NORGES HANDELSHØYSKOLE
Bergen, Spring 2011

Forecast behavior in business cycles and periods of different monetary policy regimes: A case study of private sector forecasts in Japan from 1989 to 2010

André Kyvik and Svein Ove F. Langeland

Advisor: Krisztina Molnár

Master Thesis - Major in Financial Economics

NORGES HANDELSHØYSKOLE
Norwegian School of Economics and Business Administration

Abstract

Does forecast behavior differ in business cycles and periods of different monetary policy regimes? This paper explores private sector expert forecasts in Japan from 1989 to 2010 to answer this question. The case study suggests six broad lessons. First, private sector expectations are irrational in contractions and expansions even though they are rational in the full sample. Second, consensus forecasts are more accurate than forecasts based on theoretical models. Third, forecast accuracy differs among macroeconomic variables and is higher in expansions than in contractions. Fourth, experts are not able to predict crises or turning points in the economy. Fifth, the critique of the Bank of Japan doing too little to fight deflation in the 1990s seems harsh. And sixth, the Bank of Japan's credibility may have been hurt by twice leaving their zero interest rate policy too early.

Preface

This paper is written as the final thesis of our master degree at the Norwegian School of Economics and Business Administration, within the specialization of Financial Economics. The thesis is written as part of the research program on crisis, restructuring and growth (KOV-project) on NHH, and we are grateful for being selected as research assistants to write a contribution to the program.

During a visit to London School of Economics in November 2009, Queen Elizabeth II asked a group of eminent economists why nobody had anticipated the subprime crisis, sparking a wave of debate. Why did the economists not warn us about the crisis, or is it at all possible to predict crises? We were inspired by this question and chose to study private sector expert forecasts for macroeconomic variables in Japan. Japan has experienced one crisis after another for the last two decades, which makes the country interesting for studying how expectations behave in crises and throughout business cycles. In addition, the Bank of Japan's monetary policies in the same time period have received criticism from prominent economists, which makes it interesting to explore whether or not private sector's forecasts reflect this criticism.

The unique data set facilitates studying forecast behavior for several macroeconomic variables, enabling us to make stronger arguments than previous research. However, the analysis is mainly descriptive, and several of our findings are explained by intuition and reasoning; they cannot be viewed as hard evidence. We have tried to structure the paper as we feel appropriate, and have not been bounded by the conventional way of writing master theses.

The follow up from the KOV-project has been of great assistance, and we express gratitude for helpful input from both professors and other students. We would especially like to thank our thesis advisor, assistant professor Krisztina Molnár, for helpful advice and timely recommendations. Her support, guidance and comments have been invaluable for this thesis. Finally, we are grateful for helpful comments from Alexander Klose and Anette Kyvik.

The process has been demanding and challenging at times, but we have enjoyed working on the topic and feel that we have gained important insight in forecast behavior. Several interesting areas are open for further research, and we hope that some readers are inspired to pursuit research in these fields.

Norges Handelshøyskole

Bergen, June 16th, 2011

André Kyvik

Svein Ove F. Langeland

Contents

Abstract 1

Preface 2

Introduction 7

Part 1: Background, data and methodology

Section 1: Economic history of Japan from 1986 to 2010 11

1st Business Cycle: November 1986 to October 1993 12

2nd Business Cycle: November 1993 to January 1999 14

3rd Business Cycle: February 1999 to January 2002 17

4th Business Cycle: February 2002 to March 2009 19

5th Business Cycle: April 2009 to the present 21

Historic development in macroeconomic variables for Japan 21

Section 2: The importance and formation of expectations 24

Why are expectations important? 24

The formation of expectations: Adaptive versus rational expectations 25

Section 3: Data set 27

Selected variables for the analysis 27

The survey data 27

Actual data 32

Section 4: Transforming raw data into comparable data 34

Generating 12-month fixed horizon forecasts 34

Generating 12-month fixed horizon actual values 36

Autocorrelation 37

Section 5: Methodology for evaluating forecasts 38

Forecast accuracy 38

Rationality tests 40

Part 2: Analysis of forecasts throughout business cycles

Section 6: Preliminary look at the dataset in business cycles 45

Real GDP growth for the next 12 months 46

Consumer price growth for the next 12 months 46
Unemployment rates for the next 12 months 47
3-month CDs in three and 12 months 47

Section 7: Analysis of CE forecasts from 1989 to 2010 51
Forecast accuracy differs among variables 51
Forecasts are more accurate than theoretical models 52
Forecasts overestimate actual values on average, but not significantly 53
Forecasts are weak-form rational, but strong-form irrational 54
Concluding remarks on forecasts for the full sample 56

Section 8: Forecast performance in expansions versus contractions 58
Forecasts are more accurate in expansions than in contractions 58
Disagreement is higher in different parts of the business cycle 59
Overestimation and underestimation in contractions and expansions 61
Forecasts are not rational in contractions and expansions 62

Section 9: Forecast behavior in different expansions and contractions 68
Subprime crisis has the least accurate forecasts 68
Unusually high forecast accuracy in two episodes 68
Consumer price growth is both over- and underestimated in expansions 69
Forecasts in the expansion before the Asian crisis are surprisingly rational 69

Part 3: Analysis of forecasts in periods of different monetary regimes

Section 10: Preliminary look at forecasts in periods of different monetary policies 72
Monetary policy regime from 1991 to 1994 72
Monetary policy regime from February 1999 to June 2006 72
Monetary policy regime after June 2006 73

Section 11: Forecast behavior in distinctive monetary policy periods 74
Forecast behavior in wait-and-see monetary policy period 74
Forecast behavior in zero interest rate policy period 75
Forecast behavior in the period after zero interest rate policy 77
Regime changes as an explanation of forecast inaccuracy and bias 78
Reflections from forecasts in periods of different monetary policy regimes 79

Conclusion 81

References 83

Appendix 1: Historic development in real GDP, Nikkei 225 and land prices 90

Appendix 2: Historic development of macroeconomic variables in business cycles 91

Appendix 3: Fixed horizon forecasts - CE versus WAN12 94

Appendix 4: Construction of actual values for the next 12 months 96

Appendix 5: Description of figures in section 6 98

Appendix 6: The theoretical forecast models versus CE forecasts 99

Appendix 7: Results for forecast accuracy and rationality tests in business cycles 107

Appendix 8: Tests of disagreement 112

Appendix 9: Figures for different monetary policy regimes 115

Appendix 10: Forecast accuracy and bias in different monetary policy regimes 118

Introduction

Expectations play a key role in the economy. Households, businesses, investors and authorities act and make decisions based on expectations of the future: how will the economy develop, will prices rise or fall, how easy will it be to get a job, and what will the capital costs be for investments? Forecasts of macroeconomic variables provide important information for decision makers, and accurate forecasts are therefore essential. However, some say that the only function of economic forecasting is to make astrology look respectable; indeed it is neither a simple exercise nor an exact science. Can we trust forecasters when they predict that crises will end soon? Will they give us a warning before the economy takes a turn for the worse?

