambda_l(ds'|s).$$

and infinite horizon game reduce to $G(s,\eta^\beta)$ with

$$\eta^\beta_{i,l}(s) := \beta \sum_{l=1}^{L} \int_S v^\beta_i(s')\lambda_l(ds'|s).$$

Note that

$$\begin{aligned}\sup_{s\in S}\left|\eta^{\beta,n}_{i,l}(s) - \eta^\beta_{i,l}(s)\right| &\leq \sum_{l=1}^{L}\int_S \left|v^\beta_{i,n}(s') - v^\beta_i(s')\right| \lambda_l(ds'|s) \\ &\leq \sup_{s\in S}\left|v^\beta_{i,n}(s') - v^\beta_i(s')\right| \to 0 \quad \text{as } \beta \to 1.\end{aligned}$$

Hence and by Lemma 4 we obtain uniform convergence in (24). ■

Remark. Problem of convergence of Nash equilibria and adequates equilibria payoffs with horizon tending to ∞ is also solved for symmetric m - person games in [2]. Hence, repeating the reasoning in the proof of Theorem 14 we would also obtain the same results for m - person symmetric game.

Main Theorem 15. *For $i = 1, 2$ hold*

$$\lim_{\beta\to 1} f^\beta_i(\cdot) = f^\star_i(\cdot) \tag{26}$$

and

$$\lim_{\beta\to 1} v^\beta_i(\cdot) = v^\star_i(\cdot). \tag{27}$$

and these convergences above are uniform on S (i.e. in sup-norm on S). Moreover, $f^\star := (f^\star_1, f^\star_2)$ is a Nash equilibrium in the undiscounted game and $v^\star_i$ is an equilibrium payoff adequates to $f^\star$.

Proof. First we show that the function $\beta \to v^\beta_{i,n}(s)$ is nondecreasing. For $n = 1$ this hipothesis is clear. Assume that for some $n \in N$ the function $\beta \to v^\beta_{i,n}(s)$ is nondecreasing. Consider the game $G(s, \eta^{\beta,n})$ with

$$\eta^{\beta,n}_{i,l}(s) := \beta \int_S v^\beta_{i,n}(s')\lambda_l(ds'|s).$$

Clearly $\eta^{\beta,n}$ is nondecreasing in β. Note that by Bellman equations for finite horizon game ([5, 6] or [15]) $\left(f^\beta_{1,n+1}(s), f^\beta_{2,n+1}(s)\right) = NE\, G(s, \eta^{\beta,n})$. Hence and by Lemma 6 we immediately obtain that $v^\beta_{i,n+1}$ is nondecreasing in β as well.

Since by Theorem 14 $v^\beta_i(s) = \lim_{n\to\infty} v^\beta_{i,n}(s)$, hence $v^\beta_i(s)$ is nondecreasing in β as a limit of nondecreasing functions. By Lemma 11 $v^\beta_i(s) \leq ||u|| \frac{C}{1-C}$. Hence there exists a limit, say $v_i^\star(s) := \lim_{\beta\to 1} v_i^\beta(s)$. We show that $v_i^\star(\cdot)$ is a payoff equilibrium in undiscounted infinite horizon game. By Bellman equations ([5, 6] or [15]) for $j \neq i$ we have

$$\begin{aligned} v_i^\beta(s) &= u_i(f_i^\beta(s)) + \beta\sum_{l=1}^{L}\int_S v_i^\beta(s')\lambda_l(ds'|s)g_l\left(s - f_1^\beta(s) - f_2^\beta(s)\right) \\ &\geq u_i(x_i) + \beta\sum_{l=1}^{L}\int_S v_i^\beta(s')\lambda_l(ds'|s)g_l\left(s - f_j^\beta(s) - x_i\right). \end{aligned} \tag{28}$$

Fix $s \in S$. Suppose that $x^* := (x_1^*, x_2^*)$ is a cumulation point of $f^\beta(s) := \left(f_1^\beta(s), f_2^\beta(s)\right)$. Suppose that $\beta_k \to 1$ is such sequence for which $\lim_{k\to 1} f^{\beta_k} = x^*$. Let $(x_1, x_2) \in X(s)$ be arbitrary. Now, if we put $\beta := \beta_k$ in (28), and take a limit with $k \to \infty$, then we obtain

$$\begin{aligned} v_i^\star(s) &= u_i(x_i^*(s)) + \sum_{l=1}^{L}\int_S v_i^\star(s')\lambda_l(ds'|s)g_l\left(s - x_1^*(s) - x_2^*(s)\right) \\ &\geq u_i(x_i) + \sum_{l=1}^{L}\int_S v_i^\star(s')\lambda_l(ds'|s)g_l\left(s - x_j^*(s) - x_i\right). \end{aligned} \tag{29}$$

Hence we obtain that $(x_1^*(s), x_2^*(s)) = NE\Gamma(s, v_1^\star, v_2^\star, 1)$. Proposition 1 in [3] guarantes that this definition is well. Hence for each $s \in S$, there exists a limit, say $f_i^\star(s) := \lim_{\beta\to 1} f_i^\beta(s)$. By Corrolary 12 and then by Bellman equations (see [5, 6] or [15]) we conclude that $f^\star$ is a Nash equilibrium in undiscounted infinite horizon game, and $v_i^\star(s) = \gamma_i\left(f^\star\right)(s)$.

Now we show that $f_i^\star(\cdot)$ is continous. Note that by Assumption **C2** the function $s \to \int_S v_i^\star(s')\lambda_l(ds'|s)$ is continous. Let $s_n \to s_0$. By (29) we have

$$\begin{aligned} &u_i(f_i^\star(s_n)) + \sum_{l=1}^{L}\int_S v_i^\star(s')\lambda_l(ds'|s_n)g_l\left(s_n - f_1^\star(s_n) - f_2^\star(s_n)\right) \\ &\qquad \geq u_i(x_i) + \sum_{l=1}^{L}\int_S v_i^\star(s')\lambda_l(ds'|s_n)g_l\left(s_n - f_j^\star(s_n) - x_i\right). \end{aligned} \tag{30}$$

Let $x^0 := (x_1^0, x_2^0)$ be a cumulation point of the sequence $f^\star(s_n) := (f_1^\star(s_n), f_2^\star(s_n))$. Let s_k be a subsequence of the sequence s_n (again we denote s_k instead s_{n_k} for simplify the notation), such that $x^0 = \lim_{k\to\infty} f^\star(s_k)$. Hence, if we put $s_n := s_k$ in (30) and take a limit with $k \to \infty$ we have

$$u_i(x_i^0) + \sum_{l=1}^{L} \int_S v_i^\star(s')\lambda_l(ds'|s_0) g_l \left(s_0 - x_1^0 - x_2^0\right)$$

$$\geq u_i(x_i) + \sum_{l=1}^{L} \int_S v_i^\star(s')\lambda_l(ds'|s_0) g_l \left(s_0 - x_j^0(s) - x_i\right).$$

It means that $(x_1^0, x_2^0) = NE\Gamma(s_0, v_1^\star, v_2^\star, 1)$. Hence and by uniquness of Nash equilibrium in $\Gamma(s_0, v_1^\star, v_2^\star, 1)$ we obtain that the cumulation point of the sequence $(f_1^\star(s_n), f_2^\star(s_n))$ is unique and is equal $(f_1^\star(s_0), f_2^\star(s_0))$. Hence we obtain continouity of $f_i^\star(\cdot)$ and hence also $v_i^\star(\cdot)$. Hence, because $\beta \to v_i^\beta(s)$ is monotone and $v_i^\star(\cdot)$ is continous as well, by Dini Theorem it follows that $v_i^\beta(\cdot) \to v_i^\star(\cdot)$ as $\beta \to 1$ uniformely which ends proof of the uniform convergence in (27) . Consider a game $G(s, \eta^\beta)$ with

$$\eta_{i,l}^\beta := \beta \int_S v_i^\beta(s')\lambda_l(ds'|s),$$

and game $G(s, \eta)$ with

$$\eta_{i,l} := \int_S v_i^\star(s')\lambda_l(ds'|s).$$

Clearly by Bellman equations (see [5, 6] or [15]) $f^\beta(s) = NE\,G(s, \eta^\beta)$ and $f^\star(s) = NE\,G(s, \eta)$. We show that

$$\eta_{i,l}^\beta(\cdot) \to \eta_{i,l}(\cdot)\,(\beta \to 1).$$

uniformly on S. By Lemma 11 we know that $v_i^\star(\cdot) \leq \frac{C}{1-C}\,||u_i||_\infty$. Hence we have

$$\begin{aligned}
\sup_{s\in S}\left|\eta_{i,l}^\beta(s) - \eta_{i,l}(s)\right| \quad &\leq \quad \sup_{s\in S}\left\{\int_S \left|v_i^\beta(s') - v_i^\star(s')\right| \lambda_l(ds'|s)\right\} \\
&+ \quad (1-\beta)\sup_{s\in S}\left\{\int_S v_i^\star(s')\lambda_l(ds'|s)\right\} \\
&\leq \quad \sup_{s\in S}\left|v_i^\beta(s) - v_i^\star(s)\right| + (1-\beta)\frac{C}{1-C}\,||u_i||_\infty \to 0 \quad \text{as } \beta \to 1.
\end{aligned}$$

Hence and by Lemma 4 we obtain that $f_i^\beta(\cdot) \to f_i^\star(\cdot)$ uniformely on S, hence the part (26) is proven. ■

For $i = 1, 2$ define $\pi_i^\star := (f_i^\star, f_i^\star, \ldots)$.

Main Theorem 16. *For arbitrary ε and there exist β_0 such that if $\beta > \beta_0$, stationary multi - strategy $(\pi_1^\star, \pi_2^\star)$ is ε- equilibrium in β - discounted infinite horizon game.*

Proof. Fix arbitrary $\varepsilon > 0$ and $\sigma_i \in \mathcal{S}_\rangle$. By Lemma 13 and Main Theorem 15 we conclude existence of β_1 such that for $\beta > \beta_1$ we have

$$\gamma_i^\beta(f^\star)(s) \quad \geq \quad \gamma_i^\beta(f^\beta)(s) - \frac{\epsilon}{3}. \tag{31}$$

Since f^β is a Nash equilibrium in the β discounted game, hence we obtain

$$\gamma_i^\beta(f^\beta)(s) \quad \geq \quad \gamma_i^\beta(f_{-i}^\beta, \sigma_i)(s), \tag{32}$$

By Lemma 11 there exists n_0 such that for all $n > n_0$ holds

$$\gamma_i^\beta(f_{-i}^\beta, \sigma_i)(s) \quad \geq \quad \gamma_{i,n}^\beta(f_{-i}^\beta, \sigma_i)(s) - \frac{\epsilon}{3}. \tag{33}$$

From Lemma 8 and Main Theorem 15 there exists $\beta_2 > \beta_1$ such that for $\beta > \beta_2$ we have

$$\gamma_{i,n}^\beta(f_{-i}^\beta, \sigma_i)(s) \quad \geq \quad \gamma_{i,n}^\beta(f_{-i}^*, \sigma_i)(s) - \frac{\epsilon}{3}. \tag{34}$$

Combining (31), (32), (33), (34) we obtain

$$\gamma_i^\beta(f^\star)(s) \quad \geq \quad \gamma_{i,n}^\beta(f_{-i}^\star, \sigma_i)(s) - \epsilon, \tag{35}$$

for arbitrary $\beta > \beta_2$ and $n > n_0$. By Lemma 9 if we take a limit with $n \to \infty$ in (35) we obtain

$$\gamma_i^\beta(f^\star)(s) \quad \geq \quad \gamma_i^\beta(f_{-i}^\star, \sigma_i)(s) - \epsilon. \tag{36}$$

To see that (36) is satisfied when σ_i is arbitrary strategy, we first note that for each β, there exists a stationary optimal policy on $f_i^\star$ (say σ_i^β) in the β - discounted game. If we put $\sigma_i := \sigma_i^\beta$ in (36) we immediately obtain (36) with arbitrary σ_i. To complete the proof we can take $\beta_0 := \beta_2$. ■

Example 17. Let us consider two person game in which $S = [0, 1]$, $a_1(s) = a_2(s) = s/2$, $u_1(s) = u_2(s) = \sqrt{s}$, and the transition probability is of the form

$$q\,(\cdot\,|s, x\,) = \sqrt{s - x_1 - x_2}\lambda(\cdot) + \left(1 - \sqrt{s - x_1 - x_2}\right)\delta_0(\cdot),$$

where λ a uniform distribution on $[0, 1]$. By [2] and Theorem 2 [3] there exists an unique Nash equilibrium in the β discounted finite horizon game. We see that

$$f_{1,1}^\beta(s) = f_{2,1}^\beta(s) = s/2,$$

and

$$v_{1,1}^\beta(s) = v_{2,1}^\beta(s) = \sqrt{s/2}.$$

For $i = 1, 2$ and $n \geq 1$ we obtain $\pi_i^{(n),\beta} = f_{i,n}^\beta$ for

$$f_{i,n}^\beta(s) = \frac{s}{2 + c_n^2},$$

and equilibria functions

$$v_{i,n}^{\beta}(s) = \frac{1+\beta c_n^2}{\sqrt{c_n^2+2}}\sqrt{s},$$

when c_n is define in the following way: $c_1 = 0$ and for $n \geq 1$

$$c_{n+1} = \frac{2}{3}\frac{1+\beta c_n^2}{\sqrt{c_n^2+2}}.$$

If we take a limit $n \to \infty$, we obtain

$$c^{\beta} = \sqrt{\frac{\sqrt{(18-8\beta)^2+16(9+4\beta^2)}-(18-8\beta)}{2(18-8\beta)}}.$$

From Theorem 14 we immediately conclude that

$$f_i^{\beta}(s) = \frac{s}{2+\left(c^{\beta}\right)^2},$$

and $\pi^{\beta} := \left(f_1^{\beta}, f_2^{\beta}\right)$ is a stationary Nash equilibrium in the β - discounted infinite horizon game and

$$v_i^{\beta}(s) = \frac{1+\beta\left(c^{\beta}\right)^2}{\sqrt{\left(c^{\beta}\right)^2+2}}\sqrt{s}$$

is the equilibrium function. By Main Theorem 15, taking a limit $\beta \to 1$ we obtain

$$f_i^{\star}(s) = \frac{20}{30+\sqrt{192}}s,$$

and

$$v_i^{\star}(s) = \frac{\sqrt{192}+10}{\sqrt{20\sqrt{192}+600}}\sqrt{s}.$$

By Main Theorem 16 it follows that the stationary strategy $\pi^{\star} = (f_1^{\star}, f_2^{\star})$ is a Nash equilibrium in the undiscounted stochastic game with limiting average criterion. Moreover, this strategy is ε equilibrium in β - discounted infinite horizon game for sufficiently large β.

References

[1] R. Amir, Continuous stochastic games of capital accumulation with convex transitions, Games and Economic Behavior 15 (1996), 111–131.

[2] Ł. Balbus and A.S. Nowak, Construction of Nash equilibria in symmetric stochastic games of capital accumulation, Mathematical Methods of Operations Research 60 (2004), 267–277.

[3] Ł. Balbus and A.S. Nowak, Existence of perfect equilibria in a class of multigenerational stochastic games of capital accumulation, Automatica 44 (2008), 1471–1479.

[4] Ł. Balbus and A.S. Nowak, Nash equilibria in unconstrained stochastic games of resource extractions, International Game Theory Review 10 (2008), 25–35.

[5] D.P. Bertsekas and S.E. Shreve, Stochastic Optimal Control: the Discrete-Time Case, (New York: Academic Press, 1978).

[6] D. Blackwell, Discounted dynamic programming, The Annals of Mathematical Statistics 36 (1965), 226-235.

[7] L.O. Curtat, Markov equilibria of stochastic games with complementarities, Games and Economic Behavior 17 (1996), 177–199.

[8] P.K. Dutta and R.K. Sundaram, Markovian equilibrium in a class of stochastic games: existence theorem for discounted and undiscounted models, Economic Theory 2 (1992), 197–214.

[9] A.M. Fink, Equilibrium in a stochastic n - person game, Journal of Science of the Hiroshima University 28 (1964), 89–93.

[10] D. Levhari and L. Mirman, The great fish war: an example using a dynamic Cournot-Nash solution, Bell Journal of Economics 11 (1980), 322-344.

[11] M. Majumdar and R.K. Sundaram, Symmetric stochastic games of resource extraction: the existence of non-randomized stationary equilibrium, in: Stochastic Games and Related Topics (Kluwer Academic Publishers, Dordrecht, The Netherlands, 1991), T.E.S. Raghavan et al. (eds.), Shapley Honor Volume, 175–190.

[12] A.S. Nowak, On a new class of nonzero-sum discounted stochastic games having stationary Nash equilibrium points, International Journal of Game Theory 32 (2003), 121–132.

[13] A.S. Nowak, N-person stochastic games: extensions of the finite state space case and correlation, in: Stochastic Games and Applications (Kluwer Academic Publishers, Dordrecht, The Netherlands, 2003), A. Neyman and S. Sorin (eds.), Lecture notes in NATO Science Series C, Mathematical and Physical Sciences, 93–106.

[14] A.S. Nowak and P. Szajowski, On Nash equilibria in stochastic games of capital accumulation, in: Game Theory and Applications 9 (2003), L.A. Petrosjan and V.V. Mazalov (eds.), 119–129.

[15] U. Rieder, Equilibrium plans for nonzero-sum Markov games, in: Game Theory and Related Topics (North-Holland, Amsterdam, 1979), 91-102.

[16] P. Szajowski, Constructions of Nash equilibria in stochastic games of resource extraction with additive transition structure, Mathematical Methods of Operations Research 63 (2006), 239–260.

[17] Q. Zhu, X. Guo, and Y. Dai, Unbounded cost Markov decision proces with limsup and liminf average criteria: new conditions, Mathematical Methods of Operations Research 61 (2005), 469–482.

Quantitative Methods in Economics
Vol. XI, No. 1, 2010, pp. 189–198

ON INCENTIVE COMPATIBLE DESIGNS OF FORECASTING CONTRACTS

Bogumił Kamiński

Warsaw School of Economics
Al. Niepodległości 162, 02-554 Warszawa, Polska
e-mail: bkamins@sgh.waw.pl

Abstract: In the paper the optimal design of forecasting contracts in principal-agent setting is investigated. It is assumed that the principal pays the agent (the forecaster) based on an announced forecast and an event that materializes next. Such a contract is called incentive compatible if the agent maximizes her payoff when she announces her true beliefs. This paper relaxes the assumption present in earlier works on this subject that agent's beliefs are deterministic by allowing them to be random (i.e. stemming from estimation). It is shown that for binary or nominal events the principal can learn only expected values of agent's predictions in an incentive compatible way independent of agent's signal space. Additionally it is proven that incentive compatible payment schemes give the agent a strictly positive incentive to improve the precision of her estimates.

Key words: contract design, forecasting, scoring rules

Introduction

The demand for reliable forecasts is omnipresent in private, business and academic applications. People watch weather forecasts, listen to GDP growth predictions given by central banks and are lively interested in expected rate of ozone layer depletion. The process of forecast preparation is often difficult and requires much effort and skill. Therefore it is not unusual that someone who demands a forecast relies on expert help. Proffessional approach to such *delegated* activity calls for probabilistic assessments of future events[1]. For example central bank economists are interested not only in point estimate of inflation, but also the distribution of the estimator. In everyday life people prefer to know the probability of rain next day — not only a simple

[1]Gneiting and Raftery (2007) give a review of literature in weather and macroeconomic domains focusing on probabilistic forecasting.

statement of the most probable weather state. An interesting practice of probabilistic forecasting is employed by Gartner Inc., an IT market monitoring company, who assigns probabilities to its predictions of future trends.

The widely spread practice of relying on external forecasts rises a question how to evaluate them. Such assessment can be simply used as a measure of forecaster performance. It can also be a part of a contract between the forecaster and the side that demands the forecast. Further the latter understanding will be taken in order to fix the terminology used. Following the standard conventions in economic literature, see for example Salanié (2005), we will call the forecaster an agent and the forecast demanding side a principal. We assume that the principal contracts the agent to produce a forecast. The payment scheme in this contract is based on the forecast given and the event that later occurs. Later, following the literature (Hendrickson and Buehler, 1971), agent's payoff function will be called a *scoring rule*. It is taken that neither the principal nor the agent can influence the probability distribution of the event. We will say that the scoring rule is *proper* if the agent maximizes her expected profit when she reveals her true beliefs about the probability of the forecasted event[2]. In economic literature, see Salanié (2005), proper scoring rules are said to implement an *incentive compatible* contract.

The problem of evaluation of forecasts has been studied since the work of Brier (1950), who considered a binary event prediction in weather forecasting applications and proposed one proper scoring rule for such a case. Later line of research focused on classification of proper scoring rules for binary (Schervish, 1989), nominal (Savage, 1971) and continuous (Gneiting and Raftery, 2007) random variables.

The implicit assumption that was made in the forecasting contracts literature was that the true expectations of the forecaster are certain. For example, in binary event case, that she has crisp beliefs about the probability of the event. McCarthy (1956) has considered an experimentation procedure on agent's side that would lead to a change in her beliefs but still the a priori and a posteriori beliefs of the agent were non-random. However, in real applications a far more common situation is that the true agent's expectations are represented as estimators. For example when predicting the share of votes on some candidate in elections a pool-making company gets its imprecise estimate by questioning a random sample of voters[3].

The objective of this research is to extend the existing results on proper scoring rules by adding the possibility that agent's beliefs are given as a random variable. A focus is put on the classical binary event case, as it allows for most clear presentation of the analysis. An extension of the results to nominal events is presented.

The paper is organized as follows. First a formal model of a binary event forecasting contract under non-random (classical) and random expectations assumption are presented and outline the standard results obtained in the literature for the former case are given. In the next section the properties of random beliefs model in binary case are investigated. Finally it is shown how the results can be extended to nominal variables. The paper is finished by concluding remarks.

[2]A formal definition of a proper scoring rule is given in section .

[3]Moreover, in practice such companies often quote the confidence of their predictions.

The binary event forecasting contract model

In this section we first introduce the standard model of binary forecasting contract and outline its properties. Next it is shown how it can be extended to agent's random beliefs.

Let $x \in \{0;1\}$ be the event that is predicted. For example $x = 1$ can be associated with rain and $x = 0$ with no-rain forecast[4]. It is important to distinguish between agent's belief on $\Pr(x = 1)$, which will be denoted q and the value announced by her to the principal, further denoted p.

A scoring rule is a pair of mappings $S_x : [0;1] \to \mathbf{R}$ taking announced probability p as an argument and returning agent's payoff. Notice that the mappings are indexed by the event x. The agent obtains payoff $S_1(p)$ when event $x = 1$ occurs and $S_0(p)$ when $x = 0$ happens. We assume that S_x can be set by the principal. Moreover we take that the agent wants to maximize her expected payoff, denoted $R(p)$, subject to her beliefs[5]:

$$R(p) = qS_1(p) + (1-q)S_0(p) \to \max \tag{1}$$

Using this assumption we can define proper scoring rules.

Definition 1. A scoring rule S_x is proper if and only if

$$\forall q \in [0;1] \forall [0;1] \ni x \neq q : R(q) > R(x)$$

and S_x is bounded from above and real valued (except possibly that $S_1(0)$ and $S_0(1)$ can be equal to $-\infty$).

Under this condition Gneiting and Raftery (2007) give a convenient condition characterizing the class of proper scoring rules[6].

Theorem 2 (Gneiting and Raftery, 2007)**.** *Every proper scoring rule is of the form*

$$\begin{aligned} S_1(p) &= G(p) + (1-p)G'(p) \\ S_0(p) &= G(p) - pG'(p) \end{aligned} \tag{2}$$

where $G : [0;1] \to \mathbf{R}$ is a bounded and strictly convex function and $G'(p)$ is a subgradient of G at point p, for all $p \in [0;1]$.

Notice that:

$$R(p) = q(G(p) + (1-p)G'(p)) + (1-q)(G(p) - pG'(p)) = G(p) + (q-p)G'(p)$$

so $G(q) = R(q)$.

[4]The rain/no rain probability prediction example was originally used by Brier (1950) in his pioneering research on scoring rules.

[5]This follows the standard economics assumption of agent's rationality following expectation maximization principle, see Mas-Collel et al. (1995).

[6]Gneiting and Raftery (2007) use the term *regular strictly* proper scoring rule.

Example 3. By putting $G(p) = -p(1-p)$ we get $S_1(p) = -(1-p)^2$ and $S_0(p) = -p^2$. It is a scoring rule originally proposed by Brier (1950). It is illustrated on Figure 1. Such scoring function always gives the agent negative payouts. However it can be seen from Theorem 2 that transforming $G(p)$ to $G(p) + \alpha$, where $\alpha \in \mathbf{R}$, keeps it convex while it does not affect the shape of $S_x(p)$. Therefore the principal can adjust the payouts to agent's reservation level if it exists[7]. Also notice that Brier score is a symmetric scoring function, that is $S_1(p) = S_0(1 - p)$, so it does not depend on labeling of event realizations. However in general scoring rules do not have to be symmetric. Asymmetric scoring functions could be preferred by the principal when she assigns different values to $x = 1$ and $x = 0$.

If we take Shannon's entropy[8] $G(p) = p\ln(p)+(1-p)\ln(1-p)$ as a second example then $S_1(p) = \ln(p)$ and $S_0(p) = \ln(1 - p)$. Note that in this case the payoff is $-\infty$ if the event is assigned probability 0 and it happens.

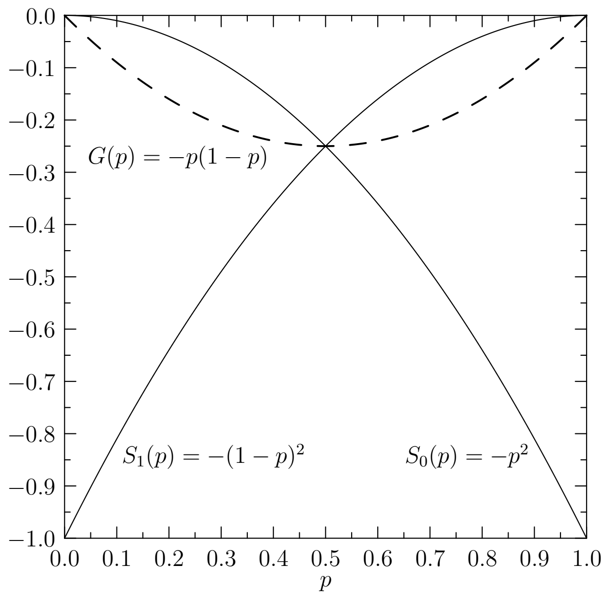


Figure 1. Brier scoring functions $S_x(p)$ and expected score $G(p)$.

Now assume that the agent does not know q exactly and Q is a random variable representing agent's uncertainty about true value of q. Denote cumulative distribution function of Q by $F : [0; 1] \to [0; 1]$.

Let us denote $\mathcal{S}$ the set of signals that can be sent by the agent. In the simplest case if $\mathcal{S} = [0; 1]$ the agent is asked to report a fraction, but in general the principal

[7]The reservation level is understood as minimal payout that the agent requires to receive.

[8]Shannon and Weaver (1949) first introduced it as a measure of amount of information contained in a message. In our case the message is what the agent tells the principal.

might want to have something else reported. For example she might ask the agent to reveal F. Based on this we define *extended scoring rules* as $S_x^E : \mathcal{S} \to \mathbf{R}$, that are natural generalization of scoring rules. Notice that in the case when $\mathcal{S} = [0;1]$ extended scoring rule reduces to scoring rule.

The calculation of agent's expected payoff takes into consideration the uncertainty of Q as follows:

$$\forall s \in \mathcal{S} : R^E(s) = \int_0^1 qS_1^E(s) + (1-q)S_0^E(s)dF(q) \tag{3}$$

In the next section the consequences of introducing uncertainty of q to agent's beliefs will be analyzed.

Uncertain expectation for binary event properties

Under uncertainty of q several natural questions arise. Firstly it will be investigated how the agent will behave when faced with classical proper scoring rules. Next it will be shown what kind of information the principal can extract from the agent by appropriately setting an extended scoring rule. Lastly a question if extended scoring rules can give an incentive to the agent to put extra effort to improve the precision of estimation of q will be answered.

Expectation revelation principle

In this subsection we will show that when the agent has uncertain beliefs about q the only information that the principal can obtain from her in an incentive compatible way is an expected value of Q — the phenomenon will be called *expectation revelation principle*. First it is shown that it is possible to construct a scoring rule under which the agent will reveal $\mathrm{E}(Q)$.

Theorem 4. *If $\mathcal{S} = [0;1]$ and $S_x^E(s)$ is a proper scoring rule then $R^E(s)$ is maximized for $s = E(q)$.*

Proof. Notice that:

$$R^E(s) = \int_0^1 qS_1^E(s) + (1-q)S_0^E(s)dF(q) = \mathrm{E}(Q)S_1^E(s) + (1-\mathrm{E}(Q)) - S_0^E(s)$$

so $R^E(s)$ equals $R(s)$ for $q = \mathrm{E}(Q)$. Therefore by Theorem 2 $R^E(s)$ is maximized for $s = \mathrm{E}(Q)$. ∎

The above theorem shows that by using proper scoring rules the principal will learn the expected value of Q. Using this one can notice that all results in existing literature concerning binary forecasting contracts can be directly extended to uncertain estimation of q by assumption that the agent reveals her expected value of Q.

In general the principal might be interested in other statistics of the random variable Q than expected value. In general we can denote the statistics as a function of cumulative distribution function of Q: $\chi : [0;1]^2 \to \mathcal{S}$. This could be for example a median of the distribution — then $\chi(F) = F^{-1}(0.5)$ or the distribution function — then $\chi(F) = F$. Analogously to to the proper scoring rules we define χ-proper extended scoring rule.

Definition 5. An extended scoring rule S_x^E is χ-proper if and only if for all cumulative distribution functions F:

$$\forall \mathcal{S} \ni s \neq \chi(F) : R^E(\chi(F)) > R^E(s)$$

and S_x^E is bounded.

The following theorem shows that the principal is unable to retrieve any significantly different information from the agent than expected value of Q.

Theorem 6. *If S_x^E is χ-proper extended scoring rule then there exists such mapping $h(\cdot)$ that $\chi(F) = h(E(Q))$.*

Proof. Assume that the converse is true and such χ-proper extended scoring rule exists. Then there exist such random variables Q_1 and Q_2 with cumulative distribution functions F_1 and F_2 for which:

$$\mathrm{E}(Q_1) = \mathrm{E}(Q_2) \wedge \chi(F_1) \neq \chi(F_2).$$

The above condition reads that χ is not a function of expected value of random variable it transforms.

But then using equation 3 and definition of χ-proper extended scoring rule we get:

$$\begin{cases} \quad \mathrm{E}(Q_1)(S_1^E(\chi(F_1)) - S_0^E(\chi(F_1))) + S_0^E(\chi(F_1)) > \\ > \mathrm{E}(Q_1)(S_1^E(\chi(F_2)) - S_0^E(\chi(F_2))) + S_0^E(\chi(F_2)) \\ \quad \mathrm{E}(Q_2)(S_1^E(\chi(F_1)) - S_0^E(\chi(F_1))) + S_0^E(\chi(F_1)) < \\ < \mathrm{E}(Q_2)(S_1^E(\chi(F_2)) - S_0^E(\chi(F_2))) + S_0^E(\chi(F_2)) \end{cases}$$

However, by assumption $\mathrm{E}(Q_1) = \mathrm{E}(Q_2)$ so the above equations are contradicting. ■

The above result shows two important properties of the forecasting contract problem. Firstly - the only information the principal can get from the agent is expected value of Q under any assumption on the structure of agent's signal. Secondly - we know that $\mathrm{E}(Q)$ is reported under classical proper scoring rules. Therefore it is obsolete to consider extended scoring rules — all the information that can be obtained by the principal from the agent can be extracted in the standard framework.

Agent's effort optimization

In this section it will be analyzed how can proper scoring rules incentivize the agent to improve the precision of her initial estimation of Q. For this assume that the agent can gather some information i coming frome the information space $\mathcal{I}$. The expected realization of gathered information i is a priori random and conditional on agent's beliefs on the distribution of Q. It is assumed that, conditional on Q, the agent can assign a probability measure to $\mathcal{I}$ and define a random variable I representing the predicted distribution of received information. We assume that after the information i is gathered the agent updates her beliefs on the distribution of Q. We will denote the updated distribution conditional on realized i by Q_i.