This paper studies the behavior of private sector forecasts in business cycles and different monetary policy regimes. We have chosen Japan as a subject of study because Japan has been an important player in the global economy. Still, forecasts for Japan have not been analyzed as comprehensively as for other nations, for instance the US. Furthermore, Japan has experienced multiple crises over the last decades, which makes it a particularly interesting case for exploring forecast behavior in business cycles. How do expectations behave in these crises and throughout the business cycles? Do forecasters make systematic mistakes in certain periods, or are they accurate and rational? Finally, the Bank of Japan's monetary policies have received criticism from Bernanke, the IMF and the OECD. Is this criticism reflected by private sector expectations?

Our analysis explores forecasts from experts in the private sector provided by Consensus Economics. Forecasts of real GDP growth, consumer price growth, unemployment rates and interest rates are studied and evaluated according to forecast accuracy and rationality. This is more macroeconomic variables than what is commonly analyzed. The forecast accuracy of private sector's forecasts is also tested against forecasts from theoretical models. In addition, disagreement among forecasters is explored. This paper adds to the literature by focusing on forecast behavior in business cycles, crises and periods of different monetary policy regimes.

When studying the whole sample, we find no evidence of bias in forecasts for real GDP growth, consumer price growth or unemployment rates. These variables also stand up to tests of weak-form efficiency. However, when dividing the sample in contractions and expansions, we find both systematic bias and inefficiency.

The idea that forecasts are clearly irrational over the business cycle is, as far as we know, a new finding and thus a valuable contribution to the literature of expectations and forecasts. Even though we find systematic mistakes and irrational behavior in forecasting, this does not necessarily mean that forecasts are rendered useless. Forecasts are more accurate than theoretical forecast models. Still, private sector forecasts do not seem to predict turning points, and at times they miss by considerable margins.

Forecasts of interest rates are both biased and inefficient for the full sample. Therefore these expectations are studied more closely in a monetary policy setting. There are three main findings from analyzing private sector forecasts in periods of different monetary policy regimes: (1) Bernanke's criticism of the Bank of Japan's monetary policy being too defensive from 1991 to 1994 seems harsh, (2) the private sector believed the zero interest rate policy from 1999 to 2006, but seems to have lost confidence in the central bank's ability to combat deflation in this period, and (3) high interest rate forecast errors after the Bank of Japan abandoned the zero interest rate policy twice in seven years, indicate that the central bank's credibility declined in the private sector.

The paper is structured as follows: Part 1 sets a frame of reference for the analysis of forecast behavior in Japan. Section 1 outlines the economic history of Japan over the last 25 years. Expectations and the rational expectations hypothesis are presented in section 2, before section 3 introduces the selected data and explains why expert consensus surveys provide the best forecasts. Section 4 describes how our raw forecast data and actual data are transformed into comparable fixed horizon time series, and section 5 explains the methods used to evaluate forecast behavior and performance.

The analysis is divided into two parts. The first part of the analysis (Part 2) examines expectations throughout the course of business cycles. We begin the analysis in section 6 with a preliminary look at the data. Section 7 analyzes forecast accuracy, disagreement and tests for rationality over the full sample. The sample is then divided in section 8 to examine the behavior of expectations in contractions and expansions. Consistent irrationality is found in both contractions and expansions, and some possible reasons for the test results are discussed. Section 9 looks further at some particular findings in specific contractions and expansions.

The second part of the analysis (Part 3) studies interest rate and inflation expectations in periods of different monetary policy. Section 10 outlines three different monetary policy regimes in Japan over the last two decades. These distinct monetary policy periods are then examined more closely in section 11 by looking at forecast accuracy and bias in order to provide probable explanations for the findings. The conclusion sums up the major findings of the paper. The appendices explain in greater detail some of the methodology as well as results from all tests of forecast accuracy, rationality (bias and efficiency tests) and disagreement.

Part 1: Background, data and methodology

This paper investigates how private sector forecasts have performed compared to the actual development of different macroeconomic variables in Japan. Before we begin our analysis, it is important to build an understanding of Japan's economic history in the relevant time period and look at the actual development in macroeconomic variables. This will be done in section 1. After we have outlined the history, we will explain the importance of studying expectations in section 2. The theory behind the formation of expectations is also included in this section. Section 3 discusses the data set of this paper and the reasoning behind choosing the data. The raw data extracted from external sources is not comparable, and we therefore describe how to transform this data into comparable data in section 4. The methodology used in the analysis is outlined in section 5, which is the end of the first part of this paper.

Section 1: Economic history of Japan from 1986 to 2010

In order to analyze forecasts of macroeconomic variables we first need a basic understanding of Japanese economic history from 1986 to 2010, which includes the time period analyzed in this paper. The history helps us understand the actual development in macroeconomic variables and makes us better able to interpret findings from the analysis in later sections. We will create a basic understanding of Japanese economic history by focusing on underlying factors that have driven change in the Japanese economy.

Many economists have discussed Japan's economic history over the last 20 years[1] and several describe the 1990s as "the lost decade" due to low economic growth in this period. Low economic growth was indeed present in this period as the annual real GDP growth was 0.71 percent from 1991 to 1999. However, one may talk about "the lost decades" after the subprime crisis for Japan since the annual real GDP growth was 0.65 percent from 1991 to 2009 (for example, Fujii and Kawai (2010) claim the lost decade started in 1991 and ended in 2005). These growth rates are slow compared to Japan's high growth from 1961 to 1980 and the growth path Japan was on from 1981 to 1985. We will argue that the slow growth was mainly triggered the instability in the Japanese financial system caused by a continuous increase in the amount of outstanding non-performing loans. Also, measures from the authorities failed to return the economy to a stable growth path after this bubble burst. (Figures 20 and 21, appendix 1)

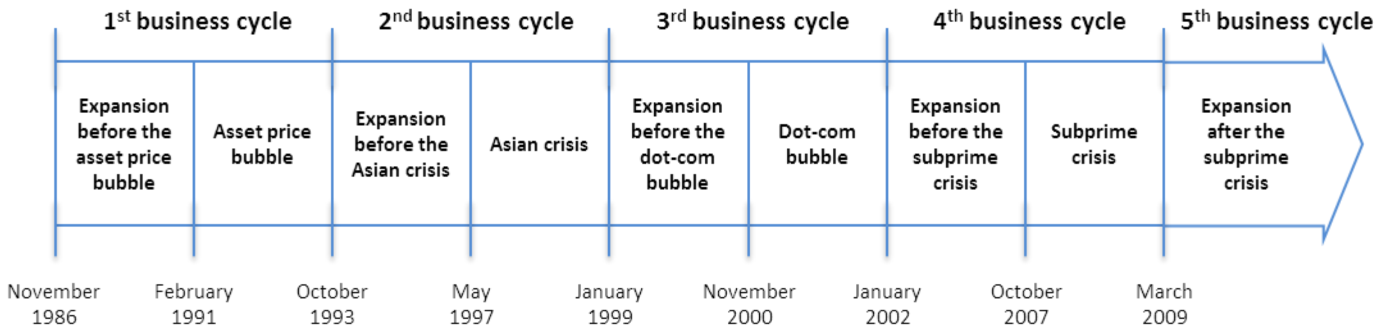
Economists differ on the timing of different business cycles. We have chosen to use the definitions from the Economic and Social Research Institute (ESRI)[2]. The reason for this choice is two-fold: 1) ESRI is a reliable source and 2) leading economists frequently use ESRI's timing of business cycles in Japan (for example, Okina et al. 2001). According to ESRI, Japan experienced four business cycles from November, 1986 to March, 2009. This section will explain the history of the four business cycles and the beginning of the fifth

[1] For example: Bernanke (1999), Fujii and Kawai (2010), Hayakawa and Meada (2000), Nakaso (2001), Oda and Ueda (2005), Okina et al. (2001), Shiratsuka (2003) and Syed et al. (2009).

[2] ESRI is a working group within the Cabinet Office's think tank that dates peaks and troughs by applying diffusion indexes and the Bry-Boschan method.

business cycle (see figure 1 for time definitions), and finally describe the historic development in macroeconomic variables.

Figure 1: Timeline of contractions and expansions



1st Business Cycle: November 1986 to October 1993

The expansion before the asset price bubble began after the Japanese economy bottomed in November 1986. The Japanese economy peaked in February 1991, and the asset price bubble lasted until October 1993. The main characteristic of the first business cycle was a surge and fall in asset prices, especially in stock and land prices. We begin by describing the drivers of the bubble, followed by a description of the burst.

Booming stock and land prices

Stock and land prices boomed in the expansion before the asset price bubble. The Nikkei 225 index increased by 132 percent from November 4th, 1986, to its all-time high in December 29th, 1989 (figure 22, appendix 1). At its peak, the Japanese stock market had a value of about USD 4 trillion, which was approximately 44 percent of the world's equity market capitalization (Stone and Ziemba 1993). From there, the stock prices plummeted, and on August 18th, 1992 the Nikkei 225 bottomed after falling 63 percent from its peak (figure 22, appendix 1). Land prices also saw tremendous growth in this expansion. In September 1990, the Urban Land Price Index was 275 percent higher than the level in September 1985 (figure 23, appendix 1), and prime property prices in the Tokyo Ginza district reached approximately USD 1 million per square meter in 1989 (PFP Wealth Management 2009). According to Stone and Ziemba (1993), Japanese land was valued at about five times that of the United States, and the land under the Emperor's Palace - which is about three-quarters of a square mile - was estimated to be worth about the same as all the land in California or in Canada. However, a long-lasting decline in land prices followed, and in September 1999 the prices were 80 percent lower than at their peak in September 1990 (figure 23, appendix 1).

Drivers of a surge in asset prices

Okina et al. (2001) describe five interrelated drivers of a surge in asset prices: (1) aggressive bank behavior, (2) protracted monetary easing, (3) taxations and regulations, (4) weak mechanisms to impose discipline on agents, and (5) self-confidence in Japan.

The first driver of the surge in asset prices was the aggressive bank behavior that emerged in the late 1980s that was largely due to financial deregulation and weak capital adequacy requirements. Financial deregulation led to higher competition among banks, and banks' lending activities became more aggressive than before, resulting in high credit growth in Japan. More property backed or property-related loans were given, and Japanese banks became more exposed to the development in the property market. A higher capital base due to higher profits, unrealized gains on stocks and equity financing enforced the loose lending in the expansion period. (Ibid.)

Protracted monetary easing was the second driver of the increase in asset prices. Okina et al. (2001) say that monetary easing from the latter half of the 1980s to early 1990s increased the pressure on the economy and supported the increase in asset prices. They point out three effects of monetary easing: (1) funding costs were reduced and made it easier for speculators to raise funds for asset purchases, (2) higher stock prices reduced capital costs and facilitated financing in capital markets, and (3) the collateral value of assets increased the funding ability for firms.

Japan's tax rates were also identified as a driver of the rise in asset prices because they had a negative effect on the supply of land. Taxes were low on holding land and high on land transactions, which made it cheap to keep the land and expensive to sell it. The negative effect was enforced by the possibility of future increases in land prices in the market, mainly in local areas, driven by the expectation that agricultural land would be converted to residential use. (Ibid.)

The fourth driver of the surge in asset prices was weak mechanisms to impose discipline on agents. Japanese banks had long been heavily regulated before deregulation started on a step-by-step basis in the early 1970s (Nakaso 2001). As the banks were deregulated, other mechanisms to impose discipline on banks seemed to be missing. In particular, the mechanisms on banks' lending activities were weak. Also, it appears that Japan lacked effective means to impose discipline on other agents such as firms, individuals and the government. For example, Japanese banks traditionally disciplined firms, but when major firms increased their funding through capital markets, this mechanism weakened (Okina et al. 2001).

Finally, self-confidence was identified as a driver of the increase in asset prices. Japan experienced euphoria where market participants' enthusiasm was not consistent with the projection of fundamentals (Shiratsuka 2003). This view is supported by rapid increase in asset prices and the fall after the bubble burst. So what caused these bullish beliefs? Okina et al. (2001) points out several factors. First, good macroeconomic conditions with high business fixed investment, large increase in housing investment and strong consumer expenditure were present in the expansion. Second, Japan played a greater role in international markets and at most had a 41 percent share of the international bank lending market (fourth quarter of 1989). Third, Japanese firms were leading in manufacturing technology and were believed to have a competitive advantage in management. Finally, there was a rush from overseas financial institutions seeking to establish branches in Tokyo. Nakaso (2001) also points out the fact that there had not been a major bank failure in the postwar period and that there was a strong belief that big banks would never fail.

The burst of the asset price bubble

When the stock and property market turned, Japanese banks' profitability declined and the banks ended up with large stocks of non-performing loans (NPLs) in their portfolios (Shiratsuka 2003). Some small financial institutions went bankrupt, and risk-taking ability declined for Japanese banks. As a consequence, a "credit crunch" arose and the channel of funds was disrupted. Fewer loans were granted for growing firms and loans were provided to unprofitable firms as banks tried to prevent losses from materializing. This disruption had a negative effect on economic activity. (Ibid.)

2nd Business Cycle: November 1993 to January 1999

After the asset price bubble ended in October 1993, a new expansion began that lasted until the Asian crisis hit Japan in June 1997. However, we will argue that much of the distress in the Japanese economy that was revealed in the Asian crisis had its origins in the asset price bubble, a view that is supported by Fujii and Kawai (2010).

The expansion period before the Asian crisis

Land prices continued to fall during the expansion after the asset price bubble. Also, stock prices saw a modest change in this period, as the Nikkei 225 only increased by one percent from October, 1993 to June, 1997. (Figures 22 and 23, appendix 1)

The continuing decline in land prices increased the number of NPLs, but the scale of NPLs was hard to assess due to weak requirements of banks' disclosure of these assets. Some institutions, such as the Bank of International Settlements, reported concern about higher levels of NPLs than what the consensus expected. Higher NPLs than reported implies that banks had buffers that were too low against credit risk and lower effective capital ratios than reported in this period. This view is supported by the increase in bank failures during this expansion, which increased in scale and complexity over time. This put pressure on the authorities, who repeatedly had to modify their handling of such failures by introducing new reforms. The authorities improved their flexibility, but the system was still not prepared for handling failures of big banks before the Asian crisis. (Nakaso 2001)

Japan's fiscal and monetary policies during the expansion after the asset price bubble have received criticism from several economists, including Ben Bernanke. Bernanke (1999) points out that fiscal and monetary policy were deflationary in this period as policies failed to ease adequately in the 1991-1994 period to stimulate aggregate demand. Also, Nakaso (2001) states that the monetary policy in Japan adopted a wait-and-see policy as authorities expected that the threat to the financial system would be eliminated by a collateral value pick-up, and asset prices would stop decreasing and start increasing again.