Example 7. Assume that initial distribution of Q is given by Beta distribution with parameters α and β (it will be denoted $B(\alpha, \beta)$)[9]. We know that $\mathrm{E}(Q) = \alpha/(\alpha+\beta)$. Now we take, that the experiment that can be made is one random draw of x from the population. We have $\mathcal{I} = \{0; 1\}$ and conditional on agent's beliefs $\Pr(I = 1) = \alpha/(\alpha+\beta)$. Note that beta distribution is a conjugate prior to Bernoulli distribution. Using standard results on Bayesian updating, see DeGroot (2004), posterior belefs of the agent will be:

$$\begin{array}{llll} Q_0 \sim B(\alpha, \beta+1) & \text{for} & i=0 & \text{with probability} \quad \beta/(\alpha+\beta) \\ Q_1 \sim B(\alpha+1, \beta) & \text{for} & i=1 & \text{with probability} \quad \alpha/(\alpha+\beta). \end{array} \tag{4}$$

Continuing the analysis we have from the Theorem 2 that a priori agent's expected profit under proper scoring rule is equal to $G(\mathrm{E}(Q))$. Having seen i she will report $\mathrm{E}(Q_i)$, further denoted X_i. A priori agent's expected profit, conditional on the decision of running the experiment, is equal to $\mathrm{E}(G(X_I))$ (notice that this expectation is taken over I). Also we can notice that $\mathrm{E}(X_I) = \mathrm{E}(Q)$, as a priori the expected beliefs must be equal. The following example follows Example 7 to illustrate this property:

Example 8. Following equations 4 we can calculate that:

$$\begin{array}{l} X_0 = \alpha/(\alpha+\beta+1) \\ X_1 = (\alpha+1)/(\alpha+\beta+1). \end{array}$$

Using this we have:

$$\mathrm{E}(X_I) = \frac{\beta}{\alpha+\beta}\alpha/(\alpha+\beta+1) + \frac{\alpha}{\alpha+\beta}(\alpha+1)/(\alpha+\beta+1) = \alpha/(\alpha+\beta) = \mathrm{E}(Q).$$

So the agent does not expect to change her expectations.

Returning to the main line of reasoning it can be seen that the change in expected profit of the agent when she decides to run the experiment is equal to:

$$V(Q, \mathcal{I}) = \mathrm{E}(G(X_I)) - G(\mathrm{E}(Q)). \tag{5}$$

It can be shown that this change is always positive:

Theorem 9. *If X_I is not certain then $V(Q, \mathcal{I}) > 0$.*

Proof. It is enough to show that $G(\mathrm{E}(Q)) < \mathrm{E}(G(X_I))$. However, $G(\mathrm{E}(Q)) = G(\mathrm{E}(X_I))$, so using the convexity of G and Jensen's inequality we get the result. ■

[9] The probability distribution function is given as $f(x, \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}$, see Gelman et al. (2004).

This means that the agent is given the incentive to improve the precision of the estimate of Q. Of course she will have to compare $V(Q,\mathcal{I})$ with her cost of running the experiment.

As a special case we can assume that the experiment gives perfect information on the value of q. In such a situation we have that cumulative distribution functions of X_I and Q are equal. Under this the expected value of perfect information (EVPI) is equal to $\mathrm{E}(G(Q)) - G(\mathrm{E}(Q))$. The above results are illustrated with the example using Brier score (see Example 3):

Example 10. Under Brier score we have $G(q) = -q(1-q)$ therefore:

$$\mathrm{E}(G(X_I)) = \mathrm{E}(-X_I(1-X_I)) = \mathrm{E}(X_I^2) - \mathrm{E}(X_I) = \mathrm{D}^2(X_I) + \mathrm{E}^2(X_I) - \mathrm{E}(X_I).$$

Remembering that $\mathrm{E}(X_I) = \mathrm{E}(Q)$ it follows that:

$$\begin{aligned} &V(Q,\mathcal{I}) = \mathrm{E}(G(X_I)) - G(\mathrm{E}(Q)) = \\ &= \big(\mathrm{D}^2(X_I) + \mathrm{E}^2(Q) - \mathrm{E}(Q)\big) - \big(\mathrm{E}^2(Q) - \mathrm{E}(Q)\big) = \mathrm{D}^2(X_I). \end{aligned}$$

In particular $\mathrm{EVPI}(Q) = \mathrm{D}^2(Q)$.

Summing up the results of this section it has been shown that the only information that the principal can get from the agent in an incentive compatible way is $\mathrm{E}(Q)$ and allowing the agent to signal other information than her probability beliefs does not change the result. Finally — the agent gets a positive incentive to improve the precision of the estimate of Q. In the next section it will be shown how to extend this result to nominal events.

Nominal event forecasting contracts

In the nominal event case we assume that $x \in \{1,\dots,n\}$ and $\mathbf{N} \ni n > 2$. That is — there is a finite number of possible outcomes. In certain beliefs case by $\mathbf{q} = (q_1,\dots,q_n)$ we will denote agent's expectations and by $\mathbf{p} = (p_1,\dots,p_n)$ announced probabilities. In this case every scoring rule S_x takes $\mathbf{p}$ as an argument. Agent's expected payoff can be calculated as $R(\mathbf{p}) = \sum_{i=1}^n q_i S_i(\mathbf{p})$. The Definition 1 of proper scoring rule is extended naturally by requiring that the strict maximum of $R(\mathbf{p})$ is attained in $\mathbf{q}$.

For uncertain beliefs we define $\mathbf{Q} = (Q_1,\dots,Q_n)$ as a random vector representing beliefs of the agent. The definition of S_x^E remains unchanged (as it relates to arbitrary set $\mathcal{S}$) and we can calculate expected payoff $R^E(s) = \mathrm{E}\left(\sum_{i=1}^n Q_i S_i^E(s)\right)$. The definition of χ-proper extended scoring rule is also unchanged (except for that now $\chi : [0;1]^{n+1} \to \mathcal{S}$). In nominal case it can be shown that the binary results can be naturally extended:

Theorem 11. *If $\mathcal{S} = [0;1]^n$ and $S_x^E(s)$ is a proper scoring rule then $R^E(s)$ is maximized for $s = E(\mathbf{q})$.*

Proof. Using the linear separability of expected value operator the definition of $R^E(s)$ in nominal case can be rewritten as $R^E(s) = \sum_{i=1}^n \mathrm{E}(Q_i) S_i^E(s)$. And because $S_x^E(s)$ is a proper scoring rule we get that in optimum $s = \mathrm{E}(\mathbf{q})$. ∎

Theorem 12. *If S_x^E is χ-proper extended scoring rule then there exists such mapping $h(\cdot)$ that $\chi(F) = h(E(\mathbf{Q}))$.*

Proof. Assume that the converse is true and such χ-proper extended scoring rule exists. Then there exist such random vectors $\mathbf{T}$ and $\mathbf{U}$ with n-variate cumulative distribution functions F_1 and F_2 for which: $\mathrm{E}(\mathbf{T}) = \mathrm{E}(\mathbf{U}) \wedge \chi(F_1) \neq \chi(F_2)$. But S_x^E χ-proper extended scoring rule so:

$$\begin{array}{c}\sum_{i=1}^{n} \mathrm{E}(T_i) S_i^E(\chi(F_1)) > \sum_{i=1}^{n} \mathrm{E}(T_i) S_i^E(\chi(F_2)) \\ \sum_{i=1}^{n} \mathrm{E}(U_i) S_i^E(\chi(F_1)) < \sum_{i=1}^{n} \mathrm{E}(U_i) S_i^E(\chi(F_2))\end{array}$$

A contradiction, so it is impossible that S_x^E is χ-proper. ■

Now we will show that Theorem 9 also holds for nominal case. For this a result first noted by McCarthy (1956) is needed:

Theorem 13 (McCarthy, 1956). *If S_x is a proper scoring rule for nominal x then $R(\mathbf{p})$ is strictly convex.*

And it can be seen that the proof of Theorem 9 remains unchanged for nominal event.

Summary

In the paper a forecasting contract design with uncertain forecaster's (agent's) beliefs was considered. It was assumed that such contract is incentive compatible if it is strictly optimal for the agent to reveal her true beliefs.

It was shown that in binary and nominal case the principal can learn only expected values of agent's predictions on event probabilities in an incentive compatible way. Payment schemes having this property, called proper scoring rules, can rely only on asking the agent to reveal her expectation. Additionally it was shown that proper scoring rules give the agent strictly positive incentive to improve the precision of her estimates.

The consequences of changing the assumption that the agent maximizes her expected payoff can be investigated in further work. Here one remark will be only made. Changing this assumption to maximization of *expected utility* (see for example Mas-Collel et al., 1995) does not change the results if we use a natural assumption that the utility function is strictly increasing. Then it is reversible and any proper (or χ-proper) scoring rule under expected utility maximization assumption must be equivalent to some proper scoring rule under expected payoff maximization.

It it also worth to investigate the problem on designing such incentive schemes that would allow the principal to learn other statistics of agent's beliefs than expected value. However, from the results presented in the paper, it can be seen that in order to achieve this goal such systems must allow for some additional information, for example: repeated measurement or comparison of several forecasters.

References

[Br] G. W. Brier, Verification of Forecasts Expressed in Terms of Probability, Monthly Weather Review, Vol. 78, No. 1 (1950), 1–3

[DeG] M.H. DeGroot, Optimal Statistical Decisions, John Wiley & Sons (2004)

[G et al] A. Gelman, J.B. Carlin, H.S. Stern and D.B. Rubin, Bayesian Data analysis, 2nd edition, Chapman & Hall / CRC (2004)

[GR] T. Gneiting and A.E. Raftery, Strictly Proper Scoring Rules, Prediction, and Estimation, Journal of the American Statistical Association, Vol. 102, No. 477. (2007), 359–378

[HB] A.D. Hendrickson and R.J. Buehler, Proper Scores for Probability Forecasters, The Annals of Mathematical Statistics, Vol. 42, No. 6 (1971), 1916–1921

[MWG] A. Mas-Collel, M.D. Whinston, J.R. Green, Microeconomic Theory, Oxford University Press (1995)

[McC] J. McCarthy, Measures of the value of information, Proceedings of the National Academy of Sciences, Vol. 42 (1956), 654–655

[Sl] B. Salanié, The Economics of Contracts, The MIT Press (2005)

[Sv] L.J. Savage, Elicitation of personal probabilities and expectations, Journal of the American Statistical Association, Vol. 66, No. 336 (1971), 783–801

[Sch] M.J. Schervish, A General Method for Comparing Probability Assessors, The Annals of Statistics, Vol. 17, No. 4 (1989), 1856–1879

[SW] C.E. Shannon and W. Weaver, The Matematical Theory of Communications, Urbana, University of Illinois Press (1949)

Quantitative Methods in Economics
Vol. XI, No. 1, 2010, pp. 199–211

STRATEGIC SUBSTITUTES AND COMPLEMENTS IN COURNOT OLIGOPOLY WITH PRODUCT DIFFERENTIATION

Małgorzata Knauff

Warsaw School of Economics
Al. Niepodleglosci 162, 02-554 Warsaw, Poland
e-mail: mknauff@sgh.waw.pl

Abstract: We consider Cournot oligopoly with differentiated product. We develop respective sufficient conditions on the inverse demand and cost function that make the oligopoly a game of strategic substitutes when goods are substitutes and a game of strategic complementarities when goods are complements. The scope of this result is illustrated by examples.
JEL classification: C72, L10, L13.

Key words: Cournot oligopoly, product differentiation, strategic complements and substitutes.

Introduction

Cournot's model of oligopolistic competition (1838) represents the starting point of formalized economic theory and game theory. Nowadays it continues to be widely used in a number of applications in economic theory. The reason for this is its tractability, when product homogeneity can be assumed, and the fact that its properties are well established. In this paper we follow the line of literature dealing with the monotonicity of reaction correspondences, which is important in particular for studying problems of equilibrium existence and comparative statics.

Increasing reaction correspondences characterizes games with strategic complementarities, while games of strategic substitutes have downward sloping best replies. This expresses the strategic relationship among actions in this kind of games. Strategic complementarities covers situations when an increase in one player action leads to an increase in other players' marginal payoffs. Strategic substitutes describe the opposite situation, when an increase in one player action causes a decrease in other players' marginal payoffs. The former kind of games always possesses pure strategy Nash equilibria. Moreover, the set of pure strategy Nash equilibria, correlated equilibria and rationalizable strategies have identical bounds (see, Milgrom and Roberts, 1990, and Milgrom and Shannon, 1994). The latter kind of games, in the case of $2-$player game, can be converted into the game of strategic complementarities by reordering of one player's action set (Milgrom and Roberts, 1990).

To establish properties of a Cournot oligopoly with heterogeneous products several simplifying assumptions have been imposed. For instance, inverse demand function have an aggregative form with respect to other players' strategies (cf. Dubey et al., 2005). Hoernig (2003) assumes that, departing from the situation when firms produce the same quantities, when one firm deviates by raising its output, other firms adjust their outputs to the level allowing for avoiding an increase in market price. He conducts comparative statics in case of market entry and increasing number of firms, extending the results of Amir and Lambson (2000) for homogeneous products.

In this paper we provide general conditions for a Cournot oligopoly with product differentiation to have monotonic reaction correspondences. These results are a generalization of the results of Amir (1996) for homogeneous products to the case of differentiated product. We give various sufficient conditions for downward and upward sloping reaction correspondences. They allow for identifying increasing best responses even in case of inverse demand being submodular, and similarly, decreasing best responses in case of supermodular inverse demand. Examples illustrating the scope of applicability of these results are provided.

This approach gives a significant value added to the problem of equilibria existence and comparative statics. The standard approach demands profit function to be quasi-concave in own quantities. This is quite restrictive, particularly because nonconcavities in costs are not uncommon and very convex demand functions cannot be ruled out (Vives, 1999). The lack of quasi-concavity of payoffs causes discontinuities in the best response correspondences of firms and makes possible the nonexistence of equilibrium. Our approach covers games with monotonic reaction correspondences and does not rely on the regularity condition.

The paper is organized as follows. Next section contains an brief overview of relevant notions from supermodular optimization and games. Section 3 presents main results. In Section 4 they are discussed. All proofs are placed in Appendix.

Supermodular games

We introduce in this section a summary of all relevant notions and results from lattice theory, supermodular optimization and supermodular games useful in the remainder. We present them in context of real decision parameter spaces, since this is sufficient for our needs.

A function $F : X \times Y \to R$ is *supermodular* if, for all $x_1 \geq x_2, y_1 \geq y_2$

$$F(x_1, y_1) - F(x_2, y_1) \geq F(x_1, y_2) - F(x_2, y_2). \tag{1}$$

A function F is *submodular* whenever $-F$ is supermodular. For twice continuously differentiable functions this notion has a useful differential characterization, namely supermodularity (submodularity) is equivalent to cross-partial derivative being positive (negative).

We say that F is *log-supermodular* (*log-submodular*) if logarithm of F is supermodular (*submodular).*

Topkis (1978) showed the following theorem on monotone optimization, central to our approach.

Theorem 1. *If F is upper semi-continuous and supermodular (submodular), then the maximal and minimal selections of*

$$\arg\max_{x\in X}\{F(x,y), y\in Y\}$$

are non-decreasing (non-increasing).

If F is strictly supermodular (submodular), then this theorem holds for every selection of $\arg\max_{x\in X}\{F(x,y), y\in Y\}$.

The property of supermodularity is of cardinal nature, in the sense that it is not preserved by monotonic transformation. Below we provide a definition of another notion, of ordinal nature, which can be treated as a generalization of supermodularity.

A function F has the (*dual*) *single-crossing property* in (x,y) if, for all $x_1 \geq x_2, y_1 \geq y_2$

$$F(x_2,y_1) - F(x_2,y_2)(\leq) \geq 0 \Rightarrow F(x_1,y_1) - F(x_1,y_2)(\leq) \geq 0. \tag{2}$$

Strict (dual) single crossing property is defined by the implication with strict inequality on the right hand side.

Obviously, (1) implies (2), but the converse does not hold.

Milgrom and Shannon (1994) generalized the result of Topkis for functions possessing the single crossing property.

Theorem 2. *If F is upper semi-continuous and satisfies the (dual) single crossing property, then the maximal and minimal selections of*

$$\arg\max_{x\in X}\{F(x,y), y\in Y\}$$

are non-decreasing (non-increasing).

If F has strict (dual) single crossing property, then this theorem holds for every selection of $\arg\max_{x\in X}\{F(x,y), y\in Y\}$.

The single crossing property has no differential characterization like supermodularity. Milgrom and Shannon (1994) showed that this property can be tested using the Spence-Mirrlees condition defined in the following theorem.

Theorem 3. *Let $F : R^3 \to R$ be continuously differentiable and $F_2(a,b,s) \neq 0$.*[1] *$F(a,h(a),s)$ satisfies the single-crossing property in (a,s) for all functions $h : R \to R$ if and only if*

$$\frac{F_1(a,b,s)}{|F_2(a,b,s)|} \tag{3}$$

[1]Subscripts denote partial derivatives with respect to the certain variable, e.g. here $F_2(a,b,s) = \frac{\partial F(a,b,s)}{\partial b}$. We keep this notation throughout the paper.

is increasing in s.

In applications, verifying (3) leads to conclusion that $F(a, h(a), s)$ satisfies the single crossing property in (a, s) for suitable choice of function h (which is often an identity function, see Milgrom and Shannon, 1994, and Amir, 2005).

A game with compact real action spaces is *supermodular* (*submodular*) if each payoff function is supermodular (submodular) and upper semi continuous in own actions.

Replacing supermodularity by the single crossing property enables to define a broader class of games: a game is *ordinally supermodular* (*submodular*) when its action space is compact, the payoff functions have the (dual) single crossing property and are upper semi continuous in own actions.

Supermodularity of the payoff functions can be interpreted as a complementarity between the players' actions, namely an increase some players' actions causes an increase in the marginal payoff of the others. Hence, this kind of games are also called *games with strategic complementarities.* Submodularity expresses the opposite phenomenon - substitutability of the players actions: increasing some players actions causes a decrease in the marginal payoff of the others, thus this kind of games are also called *games of strategic substitutabilities.*

Corollary 4. *Every $n-$player game of strategic complementarities has pure strategy Nash equilibrium.*

There is no equivalent corollary for $n-$player games of strategic substitutes. Only for the case of two players the existence of equilibrium is guaranteed.

Corollary 5. *Every $2-$player game of strategic substitutes has pure strategy Nash equilibrium.*

Conditions and examples

Consider a Cournot oligopoly game, when n firms decide simultaneously about the products quantity. Let $x_i \in X_i$ be a production quantity of firm i and $x_{-i} \in X_{-i}$ represents vector of production quantities of the other $n-1$ firms. The products are heterogeneous. Denote $P^i(x_i, x_{-i})$, $i = 1, .., n$, a system of inverse demand functions describing the market. Each of the functions is twice continuously differentiable and $P^i_j(x_i, x_{-i}) = P^j_i(x_j, x_{-j})$, $i \neq j$. Then the profit of firm i is given by

$$\Pi^i(x_i, x_{-i}) = x_i P^i(x_i, x_{-i}) - C^i(x_i) \tag{4}$$

where $C^i(\cdot)$ is a differentiable cost function. Assume that $P^i(x_i, x_{-i})$ is decreasing in own action, and $C^i(\cdot)$ is increasing.

We say that goods i and j are (strict) *substitutes*, if demand for i (strictly) rises with the increase in the price of j. Goods i and j are (strict) *complements,* if demand for i (strictly) goes down with the increase of the price of j. It can be translated in terms of inverse demand function, so that P^i is (strictly) decreasing in x_j, if goods i and j are substitutes, and the converse holds if goods i and j are complements.

Define reaction (best response) correspondence of firm i as

$$r_i(x_{-i}) = \arg\max_{x_i \in X_i} \{\Pi^i(x_i, x_{-i}), x_{-i} \in X_{-i}\}.$$

From Theorem 1 the game has increasing reaction correspondences if $\Pi^i, i = 1, .., n$, is supermodular or $\Pi^i_{ij} \geq 0, \forall j \neq i$, hence if each firm's revenue is supermodular (see also Novshek, 1985). For firm i it is equivalent to the following condition:

$$P^i_j(x_i, x_{-i}) - x_i P^i_{ij}(x_i, x_{-i}) \geq 0, \forall j \neq i.$$

But from Theorem 2 it follows that even weaker conditions can secure the monotonicity of the reaction correspondences. It is enough for the payoff function to satisfy the single crossing property.

Vives (1999) formulated independent conditions on demand and cost function in a Cournot oligopoly with product differentiation to be an ordinally supermodular game. We give them in the form of a theorem.

Theorem 6. *Assume that for $i = 1, .., n$*

1. *$P^i(\cdot)$ is log-supermodular.*
2. *C^i is strictly increasing.*
3. *Goods i and j are strict complements $\forall j \neq i$.*

Then the Cournot oligopoly, with profits given by (4) is an ordinally supermodular game.

The proof (not given by Vives, 1999), presented in Appendix, does not require differentiability of the inverse demand nor the cost function.

An analogous theorem can be formulated to provide conditions on a Cournot duopoly to have decreasing reaction correspondences.

Theorem 7. *Assume that for $i = 1, .., n$*

1. *$P^i(\cdot)$ is log-submodular.*
2. *C^i is strictly increasing.*
3. *Goods i and j are strict substitutes $\forall j \neq i$.*

Then the Cournot oligopoly, with profits given by (4) is an ordinally submodular game.

To provide some intuition of the scope of duality between these theorems for a duopoly case we can use Milgrom and Roberts (1990) action reordering argument. In case of 2−players game, changing order of the action space of one of the players converts a submodular game into a supermodular game. Take a game Γ, the Cournot duopoly and consider situation of player 1 with $P^1(x_1, x_2)$ log-submodular and x_1, x_2 substitute goods. Reordering of the player's 1 action space creates a new game $\widehat{\Gamma}$.

This is the Cournot duopoly with $P^1(\widehat{x_1}, x_2)$. Observe that whenever $P^1(\widehat{x_1}, x_2)$ is positive, it is log-supermodular, since $P^1 P^1_{12} - P^1_1 P^1_2 \leq 0$ and $\widehat{x_1}' = -1$ implies that $\widehat{x_1}' \left(P^1 P^1_{12} - P^1_1 P^1_2\right) \geq 0$ for every $x_1 \in X_1$ and $x_2 \in X_2$. Moreover, $\widehat{x_1}$ and x_2 are complements now, since relation between them changed to the reverse.[2] Therefore, there is complete duality between these two results. This kind of duality does not work if there is more than two players in the game.

The theorems given above work for any increasing cost function. But considering a specific cost function in relation with the inverse demand we may relax the assumption of log-supermodularity of the demand. It is possible to formulate a more general condition for a specific cost function, such that it guarantees increasing best response correspondence. We provide it in the theorem below. The differentiability of $P^i(x_i, x_{-i})$ and $C^i(x_i)$ is assumed here .

Theorem 8. *Assume that* $\forall j \neq i$

$$P^i_{ij}(x_i, x_{-i}) P^i(x_i, x_{-i}) - P^i_i(x_i, x_{-i}) P^i_j(x_i, x_{-i}) \geq P^i_{ij}(x_i, x_{-i}) C^{i\prime}(x_i). \tag{5}$$

Then the Cournot oligopoly, with profits given by (4) is an ordinally supermodular game.

Condition (5) can be reformulated into

$$P^i_{ij}(x_i, x_{-i})(P^i(x_i, x_{-i}) - C^{i\prime}(x_i)) - P^i_i(x_i, x_{-i}) P^i_j(x_i, x_{-i}) \geq 0. \tag{6}$$

It is straightforward that this condition is met, when the assumptions of Theorem 6 are satisfied. Moreover, for a suitable choice of C^i, this condition can be satisfied for a number of inverted demand functions, which do not satisfy log-supermodularity, as long as goods i and j remain complements. In particular, there are log-submodular functions giving rise strategic complementarity in the Cournot model. An illustration of this possibility is provided in Example 10.

When we assume fixed marginal costs, condition (5) can be interpreted as log-super-modularity of net-of-cost inverse demand functions $P^i(x_i, x_{-i}) - c^i$, $i = 1, 2$. Moreover, this condition says that the firm's i perceived net-of-cost inverse demand elasticity is increasing in firm's j output. Indeed, elasticity of net-of-cost inverse demand is given by

$$\begin{aligned}\varepsilon(x_i, x_j) \triangleq \frac{\partial\left(P^i(x_i, x_{-i}) - c^i\right)}{\partial x_i} & \frac{x_i}{P^i(x_i, x_{-i}) - c^i} \\ = P^i_i(x_i, x_{-i}) & \frac{x_i}{P^i(x_i, x_{-i}) - c^i}.\end{aligned}$$

It is increasing in x_j whenever its derivative with respect to x_j is positive:

$$\frac{\partial \varepsilon(x_i, x_j)}{\partial x_j} = x_i \frac{P^i_{ij}(x_i, x_{-i})\left(P^i(x_i, x_{-i}) - c^i\right) - P^i_i(x_i, x_{-i}) P^i_j(x_i, x_{-i})}{\left(P^i(x_i, x_{-i}) - c^i\right)^2}.$$

It holds if and only if (6) holds.

[2]Now, when price of good 2 increases, the demand for $\widehat{x_1}$ goes up.

Corollary 9. *Assume that marginal cost is constant, denoted* c^i *and* $P^i(x_i, x_{-i}) - c^i$ *is log-supermodular,* $i = 1, .., n$. *Then the Cournot oligopoly, with profits given by (4) is an ordinally supermodular game.*

We provide now an example illustrating the scope of application of Theorem 8 in the duopoly case. Our generalization enables to capture even situations when inverse demand, being submodular and/or log-submodular, gives rise, for specific cost functions, to an ordinally supermodular game.

Example 10. Consider a 2−player game. Let

$$P^1(x_1, x_2) = 1 + \frac{1}{(x_1+1)^2} + (x_2+1)\exp(-x_1).$$

It is easily verified that

$$\begin{aligned} P_1^1(x_1, x_2) &= \frac{-2}{(x_1+1)^3} - (x_2+1)\exp(-x_1) < 0 \\ P_2^1(x_1, x_2) &= \exp(-x_1) > 0, \end{aligned}$$

Hence, the goods are complements. Also

$$P_{12}^1(x_1, x_2) = -\exp(-x_1) < 0$$

and moreover

$$\left(\ln P^1(x_1, x_2)\right)_{12} = \frac{-\left(x_1^2 + 3x_1 + 4\right) x_1 \left(x_1 + 1\right) e^{-x_1}}{\left(1 + \left(x_1+1\right)^2 \left(\left(x_2+1\right) e^{-x_1} + 1\right)\right)^2} < 0.$$

Hence, the inverse demand is strictly submodular and strictly log-submodular, and thus cannot be log-supermodular. Therefore, the conditions of Theorem 6 are not satisfied. Now take the following cost function: $C^1(x_1) = 2x_1$ and check the condition (6):

$$\begin{aligned} & P_{12}^1(x_1, x_2)(P^1(x_1, x_2) - C^{1\prime}(x_1)) - P_1^1(x_1, x_2)P_2^1(x_1, x_2) \\ = \; & e^{-x_1}\left((x_2+1)\, e^{-x_1} + 2x_1 + \frac{2}{\left(x_1+1\right)^3}\right) > 0. \end{aligned}$$

It is satisfied for all positive x_1 and x_2. The verification that not all cost function satisfy this condition is left to the reader.

Concavity of firm's 1 profit can be easily verified, hence we use first order condition to find the reaction curve.

$$\Pi_1^1(x_1, x_2) = -\frac{(x_1-1)\left(x_1+1\right)^3 (x_2+1)\, e^{-x_1} + x_1\left(x_1^2 + 3x_1 + 4\right)}{\left(x_1+1\right)^3} = 0.$$

It cannot be solved explicitly for x_1, thus we present a numerical plot in Figure 1.

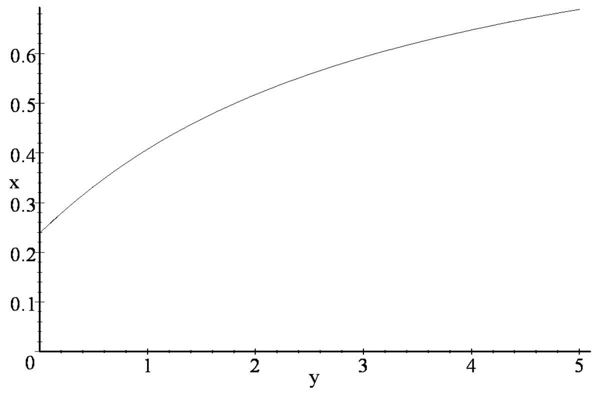


Figure 1. Reaction function of firm 1 in Example 10 is upward sloping.

Again, the dual theorem guaranteeing decreasing best responses can be formulated, and is given below.

Theorem 11. *Assume that* $\forall j \neq i$

$$P^i_{ij}(x_i, x_{-i})P^i(x_i, x_{-i}) - P^i_i(x_i, x_{-i})P^i_j(x_i, x_{-i}) \leq P^i_{ij}(x_i, x_{-i})C^{i\prime}(x_i). \tag{7}$$

Then the Cournot oligopoly, with profits given by (4) is an ordinally submodular game.

Similarly, in the case of duopoly, one can imagine this condition as derived from the reordering approach.

This condition covers all the cases, which satisfy conditions of Theorem 7 and, moreover, for specific cost functions, inverted demands, which are not log-submodular. In fact even log-supermodular inverted demands may lead to an ordinally submodular Cournot game, an illustration is provided in Example 13.

As before, we formulate a corollary for the case of constant marginal costs.

Corollary 12. *Assume that marginal cost is constant, denoted* c^i *and* $P^i(x_i, x_{-i}) - c^i$ *is log-submodular,* $i = 1, .., n$. *Then the Cournot oligopoly, with profits given by (4) is an ordinally submodular game.*

Log-submodularity of net-of-cost inverted demand can be interpreted as its elasticity being decreasing in x_j.

Usefulness of the Theorem 11 in the duopoly case is illustrated by the example below. Again, even supermodular and log-supermodular inverse demand function can give rise to an ordinally submodular game for some cost functions.

Example 13. Consider a $2-$player game and take

$$P^1(x_1, x_2) = \frac{e^{-x_1}}{(x_2 + 1)} + \frac{1}{x_1 + 1}.$$

It is easy to verify that

$$\begin{aligned} P^1_1(x_1,x_2) &= -\frac{e^{-x_1}}{(x_2+1)} - \frac{1}{(x_1+1)^2} < 0, \\ P^1_2(x_1,x_2) &= -\frac{e^{-x_1}}{(x_2+1)^2} < 0, \\ P^1_{12}(x_1,x_2) &= \frac{e^{-x_1}}{(x_2+1)^2} > 0, \end{aligned}$$

and

$$\left(\ln P^1(x_1,x_2)\right)_{12} = e^{-x_1}\frac{x_1}{\left(e^{-x_1}x_1 + e^{-x_1} + x_2 + 1\right)^2} > 0.$$

Hence, the products are substitutes and inverse demand is log-supermodular. Therefore the game does not satisfy the conditions of Theorem 7.

Take now the following cost function: $C^1(x_1) = \ln(x_1+1)$ and check condition (7),

$$\begin{aligned} P^1_{12}(x_1,x_2)(P^1(x_1,x_2) - C^{1\prime}(x_1)) - P^1_1(x_1,x_2)P^1_2(x_1,x_2) &= \\ -e^{-x_1}\frac{(x_2+1)(x_1+1)^2\ln(x_1+1) + (x_1+1)^2 e^{-x_1} + x_2 + 1}{(x_2+1)^3(x_1+1)^2} &< 0. \end{aligned}$$

It is satisfied for all positive x_1 and x_2. Therefore, the game is ordinally submodular.

From the first order condition of profit maximization we find best response curve. We verified that $\Pi^1(x_1,x_{12})$ is locally concave on the best response function. It is given by an implicit formula

$$e^{-x_1}x_1^2 - e^{-x_1}x_1 - e^{-x_1} + e^{-x_1}x_1^3 + x_2x_1 + x_1 = 0.$$

Since it cannot be explicitly solved for x_1, we present a numerical plot in Figure 2.

Discussion

Games of strategic complementarities have Nash equilibria in pure strategies. Hence, when the proper conditions are satisfied for a Cournot oligopoly with differentiated products, one can be sure that equilibria exist, moreover, in terms of strategies, there exist largest and smallest equilibria.

For games of strategic substitutes there is no equivalent result. We may only guarantee the existence of pure strategy Nash equilibria in 2−player game. For $n > 2$ one can distinguished games, in which the strategies of all opponents can be aggregated into one number. For this kind of games the existence of the pure strategy Nash equilibrium was shown by Dubey et al. (2005), see also Novshek (1985). Unfortunately, not all reasonable inverse demand functions have this property (cf. Hoernig, 2003).

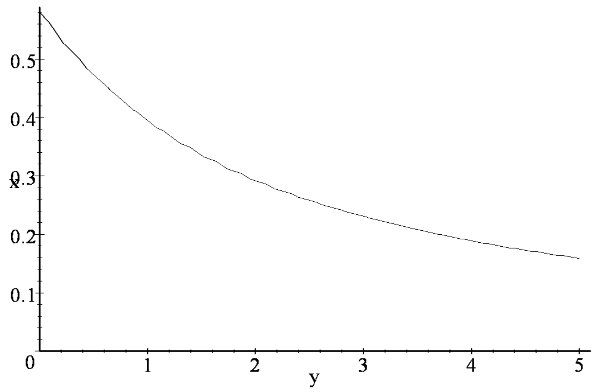


Figure 2. Downward sloping reaction function of firm 1 in Example 13.

Appendix

Proof. (of theorem 6) Consider firm i. Take $x_i^1 > x_i^2, x_{-i}^1 > x_{-i}^2$. From assumption 1 we have:

$$\ln P^i(x_i^1, x_{-i}^1) - \ln P^i(x_i^2, x_{-i}^1) \geq \ln P^i(x_i^1, x_{-i}^2) - \ln P^i(x_i^2, x_{-i}^2).$$

This implies that

$$\frac{P^i(x_i^1, x_{-i}^1)}{P^i(x_i^2, x_{-i}^1)} \geq \frac{P^i(x_i^1, x_{-i}^2)}{P^i(x_i^2, x_{-i}^2)}$$

or

$$P^i(x_i^2, x_{-i}^2)\frac{P^i(x_i^1, x_{-i}^1)}{P^i(x_i^2, x_{-i}^1)} \geq P^i(x_i^1, x_{-i}^2). \tag{8}$$

To prove that the game with log-supermodular inverse demand and increasing cost function is ordinally supermodular, it is enough to show that $\Pi^i(x_i, x_{-i})$ has the single crossing property. To this end we start from assuming that

$$\Pi^i(x_i^1, x_{-i}^2) \geq \Pi^i(x_i^2, x_{-i}^2).$$

Then

$$x_i^1 P^i(x_i^1, x_{-i}^2) - C^i(x_i^1) \geq x_i^2 P^i(x_i^2, x_{-i}^2) - C^i(x_i^2).$$

We can replace $P^1(x_i^1, x_{-i}^2)$ from (8)

$$x_i^1 P^i(x_i^2, x_{-i}^2)\frac{P^i(x_i^1, x_{-i}^1)}{P^i(x_i^2, x_{-i}^1)} - C^i(x_i^1) \geq x_i^2 P^i(x_i^2, x_{-i}^2) - C^i(x_i^2).$$

Multiplying by $\frac{P^i(x_i^2,x_{-i}^1)}{P^i(x_i^2,x_{-i}^2)}$ we get

$$x_i^1 P^i(x_i^1, x_{-i}^1) - \frac{P^i(x_i^2, x_{-i}^1)}{P^i(x_i^2, x_{-i}^2)} C^i(x_i^1) \geq x_i^2 P^i(x_i^2, x_{-i}^1) - \frac{P^i(x_i^2, x_{-i}^1)}{P^i(x_i^2, x_{-i}^2)} C^i(x_i^2).$$

$$\begin{aligned} x_i^1 P^i(x_i^1, x_{-i}^1) - x_i^2 P^i(x_i^2, x_{-i}^1) &\geq \frac{P^i(x_i^2, x_{-i}^1)}{P^i(x_i^2, x_{-i}^2)} (C^i(x_i^1) - C^i(x_i^2)) \\ &> C^i(x_i^1) - C^i(x_i^2). \end{aligned}$$

Since $P^i(x_i^2, x_{-i}^1) > P^i(x_i^2, x_{-i}^2)$ from assumption 2 and $C^i(x_i^1) > C^i(x_i^2)$ from assumption 3, it follows that

$$x_i^1 P^i(x_i^2, x_{-i}^1) - C^i(x_i^1) > x_i^2 P^i(x_i^1, x_{-i}^1) - C^i(x_i^2).$$

Hence

$$\Pi^i(x_i^1, x_{-i}^1) > \Pi^i(x_i^2, x_{-i}^1),$$

which means that (2) is satisfied. ■

Proof. (of Theorem 7) is analogous to the previous one. ■

Proof. (of Theorem 8) Consider firm i. Let the $F(a,b,s) = bP^i(a,s) - C^i(b)$. Then $F_1(a,b,s) = bP_1^i(a,s)$, $F_2(a,b,s) = P^i(a,s) - C^{i\prime}(b)$ and

$$\begin{aligned} \frac{F_1(a,b,s)}{|F_2(a,b,s)|} &= \frac{bP_1^i(a,s)}{|P^i(a,s) - C^{i\prime}(b)|} \\ &\quad \frac{\partial F_1(a,b,s)/\,|F_2(a,b,s)|}{\partial s} \\ &= b\frac{P_{12}^i(a,s)\left|P^i(a,s) - C^{i\prime}(b)\right| - P_1^i(a,s)P_2^i(a,s)}{\left(P^i(a,s) - C^{i\prime}(b)\right)^2}. \end{aligned}$$

Then, condition (3) holds when

$$P_{12}^i(a,s)\left|P^i(a,s) - C^{i\prime}(b)\right| - P_1^i(a,s)P_2^i(a,s) \geq 0.$$

Sufficient condition for this is (5). ■

Proof. (of Corollary 9) Consider firm i. Take $x_i^1 > x_i^2, x_{-i}^1 > x_{-i}^2$. Form assumption 2 we have:

$$\begin{aligned} &\ln(P^i(x_i^1, x_{-i}^1) - c^i) - \ln(P^i(x_i^2, x_{-i}^1) - c^i) \\ \geq\ &\ln(P^i(x_i^1, x_{-i}^2) - c^i) - \ln(P^i(x_i^2, x_{-i}^2) - c^i). \end{aligned}$$

This implies that

$$\frac{P^i(x_i^1, x_{-i}^1) - c^i}{P^i(x_i^2, x_{-i}^1) - c^i} \geq \frac{P^i(x_i^1, x_{-i}^2) - c^i}{P^i(x_i^2, x_{-i}^2) - c^i}$$

and

$$(P^i(x_i^2, x_{-i}^2) - c^i)\frac{P^i(x_i^1, x_{-i}^1) - c^i}{P^i(x_i^2, x_{-i}^1) - c^i} \geq P^i(x_i^1, x_{-i}^2) - c^i. \tag{9}$$

To prove that the game characterized by log-supermodular net-of cost inverse demand and constant marginal cost is ordinally supermodular, it is enough to show that $\Pi^i(x_i, x_{-i})$ has the single crossing property. To this end we start from assuming that

$$\Pi^i(x_i^1, x_{-i}^2) \geq \Pi^i(x_i^2, x_{-i}^2). \tag{10}$$

Then

$$x_i^1(P^i(x_i^1, x_{-i}^2) - c^i) \geq x_i^2(P^i(x_i^2, x_{-i}^2) - c^i).$$

We can replace $P^i(x_i^1, x_{-i}^2) - c^i$ from (9)

$$\begin{aligned} x_i^1((P^i(x_i^2, x_{-i}^2) - c^i)\frac{P^i(x_i^1, x_{-i}^1) - c^i}{P^i(x_i^2, x_{-i}^1) - c^i}) &\geq x_i^2(P^i(x_i^2, x_{-i}^2) - c^i). \\ x_i^1(\big(P^i(x_i^1, x_{-i}^1) - c^i\big)\frac{P^i(x_i^2, x_{-i}^2) - c^i}{P^i(x_i^2, x_{-i}^1) - c^i}) &\geq x_i^2(P^i(x_i^2, x_{-i}^2) - c^i). \end{aligned}$$

Dividing by $\frac{P^i(x_i^2, x_{-i}^2) - c^i}{P^i(x_i^2, x_{-i}^1) - c^i}$ we get

$$x_i^1(P^i(x_i^1, x_{-i}^1) - c^i) \geq x_i^2(P^i(x_i^2, x_{-i}^1) - c^i),$$

which means that (10) implies

$$\Pi^i(x_i^1, x_{-i}^1) \geq \Pi^i(x_i^2, x_{-i}^1),$$

which concludes the proof. ■

Proof. (of Theorem 11) is analogous to the proof of Theorem 8. ■

Proof. (of Corollary 12) is analogous to the proof of Corollary 9. ■

References

[A96] Amir, R. (1996), Cournot oligopoly and the theory of supermodular games; *Games and Economic Behavior,* 15, 132-148.

[A05] Amir, R. (2005), Supermodularity and complementarity in economics: an elementary survey, *Southern Economic Journal,* 71(3), 636-660.

[AL00] Amir, R., and V. Lambson (2000), On the effects of entry in Cournot markets, *Review of Economic Studies*, 67, 235-254.

[D05] Dubey, P., Haimanko, O. and A. Zapechelnyuk (2006), Strategic complements and substitutes, and potential games, *Games and Economic Behavior*, 54(1), 77-94.

[H03] Hoernig, S.H. (2003), Existence of equilibrium and comparative statics in differentiated goods Cournot oligopolies; *International Journal of Industrial Organization,* 21, 989-1019.

[N85] Novshek, W. (1985), On the existence of Cournot equilibria, *Review of Economic Studies,* 52, 85-98.

[MR90] Milgrom, P. and J. Roberts (1990), Rationalizability, learning, and equilibrium in games with strategic complementarities, *Econometrica,* 58, 1255-78.

[MS94] Milgrom, P. and C. Shannon (1994), Monotone comparative statics, *Econometrica*, 62, 157-180.

[T78] Topkis, D. (1978), Minimizing a submodular function on a lattice, *Operations Research,* 26, 305-321.

[V90] Vives, X. (1990), Nash equilibrium with strategic complementarities; *Journal of Mathematical Economics*, 19, 305-321.

[V99] Vives, X. (1999), *Oligopoly pricing: old ideas and new tools*, The MIT Press, Cambridge, Massachusetts.

Quantitative Methods in Economics
Vol. XI, No. 1, 2010, pp. 212–222

A DYNAMIC APPROACH TO THE STUDY OF UNEMPLOYMENT DURATION

Joanna Landmesser

Chair of Econometrics and Statistics, Warsaw University of Life Sciences
Nowoursynowska 159, 02-776 Warszawa, Poland
e-mail: joanna_landmesser@sggw.pl

Abstract: In this research work we investigate which factors influence the probability of leaving the unemployment state among people registered in the District Labor Office in Słupsk. The multiepisode hazard models with time-varying variables are suitable tools for this analysis. We introduced the changing labor market structure into the risk model. The main results achieved show that the job finding process depends on the historical time of the entry into the unemployment state and the actual historical time.
Also, the specific individual characteristics of people unemployed, such as gender, age, marital status, place of residence, education level, influence the probability of exiting the unemployment state. There is a greater tendency to leave the unemployment state when the person doesn't receive the unemployment benefit. The participation in the vocational training doesn't increase the transition rate into employment.

Key words: unemployment duration, hazard models, factor analysis

Introduction

Unemployment is a problem frequently connected with specific regions. In our research work we try to analyze in detail the situation in district Słupsk in north Poland (voivodeship Pomorskie).

The widely used unemployment measure – the unemployment rate – presents the proportion of the unemployed workforce which is seeking employment. However, such a ratio masks the dynamic nature of the labor market by failing to cover the length of time individuals are unemployed. It will be more useful to understand how the probability of exiting unemployment varies with demographic and economic characteristics. Therefore, to analyze the duration of unemployment among residents of district Słupsk, we use econometric risk models.

The main goal of our study is to introduce the changing labor market structure into the risk model in order to treat the time-dependent nature of the unemployment duration process in an adequate manner. There are two main ways in which the changing labor market affects finding job opportunities. First, people start their unemployment episodes in different labor market contexts (cohort effect, see also [B], [BGR]), and second, the labor market structure influences the opportunities of all people unemployed within the labor market at every moment (period effect).

We attempt to find measures for the macro effects on the labor market and suggest using 12 time series from official statistics, indicating the development of the labor market structure in Słupsk region. But such variables often measure similar features of a process and are highly correlated, which implies an identification problem. When we choose only uncorrelated series, it may only capture specific features of the labor market development. If time series represent aspects of an underlying regularity, it is more appropriate to look for these latent dimensions. To do this, we apply the statistical method – the exploratory factor analysis. The application of this method allows to isolate the basic principal factors from the given set of variables describing the development of the labor market in Słupsk region. The factors obtained are introduced into the multiepisode hazard model as explanatory variables.

The second goal of our analysis is to estimate the effect of the unemployment compensation system on the individual's unemployment duration. Duration models with time-varying covariates serve as proper tools for the analysis of the influence of the unemployment benefit received at the risk of leaving unemployment. We would also like to look at the means being used by the labor office to combat unemployment. Vocational activation of people belonging to risk groups in the local job market is very important for the labor office. Therefore, additionally, we intend to investigate the impact of vocational training on the unemployment duration.

Our research work is based on the data obtained from the District Labor Office in Słupsk from 1999 to 2007.

Description of the analysis method

Hazard models

The dependent variable we are interested in is the duration of time an individual spends in the state of being unemployed. Empirical data for the duration variable can take only positive values the negative duration periods do not exist. Moreover, duration of the phenomenon can be observed only temporarily (censoring problem). All this makes it impossible to apply traditional models of regression. An appropriate approach, which considers right censoring of unemployment spells, and which controls for observable personal characteristics of individuals that influence the unemployment duration, is the application of hazard models.

In case of continuous hazard models the duration variable T is a continuous non-negative random variable describing duration in any state, where t is realization. The distribution function of T is denoted F and is defined as $F(t) = Pr[T \leq t]$. The density function of the duration variable T is $f(t) = dF(t)/dt$. The probability of survival to t is given by the survivor function $S(t)$:

$$S(t) = Pr[T > t] = 1 - F(t)$$

The survival function $S(t)$, gives the probability of the surviving of the process over a certain moment t. The survival function lies between zero and one; it is equal to one at the beginning of the spell ($t = 0$); and its slope is non-positive.

Hazard models are concerned with observation of the instantaneous rate of leaving a certain state (e.g. unemployment) per unit time period at t:

$$h(t) = \frac{f(t)}{S(t)} = \lim_{dt \to 0} \frac{Pr\left[t \leq T < t + dt | T \geq t\right]}{dt}$$

The hazard function $h(t)$ is the limit of probability that the spell is completed during the interval $[t, t + dt]$, given that it has not been completed before time t, for $dt \to 0$. The hazard rates describe the intensity of transition from one state to another. A higher value of hazard function means that the transition from state A to state B follows faster.

For general surveys of hazard rate models – also called survival models or duration models – see, e.g. [KP], [CO], [Ki], [HL], [CT]. The first intensive application of duration models is the analysis of individual unemployment duration data by Lancaster [L]. A forecasting duration period of the unemployed people, over which they are without jobs, until taking up a job is a typical example of the use of hazard models (see, for example, [DK], [L]).

As far as the use of hazard models for unemployment duration in Poland is concerned, literature is modest. The first hazard models to analyze the time spent in the unemployment state have been applied by Frątczak, Jóźwiak, Paszek [FJP] and later by Malarska [M]. The other researchers performed microeconometric analysis of job market based on the logit or probit model. However, such a methodology fails to cover the individual duration of unemployment period. These researchers concentrated only on the detecting of determinants, which influence the probability of finding jobs.

Hazard models are constructed whenever there is the purpose of forecasting the moment, in which a certain event will occur. These models differ in assumptions concerning distribution of individual time in which the event T occurs. Among duration models, which allow to estimate the influence of different determinants, the following parametric hazard models can be noted: proportional hazard models (PH) and accelerated failure-time models (AFT).

In the PH models, the conditional hazard rate $h(t|X)$ can be factored into separate functions: $h(t|X) = h_0(t)\, g_0(X\beta) = h_0(t)\, exp(X\beta)$, where $h_0(t)$ is called the baseline hazard and $exp(X\beta)$ is a function of explanatory variables vector X. The characteristics of hazard function change proportionally to the influence of explanatory variables. This category of models comprises the whole range of models which show differences when it comes to assumptions concerning distribution of baseline hazard.

The most widely applied semiparametric method of analyzing the effect of covariates on the hazard rate is the proportional hazard model proposed by Cox [C]. The Cox model states that the hazard rate for the j-th subject in the data is $h(t|X_j) = h_0(t)\, exp(X_j\beta)$. Compared with the parametric approaches, the advantage of the semiparametric Cox model is that we have no need to make assumptions about baseline hazard; $h_0(t)$ is left unestimated. This model is particularly attractive when the researcher has only a weak theory supporting a specific parametric model and is only interested in the magnitude and direction of the effects of observed covariates.

Factor Analysis

Explanatory factor analysis (EFA) is a statistical technique for data reduction. The term factor analysis was first introduced by Thurstone [T]. EFA is used to uncover the latent structure of a set of variables. The main aim of this method is to get a small set of variables from a large set of variables (most of which are correlated to each other). The variability among observed variables will be described in terms of fewer unobserved variables called factors. All factors will be orthogonal to one another, meaning that they will be uncorrelated.

In the EFA the observed variables $X_i, i = 1, ..., p$ are modeled as linear combinations of the common factors $F_j, j = 1, ..., k\,(k < p)$, plus the unique factors u_i ("error" terms):

$$
\begin{aligned}
X_1 &= a_{11}F_1 + a_{12}F_2 + ... + a_{1k}F_k + u_1\\
X_2 &= a_{21}F_1 + a_{22}F_2 + ... + a_{2k}F_k + u_2\\
&\;\;\ldots\\
X_p &= a_{p1}F_1 + a_{p2}F_2 + ... + a_{pk}F_k + u_p
\end{aligned}
$$

where a_{ij} is the linear coefficient called the factor loading. Everything except the left-hand-side variables is to be estimated, which implies that the model has an infinite number of solutions.

EFA assumes that the variance in the measured variables can be decomposed into that accounted for by common factors and that accounted for by unique factors. EFA estimates how much of the variability is due to common factors ("communality").

We use EFA when we are interested in making statements about the factors that are responsible for a set of observed variables and when the goal of the analysis is to detect structure.

Advantages of EFA are the following: the reduction of the number of variables by combining two or more variables into a single factor and the identification of groups of inter-related variables. However, there is a disadvantage, too. Interpreting factor analysis is based on using a "heuristic". More than one interpretation of the same data can be made.

Subject of the research

In our analysis we use data taken from the District Labor Office in Słupsk in Poland concerning registered unemployed people in the period from 1999 to 2007.

Our selected sample consists of 1690 persons, who were registered unemployed in the labor office during the survey time at least for one day. They are residents of district Słupsk and the city with district of Słupsk status.

On the basis of the registered history of events in the labor office Puls computer system we can find out for how long was a person looking for a job every time or for how long unemployed is actually looking for a job (in days). The time spent in the unemployment state is called an episode. The episode finishes when the event occurs (finding a job etc.). The duration of a single episode is marked by the neighboring days, during which a given person has been in a given state. Unemployment spells are completed if they end up with transition from unemployment state. Otherwise, unemployment spells are treated as right censored. The last observed exit from unemployment was noted at the 3235 nd day. While our data basis contains multiple spells for 1690 persons we have got 3614 episodes.

Method of episode splitting

For survival data, however, individuals may be observed at several stages during an episode. and relevant time-varying regressors may take different values over an episode. Time-dependent covariates can be included in parametric and semiparametric transition rate models by applying the method of episode splitting. The idea of this method can be described as following: time-dependent qualitative covariates change their values only at discrete points in time. At all points in time, when at least one of the covariates changes its value, the original episode is split into subepisodes (splits).

The Table 1 shows, for example, that the individual with "id"=19 has had two unemployment episodes (numbers 26 and 27). The variable "des" serves as the censoring indicator ("des=0" for right censoring). The episode 26, for example, is divided into two subepisodes: the first one with starting time 2004-06-01 and the second one with tstart=2005-06-02. The symbol "R B" marks the registration of unemployed received unemployment benefit, "T U" means the change in the type of unemployment, "N PU" – new type: the unemployed without unemployment benefit, "W PC" means finding a job. The episode 27 starts with the registration of unemployed, who receives no unemployment benefit; contains three subepisodes; the second one concerns the vocational training.

Table 1. Records of data after episode splitting for one individual.

id	newid	des	type	tstart	type	tfin	tf	t1	benefit	training
19	26	0	R B	2004-06-01	T U	2005-06-01	365	365	1	0
19	26	1	N PU	2005-06-02	W PC	2005-06-22	20	385	0	0
19	27	0	R P	2006-04-12		2006-04-30	28	28	0	0
19	27	0	Z X	2006-05-01	O O	2006-05-15	14	42	0	1
19	27	1		2006-05-16	W PC	2006-05-26	10	52	0	0

For the whole data set with 3614 episodes we created 6751 subepisodes, 3445 of them were right censored.

Covariates in hazard models

Estimated hazard models will not only comprise present duration as a determinant for the probability of leaving the state of unemployment, but also other observable characteristics of individuals such as gender, age, marital status, place of residence, education level. The Table 2 defines the first part of covariates used to explain the joblessness duration with hazard models.

Table 2. Definitions of Variables.

Variable	Description	Change in time
sex	1 if individual is male	no
age	age of the individual at the beginning of a subepisode in years	yes
edu1	1 if individual has incomplete primary, primary, lower secondary or basic vocational education level	no
edu2	1 if individual has general secondary, vocational secondary or post-secondary education level	no
edu3	1 if individual has tertiary education level	no
marr	1 if individual is married	no
town	1 if the place of residence is town	no
language	1 if individual declares any foreign language skills	no
benefit	1 if individual receives unemployment benefit	yes
training	1 if individual takes part in the vocational training	yes

Below, we present the results of the empirical calculation in which the method of factor analysis was applied. The application of this method to the 12 time series from official statistics (Regional Data Bank) allowed to isolate the basic factors describing the development of labor market in Słupsk region in the time period 1999-2007. The suggested measures for the macro effects are shown in the Table 3.

Table 3. Suggested measures for the macro effects in Słupsk region.

Variable	Description	Data source
workpop	the proportion of the population at working age in % of total population	district słupski and city with district of Słupsk status
service	the proportion of the employed in the service sector in % of all persons employed	district słupski and city with district of Słupsk status
industry	the proportion of the employed in the industrial sector in % of all persons employed	district słupski and city with district of Słupsk status
offers	job offers per 1000 unemployed registered	district słupski and city with district of Słupsk status
entities	entities of the national economy recorded in the REGON register per 1000 of the population	district słupski and city with district of Słupsk status
dwellings	dwellings completed per 1000 of the population	district słupski and city with district of Słupsk status
budgreven	revenue of districts and cities with district status budgets per 1 inhabitant	district słupski and city with district of Słupsk status
unempl	the unemployment rate registered	subregion Słupsk
wages	average monthly gross wages and salaries	voivodeship Pomorskie
gdp	gross domestic product per capita	voivodeship Pomorskie
invest	investment outlays per capita	voivodeship Pomorskie
rd	expenditures on R&D per 1 inhabitant	voivodeship Pomorskie

We perform the factor analysis with principal factoring and the Equimax rotation. The variance accounted for by factors is summarized in the Table 4. In the second column (Eigenvalue), we find the variance on the new factors that were extracted. In the third column, these values are expressed as a percentage the total variance. The next column contains the cumulative variance extracted.

The Table 4 attempts to determine the number of orthogonal factors to be retained for further analysis. The Kaiser criterion for determining the number of factors, is the "eigenvalue greater than 1" criterion [Ka]. This entails that, unless a factor extracts at least as much as the equivalent of one original variable, we drop it. Since the first two factors were the only ones that had eigenvalues greater than 1, the final factor solution will represent only 87.649% of the variance in the 12 time series.

Table 4. Extraction of factors.

	Eigenvalues			Sum of square of loadings after extraction			Sum of square of loadings after rotation		
Factors	Eigen-value	% of variance	Cum. % of variance	Eigen-value	% of variance	Cum. % of variance	Eigen-value	% of variance	Cum. % of variance
1	6.838	56.984	56.984	6.838	56.984	56.984	6.648	55.397	55.397
2	3.680	30.665	87.649	3.680	30.665	87.649	3.870	32.251	87.649
3	0.736	6.133	93.781						
4	0.496	4.133	97.914						
5	0.141	1.179	99.093						
6	0.080	0.668	99.761						
7	0.022	0.186	99.947						
8	0.006	0.053	100.000						
9	0.000	0.000	100.000						
10	0.000	0.000	100.000						
11	0.000	0.000	100.000						
12	0.000	0.000	100.000						

A graphical method for determining the number of factors is the scree test. We can plot the eigenvalues shown above in a simple line plot (Figure 1).

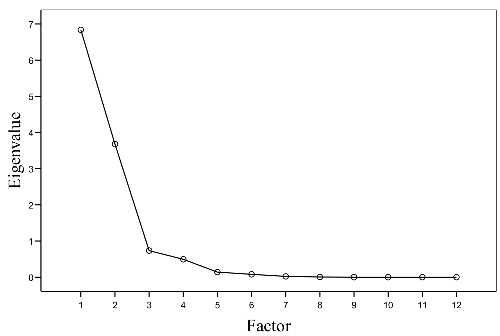


Figure 1. Plot of eigenvalues.

We have to find the place where the smooth decrease of eigenvalues appears to level off to the right of the plot. According to this criterion, we would retain two factors (principal components).

The Table 5 shows the loadings representing a correlation between the items and the factors. The loadings are distributed between the two factors. In order to obtain interpretable results the basic solution was rotated (we used Equimax rotation).

Table 5. Factor loading matrix after Equimax rotation.

Variables	Factors 1	Factors 2
gdp	**0.971**	0.234
rd	**0.959**	0.107
budgreven	**0.943**	0.278
invest	**0.927**	-0.329
wages	**0.926**	0.294
offers	**0.924**	0.088
unempl	**-0.875**	0.429
service	0.087	**0.968**
entities	0.285	**0.881**
workpop	0.566	**0.798**
industry	0.387	**-0.779**
dwellings	0.003	**0.619**

Table 6. Table of factor scores.

Variables	Factors 1	Factors 2
gdp	0.142	0.033
rd	0.144	0.000
budgreven	0.137	0.045
invest	0.152	-0.115
wages	0.134	0.050
offers	0.139	-0.004
unempl	-0.147	0.139
service	-0.016	0.253
entities	0.018	0.224
workpop	0.063	0.194
industry	0.083	-0.217
dwellings	-0.018	0.163

The Table 6 shows the factor scores for each row of the data file; we use these to obtain the values for each factor.

The first factor could be interpreted as representing the changing "level of economy development" in the whole region (voivodeship) and, the second one as a measure of changes in the "local economic activity".

As can be seen from the plots of the scores of the two factors in Figure 2, the first factor shows a trend with an increasing slope, while the "local economic activity" factor shows contrary development with downturn around 2004.

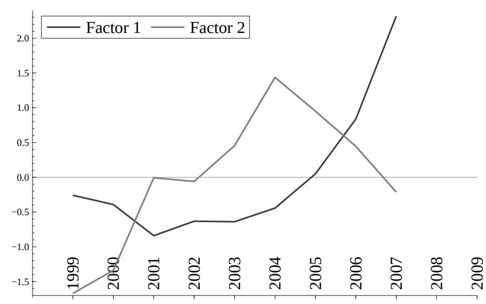


Figure 2. Plots of factor values.

Because the factors obtained are orthogonally constructed, it is possible to introduce both measures into the multiepisode hazard model as time-varying explanatory variables.

To represent the changing conditions under which cohorts enter the unemployment state, we use the factor values at the day, on which the person was registered in the labor office (variables fac1rd and fac2rd). To introduce the period effects, we use the earlier splitted episodes so that the factor values are updated in each job subepisode (variables fac1 and fac2).

Results of models estimation

Below, you can find estimated Cox proportional hazards model describing the unemployment duration. The results of the estimation of cohort and period effects and the effects of the other variables on the probability of leaving the unemployment state in the Cox model are shown in the Table 7.

Table 7. Results of Cox PH models estimation for the risk of leaving the unemployment state.

	Cox regression		
Variable	Coef.		Haz. Ratio
sex	0.490	***	1.632
age	-0.009	***	0.991
edu1	-0.194	***	0.824
edu2	-0.096		0.909
marr	0.169	***	1.184
town	0.136	***	1.146
language	0.126	***	1.134
benefit	-6.504	***	0.001
training	-0.463	**	0.629
fac1rd	-0.646	***	0.524
fac2rd	-0.504	***	0.604
fac1	0.788	***	2.199
fac2	0.566	***	1.760
	Log likelihood = -22,913.006		
	LR chi2(13) = 2,412.20		
	Prob > chi2 = 0.0000		

***,**,* - significant at 1%, 5%, 10% level respectively.
Source: own computations using Stata Statistical Software.

The semiparametric Cox model was estimated by the partial likelihood method. By interpreting the results of models parameters it can be stated that:

- the hazard of leaving the unemployment state in the case of a man is 63.2% greater than in the case of a woman,
- the age coefficient implies that older people are at a disadvantage; the one-year-older age of the individual at the beginning of a jobless subepisode leads to 0.9% decrease of chance for exiting unemployment,
- the primary, lower secondary or basic vocational education levels, in comparison with the tertiary education level, lead to the significant decrease of opportunities to break unemployment,
- the chance of leaving unemployment is greater in urban than in rural areas,
- the hazard of breaking unemployment is greater in the case of people with any foreign language skills (by 13.4%),

- there persists a lower tendency to leave the unemployment state if the person registered receives the unemployment benefit (hazard rate decreases by 99.9%),
- the participation in the vocational training significantly decreases the transition rate from unemployment.

Now we concentrate on the interpretation of the effects of changes in markets development. The parameters, which we are interested in, are statistically significant.

The "level of economy development" at the time of entry into the unemployment state has a negative effect on the chance of finding a job. The higher the "level of economy development", the smaller the attractiveness of the people newly registered in the labor office. It is less likely that these people will be further moved from unemployment. It is also true for the negative effect of the "local economic activity" level at the entry into jobless state.

Conversely, the period effect of the "economy development" is positive. The continuous developing of the regional economy leads to the increasing opportunities for all people to move up to the job. The same is true for the period effect of the "local economic activity" level; the better this level is, the more current opportunities for unemployed people to find a job.

Conclusions

Using the multiepisode hazard models with time-varying variables, we investigated which factors influence the probability of leaving the unemployment state among the people registered in the District Labor Office in Słupsk.

We introduced the changing labor market structure into the risk model in order to treat the time-dependent nature of the unemployment duration process in an adequate manner. The main results achieved show that the job finding process depends on the historical time of the entry into the unemployment state and the actual time. We found that the "level of economy development" in the whole region and the "local economic activity" level at the time of entry into the unemployment state have a negative effect on the chance of finding a job, while the period effects are positive.

Having estimated the Cox PH model for the risk of leaving the unemployment state, we can take into account three future scenarios for the regional and local economy development. The optimistic scenario A assumes a continuous development of the regional and local economy, the worst-case scenario B reflects the decline in development at the regional and local level and the base scenario C assumes a continuation of existing tendencies. The continuous development (scenario A) leads to the increasing opportunities for all registered unemployed to move up to the job. The reduction of development (scenario B) leads to fewer opportunities to find a job. In scenario C, the effects of increasing development at the voivodeship level and decline in local economic activity may be reduced.

Apart from that, the specific individual characteristics of people unemployed, such as gender, age, marital status, place of residence, education level, influence the probability of exiting the unemployment state.

The examination of the impact of the unemployment benefit received exhibits, that there is a greater tendency to leave the unemployment state when the person doesn't receive the unemployment benefit. What is more, the actual participation in the vocational training doesn't increase the transition rate into employment.

Acknowledgments

This work was supported by grant No. N N111 209436 from Polish Ministry of Science and Higher Education.

References

[B] H. P. Blossfeld, Career opportunities in the Federal Republic of Germany: A dynamic approach to the study of life-course, cohort, and period effects, European Sociological Review 2 (1986), 208–225.

[BGR] H. P. Blossfeld, K. Golsch K. and G. Rohwer, Event History Analysis with Stata, Lawrence Erlbaum Associates, Inc., Mahwah, New Jersey, 2007.

[CT] A. C. Cameron and P.K. Trivedi, Microeconometrics: Methods and Applications, Cambridge University Press, New York, 2005.

[C] D. R. Cox, Regression Models and Life Tables (with Discussion), Journal of the Royal Statistical Society, Series B 34 (1972).

[CO] D. R. Cox and D. Oakes, Analysis of Survival Data, Chapman and Hall, London, 1972.

[DK] D. Devine and N. Kiefer, Empirical Labor Economics, New York, Oxford, 1991.