The Asian crisis

The Asian crisis was triggered by the currency crisis in Thailand (Hayakawa and Meada 2000). However, Japan was not only affected by external pressure in this contraction; internal pressure was at least equally important for the development in Japan's economy in this period. These two forces tested the fragility of the financial system in Japan

(Ibid.)[3]. We will focus on describing the internal conditions as it may be argued that a troubled financial system in Japan was the main driver of the Asian crisis in Japan, even though global conditions might have triggered and amplified the crisis.

Japanese financial institutions still had large stocks of NPLs in their balance sheets, and failures of major financial institutions were soon a fact. Sanyo Securities, Hokkaido Takushoku Bank, Yamaichi Securities, Nippon Credit Bank and Long Term Credit Bank of Japan (BOJ) all failed in this period. These failures led to increased credit risk and banks soon preferred loans from the BOJ to loans from the interbank market, because they were afraid that the counterpart would fail to fulfill their obligations (Nakaso 2001). Raising funds became more difficult for financial institutions, and their capital base declined (Hayakawa and Meada 2000). The capital base was also negatively affected by the disposals of NPLs and the shock to people's confidence in Japan's financial system (Ibid.). Depositors formed long lines outside the banks to withdraw their money as rumors of more bank failures spread.

Lending ability for banks deteriorated in the Asian crisis due to lower capital base. In addition, Hayakawa and Meada (2000) points out three other factors that explains this decline in lending ability: (1) banks increased their on-hand liquidity, (2) banks became stricter in reviewing all assets to improve profitability and financial soundness, and (3) risk evaluation became more severe. Business fixed investment and consumer spending declined due to tighter credit conditions. However, consumer spending declined mainly due to worries of future and not current employment and income conditions. Weakened demand led to a higher output gap, which again led to a further decline in prices. Lower prices then continued to undermine the asset quality of Japanese financial institutions, which were constantly downgraded by rating agencies.

The Japanese economy was trapped in a vicious circle that the authorities tried to get out of by utilizing several measures. For example, several bailout packages were worked out for troubled financial institutions, new legislation was passed that improved the authorities' flexibility and financial resources, the overnight call rate was lowered and

[3] "Japan's financial crisis" or "Japan's banking crisis" might have been more appropriate names. However, we refer to this contraction by "the Asian crisis" because "Japan's financial crisis" might be confused with other contractions in this paper.

the first capital injection in the Japanese economy happened in February 1998 (Nakaso 2001). However, monetary easing was countered by increased savings rates for corporations and private consumers due to high precautionary demand for money (Hayakawa and Meada 2000).

3rd Business Cycle: February 1999 to January 2002

Another period of recovery emerged from February 1999, as a new expansion period began. The expansion continued until the dot-com bubble burst in December 2000, leading Japan into a new contraction that lasted through January 2002.

The expansion after the Asian crisis

The expansion came after Japanese authorities stimulated the economy further. The BOJ shifted to a zero-rate interest policy as the uncollateralized overnight call rate was set as low as possible in February 1999. The aim of the zero interest rate policy was to remove concerns about financial stability in Japan, have a positive effect on the struggling Japanese economy and, most importantly, end the deflationary environment. In their announcement of the zero interest rate policy the BOJ stated that it would maintain its zero interest rate policy "until deflationary concern has been dispelled". This statement was supposed to ensure the desired effects of the policy and in effect the policy meant that the BOJ would inject funds into the money market without limit whenever necessary. (Shinotsuka 2000)

Why was it so important for the Japanese authorities to end the deflationary environment? If interest rates are zero, one might think that people and firms will run to the bank for cheap loaning opportunities. However, deflation causes positive real interest rates even though the nominal interest rate is zero. Furthermore, deflation discourages consumption as consumer goods become cheaper over time. This also leads to investments being postponed. This is further enforced as loans increase in real value when prices decrease. Lower activity leads to a further downward pressure on prices, and the economy can enter a vicious circle called the deflationary spiral. The gap between supply and demand can increase, leading to surplus capacity and further cuts in prices, as well as lay-offs and higher unemployment.

Monetary policy is the first line of attack to reverse a deflationary spiral, primarily by slashing interest rates. However, stronger measures should perhaps be taken to boost the economy, and the BOJ did turn to unconventional measures in addition to lowering interest rates. Further quantitative easing was introduced when the second major capital injection in the banking system took place in March 1999. The objective of this injection was to address the undercapitalization of Japanese banks. Also, a systematic approach to dealing with bank failures was implemented and steps were taken to remove bad loans from banks' balance sheets. (Nakaso 2001)

These measures, along with improved global economic conditions, led to improved activity (Syed et al. 2009). Hayakawa and Meada (2000) state three factors that caused higher demand: (1) increased public works and housing investment due to government measures, (2) increased exports as the Asian economies recovered, and (3) lower savings rates due to lower anxiety over the financial system. Also, risk premiums started to decline (Ibid.) and credit flow improved (Hoshi and Kashyap 2008).

The BOJ broke out of the zero rate policy and raised the overnight call rate in August 2000 due to improvements in the development of the Japanese economy (Schrooten 2000). However, Japan was still fighting deflation and the IMF (2000) warned the BOJ about moving away from the zero interest rate policy. The IMF recommended that the BOJ keep its policy until a sustained period of above-potential growth and a broad-based recovery of private demand was present. Also, Japan had still not gotten rid of all their problems from the asset price bubble. The amount of outstanding NPLs continued to increase and banks constantly underestimated their NPLs (Fujii and Kawai 2010), which indicates that disclosure of NPLs was still not comprehensive enough. In addition, Hoshi and Kashyap (2004) argue that no policies had focused on closing down the insolvent banks and their zombie borrowers.

The dot-com bubble

Again, the contraction was triggered by a shock in the global economy as the dot-com bubble that originated in the United States collapsed. The shock to the economy was amplified by a still-fragile financial system in Japan. Japan saw optimism about the future state of the economy turn into pessimism, and public debt rose to almost 75 percent of GDP in net terms and a large output gap existed (Syed et al. 2009).

The zero interest rate policy was put back in place in March 2001 when the economy slowed down, and the policy lasted throughout the crisis (The Bank of Tokyo-Mitsubishi

2006). Still, this policy did not prevent the capital shortage problem from reemerging, and low credit availability was once again present due to increasingly high levels of NPLs (Hoshi and Kashyap 2008). However, there was a change in the nature of the NPLs as a higher portion of loans was given to loss-making companies (Ibid.). Lending to small and medium enterprises became important for banks, as the government required the banks that received public capital to increase lending to these businesses. NPLs were now more connected to the general state of companies and less connected to the real estate industry. Also, the lending policy suppressed lending to new businesses in sectors where subsidized firms were more prevalent (Caballero et al. 2003). For example, healthy banks held back on their lending activities because they saw no point in lending to firms that would have to compete against zombies that were kept alive by sick banks (Hoshi and Kashyap 2008).