[FJP] E. Frątczak, J. Jóźwiak and B. Paszek, Zastosowanie analizy historii zdarzeń w demografii, Szkoła Główna Handlowa, Warszawa, 1996.

[HL] D. Hosmer and S. Lemeshow, Applied Survival Analysis: Regression Modeling of Time to Event Data, John Wiley and Sons, New York, 1999.

[Ka] H. F. Kaiser, The application of electronic computers to factor analysis, Educational and Psychological Measurement 20 (1960), 141–151.

[KP] J. Kalbfleisch and R. Prentice, The Statistical Analysis of Failure Time Data, John Wiley and Sons, New York, 1980.

[Ki] N. Kiefer, Economic Duration Data and Hazard Functions, Journal of Economic Literature 26/2 (1988), 646–679

[L] T. Lancaster, Econometric Methods for the Duration of Unemployment, Econometrica 47/4 (1979), 939–956 .

[M] A. Malarska, Diagnozowanie determinantów bezrobocia w Polsce nieklasycznymi metodami statystycznymi, Wydawnictwo Biblioteka, Łódź, 2007.

[T] L. L. Thurstone, Multiple factor analysis, Psychological Review 38 (1931), 406–427.

Quantitative Methods in Economics
Vol. XI, No. 1, 2010, pp. 223–253

BUYING AND SELLING PRICE FOR RISKY LOTTERIES AND EXPECTED UTILITY THEORY WITHOUT CONSEQUENTIALISM

Michał Lewandowski

EUROPEAN UNIVERSITY INSTITUTE, Florence, Italy
e-mail: michal.lewandowski@eui.eu

Abstract: In this paper I show that within expected utility large buying and selling price gap is possible and [R] paradox may be resolved if only initial wealth is allowed to be small. It implies giving up the doctrine of consequentialism which may be reduced to requiring initial wealth to be total lifetime wealth of the decision maker. Still, even when initial wealth is allowed to be small and interpreted narrowly as gambling wealth, classic preference reversal is not possible within expected utility. I show that only another kind of reversal which I call preference reversal B is possible within expected utility. Preference reversal B occurs when buying price for one lottery is higher than for another, but the latter lottery is chosen in a direct choice. I demonstrate that classic preference reversal is susceptible to arbitrage whereas preference reversal B is not which suggests that the latter reversal is more rational.

Key words: expected utility, consequentialism, total wealth, gambling wealth, narrow framing, [R] paradox, preference reversal, WTA/WTP disparity, buying and selling price for a lottery

Introduction

Willingness-to-accept or selling price for a lottery is a minimal sure amount of money which a person is willing to accept to forego the lottery. Willingness-to-pay or buying price for a lottery on the other hand is a maximal sure amount of money which a person is willing to pay in order to play the lottery. The disparity between willingness to pay (WTP) and willingness to accept (WTA) is a well-known phenomenon that arises in experimental settings. There is a large body of evidence starting with [KS] and [T] that WTA is much higher than WTP for many types of goods. [H] is a survey which documents and analyzes results from a great number of experiments and obtains mean values of the WTA/WTP ratio for different goods. Although the mean WTA/WTP ratio for lotteries is 2.10 and it is small as compared to the same ratio for other, especially non-market goods, it means nevertheless that WTA is on average more than two times higher than WTP. There have been many attempts to give account for this phenomenon.

There is strong belief in the literature that this evidence is not consistent with expected utility theory. Along the lines of [R06] I will argue that the source of this belief lies in associating expected utility theory with the doctrine of consequentialism, according to which "the decision maker makes all decisions having in mind a preference relation over the same set of final consequences". This association is harmless when considering Constant Absolute Risk Aversion, as in this case decisions whether to accept a given lottery do not depend on wealth. However as many studies confirm people usually exhibit Decreasing Absolute Risk Aversion[1], in which case wealth effects are present.

In practice the doctrine of consequentialism means that the initial wealth underlying any decision whether to accept or reject a given lottery is assumed to be the decision maker's lifetime wealth. It follows that most lotteries under consideration are small relative to initial wealth and therefore, by [R] argument for any reasonable level of risk aversion expected utility predicts approximate risk neutrality towards such lotteries. In this case, not only is expected utility incapable of accommodating large spreads between buying and selling price, but also it is inconsistent with risk averse behavior for small gambles[2]. Instead of burying expected utility theory I propose to divorce it from the doctrine of consequentialism, i.e. relax the assumption that initial wealth underlying any decision whether to accept a gamble is total lifetime wealth of the decision maker. If initial wealth is allowed to be small, I will show that expected utility is consistent with large buying/selling price spread, i.e. that within expected utility for reasonable[3] levels of risk aversion one can obtain buying/selling price spread of the magnitude consistent with experimental results. Following this finding I will propose an alternative for consequentialism involving narrow framing. Instead of asserting that preferences are always defined over total lifetime wealth, I will assume that preferences over gambling are defined over gambling wealth, i.e this part of the decision maker's total wealth which he designates for taking gambles. The idea is taken from [FH], although the seeds of this approach, and in particular the idea of separating lifetime wealth and something else for different decision problems, are already in [R06]. I will propose several ways for testing the hypothesis of gambling wealth.

There are many papers on the disparity between willingness to accept and willingness to pay for risky lotteries. It is part of a vast literature stream on WTA and WTP valuations in general. For example, [SSS] explain WTA/WTP spread for risky lotteries using prospect theory. They propose the third-generation prospect theory, in which, unlike in the previous versions, reference point is allowed to be random. They show

[1] In this paper decreasing absolute risk aversion means strictly decreasing absolute risk aversion.
[2] "Small" here means "small relative to lifetime wealth".
[3] Consistent with experimental evidence.

that loss aversion in such model implies positive WTA/WTP gap[4]. In general, there have been many accounts for the disparity based on non-expected utility models. My aim in this paper is not to offer a better explanation. I am even convinced that specific behavioral theories will fit empirical and experimental evidence better than expected utility model which I analyze. My goal is to show, that large spreads between WTA and WTP, due entirely to wealth effects, are possible within expected utility if only wealth is interpreted narrowly as gambling wealth. The advantage of this approach is that expected utility has stronger normative appeal as compared to many behavioral models. And hence, it is useful to know that certain patterns of preferences, or in this case valuations, can be accommodated not only within behavioral models but also within expected utility.

The approach I take in this paper in general is not novel. As I mentioned before, [R06] claims that a lot of recent confusion around expected utility, which led some researchers to question it as a descriptive theory is caused by associating expected utility theory with the assumption of consequentialism - the idea that there is a single preference relation over the set of lotteries with prizes being the "final wealth levels" such that the decision maker at any wealth level W who has vNM preference relation $\succsim_W$ over the set of "wealth changes" derives that preference from $\succsim$ by $L_1 \succsim_W L_2 \iff W + L_1 \succsim W + L_2$, where L_1 and L_2 are lotteries. Also, [CS] argue that the confusion around expected utility in general, and Rabin's paradox in particular, is caused by the failure in the literature to distinguish between expected utility theories, which stands for all models based on a set of axioms among which there is independence axiom, and a specific expected utility model. They show on the basis on Rabin's argument that the expected utility of income model is capable of accommodating evidence which the expected utility of terminal model cannot accommodate. Finally, [P-HS] show that in case of Rabin's paradox, it is the assumption of rejecting small gambles over a large range of wealth levels, and not expected utility, that does not match real-world behavior. For more discussion on Rabin's paradox, see section .4.

My approach in this paper follows the lines of the aforementioned articles. The difference is that these articles focus on Rabin's paradox and I focus here on buying/selling price or WTA/WTP spread.

Related to buying/selling price disparity is the issue of preference reversal analyzed by [GP]. There are two lotteries called the \$-bet and the P-bet both of which promise some prize with some probability and nothing otherwise such that the probability of winning is higher for the P-bet but the prize is bigger for the \$-bet. Preference reversal occurs when selling price for the \$-bet is higher than that for the P-bet but the P-bet is preferred to the \$-bet in a direct choice[5]. A related possibility, which I call preference reversal B occurs when buying price for the P-bet is higher than that for the \$-bet and yet the \$-bet is chosen over the P-bet in a direct choice. I will show that traditional preference reversal is susceptible to arbitrage and is not possible within

[4]In fact, by imposing some symmetry conditions on prospect theory utility function in their model, it is possible to show that loss aversion is equivalent to positive WTA/WTP gap

[5]Experimentally, in order to confirm preference reversal one must show that the asymmetry described above occurs more often than the opposite kind of asymmetry, i.e. when the \$-bet is preferred in a direct choice but the P-bet gets higher selling price.

expected utility, whereas preference reversal B is possible within expected utility and it does not allow arbitrage. This result may suggest that traditional preference reversal is less rational that preference reversal B.
Buying and selling price for a lottery are the concepts introduced by [Rai] in the context of expected utility. More popular perhaps are the terms willingness to pay (WTP) and willingness to accept (WTA), respectively, the terms introduced primarily in the context of non-expected utility theories. Except for the fact that buying and selling price terms were introduced in a different context than WTP and WTA, these terms have the same meaning. Since I focus on expected utility model I will henceforth use the former terms.
The structure of this paper is as follows. First, I introduce the model, its assumptions, definitions of buying and selling price for a lottery and buying/selling price reversal. Then I state a couple of technical propositions which describe the shape and properties of buying and selling price for a lottery for different risk attitudes. The subsequent section contains the main theses of the paper. Focusing on constant relative risk aversion class of utility functions, I demonstrate first that expected utility with consequentialism is likely to predict risk neutral behavior towards most gambles and eventually a gap between buying and selling price becomes negligible. Second, I demonstrate that if the doctrine of consequentialism is abandoned and wealth is allowed to move over the whole domain, significant spreads between buying and selling price are possible due to income effects when wealth is sufficiently small. As a next step, I propose an alternative to consequentialism involving narrow framing. Instead of defining wealth as total lifetime wealth of the decision maker I suggest to use gambling wealth which is that part of the decision maker total wealth which he designates for the purpose of taking gambles. I discuss ways to test gambling wealth hypothesis and then I examine the possibility of what I call preference reversal B which I compare to the related concept of traditional preference reversal. I show that whereas preference reversal B is possible within expected utility framework with gambling wealth instead of total lifetime wealth and it does not allow arbitrage opportunities, preference reversal allows arbitrage opportunities and is not possible within expected utility. And finally I conclude. The appendix at the end of this paper contains proofs of the propositions.

The model

I start with basic assumptions and definitions.

Assumption 1. Preferences obey expected utility axioms. Bernoulli utility function $U : \mathbb{R} \to \mathbb{R}$ is twice continuously differentiable, strictly increasing and strictly concave.

Definition 2. A lottery $\mathbf{x}$ is a real- and finite-valued random variable with finite support. The space of all lotteries will be denoted $\mathcal{X}$. I define the maximal loss of lottery $\mathbf{x}$ as: $\min(\mathbf{x}) = \min \operatorname{supp}(\mathbf{x})$.

The typical lottery will be denoted as $\mathbf{x} \equiv (x_1, p_1; ...; x_n, p_n)$, where $x_i \in \mathbb{R},\ i \in \{1, 2, ..., n\}$ are outcomes and $p_i \in [0, 1]$ $i \in \{1, 2, ..., n\}$ the corresponding probabilities. Outcomes should be interpreted here as monetary values. Although most results that follow are true for more general lotteries, the finite support assumption is sufficient for the purposes of this paper. Now I define buying and selling price for a lottery given wealth level along the lines of [Rai]. To avoid repetitions, I will henceforth skip statements of the form: "Given utility function U satisfying assumption 1, any lottery $\mathbf{x}$ and wealth W...".

Definition 3. I define selling price and buying price for a lottery $\mathbf{x}$ at wealth W as functions denoted, respectively, $S(W, \mathbf{x})$ and $B(W, \mathbf{x})$. Provided that they exist, values of these functions will be determined by the following equations:

$$\mathrm{E}U[W + \mathbf{x}] = U[W + S(W, \mathbf{x})] \tag{1}$$

$$\mathrm{E}U[W + \mathbf{x} - B(W, \mathbf{x})] = U(W) \tag{2}$$

If utility function is defined over the whole real line as is the case for constant absolute risk aversion, buying and selling price as functions of wealth exists for any wealth level by assumption 1. If the domain of utility function is restricted to a part of real line as is the case of constant relative risk aversion utility function analyzed here, I will specify later on in the paper on which domain buying and selling price are defined as functions of wealth.

In economic terms, given an individual with initial wealth W whose preferences are represented by utility function $U(\cdot)$, $S(W, \mathbf{x})$ is the minimal amount of money which he demands for giving up lottery $\mathbf{x}$. Similarly, $B(W, \mathbf{x})$ is the maximal amount of money which he is willing to pay in order to play lottery $\mathbf{x}$. Additionally I define a concept of buying/selling price reversal.

Definition 4. Given two lotteries $\mathbf{x}$ and $\mathbf{y}$ and some wealth level W, define buying/selling price reversal as:

$$S(W, \mathbf{y}) > S(W, \mathbf{x}) \quad \text{and} \quad B(W, \mathbf{x}) > B(W, \mathbf{y})$$

This kind of preference pattern may be interpreted as follows. For a given initial wealth, an individual's certainty equivalent for lottery $\mathbf{y}$ is higher than for lottery $\mathbf{x}$, and yet he is willing to pay more to play lottery $\mathbf{x}$ than to play lottery $\mathbf{y}$. In other words, an individual exhibiting buying/selling price reversal, may prefer to buy $\mathbf{x}$ than $\mathbf{y}$ if he does not play any lottery initially. When, on the other hand, he does play the lottery initially, he would prefer to sell $\mathbf{x}$ than $\mathbf{y}$.

Buying short and selling short price for a lottery

It is possible to introduce buying short and selling short for a lottery $\mathbf{x}$ at wealth level W denoted, respectively, by $B^S(W, \mathbf{x})$ and $S^S(W, \mathbf{x})$. They satisfy the following equations:

$$\mathrm{E}U[W - \mathbf{x}] = U[W - B^S(W, \mathbf{x})] \tag{3}$$

$$\mathrm{E}U[W - \mathbf{x} + S_S(W, \mathbf{x})] = U(W) \tag{4}$$

The interpretation of these two measures is the following: $B^S(W, \mathbf{x})$ is the maximal sure amount of money which an individual would pay for not taking a short position

in lottery $\mathbf{x}$. In other words if initial position is $W - \mathbf{x}$, $B^S(W, \mathbf{x})$ is the maximal sure amount of money which an individual is willing to pay for $\mathbf{x}$. On the other hand $S^S(W, \mathbf{x})$ is the minimal sure amount of money which an individual would accept for taking a short position in $\mathbf{x}$. In other words, it is the minimal selling price for a lottery which an individual does not have initially.
Notice that buying price $B(W, \mathbf{x})$ and $S^S(W, \mathbf{x})$ are evaluated with respect to the same initial position W. Using Jensen's inequality it is easy to show that for strictly concave utility function and a nondegenerate lottery $\mathbf{x}$:

$$S^S(W, \mathbf{x}), B^S(W, \mathbf{x}) \in (\mathrm{E}[\mathbf{x}], \max(\mathbf{x}))$$

where $\max(\mathbf{x})$ denotes the maximal consequence in the support of lottery $\mathbf{x}$. As shown in proposition 1 below, classical buying and selling price for a lottery are on the other hand strictly in between $\min(\mathbf{x})$ and $\mathrm{E}[\mathbf{x}]$ for strictly concave utility function. Hence, for strictly concave utility function, both $S^S(W, \mathbf{x})$ and $B^S(W, \mathbf{x})$ are strictly greater than $S(W, \mathbf{x})$ and $B(W, \mathbf{x})$ for any wealth level.
Certain global (i.e. holding for any lottery) properties of buying and selling price as functions of wealth are "mirrored" by the corresponding global properties of buying short and selling short prices for a lottery. This is due to the simple relation which holds between these measures and which is the following:

$$\begin{aligned} S^S(W, \mathbf{x}) &= -B(W, -\mathbf{x}) \\ B^S(W, \mathbf{x}) &= -S(W, -\mathbf{x}) \end{aligned}$$

So, if for example buying and selling price for any no-degenerate lottery are strictly concave and strictly increasing in W as is the case for CRRA utility function[6], then selling short and buying short prices for any non-degenerate lottery will be strictly decreasing and strictly convex in W. If $0 < B(W, \mathbf{x}) < S(W, \mathbf{x})$ as is the case for DARA[7], then $0 < S^S(W, \mathbf{x}) < B^S(W, \mathbf{x})$.

Preliminary results

Before introducing the main point of this paper I need a couple of theoretical results which describe properties of buying and selling price for a lottery for different risk attitudes. The most basic property of buying and selling price which is true for any concave strictly increasing utility function is the following:

Proposition 1 (Concave)**.** *For any non-degenerate lottery* $\mathbf{x}$ *and any wealth* W *such that buying and selling price exist,* $S(W, \mathbf{x})$ *and* $B(W, \mathbf{x})$ *lie in the interval* $(\min(\mathbf{x}), \mathbf{E}(\mathbf{x}))$. *For a degenerate lottery* $\mathbf{x}$, $S(W, \mathbf{x}) = B(W, \mathbf{x}) = x$.

Proof. In the appendix. ∎

Below I state propositions which characterize constant and decreasing absolute risk aversion utility functions in terms of buying and selling price. Proofs of these propositions may be found for example in [ML]. Also I refer to [ML] for an extensive

[6] See results in [ML].
[7] See results in [ML].

discussion on multiplicative and nominal gambles, risk aversion notions for the two kinds of gambles, etc.

Proposition 2 (CARA)**.** *The following two statements are equivalent:*

i. Bernoulli utility function exhibits CARA

ii. Buying and selling price are independent from wealth and equal i.e.

$$B(W, \mathbf{x}) = S(W, \mathbf{x}) = C_\alpha, \ \forall W$$

where α is absolute risk aversion coefficient and C_α takes real values and depends only on α.

Proposition 3 (DARA)**.** *The following two statements are equivalent:*

i. Bernoulli utility function exhibits DARA

ii. buying and selling price are increasing in W

$$B(W, \mathbf{x}) > 0 \iff B(W, \mathbf{x}) < S(W, \mathbf{x})$$

for a non-degenerate lottery $\mathbf{x}$.

The above propositions show that in expected utility model a gap between buying and selling price can only arise due to wealth effects. Selling price is higher than buying price for a lottery for which I would be willing to pay positive amount only if absolute risk aversion decreases in wealth. Since I want to focus on CRRA utility functions which is a subclass of DARA utility functions I will additionally state one more proposition, the proof of which may also be found in [ML].

Proposition 4 (CRRA)**.** *The following two statements are equivalent:*

i. Bernoulli utility function exhibits CRRA

ii. buying and selling price for any lottery are homogeneous of degree one i.e.

$$\begin{aligned} S(\lambda W, \lambda \mathbf{x}) &= \lambda S(W, \mathbf{x}), \ \forall \lambda > 0 \\ B(\lambda W, \lambda \mathbf{x}) &= \lambda B(W, \mathbf{x}), \ \forall \lambda > 0 \end{aligned}$$

Buying/selling price spread within expected utility framework

In this section I focus on constant relative risk aversion utility class, since it is simple and empirically well validated. For convenience but without loss of generality I normalize Bernoulli utility function as follows:

$$U_\alpha(x) = \begin{cases} \frac{x^{1-\alpha}-1}{1-\alpha}, & 1 \neq \alpha > 0, \quad x > 0 \\ \log x, & \alpha = 1, \quad x > 0 \end{cases} \tag{5}$$

Parameter α is required to be bounded. I also focus on non-degenerate lotteries with non-negative values such that outcome zero gets positive probability. This restriction is a matter of convenience as the forthcoming results extend to the case of general lotteries. The following proposition is necessary to establish the domain and the range of buying and selling price for a lottery as functions of wealth for the case of CRRA functions of the above form. Before I state this proposition a couple of remarks might be useful. First, since CRRA utility function used in this section is defined only for positive real numbers I need to be sure that both sides of equations (2) and (1) defining buying and selling price are well defined. Second, notice that CRRA function of the above form is unbounded from below for $\alpha \geq 1$ and bounded from below for $0 < \alpha < 1$. This is the reason why for $0 < \alpha < 1$ the infimum of $B(W, \mathbf{x})$ and $S(W, \mathbf{x})$ cannot be equal to $\min(\mathbf{x})$, the lower bound given in proposition 1. It turns out that there is a certain threshold denoted by $W_L(\mathbf{x}) \in (0, \mathrm{E}[\mathbf{x}])$ such that the infimum of $B(W, \mathbf{x})$ and $S(W, \mathbf{x})$ is equal to $W_L(\mathbf{x}) + \min(\mathbf{x})$ which is greater than $\min(\mathbf{x})$.

Proposition 5 (CRRA2)**.** *Given the class of CRRA utility function of the form given by (5) the following holds for any non-degenerate lottery* $\mathbf{x}$*: for* $\alpha \geq 1$

- $\lim_{W \to 0} B(W, \mathbf{x}) = \min(\mathbf{x})$
- $\lim_{W \to -\min(\mathbf{x})} S(W, \mathbf{x}) = \min(\mathbf{x})$

Define $W_L(\mathbf{x}) = U^{-1}[\mathrm{E}U(-\min(\mathbf{x}) + \mathbf{x})]$*. For* $0 < \alpha < 1$

- $\lim_{W \to W_L(\mathbf{x})} B(W, \mathbf{x}) = W_L(\mathbf{x}) + \min(\mathbf{x})$,
- $\lim_{W \to -\min(\mathbf{x})} S(W, \mathbf{x}) = W_L(\mathbf{x}) + \min(\mathbf{x})$

Additionally,

$$\forall \alpha > 0 \quad \lim_{W \to \infty} B(W, \mathbf{x}) = \lim_{W \to \infty} S(W, \mathbf{x}) = \mathrm{E}[\mathbf{x}] \tag{6}$$

Proof. In the appendix. ■

The above proposition establishes the domain and the range of buying and selling price for a given lottery $\mathbf{x}$ as functions of wealth for CRRA utility functions which are defined above. Now that I introduced the necessary theoretical results, I proceed to the main message of this paper.

Expected utility and consequentialism

Consequentialism is a doctrine that says that an individual makes all decisions according to a preference relation defined over one set of final consequences. In practice it means that initial wealth taken into account when making whatever decision is interpreted as the decision maker's total lifetime wealth. Most lotteries which a person may encounter are small relative to his lifetime wealth. Especially, lotteries used in experiments have values which are small relative to total lifetime wealth of experimental subjects. Therefore to explain certain experimental results it is sufficient to focus on lotteries that have values which are negligible as compared to total lifetime wealth. To represent this fact I assert here that lotteries have bounded values and consequentialism approximately means that wealth tends to infinity. In this case the following result holds:

Proposition 6. *Expected utility with consequentialism and CRRA approximately predicts no buying/selling price spread and risk neutrality.*

Proof. The proof follows directly from equation (6) in proposition 5. To represent the fact that most lotteries are small relative to lifetime wealth, I take any lottery with bounded values and let wealth go to infinity. What happens is that both selling price and buying price tend to $\mathrm{E}[\mathbf{x}]$ and hence the gap between them vanishes. Since the distance $\mathrm{E}[\mathbf{x}] - S(W, \mathbf{x})$ measures risk aversion, it is clear that there is no risk aversion either. ■

This proposition is very similar to [R] calibration theorem confined to CRRA class of utility functions. Reasonable levels of risk aversion for big gambles give rise to risk neutral behavior towards small gambles within expected utility with consequentialism. The difference between [R] argument is that I claim after [R09] that this is due to consequentialism and not due to expected utility.
This negative result immediately rises the issue of what happens if I drop the assumption of consequentialism. To answer this question I proceed in two steps. First, I show that relaxing consequentialism is promising, i.e. large buying/selling price for a lottery for reasonable levels of risk aversion may be obtained. Second, I propose an alternative assumption which could replace the assumption of consequentialism.
In the first step I allow wealth to vary freely. I will therefore analyze buying and selling price for a lottery as functions of wealth. The goal is to see for what values of wealth is the spread between buying and selling price likely to be high. To save on notation, given a fixed lottery $\mathbf{x}$ I shall write $S(W, \mathbf{x}) = S(W)$ and $B(W, \mathbf{x}) = B(W)$. I define relative spread between buying and selling price as follows:

$$\tau(W) = \frac{S(W) - B(W)}{B(W)}$$

The following lemma can be used to infer certain properties of the relative gap between buying and selling price.

Lemma 5. *For differentiable decreasing absolute risk aversion utility function, given any non-degenerate lottery* $\mathbf{x}$ *and any wealth level* W*, the following holds:*

- $B'(W) < 1$
- $S'(W - B(W)) = \frac{B'(W)}{1-B'(W)}$ *and hence* $S'(W - B(W)) > B'(W)$
- $B'(W + S(W)) = \frac{S'(W)}{1+S'(W)}$ *and hence* $B'(W + S(W)) < S'(W)$
- $S'(W) = \frac{S(W)-B(W)}{B(W)}$ *and* $B'(W) = \frac{S(W)-B(W)}{S(W)}$ *for small positive* $S(W)$

Proof. In the appendix. ■

Observe that the slope of buying price is always smaller than one whereas the slope of selling price can be higher for small values of wealth. Before I state a proposition describing the characteristics of the relative gap between buying and selling price I need the following lemma:

Lemma 6. *For CRRA utility function, given any non-degenerate lottery* $\mathbf{x}$*,* $S(W)$ *and* $B(W)$ *are concave functions.*

Proof. See [ML]. ∎

I focus now on the case when $S(W) > B(W) > 0$. The remaining cases can be analyzed similarly. By proposition 5, to make sure that $B(W)$ is positive I require that $\min(\mathbf{x})$ cannot be lower than zero. The following proposition suggests that for CRRA utility function the lower the wealth the higher the relative gap between buying and selling price.

Proposition 7. *For CRRA utility function and any lottery* $\mathbf{x}$ *with* $\min(\mathbf{x}) \geq 0$*, the relative gap between buying and selling price* $\tau(W)$ *is strictly decreasing in* W*.*

Proof. In the appendix. ∎

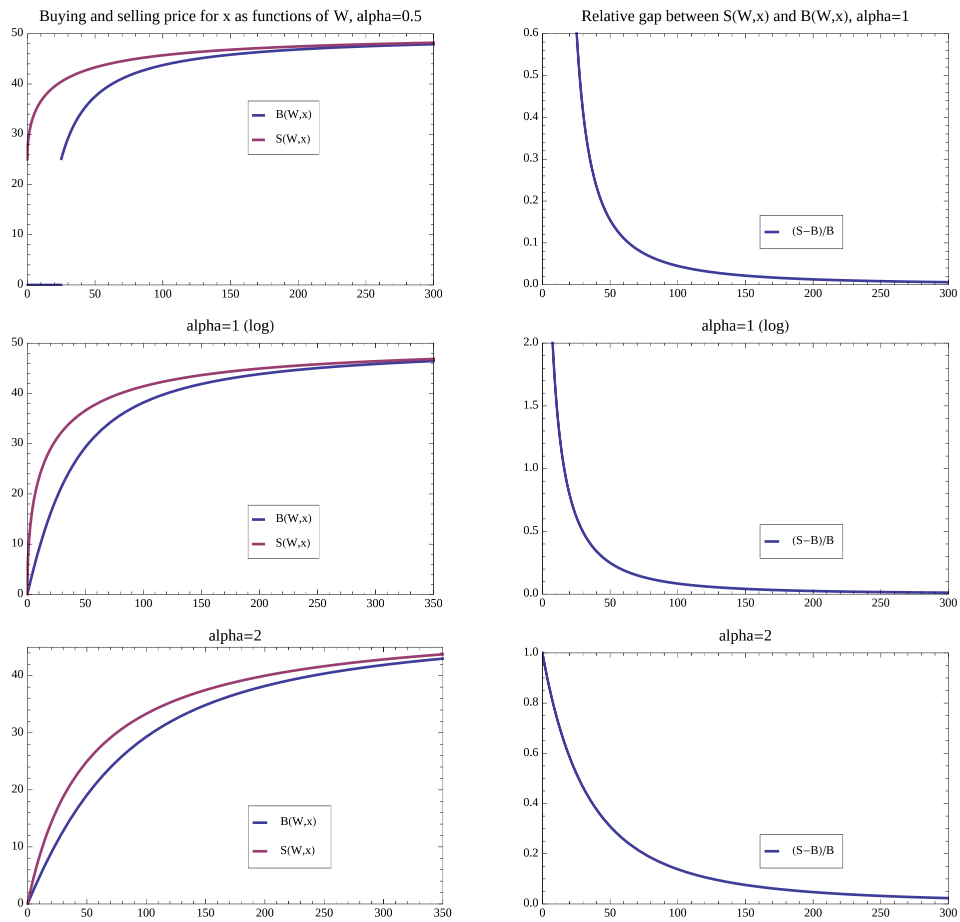
This proposition already gives an explanation of why buying/selling price gap cannot be predicted within expected utility with consequentialism for small experimental lotteries. The reason is that within expected utility, the gap between buying and selling price is the highest for small values of wealth. So if initial wealth is small, expected utility model can accommodate large buying and selling price gap. Obviously, assuming initial wealth to be total lifetime wealth of the decision maker is as far as one can go away from this possibility.
Using lemma 5 and proposition 7 it is possible to infer certain properties of buying and selling price when data on relative gap between buying and selling price is available. Also, in the opposite direction, it is possible to infer properties of the relative gap between buying and selling price when certain properties of buying and selling price are known. Here, I mention just a couple of possibilities:

- the gap is equal to the slope of selling price for small $B(W)$
- for small values of $B(W)$ the gap is equal to $\frac{B'(W)}{1-B'(W)}$ and hence
- the maximal gap depends on the slope of $B(W)$ for small values of $B(W)$

The above mathematical results can be best illustrated on the basis of an example. Let $\mathbf{x}$ be a lottery giving 100 euros or nothing with equal probabilities. The notation I use for such a lottery is $(100, \frac{1}{2}; 0, \frac{1}{2})$. Table 1 contains graphs of selling and buying price for lottery $\mathbf{x}$ on the left and relative spread between them as functions of wealth W on the right, each of them for CRRA utility function for three different coefficients of relative risk aversion: 1/2, 1 and 2[8]. Notice that as stated in propositions above

[8]The CRRA utility function is of the form given in (5).

Table 1. Buying/selling price spread for $\mathbf{x}$ for CRRA utility function



buying and selling price are between $\min(\mathbf{x})$ and $\mathrm{E}[\mathbf{x}]$ for $\alpha = 1$ and $\alpha = 2$. For $\alpha = 0.5$ I can calculate $W_L(\mathbf{x})$ as follows:

$$W_L(\mathbf{x}) = \left(\frac{1}{2}\sqrt{100} + \frac{1}{2}\sqrt{0}\right)^2 = 25$$

Hence buying and selling price for $\alpha = 0.5$ are indeed between $W_L(\mathbf{x}) + \min(\mathbf{x})$ and $\mathrm{E}[\mathbf{x}]$. Notice also that buying and selling price are increasing and strictly concave in wealth and that selling price is higher than buying price over the whole domain of buying and selling price. Finally as stated in proposition 7 the relative gap indeed is the highest for the minimal value of wealth for which both buying and selling price

are defined.
As illustrated by this simple example and stated formally in the propositions, the smaller the wealth the greater the relative gap between buying and selling price. So if wealth is small enough it is possible to obtain the gap between buying and selling price consistent with experimental evidence for reasonable levels of risk aversion. I will summarize this finding in a proposition.

Proposition 8. *For levels of risk aversion which are consistent with experimental evidence on risk attitudes there exists levels of wealth such that the expected utility model predicts high relative gap between buying and selling price.*

To illustrate the proposition consider again the above example. For instance, to obtain selling price 30 per cent higher than buying price for the lottery in consideration and for different relative risk aversion coefficients I need wealth levels which are listed in table 2.

Table 2. Selling price 30% higher than buying price

α	W
0.5	35.15
1	43.94
2	51.57

For example to obtain selling price 30% higher than buying price for the lottery $(100, \frac{1}{2}; 0, \frac{1}{2})$ for logarithmic utility function, initial wealth level of almost 44 is necessary. In the next subsection I introduce gambling wealth. If one believes that expected utility model accurately predicts behavior 44 would correspond to the calibrated gambling wealth.
Assuming that the decision maker exhibits constant relative risk aversion, one can calibrate pairs of wealth and relative risk aversion consistent with any given level of relative gap between selling and buying price for a given lottery.