4th Business Cycle: February 2002 to March 2009

In the expansion that began in February 2002 and lasted until the subprime crisis broke out in November 2007, Japan finally managed to get control over the NPLs and stabilize the financial system. The subprime crisis had its origins in the subprime market in the United States and led the global economy into turmoil. The Japanese economy was affected and the contraction lasted through March 2009. However, these dates are provisional and might be revised later.

The expansion before the subprime crisis

Favorable global economic conditions and a final cleanup of the financial system drove the expansion after the dot-com bubble. The favorable global conditions in this period are related to strong growth in emerging economies, especially in China and India. Japan benefited from this development and Syed et al. (2009) claims that net export accounted for a third of Japan's growth in this period.

The problem of NPLs remained after the dot-com crisis and the amount of outstanding NPLs reached its peak in 2002 after regulatory pressure led to a dramatic change in loan classifications (Fujii and Kawai 2010). The government then launched a new policy package, the Program for Financial Revival, with the objective of accelerating bank loan restructuring (Ibid.). Loan classification and loan loss provision were strengthened and Japan finally got a hold on its problems with NPLs, a problem that had lasted over a decade. Also, banks rebuilt their capital and reduced the level of NPLs in their portfolios, which resulted in improved operating performance due to lower loan losses and higher capital gains on stock portfolios (Hoshi and Kashyap 2008).

According to Syed et al. (2009), the above factors may have been the main drivers for the restored confidence in the banking system. In addition, lower corporate debt levels due to restructuring pressure from the government and a weak yen contributed to positive development in Japan in this period. All these factors led to a recovery with the following characteristics: increased real GDP growth, improved bank and corporate profits, improved credit flow, lower unemployment, higher stock prices and increased investment.

The long period of deflation ended in May 2006, and as soon as consumer prices reached positive terrain, the BOJ ended their zero interest rate policy with an increase in interest rate on July 14th, 2006. Again, the end of this policy was met with skepticism, and the OECD was one of the critics. The OECD states on page 1 in their economic survey of Japan in 2006: "*The Bank of Japan should be cautious in raising interest rates, given remaining deflationary pressure (...) Avoiding an early and significant rise in long-term interest rates would be beneficial to economic activity, the fiscal situation and the banking sector.*"

The subprime crisis

"*The global financial crisis has underlined in a painful way that the globalized economy is a fact*" – John Lipsky (2009). This statement characterizes the subprime crisis, which started in the United States and spread to almost every corner of the world. Japan was no exception and experienced a severe downturn in this period.

Activity in the global economy slowed due to a worldwide credit crunch that made firms and consumers cut down on their investments and spending. The drop in global demand hit Japanese exports hard because advanced manufacturing products such as cars, information technology and machinery, make up a larger share of production in Japan than in other G-7 economies. For example, exports of cars fell by 65 percent in the months after September 2008. In addition, a strong yen contributed to the fall in Japanese exports. (Sommer 2009)

According to Sommer (2009), domestic demand fell due to four factors: (1) the credit crunch led to stricter lending conditions, (2) uncertainty led to higher interest rate spreads, (3) the significant drop in stock prices decreased wealth, and (4) increased inventory ratios put downward pressure on business investments. Japanese banks had relatively small investments in subprime-related financial products and were therefore not as seriously affected by the subprime crisis until after the summer of 2008, when stock prices dropped and capital losses rose (Fujii and Kawai 2010).

The world has never seen greater measures taken by central banks and governments in countries all over the world as the measures implemented during the subprime crisis. The BOJ was no exception. Hirose and Ohyama (2009) points out monetary policy measures in three areas: (1) reductions in the policy interest rate, (2) measures to ensure stability in financial markets, and (3) steps to facilitate corporate financing.

5th Business Cycle: April 2009 to the present

The global economy improved as measures taken by central banks and governments proved effective, at least in the short to medium term (Lipsky 2009). However, it is too early to say what the effect of these measures will be in the long run. Japan benefitted from the improved global conditions and its domestic stimulus packages, and the economy entered an expansion in April 2009. The date of the beginning of the recovery is provisional as Japan is still recovering from the subprime crisis[4].

Exports have increased due to the global economic recovery, but there has been a change in demand for Japanese products from advanced economies to the fast growing emerging economies. For example, China is now the largest importer of Japanese goods as they import about 19 percent of all Japanese exports. (Berkmen et al. 2010). The deflation problem is still present even though Japan is recovering from the subprime crisis. However, enhanced growth potential and lower concerns about the future may help end deflation (Noda 2010).

Historic development in macroeconomic variables for Japan

We have now outlined the recent economic history of Japan by focusing on underlying factors for the development in the Japanese economy. However, how have these underlying factors in different periods affected key indicators? Figures 2 to 4 illustrate

[4] However, we have not studied the effect of the earthquake and tsunami that hit Japan on March 11th, 2011 (ABC News 2011).

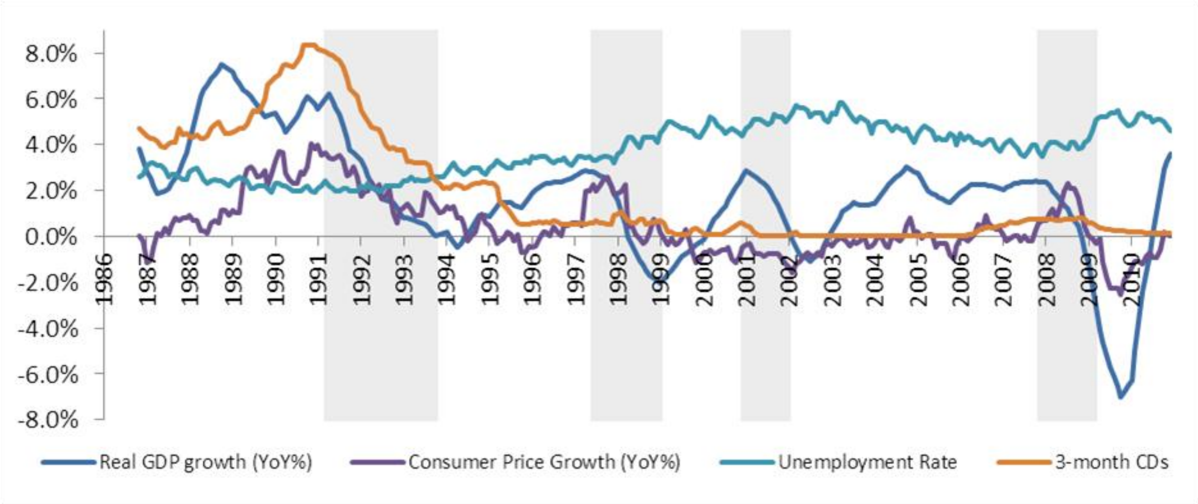
characteristics of the actual development of some macroeconomic variables in expansions and contractions.

Real GDP growth tends to increase and remain at higher levels in expansions, while it decreases in contractions. This pattern is also present for industrial production growth and real private consumption. Also, the unemployment rate usually increases in contractions and decreases in expansions. These common trends for macroeconomic variables support ESRI's definitions for the timing of different expansions and contractions.

The extraordinary deflationary environment that has been present in Japan since the asset price bubble makes it difficult to point out differences in the development of inflation and interest rates in expansions and contractions. A more thorough discussion of the historic development of macroeconomic variables in the different business cycles is included in appendix 2.

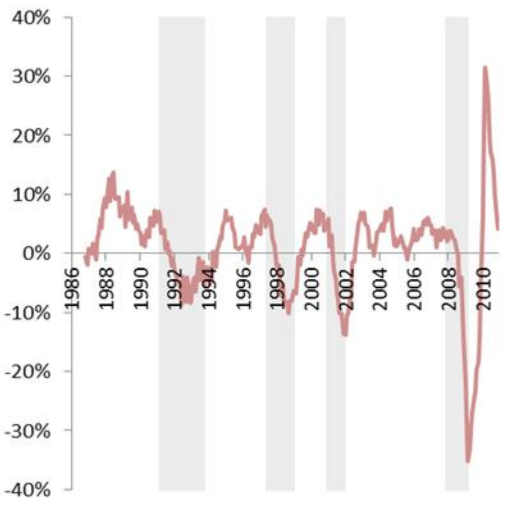
Figure 2 to 4: Historic development in Japanese macroeconomic variables from November 1986 to December 2010

Figure 2: Historic development for real GDP growth, consumer price growth, unemployment rate and 3-month CDs



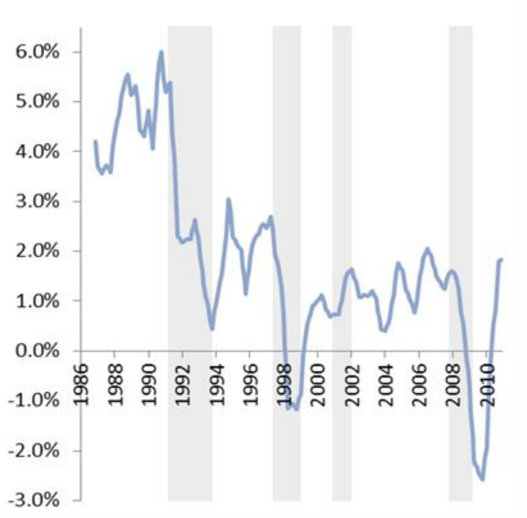
Source: Thomson Reuters Datastream

Figure 3: Industrial Production growth (YoY%)



Source: Thomson Reuters Datastream

Figure 4: Real Private Consumption growth (YoY%)



Source: Thomson Reuters Datastream

Section 2: The importance and formation of expectations

The last section gave us a basic understanding of macroeconomic development in Japan over the last 24 years. However, given this development, how have expectations for different macroeconomic variables behaved in the same period? What have economists expected to happen in the Japanese economy? Can they predict changes in macroeconomic variables, and are the forecasts rational? The rest of this paper tries to answer these and other questions related to expectations. This section will explain the importance of expectations and briefly outline the main theories behind the formation of expectations.

Why are expectations important?

Expectations are important as current decision-making depends on future prospects. This can be seen from consumers', businesses', authorities' and investors' points of view. Consumers spend and save according to their future expectations of employment and income. Businesses' future expectations of income and profitability are used in decision making about current and future business investments, production levels, employment levels, etc. Investors use expectations in their timing of buying and selling assets. The authorities use expectations when planning current and future actions. For example, expectations of future tax income and fiscal spending are considered when governments outline fiscal budgets. All these expectations are tightly linked with expectations of economic development and therefore forecasts of main macroeconomic variables. Agents make real decisions according to their predictions, and expectations will therefore inevitably affect the real economy.

Most research on expectations has been on inflation expectations because these are particularly important in monetary policy. As Gerberding (2006) states in her introduction of households versus expert forecasts on page 3: "*There is a broad consensus that inflation expectations play a key role in the transmission of monetary policy measures to aggregate output and prices. Hence, the question of how economic agents form expectations is of interest to model builders and monetary policy makers alike.*" In fact, almost every central bank that has inflation targeting studies inflation expectations surveys when forming monetary policies (Kershoff and Smit 2002). The surveys are used to forecast inflation and evaluate the credibility of inflation fighting-policies.

Changes in expectations are important. For example, consumers lower their current consumption and increase their savings rate if they suddenly expect lower future income and employment. As seen in the previous section, this was the case for Japanese consumers in the Asian crisis. Also, central banks may adjust their monetary policies if there is a sudden change in expectations.

Factors that change expectations are also important when studying expectations. New information usually changes expectations as agents adjust their beliefs according to the new information. Realized values for macroeconomic variables give agents new information if realized values differ from the expected values, probably causing agents to modify their future expectations. In other words, there is co-dependency between forecasts and actual development of macroeconomic variables.

The formation of expectations: Adaptive versus rational expectations

There are two main lines of thought within the literature on how expectations are formed: adaptive and rational expectations. According to the adaptive expectation hypothesis, expectations of future inflation are based solely on some distributed lag of past values of inflation. The same goes for other macroeconomic variables. If the nature of expectation formation were solely adaptive, then it would have been a waste of time to conduct expectation surveys, as their results would contain no additional information other than what is already provided by past values. But numerous analyses of expectations of different economic agents show that expectations can provide valuable information (for example Ang et al. (2007), Thomas (1999), Gerberding (2006) and Mestre (2007)).

On the other hand, supporters of the rational expectation hypothesis favor the view that people make use of all available information when forming expectations about future values of the variable in question. In Muth's (1961) definition of rational expectations it is assumed that the subjective expectations of economic agents match the predictions of the relevant economic theory.

If expectations are fully rational, they should exhibit two fundamental characteristics. Firstly, they should be unbiased; that is, agents should forecast correctly. Rational expectations are not about always being exactly right, however, but about getting things right on average. Secondly, forecasts should be efficient, in that forecasters should employ all relevant information when forming their expectations. It should be evident that this requires an overwhelming amount of information and processing capability from the agent.

A distinction is therefore commonly made between weak-form and strong-form rationality. Weak-form rationality implies unbiasedness and weak-form efficiency. The latter is based on the notion that information about the past history of the variable itself is costless while other information is costly, implying only that agents should have adequately considered all information contained in past values of the variable being forecasted. Strong-form rationality specifies that agents also have to be strong-form efficient, meaning that they have effectively incorporated information about all the relevant variables that a state-of-the-art model of that variable would include (Thomas 1999). Thomas (1999) proposes that this should be interpreted as agents using all information for which the marginal benefit of gathering and utilizing the information exceeds the marginal cost. However, not all agents can be economic experts or all-knowing. Agents have different marginal costs and benefits, and it is therefore impossible to exactly specify the optimal level of information to include in the information set.

Since the monetary policy implications of rational expectations are very different from those of more backward-looking models, the issue of how expectations are formed is of considerable interest to policy makers[5]. The combined insight derived from econometric studies and theoretical work tends to favor the view that neither the adaptive nor the rational hypothesis fully explains the formation of expectation. As stated by Roberts (1998), these hypotheses are two extremes and the truth probably lies somewhere in between.

[5] See Bullard and Mitra (2002), Gaspar et al. (2006), and Molnár and Santoro (2006).

Section 3: Data set

In order to answer the questions we outlined in the introduction to the previous section we need to analyze dataset that contains both actual values and forecasts of macroeconomic variables. However, several questions come into mind: Which macroeconomic variables should we analyze? Which forecasts are best suitable in our analysis, and why? Which actual data series are comparable to the forecasts series? This section will answer these questions.