Expected utility with gambling wealth

I have argued above that expected utility with total wealth interpretation of wealth predicts no gap between buying and selling price and risk neutrality for a wide range of gambles used in experiments. On the other hand I have shown that if small values of wealth are possible one can obtain large gaps between buying and selling price for a lottery for reasonable levels of risk aversion. One way to proceed would be to make wealth a free parameter of the model. Then, if one believes that expected utility is a good descriptive model of behavior, then given the data on risky choices one can calibrate which pairs of risk attitude and wealth level are consistent with the data, as I have illustrated in table 2. Unfortunately, by making wealth a free parameter, the model loses much of its predictive power. In particular, it is harder to falsify the model or design testable predictions. Another way to proceed is to give wealth a new interpretation or, even better, to develop a theory of endogenous wealth determination and then to test whether this new interpretation gives better answers than consequentialist interpretation. Since at this point I am unable to offer a theory of endogenous wealth determination, I will only propose a new interpretation of wealth and ways to test it.

.1 Gambling wealth

Consequentialism assumption implies that when making any kind of decision people consider and have in mind their lifetime wealth. I think a good alternative assumption is that people frame decisions narrowly and separate them into categories. When they engage themselves in housing decisions they think about housing budget, when they consume they think about consumption budget and when they consider gambling or whether to accept of reject an offered gamble, they consider gambling budget. Of course, personal assignment of different categories, budgets for them and time span for the budgets is a very complex subject and certainly there is plenty of factors which influence such decisions. Therefore I do not aim at a theory of endogenous budget determination. For the purposes of this paper I focus only on gambling category and a budget assigned to it, which I call gambling wealth. Gambling wealth was proposed informally by [FH]. They define gambling wealth as that part of total wealth designated only for taking gambles. Alternatively, if $\underline{W}$ is wealth designated for the purposes of living, housing and consumption, then gambling wealth is what is left over.
In the light of the results from previous subsection, one can argue that the idea of gambling wealth and more generally, the idea of separate budgets for different categories of decisions could explain a number of interesting phenomena, for example:

- Agents who gamble more, have higher gambling wealth and therefore buying and selling price gap for a given lottery is smaller than for less experienced individuals
- If an object is treated narrowly the disparity should be higher; if it is integrated into a wider set of objects the disparity should decrease ([H])
- The disparity should also be higher for artificial environments such as experiments than for a real market place.

This approach also has potential of explaining why buying/selling price gap is more pronounced when objects of choice are not monetary, e.g. coffee mugs. The more specific or narrowly defined is the object of choice the more pronounced are wealth or income effects since the value of the object is comparable with the money designated for taking such objects.
The attractive feature of all these explanations is that they are all within expected utility framework. The only novel thing is narrow framing with which expected utility model is supplemented. Naturally, a theory of endogenous wealth determination would be much appreciated to make this kind of explanations fully testable. At this point, I may suggest a couple of ways to test gambling wealth hypothesis.

I propose the following experiment design which could shed light on the validity of this approach. The first stage of experiment is to give people small amount of money for trading in gambles and then to elicit buying price and selling price for a given lottery. It is possible to use sealed bid second price auction to elicit the true buying price and [BDGM] procedure to elicit the true selling price for a lottery. In the second stage subjects are given more money for trading in gambles and again buying and selling price is elicited. Alternatively, instead of giving the subjects more money it is possible to scale down or up the lotteries being played. If subjects exhibit constant relative risk aversion it should be equivalent to increasing or decreasing initial wealth - here the characterization results from [ML] are useful. If my explanation for the gap between buying and selling price is correct then the gap should decrease when subjects are given more gambling money or if the lotteries are scaled down without changing gambling wealth.
The second experiment design is the following - I show some possible lotteries to the subjects. Then I ask them how much money maximally, they would risk playing these lotteries. Their answer would correspond to their gambling wealth. Then I again repeat the procedure as in the first experiment design.
Another way to test the approach would be to elicit buying and selling prices for objects from a very narrowly defined set, such as coffee mugs and then extend the set of objects to say all kitchen stuff and again elicit buying and selling prices. The gap between reported buying and selling price in the first case should be bigger than the one in the second case. The reason is that the money designated for trading in coffee mugs is definitely no bigger than money designated to trade in all kitchen stuff. This also would be consistent with results of [H]. He argues that buying/selling price gap should be small if there is some substitute on the market and should be bigger if there is no.
Assuming the approach is valid then I propose the following experiment for calibrating gambling wealth. The experiment should be designed to test risk attitudes[9] and at the same time to elicit selling and buying prices for lotteries. Given the data it is then easy to calculate the underlying wealth level. This is then interpreted as gambling wealth. More precisely, given observed buying and selling price for a given lottery I can calculate wealth-relative risk aversion coefficient pair which is consistent with these prices.
Gambling wealth hypothesis is promising. However, until there is no theory of gambling wealth interpretation it can not be fully testable. In the next subsection I discuss another concept which is related to gambling wealth - the concept of pocket cash by [FL]. The advantage of pocket cash idea is that there is a theory of pocket cash determination. I would like to show in what respect pocket cash and gambling wealth are similar and in what respect they differ.

.2 Pocket cash

The idea of pocket cash money in the context of gambling decisions is the following. If a small gamble is offered, an individual decides whether to take it or not on the basis

[9] Characterization results from [ML] are useful here.

of what he has in his pockets, and hence pocket cash will be the relevant wealth level for this decision. If, on the other hand, the same individual is offered a big gamble the values of which exceed significantly what he has in his pockets, the individual decides more carefully taking into account his lifetime wealth. I will introduce now some details of the model.
[FL] develop a dynamic model in which long-run self controls the series of short-run selves. In each period t there are two subperiods:

- bank subperiod
  - consumption is not possible
  - wealth y_t is divided between savings s_t, which remain in the bank, and pocket cash x_t which is carried to the nightclub
- nightclub subperiod
  - consumption $0 \leq c_t \leq x_t$ is determined and $x_t - c_t$ is returned to the bank at the end of the period
  - wealth next period is $y_{t+1} = R(s_t + x_t - c_t)$

The long-run self can implement a^*, the optimum of the problem without self-control, by simply choosing pocket cash $x_t = (1 - a^*)y_t$ to be the target consumption. In this way self-control costs might be avoided.

- At the nightclub in the first period there is a small probability the agent will be offered a choice between several lotteries.
- The model predicts then that:
  - for large gambles risk aversion is relative to wealth
  - for small gambles it is relative to pocket cash

In this way the model can explain [R] paradox and large buying and selling price gap.

.3 Gambling wealth vs. pocket cash

An interesting feature of [FL] approach is the following. [FL] estimate pocket cash to be roughly in the range of 20-100 dollars. This is very similar to the range of gambling wealth necessary to get large and consistent with the evidence buying/selling price gaps as indicated in table 2. Even if not supported by the thorough econometric analysis it is striking that two totally different approaches give rise to results of a very similar range.
In spite of the similarities, the two concepts are nevertheless different from each other. To illustrate the difference I will now discuss what testable predictions are obtained in [R] paradox according to the dual self model with pocket cash and what testable prediction are obtained according to gambling wealth approach.
[R] calibrated that expected utility model predicts the following:

- if a risk averse agent with wealth ≤ 350000 rejects the lottery $(105, 1/2; -100, 1/2)$

- then he should reject the lottery $(635000, 1/2; -4000, 1/2)$ at wealth level 340000

Denote the first of the above lotteries by lottery 1 and the second by lottery 2. According to [R] the first statement is plausible and the second is not and hence it is called a paradox.
In the dual-self model it is not true anymore that the decision maker rejects both lotteries. The first lottery is small and hence it is evaluated relative to pocket cash. The second lottery is big and therefore it is evaluated according to total wealth. Suppose that the utility function is logarithmic. Then the following is true:

- lottery 1 small - reject if pocket cash < 2100
- lottery 2 large - accept if total wealth higher than 4035

Now both statements (pocket cash less than 2100 and total wealth higher than 4035) are plausible.
Now consider gambling wealth interpretation. Suppose that utility is logarithmic. If gambling wealth is less than 2100 then the decision maker should

- reject lottery 1
- reject lottery 2

There is nothing paradoxical in rejecting the second lottery since gambling wealth in the amount of 2100 is too little to cover the loss (-4000) which occurs with probability 1/2. No matter how attractive is the second prize, the decision maker cannot afford to take lottery 2.

.4 Rabin's paradox in the literature

Although, in this paper, I adopt the lines of [R09], and focus on the assumption of consequentialism, there has been other explanations for Rabin's paradox in the literature. [P-HS] claim that it is the assumption of rejecting small gambles over a large range of wealth levels, which should be questioned as it does not match real-world behavior. In particular they show that the assumption that an expected utility maximizer turns down a given even-odds gamble with gain and loss for a given range of wealth levels implies that there exists a positive lower bound on the coefficient of absolute risk aversion which can be calculated exactly. This lower bound is an additional assumption imposed on a utility function. [P-HS] show that in Rabin's examples this lower bound turns out to be very high, which is not consistent with empirical evidence. Another paper which addresses Rabin's critique of expected utility is [CS]. They argue that the source of confusion around expected utility lies in a failure to distinguish between expected utility theories, i.e. all models based on a set of axioms with independence axiom being the key axiom, and a specific expected utility model. In a similar spirit to [R06], they claim that Rabin in fact criticizes expected utility model of terminal wealth, in which there is a single preference relation over final wealth consequences. They show that expected utility of income model, in which prizes are interpreted as changes in wealth levels, does not exhibit Rabin's paradoxical behavior. In order to enable the dependence of preference over income on initial wealth, they design an expected utility of initial wealth and income model. They demonstrate that such a model can withstand the Rabin's critique if initial wealth is not additive to income in the utility function. [SS] on the other hand point out that paradoxes of the kind considered by Rabin, are not specific to expected utility theory. They show that they can be constructed in non-expected utility theories as well.

Preference reversal versus buying/selling price reversal

Preference reversal is commonly observed in experiments. Suppose that $A \succ_C B$ denotes "A preferred to B in a direct choice". Using my notation, preference reversal is possible if:

$$S(W, \mathbf{y}) > S(W, \mathbf{x}) \text{ and } \mathbf{x} \succ_C \mathbf{y}$$

Preference reversal is not possible within expected utility framework. To see this, note that expected utility implies that $\mathbf{x} \succ_C \mathbf{y}$ which can be equivalently written as $\mathrm{E}U(W+\mathbf{x}) > \mathrm{E}U(W+\mathbf{y})$. By definition of S, this is equivalent to $U(W+S(W,\mathbf{x})) > U(W+S(W,\mathbf{y})$ and since utility function is strictly increasing: $S(W,\mathbf{x}) > S(W,\mathbf{y})$. So expected utility implies the following:

$$S(W, \mathbf{y}) > S(W, \mathbf{x}) \iff \mathbf{y} \succ_C \mathbf{x} \tag{7}$$

On the other hand buying/selling price is possible within expected utility framework:

Proposition 9. *For a given decreasing absolute risk aversion utility function and any wealth level W, buying/selling price reversal is possible.*

Proof. In the appendix. ■

By condition (7) this proposition implies that expected utility admits the possibility of the following kind of preference reversal:

$$B(W, \mathbf{y}) > B(W, \mathbf{x}) \text{ and } \mathbf{x} \succ_C \mathbf{y}$$

This kind of preference reversal will be referred to as preference reversal B. Preference reversal B is equivalent to buying/selling price reversal within expected utility framework.

Since expected utility theory imposes rather strong consistency assumptions, the result above suggests that the possibility of preference reversal is less rational than the related possibility of buying/selling price reversal or preference reversal B. The following two propositions clarify the meaning of "less rational" beyond the strength of consistency requirements argument.

Proposition 10. *Suppose that preferences of the decision maker are continuous, monotonic and that preference reversal pattern is fixed for the range of wealth $W \in [\underline{w}, \bar{w}]$. Then arbitrage opportunities exist.*

Proof. In the appendix. ■

Hence preference reversal allows arbitrage. On the other hand buying/selling price reversal or preference reversal B does not allow arbitrage.

Proposition 11. *Buying/selling price reversal does not allow arbitrage.*

Proof. In the appendix. ■

The analysis shows that buying/selling price reversal or preference reversal B is more rational than traditional preference reversal in two respects - it is consistent with expected utility and it does not allow arbitrage.
Preference reversal B or buying/selling price reversal occur within expected utility theory. However it does not mean that they have to be meaningful. If buying/selling price gap is small, then these two reversals are not meaningful i.e. they can occur theoretically but the scope for their occurrence is negligible. For these reversals to be meaningful, it is necessary for buying/selling price gap to be non-negligible. Testing of preference reversal B might be therefore relevant only if wealth is interpreted narrowly, either as gambling wealth or pocket cash. It is not relevant if the doctrine of consequentialism is maintained. I will illustrate this fact in the following example.

Example 7. Suppose utility function is CRRA with relative risk aversion coefficient of 2, the \$-bet (denote it by $\mathbf{x}$) gives \$100 or \$0 with equal probabilities and the P-bet (denote it by $\mathbf{y}$) gives \$40 with probability 3/4 and \$0 otherwise. The picture below graphs buying and selling prices for these two lotteries as functions of wealth:

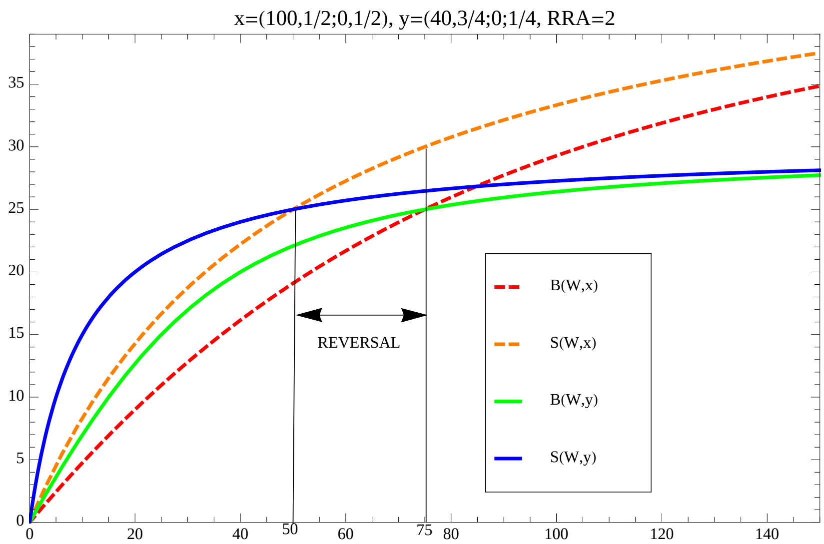


Figure 1. Wealth region for buying/seeling price reversal

In the above example, there is an interval $(50, 75)$ of wealth for which buying/selling price reversal (and hence also preference reversal B) occurs[10]. This is the common pattern that buying/selling price reversal occurs only at small wealth and only in the limited interval of wealth. The reason is that for such reversal to occur the $-bet has to have higher variance and higher expected value. Then since as wealth becomes large, buying and selling price approach expected value of a lottery, these prices for the $-bet have to increase above those of the P-bet. For smaller values of wealth, the CRRA decision maker would become very risk averse, so he will be solely preoccupied by the gamble's variance. Therefore, both selling and buying price for the $-bet are below those of the P-bet. Technically speaking, notice that if W^* denotes wealth level at which selling price of $\mathbf{x}$ and $\mathbf{y}$ are equal, i.e. $S(W^*, \mathbf{x}) = S(W^*, \mathbf{y}) = S^*$, then from definition it also holds that $S^* = B(W^* + S^*, \mathbf{x}) = B(W^* + S^*, \mathbf{y})$, so that $B(W, \mathbf{x})$ crosses $B(W, \mathbf{y})$ at $W = W^* + S^*$. Hence the interval for which buying/selling price reversal occurs is of length S^* exactly.

Further discussion

I argued above that dropping the assumption of consequentialism and interpreting initial wealth narrowly as something much smaller than total wealth can explain large buying and selling price gap but cannot explain traditional preference reversal involving WTA valuations. But in my approach, the consequentialist view, while dropped, is replaced with something different only in quantitative and not qualitative terms. I allow initial wealth to be small enough with a claim that it changes the interpretation of initial wealth. Instead of total wealth which is usually much higher than the value of consequences of a lottery in question, I propose to use a somewhat vague[11] notion of gambling wealth which, in order to explain large gap between buying and selling price, should be small enough and in particular of the magnitude similar to the lottery's consequences. This approach changes things only slightly in the following sense: While in case of total wealth interpretation of initial wealth, an individual's preferences over wealth changes are induced from his preferences over final wealth levels, in case of gambling wealth, preferences over changes in gambling wealth are induced from preferences over gambling wealth levels. In quantitative terms, these two situations may be very different. However in a structural or qualitative sense, these two situations differ only marginally. A much more fundamental departure from consequentialism assumption would be the following - suppose that the decision maker derives preference over lotteries not from preferences over the resulting total position of whatever budget he might consider but directly from changes to this budget implied by accepting the lottery in question. This approach was first undertaken by [KT] in their seminal contribution. In prospect theory which was then proposed as an alternative to expected utility theory the decision maker has preferences directly over wealth changes relative to a reference point. [R09] suggests that a similar approach is possible without having to depart from expected utility theory by reinterpreting lottery prizes as monetary change and not as total position. In what follows I will

[10] The following holds $S(50, \mathbf{y}) = S(50, \mathbf{x})$ and $B(75, \mathbf{y}) = B(75, \mathbf{x})$.

[11] Or at least hard to measure using existing data.

illustrate formally that in such an approach both large buying and selling price gap and preference reversal is possible for a wide class of utility functions defined over wealth changes.

WTA/WTP disparity for preferences over wealth changes

Suppose $U(\cdot)$ is a utility function defined over changes in wealth, and hence over the whole real line. It is assumed to be strictly increasing, continuous and that $U(0) = 0$. Willingness to accept for lottery $\mathbf{x}$, denoted by $WTA(\mathbf{x})$, and willingness to pay for lottery $\mathbf{x}$, denoted by $WTP(\mathbf{x})$, will be the notions used here in place of selling and buying price, respectively. They are defined as follows:

$$\mathrm{E}U(WTA(\mathbf{x}) - \mathbf{x}) = 0 \tag{8}$$

$$\mathrm{E}U(\mathbf{x} - WTP(\mathbf{x})) = 0 \tag{9}$$

The interpretation of these two terms is equivalent to the interpretation of selling and buying price, respectively, which was given earlier.

Proposition 12. *Suppose that $U(x) < -U(-x)$ for all $x > 0$. Then given any nondegenerate lottery $\mathbf{x}$, the following holds: $WTA(\mathbf{x}) > WTP(\mathbf{x})$.*

Proof. Denote $A \equiv WTA(\mathbf{x})$ to save on notation.

$$\begin{aligned}
&\mathrm{E}[U(\mathbf{x} - A) - U(A - \mathbf{x})] \\
&= \sum_{i:x_i > A} p_i[U(x_i - A) - U(-(x_i - A))] - \sum_{i:x_i \le A} p_i[U(A - x_i) - U(-(A - x_i))] \\
&< -2 \sum_{i:x_i > A} p_i U(-(x_i - A)) - 2 \sum_{i:x_i \le A} p_i U(A - x_i) \\
&= 0
\end{aligned}$$

The claim follows by monotonicity of U. ■

The above proposition is quite general. In what follows I will analyze two special cases. In the original version of prospect theory ([KT]) reference point was required to be constant. Hence it would be impossible to define willingness to accept in this formulation. [SSS] proposed a so called "third-generation prospect theory" in which reference point is allowed to be random. They defined willingness to accept and willingness to pay essentially as in (8) and (9) and showed first that willingness to pay and willingness to accept disparity is possible in prospect theory mainly due to loss aversion. Below I will show my version of their result which shows that under certain symmetry conditions WTA/WTP gap occurs solely due to loss aversion. Consider prospect theory utility function with an imposed symmetry condition of the following form:

Assumption 8. Utility function for outcomes is of the following form:

$$U(x) = \begin{cases} u(x) & \text{if } x \ge 0 \\ -\lambda u(-x) & \text{if } x < 0 \end{cases}$$

where $\lambda > 0$ denotes the loss attitude parameter. If $\lambda > 1$, there is loss aversion. Furthermore, a function $u(\cdot)$ is absolutely continuous, bounded, strictly increasing with $u(0) = 0$ and concave on its domain.

The above utility function is concave for gains, convex for losses and for $\lambda = 1$, it is symmetric around $(0, 0)$, meaning that risk loving for losses is of the same magnitude as risk aversion for gains. For $\lambda > 1$, there is loss aversion, which means that a given gain brings less satisfaction, than the dissatisfaction from the same loss.
Observe that this utility function satisfies $U(x) < -U(-x)$ for all $x > 0$ if and only if $\lambda > 1$. Hence for λ greater than one, meaning that there is loss aversion, willingness to accept for any nondegenerate lottery exceeds willingness to pay by proposition 12. In fact for a utility function of the above form an even stronger result holds, which I state below:

Proposition 13. *For a nondegenerate lottery* $\mathbf{x}$ *and utility function of the form defined by assumption 8, the following holds:*

$$\lambda > 1 \iff WTA(\mathbf{x}) > WTP(\mathbf{x})$$

Proof. Define $A \equiv WTA(\mathbf{x})$ and $P \equiv WTP(\mathbf{x})$ to save on notation. From definitions:

$$\begin{aligned} \lambda \sum_{i:x_i > A} u(x_i - A) &= \sum_{i:x_i \le A} u(A - x_i) \\ \sum_{i:x_i > P} u(x_i - P) &= \lambda \sum_{i:x_i \le P} u(P - x_i) \end{aligned}$$

First notice that $\lambda = 1$ if and only if A is equal to P. Now observe that $\lambda > 1$ if and only if

$$\sum_{i:x_i > A} u(x_i - A) \quad < \quad \lambda \sum_{i:x_i > A} u(x_i - A) = \sum_{i:x_i \le A} u(A - x_i) < \lambda \sum_{i:x_i \le A} u(A - x_i)$$

And by monotonicity of u it follows immediately that $A > P$ and hence $WTA(\mathbf{x}) > WTP(\mathbf{x})$. ■

The above proposition shows that in prospect theory willingness to accept/willingness to pay disparity may be explained solely by loss aversion.
An even simpler version of this result obtains in case of the prospect theory utility function without risk aversion.

Assumption 9. Utility function for outcomes is of the following form:

$$U(x) = \begin{cases} x & \text{if } x \ge 0 \\ \lambda x & \text{if } x < 0 \end{cases} \tag{10}$$

Proposition 14. *For a nondegenerate lottery* $\mathbf{x}$ *and utility function defined by assumption 9*

$$\lambda > 1 \iff WTA(\mathbf{x}) > \mathrm{E}[\mathbf{x}] > WTP(\mathbf{x})$$

Proof. Define $A \equiv WTA(\mathbf{x})$ to save on notation. From definition:

$$\begin{aligned} 0 &= -\lambda \sum_{i:x_i>A} p_i(x_i - A) + \sum_{i:x_i \le A} p_i(A - x_i) \\ &= (1-\lambda) \sum_{i:x_i>A} p_i(x_i - A) + \sum_{i=1}^{n} p_i(A - x_i) \\ &= (1-\lambda) \sum_{i:x_i>A} p_i(x_i - A) + A - \mathrm{E}[\mathbf{x}] \\ &< A - \mathrm{E}[\mathbf{x}] \end{aligned}$$

The proof that $\mathrm{E}[\mathbf{x}] > WTP(\mathbf{x})$ is similar and hence omitted. ■

In fact utility function 9 is a special case of an overall concave utility function for which similar result holds:

Proposition 15. *For a nondegenerate lottery* $\mathbf{x}$*, given utility function* $u : \mathbb{R} \to \mathbb{R}$ *that is strictly increasing, continuous and bounded with* $u(0) = 0$*, the following holds:*

$$u(\cdot) \textit{ is concave} \iff WTA(\mathbf{x}) \ge \mathrm{E}[\mathbf{x}] \ge WTP(\mathbf{x})$$

Proof. By Jensen's inequality:

$$\begin{aligned} 0 &= \mathrm{E}u(\mathbf{x} - WTP(\mathbf{x})) \le u(\mathrm{E}[\mathbf{x}] - WTP(\mathbf{x})) \\ 0 &= \mathrm{E}u(WTA(\mathbf{x}) - \mathbf{x}) \le u(WTA(\mathbf{x}) - \mathrm{E}[\mathbf{x}]) \end{aligned}$$

Since $u(0) = 0$ and u is strictly increasing, the conclusion follows. ■

The conclusion of this section is that the gap between willingness to accept and willingness to pay in case of preferences defined over wealth changes and not wealth levels may be explained by a kind of a general loss aversion, which is defined by the requirement: $u(x) < -u(-x)$, for all $x > 0$. This requirement defines a wide class of available utility functions and in particular, an S shaped utility function with sufficient level of loss aversion as well as a traditional overall concave utility function over the whole real line satisfies this condition.

Preference reversal for preferences over wealth changes

First, recall that traditional preference reversal is not possible within expected utility when preferences are defined over wealth levels, irrespective of whether these wealth levels are interpreted narrowly as levels of gambling wealth for instance or whether they are interpreted traditionally as total wealth levels. On the other hand, when preferences are defined over wealth changes, it turns out that traditional preference reversal is possible. [SSS] shows that preference reversal may occur in third generation prospect theory. They calibrate for which values of parameters, a very general but parametrized version of prospect theory is compatible with preference reversal . Below I will show that for a very simple version of third generation prospect theory, preference reversal is obtained as a generic element.

Let $\mathbf{x} \equiv (x, p)$ and $\mathbf{y} \equiv (y, q)$ be two prospects such that $y > x > 0$ and $1 > p > q > 0$. Lottery $\mathbf{x}$ will be called the P-bet and lottery $\mathbf{y}$ will be called the \$-bet. In what follows I want to demonstrate that preference reversal is possible.

Lemma 10. *For a utility function satisfying assumption 9, the following holds: If $px = qy$, so that the decision maker is indifferent between lottery $\mathbf{x}$ and $\mathbf{y}$ in a direct choice, then*

$$\lambda > 1 \iff WTA(\mathbf{y}) > WTA(\mathbf{x}$$

Proof. From definitions I calculate that:

$$WTA(\mathbf{x}) = \frac{\lambda p}{1 - p + \lambda p} x, \quad WTA(\mathbf{y}) = \frac{\lambda q}{1 - q + \lambda q} y$$

Using the fact that $px = qy$, I obtain:

$$WTA(\mathbf{y}) - WTA(\mathbf{x}) = WTA(\mathbf{y}) \frac{p - q}{1 - p + \lambda p} (\lambda - 1)$$

The claim immediately follows. ■

Proposition 16. *Given a utility function satisfying assumption 9, preference reversal is possible if and only if $\lambda > 1$*

Proof. Preference reversal occurs when the decision maker chooses the P-bet in a direct choice but assigns higher willingness to accept to the \$-bet. In terms of the model, preference reversal occurs when $px > qy$ and $WTA(\mathbf{y}) > WTA(\mathbf{x})$. By lemma 10, I know that if $px = qy$ then $WTA(\mathbf{y}) > WTA(\mathbf{x}) \iff \lambda > 1$. Since utility function $u(\cdot)$ is continuous, it follows that willingness to accept as a function of a given lottery is also continuous. Hence changing the lottery slightly changes willingness to accept for it slightly. It follows that if initially $px = qy$ and I increase p or x slightly or decrease q or y slightly, the new lottery $\mathbf{x}$ will be preferred to a new lottery $\mathbf{y}$ in a direct choice and yet it will remain true that willingness to accept for a new lottery $\mathbf{y}$ will still be higher than willingness to accept for a new lottery $\mathbf{x}$. ■

Again, a more general result for concave functions is possible:

Lemma 11. *Suppose that $px = qy$. Given utility function $u : \mathbb{R} \to \mathbb{R}$ that is strictly increasing, continuous and bounded with $u(0) = 0$, the following holds:*

$$u(\cdot) \textit{ is concave} \iff WTA(\mathbf{y}) \geq WTA(\mathbf{x})$$

Proof. Define $A \equiv WTA(\mathbf{x})$ to save on notation. A satisfies the following equation:

$$pu(A - x) + (1 - p)u(A) = 0 \tag{11}$$

The following is the so called three-strings lemma for concave functions:

Lemma 12 (Three strings lemma)**.** *Utility function $u(\cdot)$ is concave if and only if for $a > b > c$ the following holds:*

$$\frac{u(a)}{a} < \frac{u(b)}{b} < \frac{u(c)}{c} \tag{12}$$

Hence $u(\cdot)$ is concave if and only if

$$\begin{aligned}
qu(A-y)+(1-q)u(A) &\leq q\tfrac{A-y}{A-x}u(A-x)+(1-q)u(A) && \text{[by (12)]}\\
&= u(A-x)\left[q\tfrac{A-y}{A-x}-p\right]+(p-q)u(A) && \text{[by (11)]}\\
&= u(A-x)\tfrac{qA-qy-pA+px}{A-x}+(p-q)u(A) &&\\
&= A(p-q)\left[\tfrac{u(A)}{A}-\tfrac{u(A-x)}{A-x}\right] && [px=qy]\\
&\leq 0 && \text{[by (12)]}
\end{aligned}$$

■

Proposition 17. *Suppose that $px = qy$. Preference reversal occurs if $u(\cdot)$*[12] *is strictly concave.*

Proof. Suppose that $u(\cdot)$ is strictly concave. By lemma 11, $WTA(\mathbf{y}) > WTA(\mathbf{x})$. Since $x < y$, $\frac{u(x)}{x} > \frac{u(y)}{y}$, by the three strings lemma for concave function $u(\cdot)$. Hence the following holds:

$$\mathrm{E}u(\mathbf{x}) = pu(x) > p\frac{x}{y}u(y) = qu(y) = \mathrm{E}u(\mathbf{y})$$

So lottery $\mathbf{x}$ or a P-bet is chosen over lottery $\mathbf{y}$ or a \$-bet in a direct choice and yet $WTA(\mathbf{y}) > WTA(\mathbf{x})$ as required. ■

Concavity of a utility function is sufficient for preference reversal in the above example. However it is not necessary. In particular, [SSS] show that preference reversal is possible with an S-shaped prospect utility function, which is convex for losses. If one wants to obtain a possibility of preference reversal for specific lotteries and not as a generic feature of the model, then the following requirement, which is weaker than the overall concavity of utility function, may be imposed: For a given P-bet and a given \$-bet and utility function $u(\cdot)$, define $A = WTA(\mathbf{x})$. Then:

$$\frac{u(A-y)}{A-y} > \frac{u(A-x)}{A-x} > \frac{u(A)}{A}$$

Concluding remarks

Expected utility theory by [vNM] imposes a set of consistency assumptions on choices among lotteries. The theory is used in a large part of economic theory, including the famous Nash existence theorem. However there is a lot of mainly experimental evidence that people often violate [vNM] axioms, in particular the most crucial among them - independence. In response to this evidence economists started to question expected utility theory and investigate other models of choice which describe human behavior better. However, since these new theories usually have lower consistency requirements being imposed on the admissible choice, they necessarily also have lower prediction power and less scope for testable predictions. Moreover, they also have

[12] Strictly increasing, continuous and $u(0) = 0$.

weaker normative appeal, since the decision makers violating expected utility axioms are vulnerable to money pumps. It is therefore an important issue to identify patterns of choices and behavior which are consistent with expected utility and contrast them with those which are impossible within expected utility. In order to perform this task it is important to identify expected utility theory in its bare form and in particular separate it from the doctrine of consequentialism. More precisely, it is necessary to abandon the common practice of interpreting wealth variable as total wealth position common to all decisions.

If one is willing to accept that wealth underlying gambling decisions is separated from total wealth so that gambling decisions are framed narrowly, important implications can be derived. If gambling wealth is small enough, which should be tested in an experiment, then selling price for a lottery can be significantly greater than buying price without going beyond expected utility model and the extent of this difference can be as high as the one found in experiments. Also, the famous [R] paradox can be resolved, suggesting that expected utility is not guilty here, but rather the doctrine of consequentialism.

Still, traditional preference reversal is not possible even if wealth is allowed to be small. If expected utility is to be regarded as a positive theory, it is definitely a negative result. However, if one is willing to accept expected utility as a good normative theory, then the same result is very useful. It informs us then, that preference reversal is not rational. It is confirmed further by the result proved in the paper, that individuals exhibiting preference reversal are susceptible to arbitrage under certain mild conditions. The same kind of arbitrage, which I prefer to call strong arbitrage, is not possible within expected utility. What might be interesting is that another kind of preference reversal, which I call preference reversal B and which involves buying price in place of selling price and otherwise is the same as the traditional preference reversal, is possible within expected utility and is not vulnerable to arbitrage as shown in the paper. What it could suggest if one is willing to treat expected utility as a good normative theory, is that preference reversal B is perhaps "more rational" than traditional preference reversal. An interesting thing to do in the future would be to check whether people exhibit preference reversal B as frequently as they exhibit the traditional preference reversal and if not, then check why this is so.