Selected variables for the analysis

This paper compares actual development of macroeconomic variables with forecasts for the same variables. Data, time and space limitations make it difficult to look at all the variables identified in section 1, and we have therefore chosen to focus on four variables. The variables in comparison are real GDP growth, consumer price growth (inflation), unemployment rates and 3-month CDs (interest rates).

Real GDP growth is preferred to private consumption growth and industrial production, because the variable gives a broader view of the overall activity in the economy. Also, industrial production fluctuates more than real GDP growth and is therefore expected to be harder to predict.

The survey data

The specific survey chosen is the consensus forecasts from the renowned firm Consensus Economics (CE). For each variable CE reports the forecast of each individual forecaster as well as the arithmetic average, or the "consensus forecast" for that variable. Since we obtained the data electronically, we do not have the individual forecasts. However, we do have an overview of the mean, the highest forecast, the lowest forecast, and the standard deviation for each variable for each month from October 1989 to January 2010. An introduction of Consensus Economics and an elaboration of why professional consensus survey forecasts are chosen follow below.

Consensus Economics

Consensus Economics is the world's leading international economic survey organization and polls more than 700 economists and institutions each month to obtain their forecasts and views (Consensus Economics 2011). Founded in the autumn of 1989, this

London-based firm conducts monthly surveys each month in which renowned experts from leading economic firms are asked to give their forecast for the development of a range of important macroeconomic variables. These experts work in major banks, investment firms, economic research institutes and other business services. For each country, 10 – 30 forecasters are asked, initially only for the G7 countries, but around 80 other economies have been added subsequently.

The survey is then published in the middle of each month. Virtually all panelists in Consensus Economics (CE) are based in the country they forecast. Several variables are included in the survey: GDP growth, consumer price growth, industrial production, business investments, personal consumption, producer prices, unemployment rate, current account, wages, 3-month interest rate and 10 year government yield.

The reasoning behind choosing mean expert forecasts

The advantage of consensus survey forecasts from professional agents is the following: (1) survey forecasts yield direct observations of expectations without depending on a priori assumptions, (2) expert surveys are superior to business and household surveys, and (3) consensus forecasts outperform individual forecasts.

Survey forecasts yield direct observations of expectations without depending on a priori assumptions

One approach to reveal expectations is to derive them from financial asset prices. This can mainly be done for inflation expectations, where the term structure of interest rates and the prices of interest-rate derivatives (e.g. TIPS) can be useful for extracting information on expectations. Alternatively, one can do surveys and ask market participants directly what their expectations are over a certain time horizon. The main benefit of surveys is that they yield direct observations of expectations, without depending on a priori assumptions, for example regarding liquidity premium and risk in interest rates. Galati et al. (2010), however, point to some shortcomings of survey measures. While market data can be provided on a daily, or even hourly, basis, surveys have a relatively low frequency. This makes them less suited for identifying the existence and timing of breaks in expectation formation over short horizons.

In a research program of rational expectations in the 1980s, economists tested the forecasts of inflation from surveys in the U.S. (the US Livingston Survey and the Survey of Professional Forecasters). They found a disturbing result: Forecasts exhibited both bias and inefficiency. The forecasters seemed to make systematic errors. The researchers concluded that maybe macroeconomic forecasters were irrational, or perhaps the surveys were poor measures of inflation expectations (see for example Akhtar et al. (1983) and Pearce (1979)). The consequence was that forecast surveys developed a poor reputation that lasted for a long time. However, Croushore (2010) argues that the researchers were hasty in their condemnation of the surveys. He discovers that the results were particular to the data sample of the time and an artifact of the data they were using. His paper also highlights several episodes in which forecasters made persistent forecast errors, but points to the fact that the episodes are so short that by the time they can be identified, they have already nearly disappeared. Croushore also tries to improve the survey forecasts in real time, but uncovers that this only leads to increased forecast errors. This result is in accordance with Ang et al. (2007), who find that surveys are difficult to beat in real time. We therefore proceed to use surveys as a measure of expectations.

Expert surveys are superior to business and households surveys

Surveys can be directed towards different economic agents, mainly households, businesses and professionals/experts. Furthermore, surveys can be carried out qualitative or quantitative. For households, there are arguments in favor of gathering qualitative data since they are more likely to have an opinion on the expected direction of future price changes than they are to give precise forecasts for a certain time horizon (Gerberding 2006). There are methods that can be used to transform qualitative surveys into quantitative (see for instance Mestre (2007)), but this will inevitably give some uncertainty in the data. Professional forecasters, however, should be able to forecast quantitatively. Furthermore, a potential weakness of consumer and business surveys is that there may be little economic incentive for the respondents to state their expectations correctly. For professionals, however, forecasting is part of their jobs, and since forecasts are sold on the market, correct forecasts will be connected to their own compensation and reputation, as well as their firm's performance and reputation[6]. As

[6] Galati et al. (2010) point to the fact that market data has the advantage of market participants putting their money where their mouths are. This is a valid point against household and business surveys, but for professional forecasters, this should not apply to a large degree, for the reasons given above.

this is part of their job and education, they will probably also have advantages in collecting and processing relevant information. This line of thought is in accordance with Keane and Runkle (1990) who argue that only the professional forecast surveys are truly reliable, because they survey people whose incomes depend on the quality of their forecasts.

The intuitive arguments above are mostly in favor of choosing professional surveys carried out quantitatively. So what does the empirical research tell us? Several articles have found professional forecasts to be superior to other agents; both in rationality and quality of forecasts, for example Mestre (2007) and Ang et al. (2007). This is not to say that business surveys or household surveys cannot contain valuable information (Mestre 2007), but all in all, *"expert forecasts are substantially more precise than the household expectations"*, as stated in Gerberding (2006) page 9. We therefore proceed in using professional forecasts.

Consensus forecasts outperform individual forecasts

The next question is whether to use individual forecasts or to use mean or "consensus" forecasts. Batchelor and Dua (1995) state that individual forecasts may be subject to various behavioral biases, but many of these biases are likely to be eliminated by pooling forecasts from several forecasters. Capistrán and Timmermann (2006) also find that although individuals' forecasts are biased, in averaging, the biases offset each other when averaged so that the mean forecast is unbiased. Bates and Grangers (1969) first formalized this point, showing that just as spreading investments over many assets will reduce risk, so will averaging forecasts across different forecasters reduce the size of the expected error. Since then, a large literature on the benefits of pooling forecasts has developed, with over 200 articles cited in a survey by Clemen (1989). After studying predictions by a panel of economists, Zarnowitz (1982: 20) concluded that: *"The group mean forecasts from a series of surveys are on average over time more accurate than most of the individual projections. This is a strong conclusion, which applies to all variables and predictive horizons covered and is consistent with evidence for different periods from other studies."* McNees (1987) later stated that for US macroeconomic forecasters *"..consensus forecasts are more accurate than most, sometimes virtually all, of the individual forecasters that constitute the consensus."* (From Batchelor (2000: 5))

This is not to say that mean forecasts are always the best, just like an index fund will not beat every individual investor. It would of course be helpful to identify beforehand which forecasters outperform the consensus, but Batchelor (1990) showed that there is typically no consistency in individual accuracy rankings from year to year, and thus no way to pick the best individual forecasters ex ante (in Bachelor 2000). The empirical evidence therefore strongly points to using consensus forecasts.