Appendix

In what follows I will need the following lemma:

Lemma 13. *For any lottery* $\mathbf{x}$ *and any wealth level* W*, the following holds:*

$$S[W, \mathbf{x} - B(W, \mathbf{x})] = 0 \tag{13}$$

$$S[W - B(W, \mathbf{x}), \mathbf{x}] = B(W, \mathbf{x}) \tag{14}$$

$$B[W + S(W, \mathbf{x}), \mathbf{x}] = S(W, \mathbf{x}) \tag{15}$$

The proof is directly from definitions. For details, see [ML].

Proof of proposition 1

Proposition 18 (Concave)**.** *For any concave and strictly increasing utility function and a non-degenerate lottery* $\mathbf{x}$ *the following holds:*

$$\min(\mathbf{x}) < B(W,\mathbf{x}) < \mathrm{E}[\mathbf{x}]$$
$$\min(\mathbf{x}) < S(W,\mathbf{x}) < \mathrm{E}[\mathbf{x}]$$

Proof. Notice first, that for degenerate lottery $\mathbf{x} = x$, equations (1) and (2)imply the following:

$$\begin{aligned} W + S(W,x) &= W + x \\ W + x - B(W,x) &= W \end{aligned}$$

And so $S(W,x) = B(W,x) = x$. From now on I will focus on a non-degenerate lottery $\mathbf{x}$. I will prove the proposition only for the case of selling price. For buying price the proof is similar. I define $S \equiv S(W,\mathbf{x})$. Suppose $\min_{i\in\{1,\ldots,n\}} x_i \geq S$. Then notice that:

$$U(W + x_i) \geq U\left(W + \min_{i\in\{1,\ldots,n\}} x_i\right) \geq U(W+S)$$

with strict inequality for any $x_i \neq \min_{i\in\{1,\ldots,n\}} x_i$. Since lottery $\mathbf{x}$ is non-degenerate there exists at least one $x_i \neq \min_{i\in\{1,\ldots,n\}} x_i$ Hence

$$\sum_{i=1}^{n} p_i U(W + x_i) > U(W+S)$$

So S cannot be the selling price - a contradiction.
Suppose now that $S \geq \mathrm{E}[\mathbf{x}]$. By strict Jensen's inequality

$$\mathrm{E}U[W+\mathbf{x}] < U[W + \mathrm{E}[\mathbf{x}]] \leq U(W+S)$$

So S cannot be the selling price - a contradiction. So I have shown that indeed $\min_{i\in\{1,\ldots,n\}} x_i < S(W,\mathbf{x}) < \mathrm{E}[\mathbf{x}]$.

■

Hence for lotteries with bounded values buying and selling price are bounded below by the minimal prize of the lottery and bounded above by the expected value of the lottery.

Proof of proposition 5

Note first that U_α is unbounded from below if $\alpha \geq 1$ and bounded from below if $\alpha < 1$.

$$\lim_{x\to 0} U_\alpha(x) = \begin{cases} -\frac{1}{1-\alpha}, & 0<\alpha<1 \\ -\infty, & \alpha \geq 1 \end{cases} \tag{16}$$

By proposition 1 buying and selling price are necessarily greater than $\min(\mathbf{x})$. For $\alpha \geq 1$ the utility function is unbounded from below, therefore from the definition it follows that: $\lim_{W\to 0} B(W,\mathbf{x}) = \min(\mathbf{x})$ and $\lim_{W\to -\min(\mathbf{x})} S(W,\mathbf{x}) = \min(\mathbf{x})$. On the other hand for $0 < \alpha < 1$ the utility function is bounded from below. Additionally, $W - B(W,\mathbf{x})$ is strictly increasing in W since $\frac{\partial B(W,\mathbf{x})}{\partial W} < 1$. Therefore the lower bound for the domain of $B(W,\mathbf{x})$ as a function of W is given by $W_L(\mathbf{x})$ such that:

$$\mathrm{E}U(-\min(\mathbf{x}) + \mathbf{x}) = U(W_L(\mathbf{x}))$$

It follows that $\lim_{W\to W_L(\mathbf{x})} B(W,\mathbf{x}) = W_L(\mathbf{x}) + \min(\mathbf{x})$. Similarly, the lower bound for the domain of $S(W,\mathbf{x})$ as a function of W is $-\min(\mathbf{x})$ and hence: $\lim_{W\to -\min(\mathbf{x})} S(W,\mathbf{x}) = W_L(\mathbf{x}) + \min(\mathbf{x})$ since

$$\begin{aligned}\mathrm{E}U(-\min(\mathbf{x}) + \mathbf{x}) &= U(-\min(\mathbf{x}) + S(-\min(\mathbf{x}),\mathbf{x})) \\ &= U(-\min(\mathbf{x}) + \min(\mathbf{x}) + U^{-1}(\mathrm{E}U(-\min(\mathbf{x}) + \mathbf{x})))\end{aligned}$$

Proof. Now I prove the following statement:

$$\forall \alpha > 0, \lim_{W\to\infty} B(W,\mathbf{x}) = \lim_{W\to\infty} S(W,\mathbf{x}) = \mathrm{E}[\mathbf{x}]$$

Note that the Absolute Risk Aversion for CRRA utility function has the form $A_\alpha(W) = \frac{\alpha}{W}$. Hence as W goes to infinity and α is bounded (no extreme risk aversion) $A_\alpha(W)$ tends to zero. This implies risk neutrality and hence $\lim_{W\to\infty} S(W,\mathbf{x}) = \lim_{W\to\infty} B(W,\mathbf{x}) = \mathrm{E}[\mathbf{x}]$ irrespective of relative risk aversion coefficient. ■

Proof of lemma 5

I prove first that $B'(W) < 1$. From the definition of buying price using implicit function formula:

$$\frac{dB}{dW} = 1 - \frac{U'(W)}{\mathrm{E}U'(W + \mathbf{x} - B(W,\mathbf{x}))}$$

Since utility function is strictly increasing it must be that $\frac{dB}{dW} < 1$.
Now I prove that $S'(W - B(W)) = \frac{B'(W)}{1-B'(W)}$ and $S'(W - B(W)) > B'(W)$
From lemma 13 equation (14), using chain rule of differentiation, I have $B'(W) = S'(W - B(W))(1 - B'(W))$ Rearranging gives

$$S'(W - B(W)) = \frac{B'(W)}{1 - B'(W)}$$

Since $0 < B'(W) < 1$ by the above argument and proposition 3, I obtain $S'(W - B(W)) > B'(W)$.
Similarly I prove that $B'(W + S(W)) = \frac{S'(W)}{1+S'(W)}$ and $B'(W + S(W)) < S'(W)$.
Using equation (15) from lemma 13, I have $S'(W) = B'(W + S(W))(1 + S'(W))$ and hence

$$B'(W + S(W)) = \frac{S'(W)}{1 + S'(W)}$$

Since $S'(W) > 0$ by proposition 3, I get $B'(W + S(W)) < S'(W)$
Now I will prove that $S'(W) = \frac{S(W)-B(W)}{B(W)}$ for small positive $S(W)$.

And by proposition 3 $S(W) > B(W) > 0$. So when $S(W)$ is small and positive, then also $B(W)$ is small and positive. By lemma 13 equation (14), $S(W - B(W)) = B(W)$. For small $B(W)$ using first order Taylor expansion $B(W) = S(W) - \frac{dS}{dW}B(W)$ and hence it follows that

$$S'(W) = \frac{S(W) - B(W)}{B(W)}$$

Similarly, by lemma 13 equation (15), $B(W + S(W)) = S(W)$. Hence, for small $S(W)$ using first order Taylor expansion $S(W) = B(W) + \frac{dB}{dW}S(W)$ and it follows that

$$B'(W) = \frac{S(W) - B(W)}{S(W)}$$

Proof of proposition 7

Without loss of generality I assume that $\min(\mathbf{x}) = 0$. Fix $\mathbf{x}$ such that $\min(\mathbf{x}) = 0$. By proposition 5, $B(W)$ and $S(W)$ are positive and hence by proposition 3 $\tau(W)$ is positive over the whole range. Notice that range of $\tau(W)$ is determined by proposition 5. If the domain of $S(W)$ is denoted D_S and the domain of $B(W)$ is denoted D_B, then the domain of $\tau(W)$ is just $D_S \cap D_B = D_B$. In particular, for $\alpha \geq 1$ the domain of $\tau(W)$ is the interval $(0, \infty)$ and for $\alpha \in (0, 1)$, the domain is the interval $(W_L(\mathbf{x}), \infty)$, where $W_L(\mathbf{x})$ is defined as in proposition 5. To prove the proposition I have to check whether the following expression is negative:

$$\tau'(W) = \frac{S(W)}{B(W)} \left[\frac{S'(W)}{S(W)} - \frac{B'(W)}{B(W)} \right] \tag{17}$$

From lemma 13 I have the following equations:

$$\begin{aligned} B(W) &= S(W - B(W)) \\ S(W) &= B(W + S(W)) \end{aligned}$$

For the proof first order effects are not sufficient, but it turns out second order effects are. Therefore, by Taylor expansion of the second order I get from the above equations:

$$\begin{aligned} B(W) &= S(W) - S'(W)B(W) + S''(W)B^2(W) \\ S(W) &= B(W) + B'(W)S(W) + B''(W)S^2(W) \end{aligned}$$

I only need to check the difference from equation (17) which I can rewrite as follows using the above Taylor expansions:

$$\begin{aligned} \frac{S'(W)}{S(W)} - \frac{B'(W)}{B(W)} &= \frac{\frac{S(W)-B(W)}{B(W)} + S''(W)B(W)}{S(W)} - \frac{\frac{S(W)-B(W)}{S(W)} - B''(W)S(W)}{B(W)} \\ &= \frac{S''(W)B^2(W) + B''(W)S^2(W)}{S(W)B(W)} < 0 \end{aligned}$$

where the last inequality follows from the fact that both $B(W)$ and $S(W)$ are concave (by lemma 6) and nonnegative (by proposition 5).

Proof of proposition 9

Take any non-degenerate lottery $\mathbf{y}$ with $S(W,\mathbf{y}) > B(W,\mathbf{y})$. Such a lottery exists by proposition 3. I can find a sequence of real numbers which all are greater than $B(W,\mathbf{y})$ and smaller than $S(W,\mathbf{y})$. I can then treat these numbers as a support for a new lottery $\mathbf{x}$. I assign probabilities to each of this numbers such that they sum to one and are positive for at least two of these numbers (such that the resulting lottery is non-degenerate). Suppose I choose n such numbers. By proposition 1 I can now conclude that:

$$S(W,\mathbf{y}) > \max_{i\in\{1,\ldots,n\}} x_i > \mathrm{E}[\mathbf{x}] > S(W,\mathbf{x}) > B(W,\mathbf{x}) > \min_{i\in\{1,\ldots,n\}} x_i > B(W,\mathbf{y})$$

And hence the result is proved.

Proof of proposition 10

Suppose that at any wealth $W \in [\underline{w}, \bar{w}]$ the decision maker prefers lottery $\mathbf{x}$ to lottery $\mathbf{y}$ in a direct choice but assigns higher certainty equivalent to lottery $\mathbf{y}$. Given such pattern of preferences it is easy to design an arbitrage strategy that extracts at least $W - \underline{w}$ from this decision maker. Suppose $W \in [\underline{w}, \bar{w}]$ is an initial wealth. Construct a sequence $W_i,\ i \in \{1, 2, ..., n\}$ such that:

- $W_0 = W$
- $W_i = W_0 - \sum_{k=1}^{i} \epsilon_k, \quad \epsilon_i > 0 \quad i \in 1, 2, ..., n$
- $W_n \geq \underline{w}$, and $W_{n+1} < \underline{w}$
- for i even (including 0) $W_{i+1} + \mathbf{x} \succ W_i + \mathbf{y}$
- for i odd: $CE(W_i + \mathbf{x}) < CE(W_{i+1} + \mathbf{y})$

Notice that such a sequence exists by monotonicity and continuity of preferences and by properties of real numbers. Assume w.l.o.g. that $W_0 + \mathbf{y} \succ W_0$. The arbitrage strategy is now the following:

0) Take $\mathbf{y}$
1) Exchange $\mathbf{y}$ for $\mathbf{x}$ and pay me ϵ_1
2) Exchange $\mathbf{x}$ for $CE(W_1 + \mathbf{x}) - W_1$
3) Exchange $CE(W_1 + \mathbf{x})$ for $CE(W_1 + \mathbf{y})$ and pay me ϵ_2
4) Exchange $CE(W_2 + \mathbf{y}) - W_2$ for $\mathbf{y}$
5) Exchange $\mathbf{y}$ for $\mathbf{x}$ and pay me ϵ_3
6) Exchange $\mathbf{x}$ for $CE(W_3 + \mathbf{x}) - W_3$
7) Exchange $CE(W_3 + \mathbf{x})$ for $CE(W_3 + \mathbf{y})$ and pay me ϵ_4

8) Exchange $CE(W_4 + \mathbf{y}) - W_4$ for $\mathbf{y}$

....

The above arbitrage strategy extracts the amount of wealth equal to $W - \underline{w}$ from the decision maker.

Proof of proposition 11

In what follows I will try to construct an arbitrage strategy to exploit the decision maker and show that it is not possible. Given DARA utility function U, take $\mathbf{x}$ such that $B(W, \mathbf{x}) < 0$. I will examine only this case since in the other cases the proof is trivial.
Suppose first, the decision maker initially has non-random position W. If the price b for the lottery is bigger than $B(W, \mathbf{x})$, the decision maker will not buy it. Hence, a price which is a part of an arbitrage strategy must be smaller than $B(W, \mathbf{x})$. Given such price b, the decision maker buys the lottery. His new position is $W + \mathbf{x} - b$. If the price s is smaller than $S(W - b, \mathbf{x})$, then the decision maker does not want to sell. Hence a price which is a part of an arbitrage strategy must be bigger than $S(W - b, \mathbf{x})$. By proposition 3, I know that S is strictly increasing and $b < B(W, \mathbf{x})$. Therefore:

$$s > S(W - b, \mathbf{x}) > S(W - B(W, \mathbf{x}), \mathbf{x}) = B(W, \mathbf{x}) > b$$

where the equality follows from lemma 13 equation (14).
Suppose now, that the decision maker initially has a random position $W + \mathbf{x}$. By the same argument as above the price s, which is a part of an arbitrage strategy has to be greater that $S(W, \mathbf{x})$, otherwise the decision maker would not sell the lottery $\mathbf{x}$. After selling the lottery, the decision maker's new position is $W + s$. The price b which is a part of an arbitrage strategy has to be smaller than $B(W + s, \mathbf{x})$. By lemma 5, I know that $\frac{\partial B(W,\mathbf{x})}{\partial W} \leq 1$ for all $W \geq 0$. Hence:

$$s - S(W, \mathbf{x}) > B(W + s, \mathbf{x}) - B(W + S(W, \mathbf{x}), \mathbf{x})$$

By lemma 13 equation (15), I know that $B(W + S(W, \mathbf{x}), \mathbf{x}) = S(W, \mathbf{x})$, and hence:

$$s > B(W + s, \mathbf{x}) > b$$

That proves that with decision maker's initial position equal to either W or $W + \mathbf{x}$, all arbitrage strategies have the property that $s > b$. However, this cannot be an arbitrage strategy since it makes negative profit equal to $b - s$. This proves that there are no arbitrage strategies.

References

[BDGM] Becker, G. M., M. H. DeGroot, and J. Marschak (1964). Measuring utility by a single-response sequential method. Behav Sci 9, 226–32.

[CS] Cox, J. C. and V. Sadiraj (2006). Small- and large-stakes risk aversion: Implications of concavity calibration for decision theory. Games and Economic Behavior 56, 45–60.

[FH] Foster, D. and S. Hart (2007, July). An operational measure of riskiness.

[FL] Fudenberg, D. and D. K. Levine (2006). A dual-self model of impulse control. American Economic Review 96, 1449–1476.

[GP] Grether, D. M. and C. R. Plott (1979). Economic theory of choice and the preference reversal phenomenon. The American Economic Review 69, 623–638.

[H] Hanemann, W. M. (1991). Willingness to pay and willingness to accept: How much can they differ? American Economic Review 81, 635–647.

[H] Horowitz, J. K. and K. E. McConnell (2002). A review of WTA/WTP studies. Journal of Environmental Economics and Management 44, 426–447.

[KT] Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. Econometrica 47, 263–292.

[KS] Knetsch, J. L. and J. A. Sinden (1984). Willingness to pay and compensation demanded: Experimental evidence of an unexpected disparity in measures of value. The Quarterly Journal of Economics 99, 507–521.

[ML] Lewandowski, M. (2009, June). Risk attitudes, buying and selling price for a lottery and simple strategies, unpublished manuscript.

[P-HS] Palacios-Huerta, I. and R. Serrano (2006). Rejecting small gambles under expected utility. Economics Letters 91, 250–259.

[R] Rabin, M. (2000). Risk aversion and expected-utility theory: A calibration theorem. Econometrica 68 (5), 1281–1292.

[Rai] Raiffa, H. (1968). Decision Analysis: Introductory Lectures on Choices Under Uncertainty. Addison-Wesley.

[R06] Rubinstein, A. (2002). Comments on the risk and time preferences in economics. Tel Aviv University Working Paper.

[R09] Rubinstein, A. (2009). Lecture Notes in Microeconomic Theory: The Economic Agent. Princeton University Press.

[SS] Safra, Z. and U. Segal (2008). Calibration results for non-expected utility theories. Econometrica 76, 1143–1166.

[SSS] Schmidt, U., C. Starmer, and R. Sugden (2008). Third-generation prospect theory. Jour- nal of Risk and Uncertainty 36, 203–223.

[T] Thaler, R. (1980). Toward a positive theory of consumer choice. Journal of Economic Behavior and Organization 1 (1), 39–60.

[vNM] von Neumann, J. and O. Morgenstern (1944). Theory of games and economic behavior. Princeton University Press.

Quantitative Methods in Economics
Vol. XI, No. 1, 2010, pp. 254–275

GAMES WITH DISTORTED INFORMATION AND SELF-VERIFICATION OF BELIEFS WITH APPLICATIONS TO FINANCIAL MARKETS

Agnieszka Wiszniewska-Matyszkiel

Institute of Applied Mathematics and Mechanics University of Warsaw
e-mail: agnese@mimuw.edu.pl

Abstract: In the paper we examine discrete time dynamic games in which the global state variable changes in response to a certain function of the profile of players' decisions, called statistic, while the players form some expectations about its future values based on the history. Besides, there are also players' private state variables.
A general model is built, encompassing both games with finitely many players as well as games with infinitely many players. This model extends the class of games with distorted information considered by the author in [20], in which there were no private state variables and there were much stronger assumptions about the statistic of players' decisions considered. The notions of pre-belief distorted Nash equilibrium (pre-BDNE), self-verification and belief distorted Nash equilibrium (BDNE), defined already in [20], are applied to our wider class of games. The relations between Nash equilibria, pre-BDNE and BDNE are examined as well as the existence and properties of pre-BDNE.
A model of a financial market – a simplified stock exchange – is presented as an example. Pre-BDNE using threshold prices are proposed. One of further results in this example is potential self-verification of fundamental beliefs and beliefs in infinite speculative bubbles.

Key words: games with continuum of players, n-player dynamic games, Nash equilibrium, pre-belief distorted Nash equilibrium, subjective equilibrium, self-verification of beliefs, financial markets.

Introduction

Motivation

The starting point of the research on games with distorted information were two phenomena arising from the problems examined in earlier papers of the author: exploitation of common renewable resources, especially by a large groups of users, and modelling financial markets.

The problem of a common ecosystem in the context of pre-belief distorted Nash equilibria (pre-BDNE) and self-verification of beliefs was examined by the author in Wiszniewska-Matyszkiel [20]. A concept of belief distorted Nash equilibria (BDNE) encompassing both pre-BDNE and self-verification of beliefs during the game was also introduced.

An example which is especially illustrative for the class of games considered in this paper, and which in fact forced extension of the class of games considered in Wiszniewska-Matyszkiel [20], is a stock exchange.

In such an institution prices of shares are calculated by the equilibrating mechanism given only a profile of players' orders. However, players usually formulate various prognostic techniques, often ridiculed by economists. Can such a behaviour with no scientific explanation turn out to be rational? Why people still believe in things like technical analysis?

Imperfect information, beliefs and game theory

There were various concepts taking beliefs into account, usually in games with stochastic environment or randomness caused by using mixed strategies.

The first were Bayesian equilibria, introduced by Harsanyi [7].

That approach was continued by e.g. Battigalli and Siniscalchi [5] using the concept of Δ-rationalizability being an iterative procedure of eliminating type-strategy pairs in which the strategy is strictly dominated according to player's beliefs or which are contradicted by a history of play.

Another concept – subjective equilibria were introduced by Kalai and Lehrer [8] and [9] (although the very idea of subjectivity in games appeared already in Aumann [2] and [3]). They apply to repeated games. In those concepts every player maximizes his expected payoff assuming some environment response function (so, in fact, calculation of equilibrium is decomposed into separate decision making problems) and non-falsification of his assumption during the game was added.

A comparison of these absolutely different concepts to the concepts of pre-BDNE and BDNE, can be found in Wiszniewska-Matyszkiel [20].

Games with a measure space of players

The term *games with a measure space of players* is usually perceived as a synonym of *games with infinitely many players* called also *large games*. In order to make it possible to evaluate the influence of the infinite set of players on aggregate variables, a measure is introduced on a σ-field of subsets of the set of players. However, the notion *games with a measure space of players* encompasses also games with finitely many players, where e.g. the counting measure on the power set may be considered.

Large games illustrate situations where the number of agents is large enough to make a single agent from a subset of the set of players (possibly the whole set) insignificant – *negligible* – when we consider the impact of his action on aggregate variables while joint action of a set of such negligible players is not negligible. This happens in many real situations: at competitive markets, stock exchange, or while we consider emission of greenhouse gases and similar global effects of exploitation of the common global ecosystem.

Although it is possible to construct models with countably many players illustrating the phenomenon of this negligibility, they are very inconvenient to cope with. Therefore simplest examples of large games are so called *games with continuum of players*, where players constitute a nonatomic measure space, usually unit interval with the Lebesgue measure.

The first attempts to use models with continuum of players are contained in Aumann [1] and Vind [13].

Some theoretical works on large games are Schmeidler [11], Mas-Colell [10], Balder [4] and Wiszniewska-Matyszkiel [14].

The general theory of dynamic games with continuum of players is still being developed, mainly by the author in [15] and [16].

Introducing a continuum of players instead of a finite number, however large, can change essentially properties of equilibria and the way of calculating them even if the measure of the space of players is preserved in order to make the results comparable. Such comparisons were made by the author in [17] and [18].

Formulation of the model

A *game with distorted information* $\mathfrak{G}$ is a tuple of the following objects $((\mathbb{I},\Im,\lambda),\mathbb{T},\mathbb{X},\{\mathbb{W}_i\}_{i\in\mathbb{I}},(\mathbb{D},\mathcal{D}),\{D_i\}_{i\in\mathbb{I}},U,\phi,\{\kappa_i\}_{i\in\mathbb{I}},\{P_i\}_{i\in\mathbb{I}},\{G_i\}_{i\in\mathbb{I}},\{B_i\}_{i\in\mathbb{I}},\{r_i\}_{i\in\mathbb{I}})$, whose interpretation and properties will be defined in the sequel.

The *set of players* is denoted by $\mathbb{I}$. In order that the definitions of the paper encompassed both games with finitely many players as well as games with infinitely many players we introduce a structure on $\mathbb{I}$ consisting of a σ-field $\Im$ of its subsets and a measure λ on it.

The game is dynamic, played over a discrete *time set* $\mathbb{T}$, without loss of generality $\mathbb{T}=\{t_0,t_0+1,\ldots,T\}$ or $\mathbb{T}=\{t_0,t_0+1,\ldots\}$, which, for uniformity of notation, will be treated as $T=+\infty$. We introduce also the symbol $\overline{\mathbb{T}}$ denoting $\{t_0,t_0+1,\ldots,T+1\}$ for finite T and equal to $\mathbb{T}$ in the opposite case.

The game is played in a *global system* with the *set of states* $\mathbb{X}$. The state of the global system (*state* for short) changes over time in response to players' decisions, constituting a trajectory X, whose equation will be stated in the sequel. The set of all potential *trajectories* – functions $X:\overline{\mathbb{T}}\to\mathbb{X}$ – will be denoted by $\mathfrak{X}$.

Besides the global system, player i has his *private state* variable with values in a *set of private states of player* i denoted by $\mathbb{W}_i$. By $\mathbb{W}$ we denote a superset containing all $\mathbb{W}_i$. The vector of private state variables also changes in response to players' decisions, constituting a trajectory $W:\overline{\mathbb{T}}\to\mathbb{W}^{\mathbb{I}}$, whose equation will be stated in the sequel.

At each time t given state x and his private state w_i player i chooses a decision from his *available decision set* $D_i(t,x,w_i)\subset\mathbb{D}$ – the *set of* (potential) *actions*. These available decision sets of player i constitute the *correspondence of available decision sets of player* i $D_i:\mathbb{T}\times\mathbb{X}\times\mathbb{W}\multimap\mathbb{D}$, while all available decision sets constitute a *correspondence of available decision sets* $D:\mathbb{I}\times\mathbb{T}\times\mathbb{X}\times\mathbb{W}\multimap\mathbb{D}$ with nonempty values. We shall also need a σ-field of subsets of $\mathbb{D}$, denoted by $\mathcal{D}$.

For any time t, state x and a vector of private states $w \in \mathbb{W}^{\mathbb{I}}$ we call any $\mathcal{D}$-measurable selection δ from the correspondence $i \multimap D(i,t,x,w_i)$ a *static profile available at t, x and w*. The set of all static profiles available at t, x and w is denoted by $\Sigma(t,x,w)$. The union of all the sets of static profiles available at various t, x and w is denoted by Σ.

The definitions of a strategy (dynamic strategy) and a profile (dynamic profile) will appear in the sequel, since first we have to define the domains of these functions.

The influence of a static profile on the global state variable is via its statistic. Without loss of generality the same statistic is the only parameter besides player's own strategy influencing the evolution of his private state variable and the value of his payoff. Formally, a *statistic* is a function $U : \Sigma \times \mathbb{X} \overset{\text{onto}}{\to} \mathbb{U}$ for a set $\mathbb{U}$ called the *set of profile statistics* and such that $U(\delta,x) = \gamma\left(\left[\int_{\mathbb{I}} g_k(i,\delta(i),x)d\lambda(i)\right]_{k\in\mathbb{K}}\right)$ for a collection of functions $\{g_k : \mathbb{I}\times\mathbb{D}\times\mathbb{X}\to\mathbb{R}\}_{k\in\mathbb{K}}$ which are $\Im\otimes\mathcal{D}$-measurable for every $x \in \mathbb{X}$ and for every k in the set of indices $\mathbb{K}$, and a function $\gamma : \mathbb{R}^{\mathbb{K}} \to \mathbb{U}$. If $\Delta : \mathbb{T} \to \Sigma$ represents choices of profiles at various time instants and X is a trajectory of the global system, then by $U(\Delta, X)$ we denote the function $u : \mathbb{T} \to \mathbb{U}$ such that $u(t) = U(\Delta(t), X(t))$. The set of all such functions will be denoted by $\mathfrak{U}$.

We do not assume any kind of continuity of the function γ. In the case of modelling financial markets it is by assumption discontinuous – it is defined, as in section , as a point at which a maximum of a kind of lexicographic ordering on a subset of $\mathbb{R}^{\mathbb{K}}$ is attained.

This class of statistic functions is a generalization of the class of statistic functions used in the previous paper Wiszniewska-Matyszkiel [20], in which $\mathbb{K}$ was finite and γ was absent.

Obviously, in the case of games with finitely many players with finite dimensional strategy set, the statistic can be the profile itself.

In a model of stock exchange an obvious candidate for such a statistic is the market price of the asset considered (in the example we shall see that another two coordinates will be useful).

Given a function $u : \mathbb{T} \to \mathbb{U}$, representing the statistics of profiles chosen at various time instants, the global system evolves according to the equation $X(t+1) = \phi(X(t), u(t))$ with the initial condition $X(t_0) = \bar{x}$. We call such a trajectory *corresponding to* u and denote it by X^u. If $u = U(\Delta, X^u)$, where $\Delta : \mathbb{T} \to \Sigma$ represents a choice of static profiles at various time instants, then, by a slight abuse of notation, we shall denote the trajectory corresponding to u by X^Δ and call it corresponding to Δ and instead of $U(\Delta, X^\Delta)$ we write $U(\Delta)$ – the statistic of Δ.

Given functions $u : \mathbb{T} \to \mathbb{U}$ and $d : \mathbb{T} \to \mathbb{D}$, representing, correspondingly, the subsequent statistics of static profiles chosen and subsequent decisions of player i, the private system of player i evolves according to the equation $W_i(t+1) = \kappa_i(W_i(t), d(t), X^u(t), u(t))$ with the initial condition $W_i(t_0) = \bar{w}_i$. We call such a trajectory of private system *corresponding to d and u* and denote it by $W_i^{d,u}$. If a function $\Delta : \mathbb{T} \to \Sigma$ represents choices of profiles at various time instants, then the trajectory of private state variables W^Δ defined by $\left(W^\Delta\right)_i = W_i^{\Delta_i, U(\Delta)}$ is called *corresponding to* Δ. This is another generalization of [20] in which only the global state variable was considered.

At each time t and the state of the global system x and the vector of players' private states w players get instantaneous payoffs. The *instantaneous payoff* of player i is a function $P_i : \mathbb{D} \times \mathbb{U} \times \mathbb{X} \times \mathbb{W} \to \mathbb{R} \cup \{-\infty\}$.

Besides, in the case of finite time horizon players get also *terminal payoffs* (after termination of the game) defined by the functions $G_i : \mathbb{X} \times \mathbb{W} \to \mathbb{R} \cup \{-\infty\}$. For uniformity of notation we take $G_i \equiv 0$ in the case of infinite time horizon.

Players observe some histories of the game, but not the whole profiles. At time t they observe the states $X(s)$ for $s \leq t$ and the statistics $u(s)$ of chosen static profiles for time instants $s < t$. Therefore the set of histories at time t equals $\mathbb{X}^{t-t_0+1} \times \mathbb{U}^{t-t_0}$. In order to simplify notation we introduce the set of all, possibly infinite, histories of the game $\mathbb{H} = \mathbb{X}^{T-t_0+2} \times \mathbb{U}^{T-t_0+1}$ and for such a history $H \in \mathbb{H}$ we denote by $H|_t$ the actual history at time t, while by $H(t)$ the pair $(X(t-1), u(t))$.

Given a history observed at time t, $H|_t$, players formulate their suppositions about future values of u and X, depending on their decision a made at time t. This is formalized as a multivalued *correspondence of belief of player* i, $B_i : \mathbb{T} \times \mathbb{D} \times \mathbb{H} \multimap \mathbb{H}$ with nonempty values. To reflect the fact that beliefs are based only on observed history, we assume that beliefs $B_i(t, a, H)$ are identical for histories H with identical $H|_t$ and that for all $H' \in B_i(t, a, H)$ we have $H'|_t = H|_t$. For simplicity of some further notation we also assume that for every history $H' \in B_i(t, a, H)$ and $H'' \in \mathbb{H}$ differing from H' only by $u'(t) \neq u''(t)$ we have also $H'' \in B_i(t, a, H)$, which means that the belief correspondence codes no information about the current value of u.

In our problem we allow players to have very compound closed loop strategies – dependent on time instant, state, private state and belief at the actual history of the game at this time instant. Formally, a (*dynamic*) *strategy of player* i is a function $S_i : \mathbb{T} \times \mathbb{X} \times \mathbb{W} \times \mathbb{H} \to \mathbb{D}$ such that for each time t, state x, a private state w_i and history H we have $S_i(t, x, w_i, H) \in D_i(t, x, w_i)$.

Such choices of players' strategies constitute a function
$S : \mathbb{I} \times \mathbb{T} \times \mathbb{X} \times \mathbb{W} \times \mathbb{H} \to \mathbb{D}$. The set of all strategies of player i will be denoted by $\mathfrak{S}_i$.

For simplicity of further notation, for a choice of strategies $S = \{S_i\}_{i \in \mathbb{I}}$ we can consider the open loop form of it $S^{OL} : \mathbb{T} \to \Sigma$, defined by
$S_i^{OL}(t) = S_i(t, X(t), W_i(t), H)$, where H is the history of the game resulting from choosing S. It is well defined, whenever the history is well defined (note that the statistic was defined only for measurable selections from players' strategy sets, therefore the statistics at time t is well defined if the function $S_{\cdot}^{OL}(t-1)$ is measurable). Therefore, we restrict the notion (*dynamic*) *profile* (*of players' strategies*) to choices of strategies such that for every t the function $S_{\cdot}^{OL}(t)$ is a static profile available at $t \in \mathbb{T}$, $X^{S^{OL}}(t)$ and $W^{S^{OL}}(t)$. The set of all dynamic profiles will be denoted by $\mathbf{\Sigma}$. Since the choice of a dynamic profile S determines the history of the game, we shall denote this history by H^S.

If the players choose a dynamic profile S, then the actual *payoff of player* i $\Pi_i : \mathbf{\Sigma} \to \overline{\mathbb{R}}$ in the game depends only on the actions actually chosen by players at subsequent time instants, i.e. the open loop form of the profile, and it is equal to

$$\begin{array}{rl}\Pi_i(S) = & \sum_{t=t_0}^{T} P_i\left(S_i^{OL}(t), U\left(S_i^{OL}(t)\right), X^{S^{OL}}(t), W_i^{S^{OL}}(t)\right)\cdot\left(\frac{1}{1+r_i}\right)^{t-t_0} + \\ & +G_i\left(X^{S^{OL}}(T+1), W_i^{S^{OL}}(T+1)\right)\cdot\left(\frac{1}{1+r_i}\right)^{T+1-t_0}\end{array},$$

where $r_i > 0$ is called a *discount rate.*

However, players do not know the whole profile, therefore instead of the actual payoff at each future time instant they can use in their calculations the *anticipated payoff functions* $\Pi_i^e : \mathbb{T} \times \mathbf{\Sigma} \to \overline{\mathbb{R}}$ corresponding to their beliefs at the corresponding time instants (the world "anticipated" is used in the colloquial meaning of "expected", while the world "expected" is not used in order not to cause associations with expected value with respect to some probability distribution).
This function for player i is defined by

$$\begin{array}{rl}\Pi_i^e(t,S) = & P_i\left(S_i^{OL}(t), U\left(S_i^{OL}(t)\right), X^{S^{OL}}(t), W_i^{S^{OL}}(t)\right) + \\ & +V_i(t+1, W_i^{S^{OL}}(t+1), B_i(t, S_i^{OL}(t), H^S))\cdot\frac{1}{1+r_i},\end{array}$$

where $V_i : \overline{\mathbb{T}} \times \mathbb{W}_i \times (\mathfrak{P}(\mathbb{H})\backslash\emptyset)) \to \overline{\mathbb{R}}$, (the *function of guaranteed anticipated value*) represents the present value of the minimal future payoff given his belief correspondence and assuming player i chooses optimally in the future.

Formally, for time t, private state w_i and belief $\mathbb{B} \in \mathfrak{P}(\mathbb{H})\backslash\emptyset$ we define

$$V_i(t, w_i, \mathbb{B}) = \inf_{H\in\mathbb{B}} v_i(t, w_i, H),$$

where the function $v_i : \overline{\mathbb{T}}\times\mathbb{W}_i \times \mathbb{H} \to \overline{\mathbb{R}}$ is the present value of the future payoff of player i along a history assuming he chooses optimally in the future: for $t \in \mathbb{T}$ we define it by

$$\begin{array}{rl}v_i(t, w_i, (X,u)) = & \sup_{d:\mathbb{T}\to\mathbb{D}\ d(\tau)\in D_i(\tau,X(\tau),W_i(\tau))\text{ for }\tau\geq t} \\ & \sum_{\tau=t}^{T} \frac{P_i(\tau,d(\tau),u(\tau),X(\tau),W_i(\tau))}{(1+r_i)^{\tau-t}} + \frac{G_i(X(T+1),W_i(T+1))}{(1+r_i)^{T+1-t}}\end{array}$$

for W_i defined by $W_i(t) = w_i$ and
$W_i(\tau+1) = \kappa(W_i(\tau), d(\tau), X(\tau), u(\tau))$ for $\tau < T$;
$v_i(T+1, w_i, H) = G_i\left(X(T+1), w_i\right)$.

Note that such a definition of anticipated payoff is inspired by the Bellman equation for calculation of best responses of players to the strategies of the others. For various versions of this equation see e.g. Blackwell [6] or Stokey and Lucas [12].

We also introduce the symbol $\mathfrak{G}_{t,H,w}$ (for $H = (X,u)$), called *subgame with distorted information at* t, x and w and denoting the game with the set of players $\mathbb{I}$, the sets of their strategies $D_i(t, X(t), w_i)$ and the payoff functions $\bar{\Pi}_i^e(t, H, \cdot)$ defined by $\bar{\Pi}_i^e(t, H, \delta) = \Pi_i^e(t, S)$ for a profile S such that $S(t) = \delta$ and $H^S|_t = H|_t$ (note that the dependence of $\Pi_i^e(t,\cdot)$ on the profile is restricted to its static profile at time t only and the history $H|_t$, therefore the definition does not depend on the choice of specific S from this class).

Nash equilibria and pre-belief distorted Nash equilibria

One of the basic concepts in game theory, Nash equilibrium, assumes that every player (almost every in case of large games with a measure space of players) chooses a strategy which maximizes his payoff given the strategies of the remaining players.

In order to simplify the notation we shall need the following abbreviation: for a profile S and a dynamic strategy d of player i the symbol $S^{i,d}$ denotes the dynamic profile such that $S_i^{i,d} = d$ and $S_j^{i,d} = S_j$ for $j \neq i$.

Definition 1. A profile S is a *Nash equilibrium* if for a.e. $i \in \mathbb{I}$ and for every strategy d of player i we have $\Pi_i(S) \geq \Pi_i(S^{i,d})$.

However, the assumption that a player knows the strategies of the remaining players or at least the statistic of these strategies which influences his payoff, is usually not fulfilled in real life situations. Moreover, even details of the other players' payoff functions or available strategy sets are sometimes not known precisely, while the other players' information at a specific situation is usually unknown. This is especially visible at financial markets.

Therefore, given their beliefs, players maximize their anticipated payoffs.

Definition 2. A profile S is a *pre-belief distorted Nash equilibrium* (*pre-BDNE* for short) if for a.e. $i \in \mathbb{I}$, and every strategy d of player i and every $t \in \mathbb{T}$ we have $\Pi_i^e(t, S) \geq \Pi_i^e(t, S^{i,d})$.

If we use the notation introduced in the formulation of the model, then a profile S is a pre-BDNE in $\mathfrak{G}$ if at every time t the static profile $S^{OL}(t)$ is a Nash equilibrium in $\mathfrak{G}_{t,H^S,W^S(t)}$.

In order to state an existence result for games with a nonatomic space of players we have to introduce the following notation.
Let $e(\delta, x) = \left[\int_{\mathbb{I}} g_k(i, \delta(i), x) d\lambda(i)\right]_{k \in \mathbb{K}}$. The functions $\widetilde{P}_i(a, e, x, w_i)$, $\widetilde{\kappa}_i(w_i, a, x, e)$ and $\widetilde{\phi}(x, e)$ will denote $P_i(a, \gamma(e), x, w_i)$, $\kappa_i(w_i, a, x, \gamma(e))$ and $\phi(x, \gamma(e))$, respectively.

Theorem 3. ***Existence of pre-BDNE***

Let $(\mathbb{I}, \Im, \lambda)$ be a nonatomic measure space and let $\mathbb{K}$ be finite, $\mathbb{D} = \mathbb{R}^n$, $\mathbb{W} = \mathbb{R}^m$ with the σ-fields of Borel subsets and let the function $\mathbb{I} \ni i \mapsto \bar{w}_i$ be measurable. Assume that for every t, x, w_i, H and for almost every i the following continuity-compactness assumptions *hold: the sets $D_i(t, x, w_i)$ are compact, the functions $\widetilde{P}_i(a, e, x, w_i)$ and*
$V_i(t, \widetilde{\kappa}_i(w_i, a, x, e), B_i(t, a, H))$ are upper semicontinuous in (a, e) jointly while for every a they are continuous in e and for all k the functions $g_k(i, a, x)$ are continuous in a for $a \in D_i(t, x, w_i)$ and assume that for every t, x, e, H the following measurability assumptions *hold: the graph of $D_\cdot(t, x, \cdot)$ is measurable and the following functions defined on $\mathbb{I} \times \mathbb{W} \times \mathbb{D}$ are measurable $(i, w, a) \mapsto \widetilde{P}_i(a, e, x, w)$, r_i, $V_i(t, \widetilde{\kappa}_i(w, a, x, e), B_i(t, a, H))$, $\widetilde{\kappa}_i(w, a, x, e)$ and $g_k(i, a, x)$ for every k. Moreover, assume that for every k and x there exists an integrable function $\Gamma : \mathbb{I} \to \mathbb{R}$ such that for every w and every $a \in D_i(t, x, w)$ $|g_k(i, a, x)| \leq \Gamma(i)$. Under these assumptions there exists a pre-BDNE for B.*

Proof. It is a conclusion from one of theorems on the existence of pure strategy Nash equilibria in games with continuum of players: Wiszniewska-Matyszkiel [14] theorem 3.1 or Balder [4] theorem 3.4.1 applied to the sequence of games $\mathfrak{G}_{t,H,w}$ for any history H such that $H|_t$ is the actual history of the game observed at time t while w describes the private states of players at time t.

In order to apply one of those theorems we first have to prove measurability of the function $i \mapsto w_i(t)$ given a measurable initial data and a dynamic profile, which is immediate. It implies that the graph of $D_.(t, x, w_.)$ (in $\mathbb{I} \times \mathbb{D}$) is measurable and the following functions defined on $\mathbb{I} \times \mathbb{D}$ are measurable: $(i, a) \mapsto \widetilde{P}_i(a, e, x, w_i)$, r_i, $V_i(t, \widetilde{\kappa}_i(w_i, a, x, e), B_i(t, a, H))$. ■

Now we return to show some properties of pre-BDNE for a special kind of belief correspondence – the perfect foresight.

Definition 4. A belief correspondence B_i of player i is the *perfect foresight* at a profile S, if for every t, $B_i(t, S_i^{OL}(t), H^S) = \{H^S\}$.

Theorem 5. ***Equivalence between pre-BDNE for perfect foresight and Nash equilibria***

Let $(\mathbb{I}, \Im, \lambda)$ *be a nonatomic measure space and let* $\sup_{S \in \mathbf{\Sigma}} \Pi^e(t, S)$ *and* $\inf_{S \in \mathbf{\Sigma}} \Pi^e(t, S)$ *be finite for every* t.

a) Let $\bar{S}$ *be a Nash equilibrium profile. If* B *is the perfect foresight at a profile* $\bar{S}$ *and the profiles* $\bar{S}^{i,d}$ *for a.e.* i *and every strategy* d *of player* i*, then for every* t $\bar{S}^{OL}(t) \in \mathrm{Argmax}_{a \in D_i(t, X(t), W_i(t))} \bar{\Pi}_i^e(t, H^{\bar{S}}, (\bar{S}^{OL}(t))^{i,a})$*, where the symbol* $\delta^{i,a}$ *for a static profile* δ *denotes the profile* δ *with strategy of player* i *changed to* a*, and* $\bar{S}_i^{OL}|\{t+1, \ldots, T\}$ *are consistent with the results of the player's optimizations used in the definition of* v_i*, i.e.* $\bar{S}_i^{OL}$ *is an element of the set*

$$\mathrm{Argmax}_{d:\mathbb{T} \to \mathbb{D}\ d(\tau) \in D_i(\tau, X(\tau), W_i(\tau))\ for\ \tau \geq t} \sum_{\tau=t}^{T} P_i(d(\tau), u(\tau), X(\tau), W_i(\tau)) \cdot \left(\frac{1}{1+r_i}\right)^{\tau - t} + G_i\left(X(T+1), W_i(T+1)\right) \cdot \left(\frac{1}{1+r_i}\right)^{T+1-t}.$$

b) Every Nash equilibrium profile $\bar{S}$ *is a pre-BDNE for a belief correspondence being the perfect foresight at* $\bar{S}$ *and the profile* $\bar{S}^{i,d}$ *for a.e.* i *and every strategy* d *of player* i*.*

c) Let $\bar{S}$ *be a pre-BDNE for a belief* B*. If* B *is the perfect foresight at* $\bar{S}$ *and the profiles* $\bar{S}^{i,d}$ *for a.e.* i *and every strategy* d *of player* i*, then choices of players are consistent with the results of their optimizations used in definition of* v_i*.*

d) If a profile $\bar{S}$ *is a pre-BDNE for a belief* B *being the perfect foresight at this* $\bar{S}$ *and* $\bar{S}^{i,d}$ *for a.e. player* i *and his every strategy* d*, then it is a Nash equilibrium.*

Proof. The proof is similar to analogous result in [20]. Nevertheless, introduction of private state variables makes it much more complicated.

First note that if the measure λ is nonatomic, then a choice of strategy by a single player influences neither u nor X. It only influences player's private state variable. Therefore, instead of looking for the best response to the profiles of the remaining players' strategies it is enough to look for best responses to the statistic of this profile, equal to the statistic of the whole profile.

In all the cases we shall consider player i outside the set of measure 0 of players for whom the condition of maximizing payoff or expected payoff does not hold.

a) We shall prove that along the perfect foresight path the equation for the expected payoff of player i becomes the Bellman equation for optimization of the actual payoff by player i and V_i coincides with the value function.

Formally, given the profile of the strategies of the remaining players coinciding with $\bar{S}$, with the statistic u and trajectory of the global system X, let us define the value function for the decision making problem of player i, $\widetilde{V}_i : \mathbb{T} \times \mathbb{W} \to \overline{\mathbb{R}}$.

$\widetilde{V}_i(t, w_i) =$

$= \sup_{d:\mathbb{T}\to\mathbb{D}\ d(\tau)\in D_i(\tau,X(\tau),W_i(\tau)) \text{ for } \tau\geq t} \sum_{\tau=t}^{T} P_i(d(\tau), u(\tau), X(\tau), W_i(\tau)) \cdot \left(\frac{1}{1+r_i}\right)^{\tau-t} +$

$+ G_i\left(X(T+1), W_i(T+1)\right) \cdot \left(\frac{1}{1+r_i}\right)^{T+1-t}$,

where W_i is defined recursively by $W_i(t) = w_i$ and $W_i(\tau{+}1) = \kappa(W_i(\tau), d(\tau), X(\tau), u(\tau))$.

In the finite horizon case $\widetilde{V}_i$ fulfills the Bellman equation

$\widetilde{V}_i(t, w_i) = \sup_{a\in D_i(t,X(t),w_i)} P_i(a, u(t), X(t), w_i) + \widetilde{V}_i(t{+}1, \kappa_i(w_i, a, X(t), u(t))) \cdot \left(\frac{1}{1+r_i}\right)$

with the terminal condition

$\widetilde{V}_i(T+1, w_i) = G_i\left(X(T+1), w_i\right)$.

In the infinite horizon case $\widetilde{V}_i$ also fulfills the Bellman equation, but the terminal condition sufficient for the solution of the Bellman equation to be the value function is different. The simplest one is $\lim_{t\to\infty} \widetilde{V}_i(t, W_i(t)) \cdot \left(\frac{1}{1+r_i}\right)^{t-t_0} = 0$ for every admissible W_i (see e.g. Blackwell [6] or Stokey and Lucas [12]). In this paper it holds by the assumption that the payoffs are bounded.

If we write the formula for $\widetilde{V}_i$ from the definition in the r.h.s. of the Bellman equation, then we get

$\widetilde{V}_i(t, w_i) = \sup_{a\in D_i(t,X(t),w_i)} P_i(a, u(t), X(t), w_i) + \left(\frac{1}{1+r_i}\right) \cdot$

$\sup_{d:\mathbb{T}\to\mathbb{D}\ d(\tau)\in D_i(\tau,X(\tau),W_i(\tau)) \text{ for } \tau\geq t+1} \sum_{\tau=t+1}^{T} P_i(d(\tau), u(\tau), X(\tau), W_i(\tau)) \cdot \left(\frac{1}{1+r_i}\right)^{\tau-t-1}$

$+ G_i\left(X(T+1), W_i(T+1)\right) \cdot \left(\frac{1}{1+r_i}\right)^{T+1-t}$ subject to

$W_i(t) = w_i$, $W_i(t{+}1) = \kappa(W_i(t), a, X(t), u(t))$ and $W_i(\tau{+}1) = \kappa(W_i(\tau), d(\tau), X(\tau), u(\tau))$ for $\tau > t$.

Note that the last supremum is equal to $\widetilde{V}_i(t+1, \kappa_i(w_i, a, X(t), u(t)))$, but also to $v_i(t+1, \kappa(w_i, a, X(t), u(t)), (X, u))$.

Since (X, u) is the only history in the belief correspondence along both $\bar{S}$ and all profiles $\bar{S}^{i,d}$, it is also equal to $V_i(t+1, \kappa(w_i, a, X(t), u(t)), \{(X, u)\})$.

Therefore $\widetilde{V}_i(t, w_i) = \sup_{a\in D_i(t,X(t),w_i)} P_i(a, u(t), X(t), w_i) +$

$+ \left(\frac{1}{1+r_i}\right) \cdot V_i(t+1, \kappa(w_i, a, X(t), u(t)), B_i(t, \bar{S}_i^{OL}(t), H^{\bar{S}})) =$

$= \sup_{a\in D_i(t,X(t),w_i)} \Pi_i^e(t, \bar{S}^{i,d_{t,a}})$, where by $d_{t,a}$ we denote such a strategy of player i that $d(t, X(t), w_i, H^{\bar{S}}) = a$ and at any other point of the domain it coincides with $\bar{S}_i$.

Let us note that for all t the set

$\operatorname{Argmax}_{d:\mathbb{T}\to\mathbb{D}\ d(\tau)\in D_i(\tau,X(\tau),W_i(\tau)) \text{ for } \tau\geq t} \sum_{\tau=t}^{T} P_i(d(\tau), u(\tau), X(\tau), W_i(\tau)) \cdot \left(\frac{1}{1+r_i}\right)^{\tau-t} +$

$+G_i\left(X(T+1),W_i(T+1)\right)\cdot\left(\frac{1}{1+r_i}\right)^{T+1-t}$ is both the set of open loop forms of strategies of player i being best responses of player i to the strategies of the remaining players along the profiles $\bar{S}$ and $\bar{S}^{i,d}$ and the set at which the supremum in the definition of the function v_i for a history being the perfect foresight along $\bar{S}$ and $\bar{S}^{i,d}$ is attained. We only have to show that $\bar{S}_i(t)\in \mathrm{Argmax}_{a\in D_i(t,X(t),w_i)}\bar{\Pi}_i^e(t,H^{\bar{S}},\bar{S}^{OL}(t)^{i,a})$.
By the definition this set is equal to
$\mathrm{Argmax}_{a\in D_i(t,X(t),w_i)}\ P_i(a,u(t),X(t),w_i)+$
$+\left(\frac{1}{1+r_i}\right)\cdot V_i(t+1,W_i^{\bar{S}^{OL}}(t+1),B_i(t,\bar{S}_i^{OL}(t),H^{\bar{S}}))=$
$=\mathrm{Argmax}_{a\in D_i(t,X(t),w_i)}\ P_i(t,a,u(t),X(t),w_i)+\left(\frac{1}{1+r_i}\right)\cdot\widetilde{V}_i(t+1,W_i^{\bar{S}^{OL}}(t+1))$, which, by the Bellman equation, defines the value of the best response at time t, which contains $\bar{S}_i(t)$, since $\bar{S}$ is an equilibrium profile.

b) An immediate conclusion from a)

d) Given $\bar{S}$, we consider $\widetilde{V}_i$ defined as in the proof of a).
By the definition of pre-BDNE
$\bar{S}_i(t)\in \mathrm{Argmax}_{a\in D_i(t,X(t),w_i)}\bar{\Pi}^e(t,H^{\bar{S}},\bar{S}^{OL}(t)^{i,a})=$
$=\mathrm{Argmax}_{a\in D_i(t,X(t),w_i)}\ P_i(a,u(t),X(t),w_i)+$
$+\left(\frac{1}{1+r_i}\right)\cdot V_i(t+1,\kappa(w_i,a,X(t),u(t)),B_i(t,\bar{S}_i^{OL}(t),H^{\bar{S}}))=$
$=\mathrm{Argmax}_{a\in D_i(t,X(t),w_i)}\ P_i(a,u(t),X(t),w_i)+$
$+\left(\frac{1}{1+r_i}\right)\cdot V_i(t+1,\kappa(w_i,a,X(t),u(t)),\{(X,u)\})=$
$=\mathrm{Argmax}_{a\in D_i(t,X(t),w_i)}\ P_i(t,a,u(t),X(t),w_i)+$
$+\left(\frac{1}{1+r_i}\right)\cdot\max_{d:\mathbb{T}\to\mathbb{D}\ d(\tau)\in D_i(\tau,X(\tau),W_i(\tau))\text{ for }\tau\ge t+1}\sum_{\tau=t+1}^{T}P_i(d(\tau),u(\tau),X(\tau),W_i(\tau))$
$\cdot\left(\frac{1}{1+r_i}\right)^{\tau-t}+G_i\left(X(T+1),W_i(T+1)\right)\cdot\left(\frac{1}{1+r_i}\right)^{T+1-t}$ (for W_i defined by $W_i(t)=w_i$ and $W_i(\tau+1)=\kappa(W_i(\tau),d(\tau),X(\tau),u(\tau))$ for $\tau>t$)
$=\mathrm{Argmax}_{a\in D_i(t,X(t),w_i)}\ P_i(a,u(t),X(t),w_i)+\left(\frac{1}{1+r_i}\right)\cdot\widetilde{V}_i(t+1,\kappa(w_i,a,X(t),u(t)))$.

If we add the fact that
$\widetilde{V}_i(t,w_i)=\max_{a\in D_i(t,X(t),w_i)}P_i(a,u(t),X(t),w_i)+\left(\frac{1}{1+r_i}\right)\cdot\widetilde{V}_i(t+1,\kappa(w_i,a,X(t),u(t)))$, then, by the Bellman condition, the set
$\mathrm{Argmax}_{a\in D_i(t,X(t),w_i)}\ P_i(a,u(t),X(t),w_i)+\left(\frac{1}{1+r_i}\right)\cdot\widetilde{V}_i(t+1,\kappa(w_i,a,X(t),u(t)))$ represents the value of the optimal choice of player i at time t, given X, u and w_i. Since we have this property for a.e. i, the profile defined in this way is a Nash equilibrium.

c) By d) and a). ■

We can also prove an equivalence theorem for repeated games – dynamic games without state variables, which can be modelled as both global and private state sets being singletons.
Not surprisingly, analogous theorem from Wiszniewska-Matyszkiel [20] can be cited and the proof does not change.

Theorem 6. *Let $\mathfrak{G}$ be a repeated game with a belief correspondence not dependent on players' own strategies in which payoffs and anticipated payoffs are bounded for a.e. player.*

a) If $(\mathbb{I}, \Im, \lambda)$ is a nonatomic measure space, then a profile S is a pre-BDNE if and only if it is a Nash equilibrium.

b) Every profile S with strategies of a.e. player being independent of histories is a pre-BDNE if and only if it is a Nash equilibrium.

Self-verification of beliefs and BDNE

The concept of pre-BDNE lacks a kind of guarantee that after some stage of the game players can still have the same beliefs, i.e. that beliefs are not contradicted by players' observations.

First we state what we mean by potential and perfect self-verification.

Definition 7. a) A collection of beliefs $\{B_i\}_{i\in\mathbb{I}}$ is *perfectly self-verifying* if for every pre-BDNE $\bar{S}$ for B for a.e. $i \in \mathbb{I}$ we have $H^{\bar{S}} \in B_i(t, \bar{S}_i^{OL}(t), H^{\bar{S}})$.

b) A collection of beliefs $\{B_i\}_{i\in\mathbb{I}}$ is *potentially self-verifying* if there exists a pre-BDNE $\bar{S}$ for B such that for a.e. $i \in \mathbb{I}$ we have $H^{\bar{S}} \in B_i(t, \bar{S}_i^{OL}(t), H^{\bar{S}})$.

c) A collection of beliefs $\{B_i\}_{i\in\mathbb{J}}$ of a set of players $\mathbb{J}$ is *perfectly self-verifying* against beliefs of the other players $\{B_i\}_{i\in\setminus\mathbb{J}}$ if for every pre-BDNE $\bar{S}$ for $\{B_i\}_{i\in\mathbb{I}}$ for a.e. $i \in \mathbb{J}$ we have $H^{\bar{S}} \in B_i(t, \bar{S}_i^{OL}(t), H^{\bar{S}})$.

d) A collection of beliefs $\{B_i\}_{i\in\mathbb{J}}$ of a set of players $\mathbb{J}$ is *potentially self-verifying* against beliefs of the other players $\{B_i\}_{i\in\setminus\mathbb{J}}$ if there exists a pre-BDNE $\bar{S}$ for $\{B_i\}_{i\in\mathbb{I}}$ such that for a.e. $i \in \mathbb{J}$ we have $H^{\bar{S}} \in B_i(t, \bar{S}_i^{OL}(t), H^{\bar{S}})$.

And, finally, the notion of BDNE.

Definition 8. a) A profile $\bar{S}$ is a *belief-distorted Nash equilibrium (BDNE) for* a collection of beliefs $B = \{B_i\}_{i\in\mathbb{I}}$ if $\bar{S}$ is a pre-BDNE for B and for a.e. i and every t $H^{\bar{S}} \in B_i(t, \bar{S}_i^{OL}(t), H^{\bar{S}})$

b) A profile $\bar{S}$ is a *belief-distorted Nash equilibrium* (*BDNE*) if there exists a collection of beliefs $B = \{B_i\}_{i\in\mathbb{I}}$ such that $\bar{S}$ is a BDNE for B.

Remark 9. Theorems 5 and 6 remain valid with pre-BDNE replaced by BDNE. ■

Simplified stock exchange

Here we present an example of games with distorted information – a model of stock exchange simplified in order to avoid enormous complexity inherent to real world large systems. More complex models of stock exchange with more than one asset sold, and coping with the problem of distorted information and self-verification of various prognostic techniques against another prognostic techniques, including random ones, were considered by the author in [19].

We examine a model of stock exchange with two assets: money m of interest rate r and a share s traded at the stock exchange. This share pays a deterministic dividend $\{A_t\}_{t\in\mathbb{N}}$. Both assets are infinitely divisible. The transaction cost is linear with a rate $C > 0$. The horizon of players' optimization is $T = +\infty$.

The set of players is either the unit interval with the Lebesgue measure or a large finite set with the normalized counting measure.

Each player i has an initial portfolio of assets $(\bar{m}_i, \bar{s}_i) \in \mathbb{R}^2_+$ with at least one coordinate strictly positive. Portfolios constitute private states of players $w_i = (m_i, s_i) \in \mathbb{R}^2_+$. Prices are, for simplicity of notation, positive integers.

We define the state of the global system being the price at the previous period $X(t) = p(t-1)$, with $p(t_0 - 1) > 0$ given. This auxiliary state variable is introduced, since price at time t, as we shall see in the sequel, depends on previous price.

At each stage players state their bid and ask prices $p_i^B, p_i^S \in \mathbb{N}$ – at which they want to buy or sell, respectively, and which constitute their decisions with the available decisions correspondence $D_i \equiv \mathbb{D} = \mathbb{N}^2$, where the first coordinate denotes p_i^B while the second p_i^S. These, together with players' actual portfolios, constitute orders – in their orders players buy or sell as much as they can. There are no other constraints for orders. The statistic function will be three dimensional with the actual market price as the first coordinate. The market price p is a function of players' decisions as follows.

Given a profile of players strategies at a fixed time instant, first the aggregate supply and demand functions $AS : \mathbb{N} \to \mathbb{R}_+$ and $AD : \mathbb{N} \to \overline{\mathbb{R}}_+$ are calculated. The *aggregate* (or *market*) *supply* at a price q is defined by

$$AS(q) = \int_{\mathbb{I}} s_i(t) \cdot \mathbf{1}_{p_i^S(t) \le q} d\lambda(i),$$

while the *aggregate* (or *market*) *demand* at a price q is defined by

$$\begin{aligned} AD(q) &= \int_{\mathbb{I}} \frac{m_i}{q \cdot (1+C)} \cdot \mathbf{1}_{p_i^B(t) \ge q} d\lambda(i) \text{ for } q > 0, \\ AD(0) &= +\infty, \end{aligned}$$

where the symbol $\mathbf{1}_{condition}$ denotes 1 when the condition is fulfilled and 0 otherwise. If we consider a finite set of players with the counting measure, then the integral is simply the sum.

The market mechanism considered in the paper at each time $t \ge t_0$ returns a price, called the *market price*, contained in the set $[(1-h)\cdot p(t-1), (1+h)\cdot p(t-1)] \cap \mathbb{N}\backslash\{0\}$, where $h > 0$, being a constraint for variability of prices, is large compared to r.

The procedure of calculating the market price operates as follows.
First we look for a strictly positive price maximizing a lexicographic order with criteria, starting from the most important:
1. *maximizing volume* i.e. the function $Vol(q) = \min(AS(q), AD(q))$;
2. *minimizing disequilibrium* i.e. the function $Dis(q) = |AS(q) - AD(q)|$;
3. *minimizing the number of shares in selling orders with price limit less then q and buying orders with price limits higher than q* i.e. the function $N(q) = (AS(q-1) - AD(q))^+ + (AD(q+1) - AS(q))^+$, where the symbol $(x)^+ = \max(0, x)$;
4. *minimizing the absolute value of the difference between the q and the previous price* i.e. $|q - p(t-1)|$.
If the resulting q is not in the interval $[(1-h)\cdot p(t-1), (1+h)\cdot p(t-1)]$, then we project it on $[(1-h)\cdot p(t-1), (1+h)\cdot p(t-1)] \cap \mathbb{N}\backslash\{0\}$.
The resulting unique price constitutes the first of three coordinates of the statistic function.

The procedure is copied from the regulations of the Warsaw Stock Exchange for the single price auction system, but it is used also at other stock exchanges.

In order that the model was complete, in case when the supply at the market price is not equal to the demand, each of the orders on the excess side is reduced at the same rate.

The buying orders are multiplied by

$$BR(t)=\begin{cases}1 & \text{if } AD(p(t))\leq AS(p(t)),\\ \frac{AS(p(t))}{AD(p(t))} & \text{otherwise,}\end{cases}$$

while the selling orders by

$$SR(t)=\begin{cases}1 & \text{if } AD(p(t))\geq AS(p(t)),\\ \frac{AD(p(t))}{AS(p(t))} & \text{otherwise.}\end{cases}$$

These BR and SR constitute the remaining two coordinates of the statistic.

The instantaneous payoff is equal to the change of money stock exactly at the time instant considered: $P_i(t,(p_i^B,p_i^S),(p,BR,SR),x,(m_i,s_i))=$
$= SR\cdot s_i\cdot p\cdot(1-C)\cdot \mathbf{1}_{p_i^S\leq p}-BR\cdot m_i\cdot \mathbf{1}_{p_i^B\geq p}+A_t\cdot s_i$.

The payoff at a profile $S=\left\{(p_i^B,p_i^S):\mathbb{T}\to\mathbb{N}^2\right\}_{i\in\mathbb{I}}$ is, therefore, equal to

$$\Pi_i(S)=\sum_{t=t_0}^{\infty}\frac{SR(t)\cdot s_i(t)\cdot p(t)\cdot(1-C)\cdot\mathbf{1}_{p_i^S(t)\leq p(t)}-BR(t)\cdot m_i(t)\cdot\mathbf{1}_{p_i^B(t)\geq p(t)}+A_t\cdot s_i(t)}{(1+r)^t}.$$

Now we can complete the definition of our game by defining the behaviour of the state variables.

The global system fulfills $X(t+1)=p(t)$, which determines ϕ in the obvious way, while the private state variables change as follows.

Money fulfil $m_i(t+1)=(1+r)\cdot\Big(m_i(t)(1-BR(t)\cdot\mathbf{1}_{p_i^B(t)\geq p(t)})+$
$+s_i(t)\cdot(SR(t)\cdot p(t)\cdot(1-C)\cdot\mathbf{1}_{p_i^S(t)\leq p(t)}+A_t)\Big)$,

while shares $s_i(t+1)=s_i(t)-SR(t)\cdot s_i(t)\cdot\mathbf{1}_{p_i^S(t)\leq p(t)}+BR(t)\cdot\frac{m_i(t)}{(1+C)\cdot p(t)}\cdot\mathbf{1}_{p_i^B(t)\geq p(t)}$, which defines the functions κ_i.

In this game we shall consider various belief correspondences.

Threshold prices and pre-BDNE for simplified stock exchange

For a history H and a time instant t we define $\bar{p}_i^S(t)\in\mathbb{N}$, called the *selling threshold price of player* i who has $s_i(t)>0$, as the minimum of the prices such that for all histories in the current belief the anticipated payoff for selling at this price is not less then the anticipated payoff for not selling at all at time t, and $\bar{p}_i^B(t)\in\mathbb{N}\cup\{+\infty\}$, called the *buying threshold price of player* i who has $m_i(t)>0$, as the maximum of the prices such that for all histories in the current belief the anticipated payoff for buying at this price is not less then the anticipated payoff for not buying at all at time t.

Formally, we have the following definition.

Definition 10. Threshold prices

Let us consider a time instant t, a history H and player i who has $w_i=W_i(t)$ resulting from some previous realization of the profile. We introduce an auxiliary anticipated payoff $\hat{\Pi}_i^e$ which does not take into account player's influence on the

current statistic of the profile $\hat{\Pi}_i^e : \mathbb{T} \times \mathbb{H} \times \mathbb{N} \times \mathbb{D} \times \mathbb{W} \to \overline{\mathbb{R}}$ by
$\hat{\Pi}_i^e(t, H, p, a, w_i) = P_i(t, a, (p, 1, 1), x, w_i) + \frac{V_i(t+1, \kappa_i(w_i, a, x, (p,1,1)), B_i(t,a,H))}{1+r}$ for some x. Note that in our case the functions considered are independent on x, so the definition is correct.

If player i has $s_i(t) > 0$, then the price $\bar{p}_i^S(t) \in \mathbb{N} \cup \{+\infty\}$ defined by
$\bar{p}_i^S(t) = \min\{p_i^S \in \mathbb{N}\backslash\{0\} : \hat{\Pi}_i^e\left(t, H, p, \left(p_i^B, p_i^S\right), w_i\right) \geq \hat{\Pi}_i^e\left(t, H, p, \left(p_i^B, p+1\right), w_i\right)$ for every $p_i^B, p \in \mathbb{N}\}$ if this set is nonempty,
$\bar{p}_i^S(t) = +\infty$ otherwise,
is called the *selling threshold price of player i*.

If player i has $m_i(t) > 0$, then the price $\bar{p}_i^B(t) \in \mathbb{N} \cup \{+\infty\}$ defined by
$\bar{p}_i^B(t) = \max\{p_i^B \in \mathbb{N} : \hat{\Pi}_i^e\left(t, H, p, \left(p_i^B, p_i^S\right), w_i\right) \geq \hat{\Pi}_i^e\left(t, H, p, \left(p-1, p_i^S\right), w_i\right)$ for every $p_i^S, p \in \mathbb{N}\}$
is called the *buying threshold price of player i*. This definition is correct because the set is nonempty – since the inequality holds for $p_i^B = 0$.

We are interested in the existence and properties of the threshold prices.

Proposition 11. *Consider a belief B independent on players' own decisions and the game $\mathfrak{G}_{t,H^S,W(t)}$ with $W(t)$ such that for player i $m_i(t), s_i(t) > 0$ and $\int_{\mathbb{I}\backslash\{i\}} m_j(t)d\lambda(j), \int_{\mathbb{I}\backslash\{i\}} s_j(t)d\lambda(j) > 0$.*

a) Both $\bar{p}_i^B(t)$ and $\bar{p}_i^S(t)$ are well defined and either $\bar{p}_i^B(t) < \bar{p}_i^S(t)$ or they are both $+\infty$.

b) Consider profiles of the form $\delta = \left\{(p_i^B(t), p_i^S(t))\right\}_{i\in\mathbb{I}}$.

(i) If $p_i^S(t) < \bar{p}_i^S(t)$, $\bar{p}_i^S(t) \geq (1-h)\cdot p(t-1)$ and $p_i^S(t) \leq (1+h)\cdot p(t-1)$, then $\bar{\Pi}_i^e\left(t, H^S, \delta\right) \leq \bar{\Pi}_i^e\left(t, H^S, \delta^{i,(p_i^B(t), \bar{p}_i^S(t))}\right)$, with strict inequality for at least one δ.

(ii) If $p_i^B(t) > \bar{p}_i^B(t)$, $\bar{p}_i^B(t) \leq (1+h)\cdot p(t-1)$ and $p_i^B(t) \geq (1-h)\cdot p(t-1)$, then $\bar{\Pi}_i^e\left(t, H^S, \delta\right) \leq \bar{\Pi}_i^e\left(t, H^S, \delta^{i,(\bar{p}_i^B(t), p_i^S(t))}\right)$, with strict inequality for at least one δ.

c) Assume a continuum of players and profiles of the form $\delta = \left\{(p_i^B(t), p_i^S(t))\right\}_{i\in\mathbb{I}}$. For every δ we have $\bar{\Pi}_i^e\left(t, H^S, \delta\right) \leq \bar{\Pi}_i^e\left(t, H^S, \delta^{i,(p_i^B(t), \bar{p}_i^S(t))}\right)$ and $\bar{\Pi}_i^e\left(t, H^S, \delta\right) \leq \bar{\Pi}_i^e\left(t, H^S, \delta^{i,(\bar{p}_i^B(t), p_i^S(t))}\right)$.

Sketch of the proof a) First note that V_i is strictly increasing in both s_i and m_i.

Assume that the inequality does not hold at time t. This means that there exists a price $\bar{p} \in [\bar{p}_i^S(t), \bar{p}_i^B(t)]$.

Let player i choose his threshold prices. If the market price is $\bar{p}$, then player i both buys and sells at time t and pays commission twice. So he can increase his anticipated payoff $\hat{\Pi}^e$ by either inreasing $p_i^S(t)$ or by decreasing $p_i^B(t)$. To prove this, we can increase the set of possible strategies allowing players to buy or sell using only a fraction of their assets and change the model respectively. Note that the actual strategies in our game can be viewed as extremal points of such an enlarged strategy set corresponding to fractions 0 and 1.

After this change the fact that buying and selling nontrivially at the same time instant decreases payoff is immediate. Because of linearity of the model, if there exists an optimal strategy, then there exists an optimal strategy consisting of extreme points.

b) (i) The inequality "$\leq$" is from definition and by the fact that choosing $p_i^S(t)$ lower than $\bar{p}_i^S(t)$ can influence price or reductions only if the resulting price is less than $\bar{p}_i^S(t)$ and resulting SR is positive.
Tedious analysis of possible cases shows that a situation in which player who wants to buy shares at lower price will not gain using such a strategy.

The strict inequality holds for a profile with $p(t) = \bar{p}_i^S(t)$ and $SR(t) \neq 0$ (which is admissible);

(ii) analogously.

c) In games with continuum of players chosing a strategy by a player does not influence neither price, nor reductions.

The cases which are not covered by b) can be proven by the fact that either nothing changes or in such cases player can only decrease his payoff because of not buying or not selling at a price at which it leads to increase of anticipated payoff. ■

Corollary 12. *In games with continuum of players and beliefs independent of player's own choice a profile consisting of pairs of respective threshold prices is a pre-BDNE.* ■

.1 Fundamental beliefs

In the "obvious" Nash equilibria, in which players believe that prices are equal to the fundamental value $F_t = \sum_{s=t+1}^{T} A_s \cdot (\frac{1}{1+r})^{s-t}$, prices are close to the fundamental value as follows.

First, lets us formally define what we mean by fundamental beliefs.

Definition 13. A belief B_i is called a *fundamental belief* if for every t, every H and every $a \in \mathbb{D}$
$B_i(t, a, H) \subset \{(X, (p, BR, SR)) : p(t) = F_t, X(t) = p(t-1) \text{ for all } t\}$.

Of course, such beliefs only make sense if for every t $F_t \in \mathbb{N}$. Otherwise, we can substitute $p(t) = F_t$ by $|p(t) - F_t| \leq \frac{1}{2}$ and we shall obtain similar results with similar proofs.

Formally, we state the following proposition.

Proposition 14. ***Beliefs based on fundamental analysis***

Assume that a.e. player has fundamental beliefs.

For every time instant t such that
$\left(\left\lfloor \frac{F_t}{1+C} \right\rfloor, \left\lceil \frac{F_t}{1-C} \right\rceil\right) \cap [(1-h) \cdot p(t-1), (1+h) \cdot p(t-1)] \cap \mathbb{N}\backslash\{0\} \neq \emptyset$ *and such that the set of players with positive $m_i(t)$ and the set of players of positive $s_i(t)$ are of positive measure we have*

a) The threshold prices are $\bar{p}_i^B(t) = \left\lfloor \frac{F_t}{1+C} \right\rfloor$ and $\bar{p}_i^S(t) = \left\lceil \frac{F_t}{1-C} \right\rceil$.

b) Every profile such that $p_i^B(t) \le \left\lceil \frac{F_t}{1-C} \right\rceil$ *and* $p_i^S(t) \ge \left\lfloor \frac{F_t}{1+C} \right\rfloor$ *for a.e. player* i *is a pre-BDNE for fundamental beliefs.*
At this profile $p(t) \in \left(\left\lfloor \frac{F_t}{1+C} \right\rfloor, \left\lceil \frac{F_t}{1-C} \right\rceil \right)$ *and there is no trade (i.e.* $Vol(p(t)) = 0$*).*

Proof. Take any history H and player i for whom both $m_i(t), s_i(t) > 0$.

a) We shall prove $\bar{p}_i^B(t) = \left\lfloor \frac{F_t}{1+C} \right\rfloor$ and $\bar{p}_i^S(t) = \left\lceil \frac{F_t}{1-C} \right\rceil$ are the threshold prices.

A simple comparison between trade and no trade strategies shows that for every decision a $V_i(t+1,(m_i(t+1),s_i(t+1)),B_i(t,a,H)) = s_i(t+1)\cdot(1+r)\cdot F_t$ and the maximum is attained for a strategy at which player i does not trade.
Therefore, in order to calculate the threshold prices we have to maximize

$$SR(t)\cdot s_i(t)\cdot p(t)\cdot(1-C)\cdot \mathbf{1}_{p_i^S(t)\le p(t)} - BR(t)\cdot m_i(t)\cdot \mathbf{1}_{p_i^B(t)\ge p(t)} + A_t\cdot s_i(t) +$$
$$+\frac{1}{1+r}\cdot s_i(t+1)\cdot(1+r)\cdot F_t =$$
$$= SR(t)\cdot s_i(t)\cdot p(t)\cdot(1-C)\cdot \mathbf{1}_{p_i^S(t)\le p(t)} - BR(t)\cdot m_i(t)\cdot \mathbf{1}_{p_i^B(t)\ge p(t)} + A_t\cdot s_i(t)+$$
$$+\left(s_i(t) - SR(t)\cdot s_i(t)\cdot \mathbf{1}_{p_i^S(t)\le p(t)} + BR(t)\cdot \frac{m_i(t)}{p(t)\cdot(1+C)}\cdot \mathbf{1}_{p_i^B(t)\ge p(t)}\right)\cdot F_t =$$
$$= SR(t)\cdot s_i(t)\cdot \mathbf{1}_{p_i^S(t)\le p(t)}\cdot(p(t)\cdot(1-C)-F_t) +$$
$$BR(t)\cdot m_i(t)\cdot \mathbf{1}_{p_i^B(t)\ge p(t)}\cdot\left(\frac{F_t}{p(t)\cdot(1+C)}-1\right) + A_t\cdot s_i(t) + s_i(t)\cdot F_t$$

for $BR(t) = SR(t) = 1$.

The last two terms of the maximized function are independent of player's own decision at time t, therefore we can equivalently take into account maximization of the function

$$f(p_i^B(t),p_i^S(t)) = 1\cdot s_i(t)\cdot \mathbf{1}_{p_i^S(t)\le p(t)}\cdot(p(t)\cdot(1-C)-F_t) +$$
$$+\left(1\cdot m_i(t)\cdot \mathbf{1}_{p_i^B(t)\ge p(t)}\cdot\left(\frac{F_t}{p(t)\cdot(1+C)}-1\right)\right).$$

If $p(t) < \frac{F_t}{(1-C)}$, then $p(t)\cdot(1-C)-F_t < 0$, therefore the first term of f is at most 0 and the maximal value is attained if player i chooses a price at which he does not sell i.e. $p_i^S(t) > p(t)$. Since it holds for all $p(t) < \frac{F_t}{(1-C)}$, we get $p_i^S(t) \ge \frac{F_t}{(1-C)}$. Since prices are integers, we have $p_i^S(t) \ge \left\lceil \frac{F_t}{(1-C)} \right\rceil$.
If $p(t) \ge \frac{F_t}{(1-C)}$, then $p(t)\cdot(1-C)-F_t \ge 0$, therefore the first term of f is nonnegative. It is 0 if $p_i^S(t) > p(t)$ and for $p(t) > \frac{F_t}{(1-C)}$ it is strictly positive whenever $p_i^S(t) \le p(t)$. Therefore $\bar{p}_i^S(t) = \left\lceil \frac{F_t}{(1-C)} \right\rceil$.

An analogous reasoning applies to the second term and it proves that $\bar{p}_i^B(t) = \left\lfloor \frac{F_t}{(1+C)} \right\rfloor$.

b) We have to prove that such a strategy of player i is a best response to the strategies of the remaining players.

Assume that all players besides i choose $\left(\left\lfloor \frac{F_t}{1+C} \right\rfloor, \left\lceil \frac{F_t}{1-C} \right\rceil \right)$. If player i also chooses this strategy, then he does not trade at time t, and his anticipated payoff is $s_i(t)\cdot(A_t+F_t)$. If he chooses any strategy with $p_i^B(t) < \left\lceil \frac{F_t}{1-C} \right\rceil$ and $p_i^S(t) > \left\lfloor \frac{F_t}{1+C} \right\rfloor$, he does not change his payoff since there is still no trade. If he chooses any strategy with $p_i^S(t) \le \left\lfloor \frac{F_t}{1+C} \right\rfloor$ then the first term of f is negative and if he chooses $p_i^B(t) \geq \left\lceil \frac{F_t}{1-C} \right\rceil$, then the second term of f is negative, while the other term is, as we have just proven, at most 0, therefore he only decreases his payoff.

We shall calculate the price at the profile at which players choose $\left(\left\lfloor \frac{F_t}{1+C} \right\rfloor, \left\lceil \frac{F_t}{1-C} \right\rceil \right)$: the market price $p(t)$ will be the price within the interval $\left(\left\lfloor \frac{F_t}{1+C} \right\rfloor, \left\lceil \frac{F_t}{1-C} \right\rceil \right)$ closest to $p(t-1)$ if this interval has a nonempty intersection with $[(1-h) \cdot p(t-1), (1+h) \cdot p(t-1)] \cap \mathbb{N} \backslash \{0\}$. For this price there is no trade. ■

.2 Speculative bubbles

However, we are interested in more counter-intuitive pre-BDNE, which correspond to speculative bubbles. They usually reflect strong trends in technical analysis. In the sequel we shall prove that belief in strong trends may be self verifying.

Proposition 15. ***Beliefs in strong trends***

Assume B is such that for every t at which a nonnegligible set of players i has $m_i(t) > 0$, and such that for every action a of a.e. player i with positive $m_i(t)$ or $s_i(t)$, every history H and every history in $B_i(t, a, H)$,

1. there exists $\bar{\tau} > t$ such that $\frac{p(\bar{\tau}) \cdot (1-C)}{(1+r)^{\bar{\tau}-t}} > (\lfloor p(t-1) \cdot (1+h) \rfloor + 1) \cdot (1+C)$ and $SR(\bar{\tau}) = \{1\}$ and

2. for every $t < \tau \leq \bar{\tau}$ $p(\tau) \geq p(\tau - 1) \cdot (1+r)$.

Assume also that for this B the anticipated payoff is always finite.

Then

a) The threshold prices fulfil $\bar{p}_i^B(t), \bar{p}_i^S(t) > \lfloor (1+h) \cdot p(t-1) \rfloor$.

b) Every profile such that $p_i^B(t) \geq \lfloor (1+h) \cdot p(t-1) \rfloor$ and $p_i^S(t) > \lfloor (1+h) \cdot p(t-1) \rfloor$ is a pre-BDNE for B.
At this pre-BDNE $p(t) = \lfloor (1+h) \cdot p(t-1) \rfloor$, $BR(t) = 0$ and $SR(t) = 1$ for every t.

c) At every pre-BDNE we have for every t and a.e. i with positive $s_i(t)$ $p_i^S(t) > p(t)$ and there is no trade.

d) In the case of continuum of players at every pre-BDNE we have also for every t
$p(t) = \min \left(\lfloor (1+h) \cdot p(t-1) \rfloor, \operatorname{essinf}_{s_i(t)>0} p_i^S(t) - 1 \right)$, $BR(t) = 0$, $SR(t) = 1$.

Proof. a) The threshold prices for these beliefs fulfil $\bar{p}^B(t), \bar{p}_i^S(t) > \lfloor (1+h) \cdot p(t-1) \rfloor$. It is enough to prove this for $\bar{p}^B(t)$, since the second inequality results from the first one by proposition 11.

We prove this by comparing a non-trade strategy with a strategy at which player i buys at the market price at time t and he sells at time $\bar{\tau} > t$ for which $\inf_{(X,(p,BR,SR)) \in B_i(t,a,H): SR(\bar{\tau})=1} \frac{p(\bar{\tau})}{(1+r)^{\bar{\tau}-t}}$ is maximal. This maximum is attained, since otherwise the anticipated payoff is infinite, and it constitutes the optimal d in the definition of v_i.
We get the required inequality, whatever are $p(t)$, $SR(t)$, $BR(t)$ and the other price limit at time t.

b) If we consider such a profile, then any strategy with $p_i^B(t) \geq \lfloor (1+h) \cdot p(t-1) \rfloor$ and $p_i^S(t) > \lfloor (1+h) \cdot p(t-1) \rfloor$ maximizes the anticipated payoff of i.

For such a profile of strategies the market mechanism returns the price $\lfloor (1+h)\cdot p(t-1)\rfloor$, $BR(t)=0$ and $SR(t)=1$ for every t.

c) Assume that $p_i^S(t) \le \lfloor (1+h)\cdot p(t-1)\rfloor$ for i in a subset of positive measure of the set of players for which $s_i(t)>0$.

Then if $p(t) \in \left[p_i^S(t), \lfloor (1+h)\cdot p(t-1)\rfloor\right]$ and $SR(t)>0$, player i can increase his anticipated payoff by changing $p_i^S(t)$ to $\lfloor (1+h)\cdot p(t-1)\rfloor+1$.

Let us check, whether $SR(t)=0$ is possible. In such a case $AS(p(t))>0$ while $AD(p(t))=0$, which implies that for almost every i for whom $m_i(t)>0$ we have $p_i^B(t)<p(t)$. Let us consider any player with positive $m_i(t)$ from this set. By changing his $p_i^B(t)$ to $p(t)$, player i does not change the price above $\lfloor (1+h)\cdot p(t-1)\rfloor$ (since $p(t)\le \lfloor (1+h)\cdot p(t-1)\rfloor$) and he does not decrease $BR(t)$ to 0 (since $AS(p)>0$ for $p\ge p(t)$), so he can buy at a price below his threshold buying price, and, consequently, he increases his anticipated payoff, which contradicts the fact that the profile was a pre-BDNE.

Therefore, in such a case $p(t)<p_i^S(t)$ for a.e. player, which implies no trade.

d) Let us consider a profile being a pre-BDNE and a time instant t. By c) we know that $p(t)<p_i^S(t)$ for a.e. i such that $s_i(t)>0$.

First we shall prove that in the continuum of players case a subset of i for whom $m_i(t)>0$ of positive measure has $p_i^B(t)\ge p(t)$.

Assume the converse. We know that $p(t)\le \lfloor (1+h)\cdot p(t-1)\rfloor < \bar{p}_i^B(t)$. So if $BR(t)>0$, then, by nonatomicity of the space of players, every player i for whom $m_i(t)>0$ and $p_i^B(t)<p(t)$ can increase his payoff by choosing some $p_i^B(t)\ge p(t)$ (and he does not affect nor p nor BR). This contradicts the pre-BDNE condition.

By c) we have that $AS(p(t))=0$. Therefore we have $SR(t)=1$. We have just excluded $BR(t)>0$.

So now let us consider $BR(t)=0$, which means that $AD(p(t))>0$, i.e. for a set of players of positive measure with positive $m_i(t)$ we have $p_i^B(t)\ge p(t)$.

We shall prove that this implies $p(t)=\min\left(\lfloor (1+h)\cdot p(t-1)\rfloor, \operatorname{essinf}_{s_i(t)>0} p_i^S(t)-1\right)$.

Indeed, either $Vol(q)=0$ for every q or it is positive for some q – by the form of p_i^S it is obtained for some $q>\lfloor (1+h)\cdot p(t-1)\rfloor$.

In the latter case $Vol(p(t))=0$ and $p(t)$ is obtained from a value greater than $\lfloor (1+h)\cdot p(t-1)\rfloor$ by projection on the set $[\lceil (1-h)\cdot p(t-1)\rceil, \lfloor (1+h)\cdot p(t-1)\rfloor]\cap(\mathbb{N}\setminus\{0\})$, since otherwise the equilibrating mechanism chooses q maximizing $Vol(q)$. In such a case $p(t)=\lfloor (1+h)\cdot p(t-1)\rfloor$.

If $Vol\equiv 0$ and $p(t)<\lfloor (1+h)\cdot p(t-1)\rfloor$,
then $\operatorname{essinf}_{s_i(t)>0} p_i^S(t) > \operatorname{esssup}_{m_i(t)>0} p_i^B(t)$.
Let us assume that for i in a subset of positive measure of the set of players with positive $m_i(t)$ we have $p_i^B(t)<\lfloor (1+h)\cdot p(t-1)\rfloor$.

As we have proven, $BR(t)=0$, $SR(t)=1$, $AD(p(t))>0$ and $AS(p(t))=0$.

Let us check what is the market price in such a situation.

The second condition restricting the price interval is minimization of disequilibrium. Let us assume that $\bar{p}<\operatorname{essinf}_{s_i(t)>0} p_i^S(t)-1$ is the market price. By the condition of minimization of disequilibrium and the facts that $AD(q)$ is a nonincreasing function of q and that in our case disequilibrium for $q<\operatorname{essinf}_{s_i(t)>0} p_i^S(t)$ is equal to

$AD(q)$, we get the result that also every price q in the interval $\left[\bar{p}, \operatorname{essinf}_{s_i(t)>0} p_i^S(t) - 1\right]$ minimizes $Dis(q)$.

The function $N(q)$ is either always zero at this set or it attains its minimum over this set at $\operatorname{essinf}_{s_i(t)>0} p_i^S(t) - 1$. In the latter case the market price cannot be less than $\operatorname{essinf}_{s_i(t)>0} p_i^S(t) - 1$. In the former case we have $AD(\bar{p}) = 0$, which contradicts our assumption.

Thus we have proven that $p(t) \geq \min\left(\lfloor (1+h) \cdot p(t-1) \rfloor, \operatorname{essinf}_{s_i(t)>0} p_i^S(t) - 1\right)$, $AS(p(t)) = 0$ while $AD(p(t)) > 0$.
Therefore, $SR(t) = 1$ and $BR(t) = 0$ at every pre-BDNE.
On the other hand, we have $p(t) < \operatorname{essinf}_{s_i(t)>0} p_i^S(t)$ and $p(t) \leq \lfloor (1+h) \cdot p(t-1) \rfloor$, which ends the proof. ■

Note that we made no comparison with fundamental values – the prices can grow because of players beliefs, whatever the fundamental values are, even if we have no uncertainty about fundamental values – speculative bubbles may happen even at a world with no external uncertainty.

This pre-BDNE profile is not a Nash equilibrium, since any deviating player who decides to sell his shares at any time t such that $p(t) > \left\lceil \frac{F_t}{1+C} \right\rceil$ and $SR(t) > 0$ increases his payoff. We shall check whether it can be a BDNE.

Simplified stock exchange and self-verification of beliefs

Now we screen the beliefs considered before to check self-verification and look for BDNE.

We can expect that fundamental beliefs of investors are self-verifying in a word without uncertainty. This is not obvious for the beliefs in strong trends.

Proposition 16. ***Self-verification***

a) If for every t $F_{t+1} \subset \left[\left\lceil (1-h) \cdot \left\lceil \frac{F_t}{1-C} \right\rceil \right\rceil, \left\lfloor (1+h) \cdot \left\lfloor \frac{F_t}{1+C} \right\rfloor \right\rfloor\right]$, $\left[\left\lceil (1-h) \cdot \left\lceil \frac{F_t}{1-C} \right\rceil \right\rceil, \left\lfloor (1+h) \cdot \left\lfloor \frac{F_t}{1+C} \right\rfloor \right\rfloor\right] \cap \mathbb{N} \backslash \{0\} \neq \emptyset$ and $F_t \in \left(\left\lfloor \frac{F_t}{1+C} \right\rfloor, \left\lceil \frac{F_t}{1-C} \right\rceil\right) \cap \mathbb{N} \backslash \{0\}$, then fundamental beliefs that contain a history with $SR(s) = BR(s) = 1$ for every $s > t$ are potentially self-verifying and the profile defined by $p_i^S(t) = F_t + 1$ and $p_i^B(t) = F_t - 1$ is a BDNE for these beliefs.

b) There exist beliefs in strong trends as defined in proposition 15 which are potentially self-verifying and the profile consisting of the corresponding threshold prices is a BDNE for these beliefs.

Proof. a) In our case the threshold strategies, as we proved in proposition 14a), are $(\bar{p}_i^B(t), \bar{p}_i^S(t)) = \left(\left\lfloor \frac{F_t}{1+C} \right\rfloor, \left\lceil \frac{F_t}{1-C} \right\rceil\right)$

As we have proven in proposition 14b), the profile fulfilling
$\bar{S}_i^{OL}(t) = \left(\left\lfloor \frac{F_t}{1+C} \right\rfloor, \left\lceil \frac{F_t}{1-C} \right\rceil\right)$ is a pre-BDNE.

However, the profile $S_i^{OL}(t) = (F_t - 1, F_t + 1)$ admits no trade and is also a pre-BDNE, since any individual deviation resulting in trade decreases the anticipated

payoff of the deviating player, which we prove in the same way as we have proven proposition 14b).
At this pre-BDNE the price is F_t, since it is the only price in the interval $(F_t - 1, F_t + 1)$. The reductions $SR(t) = BR(t) = 1$, since $AS(p(t)) = AD(p(t)) = 0$. Therefore, H^S belongs to $B_i(t, S_i^{OL}(t), H^S)$.

b) An example of such belief is
$B_i(t, a, H) = \{(X, (p, BR, SR)) : p(\tau) \geq \varepsilon \cdot p(\tau - 1), SR(\tau) = 1, BR(\tau) = 0,$
$X(\tau) = p(\tau-1)$ for every $t < \tau < \bar{\tau}$ and $(X, (p, BR, SR))|_t = H|_t\}$ for $\varepsilon \in [1+r, 1+h]$ such that for every $n \geq 1$ $\lfloor (1 + h) \cdot n \rfloor \geq \varepsilon \cdot n$ (such an ε exists since h is large compared to r) and $\bar{\tau} - t$ large enough.

Obviously, for every t we have $p(t) = \lfloor (1 + h) \cdot p(t - 1) \rfloor \geq \varepsilon \cdot p(t - 1)$, $SR(t) = 1$ and $BR(t) = 0$ for the pre-BDNE defined in proposition 15b), which is, therefore, a BDNE. ■

Note that, although we have potential self-verification and the profiles considered are BDNE, the stock exchange, in fact, cannot operate. The problem of self-verification in a more compound but less formal model of a stock exchange was considered by the author in [19], including also self-verification of various beliefs against beliefs of other players, which included also dependence on a random factor. In this paper some beliefs have the property of approximately perfect self-verification against a small group of players with random beliefs if the measure of the set of players with the beliefs under consideration was large enough. However, there were also classes of beliefs that were self-falsifying in such a case – they caused changes of prices with signs opposite to the anticipated ones.

Conclusions

In this paper new notions of equilibria in deterministic dynamic games with distorted information – *pre-belief distorted Nash equilibrium* (*pre-BDNE*) and *belief distorted Nash equilibrium* (*BDNE*) – together with concepts of self-verification of beliefs, first defined in Wiszniewska-Matyszkiel [20], were extended to a wider class of games that can encompass also models of financial markets with complicated market clearing conditions and private state variables of players besides the global state variable. These notions are especially applicable to dynamic games but they can be applied also to repeated games. In one stage games each of these concepts of equilibria is equivalent to Nash equilibrium. In games with a continuum of players also in this extended class of games the set of pre-BDNE for the perfect foresight is equal to the set of BDNE for these beliefs and to the set of Nash equilibria.

Existence and equivalence theorems were extended to the games considered in this paper.

The theoretical results were also illustrated by an example of a simplified model of a stock exchange. There were also some results proven that apply only to this model, among others no trade properties of pre-BDNE for fundamental beliefs and beliefs in strong trends, as well as potential self-verification of these beliefs. Among others, this proves that some, even very counterintuitive techniques of foreseeing prices of shares can be regarded as rational.

Acknowledgements

Research work financed by funds for science in years 2005-2007, grant no. 1 H02B 010 28.

References

[1] R. J. Aumann, 1964, *Markets with a Continuum of Traders*, Econometrica **32**, 39-50.

[2] R. J. Aumann, 1974, *Subjectivity and Correlation in Randomized Strategies*, Journal of Mathematical Economics **1**, 67-96.

[3] R. J. Aumann, 1987, *Correlated Equilibrium as an Expression of Bayesian Rationality*, Econometrica **55**, 1-19.

[4] E. Balder, 1995, *A Unifying Approach to Existence of Nash Equilibria*, International Journal of Game Theory **24**, 79-94.

[5] P. Battigalli, M. Siniscalchi, *Rationalization and Incomplete Information*, Advances in Theoretical Economics **3**, no. 1.

[6] D. Blackwell, 1965, *Discounted Dynamic Programming*, Annals of Mathematical Statistics **36**, 226-235.

[7] J.C. Harsanyi, 1967, *Games with Incomplete Information Played by Bayesian Players, Part I*, Management Science **14**, 159-182.

[8] E. Kalai, E. Lehrer, 1993, *Subjective Equilibrium in Repeated Games*, Econometrica **61**, 1231-1240.

[9] E. Kalai, E. Lehrer, 1995, *Subjective Games and Equilibria*, Games and Economic Behavior **8**, 123-163.

[10] A. Mas-Colell, 1984, *On the Theorem of Schmeidler*, Journal of Mathematical Economics **13**, 201-206.

[11] D. Schmeidler, 1973, *Equilibrium Points of Nonatomic Games*, Journal of Statistical Physics **17**, 295-300.

[12] N.L. Stokey, R.E. Lucas Jr. with E.C. Prescott, 1989, *Recursive Methods in Economic Dynamics*, Harvard University Press.

[13] Vind, 1964, *Edgeworth-Allocations is an Exchange Economy with Many Traders*, International Economic Review **5**, 165-177.

[14] A. Wiszniewska-Matyszkiel, 2000, *Existence of Pure Equilibria in Games with Continuum of Players*, Topological Methods in Nonlinear Analysis **16**, 339-349.

[15] A. Wiszniewska-Matyszkiel, 2002, *Static and Dynamic Equilibria in Games with Continuum of Players*, Positivity **6**, 433-453.

[16] A. Wiszniewska-Matyszkiel, 2003, *Static and Dynamic Equilibria in Stochastic Games with Continuum of Players*, Control and Cybernetics **32**, 103-126.

[17] A. Wiszniewska-Matyszkiel, 2005, *A Dynamic Game with Continuum of Players and its Counterpart with Finitely Many Players*, Annals of the International Society of Dynamic Games **7**, A. S. Nowak, K. Szajowski (eds.), Birkhäuser, 455-469.

[18] A. Wiszniewska-Matyszkiel, 2008, *Common Resources, Optimality and Taxes in Dynamic Games with Increasing Number of Players*, Journal of Mathematical Analysis and Applications **337**, 840-841.

[19] A. Wiszniewska-Matyszkiel, 2008, *Stock Exchange as a Game with Continuum of Players*, Control and Cybernetics **37** No. 3, 617-647.

[20] A. Wiszniewska-Matyszkiel, 2006, *Games with Distorted Information and Self-Verification of Beliefs with Applications to Exploitation of Common Resources and Minority Games*, preprint 159/2006 Institute of Applied Mathematics and Mechanics, Warsaw University, available at http://www.mimuw.edu.pl/english/research/reports/imsm/; changed and submitted as *Belief Distorted Nash Equilibria – introduction of a new kind of equilibrium in dynamic games with distorted information.*