Batchelor (2000) compares the accuracy and information content of macroeconomic forecasts for G7 countries made in the 1990s by the OECD and the IMF, and uses Consensus Economics (CE) forecasts as the benchmark. With few exceptions, he finds that CE forecasts are less biased and more accurate than both the OECD and the IMF. Blix et al. (2001) find the same, and also state that even though the mean does not show superb performance, it is always among the best performers. They consider it somewhat disturbing that organizations renowned for the high caliber of their economic analysis (OECD and IMF) do not have better forecasts, especially since the media gives so much credence to them.

Some still prefer to use individual forecasts, and many feel that big multinational agencies are trustworthy. In the media, for example, forecasts by the IMF or the OECD are often very popular and receive a lot of attention. A priori it is not clear whether these forecasts or forecasts done by governments should be more or less accurate than forecasts done in the private sector by banks, business corporations or independent consultants. Governments and multinational agencies have certain informational advantages that could help improve their relative accuracy. They have complete and timely knowledge of official statistics, and may also have some insight into future plans and reactions to events. However, for the multinational agencies, the drawback is that they are often not based in the specific countries. They may thus not have full access to the pieces of information or rumors that are available in the environment of home-country analysts. In addition, both governments and intergovernmental agencies may be subject to political pressures. There may also be bureaucratic delays of publications. More seriously, governments can be tempted to massage official statistics and forecasts to cast a favorable light on the current policy, or to justify future politics and actions. This problem was illustrated by the controversies in France and Germany over the creative accounting measures taken to bring their reported budget deficits in 1997-8 closer to the Maastricht targets (Ibid.), and later in the skepticism towards statistical figures from the People's Republic of China. However, this is not only problematic for forecasts, but also for actual figures, and can thus influence any forecast through less realistic forecast errors.

All in all, this convinces us to use consensus survey data in our research. There are not many agencies that provide this for Japan, but Consensus Economics does. We mentioned above that CE is the world's leading international economic survey organization. In addition, Batchelor (2000) finds that CE outperforms the IMF and the OECD forecasts most markedly and consistently for the two largest economies in the sample – the US and Japan. Finally, Blix et al. (2001) find that in Japan, the CE mean is the best overall trade-off for forecasts done in the 1990s.

Actual data

The analysis in this paper centers on comparing forecasts with actual values, and now that forecasts have been discussed, the focus turns to comparable "actual" or realized data. This is less trivial than it might seem due to three reasons. First, it is not obvious which variable is being asked for in surveys. For example, does CPI refer to the general consumer price index, or does the CPI exclude fresh food and/or energy prices? The "core CPI" is often used as a target for central banks, but even the definition of core CPI can vary between countries.

Second, variables in the surveys are sometimes changed. CE recently did this for interest rates, when the 3-month CD was switched to a 3-month TIBOR in June 2010. Other variables may also have changed basis over the years.[7]

Third, initial estimates of actual values are often revised several times. This is especially the case for GDP growth, which can be revised considerably. In addition, the weights in the consumption basket might be changed over time, for example as standard of living increases, which might affect the CPI years later. The standard in the literature is to use the actual data of today, but several researchers claim that one should use real-time data (the first data released for the year the forecast was done) as a comparison (Croushore 2010). However, shouldn't forecasts aim to forecast the final actual data, not the

[7] Unfortunately, CE has provided limited information of these changes to us.

preliminary data in need of revisions? If that is the case, then we should use the final revised data instead. There is no quick fix here, but because of data availability, we have chosen to use today's official values. However, we encourage future research on expectations in Japan to test using real-time data, as Croushore (2010) and Capistrán and Timmermann (2006) have done for US data.

Actual real GDP growth and 3-month CDs are provided by the Cabinet Office in Japan, actual consumer prices growth by the IMF and unemployment rates by the Statistics Bureau of Ministry of Internal Affairs and Communication[8]. The actual time series are presented in figure 2 in section 1. Even though the actual data is taken from reliable sources, the problems stated above can induce errors in our analysis and conclusions, and represent a limitation of this paper.

[8] These sources are extracted from Thomson Reuters Datastream.

Section 4: Transforming raw data into comparable data

Although we have now discussed the raw data for both forecasts and actual values, the two sets of data are not yet comparable. Forecasts for real GDP growth, inflation and unemployment are done for current year and next year. The forecasts are therefore aimed at a fixed date, but not over a fixed horizon. Also, the actual data is not comparable to the forecasts because we only have quarterly data for actual real GDP and monthly observations for actual consumer price growth and unemployment rate[9].

This section explains how forecasts and actual values for the next 12 months are generated for real GDP growth[10], consumer price growth and unemployment rates. Forecasts and actual data will thus match, and forecast errors can be evaluated at the time forecasts were made. Furthermore, the time series are auto correlated (explained in detail below), which means that we cannot use OLS regressions in statistical tests. The last part of this section discusses options for statistical tests that can be used instead of OLS regression.

Generating 12-month fixed horizon forecasts

The survey sent out by Consensus Economics asks forecasters to forecast macro variables for current year and next year. The forecasts are therefore aimed at a fixed date, not for a fixed horizon, as mentioned above. In January, the forecast for the current year will be 12 months ahead, and the forecast for next year will be 24 months ahead. In February, the forecast for the current year will be for 11 months, and the forecast for next year will be for 23 months. In December, the forecast is only 1 month and 13 months ahead, for current year and next year, respectively. Furthermore, by December, much of the real data for the year is already known, probably creating more accurate forecasts towards the end of the year. This generates an obstacle, because for our tests, we want fixed horizon forecasts. Some researchers compared all January forecasts, then all February forecasts, etc., so that we have twenty 12-month and 24-month forecasts

[9] The only directly comparable time series are forecasts and actual values for 3-month CDs, as these are fixed horizon three- and twelve-month forecasts and actual values for 3-month CDs are given each month.

[10] For example, actual real GDP growth for the next 12 months in December 2000 will represent the actual real GDP growth from December 2000 to December 2001.

(January), twenty 11-month and 23-month forecasts (February), etc. (see for instance Mestre (2007) where the results are shown in charts instead of tables due to the lack of monthly observations).

However, this paper applies a simple method and creates fixed 12-month horizon forecasts from current and next year forecasts. This method is also used by Sturm and Haan (2009), and is a weighted average of current year and next year forecasts. The simple formula is:

*Forecast for the next 12 months = Current Year Forecast * Remaining months in Current Year/12 + Next Year Forecast *(12 – Remaining Months in Current Year)/12*

This formula will from this point forward be referred to as the simple weighted average formula for the next 12 months (WAN12). WAN12 is best explained by an example: In October 1989, the current year and the next year mean forecasts for real GDP growth are 4.87 and 4.30 percent. These forecasts are published on October 15th; however, we simplify and assume that there are only 2 remaining months of the current year, 1989[11]. The forecast for real GDP growth for the next 12 months is then equal to 4.40 percent[12].

The method seems intuitive, but it is not 100% scientific and might induce some uncertainty in the numbers. However, we consider this a better compromise than to lose observations and significance.

For the last few years, CE has asked the survey participants to make forecasts every quarter for a fixed quarterly horizon for the next 6-7 quarters. This gives us an opportunity to compare fixed horizon forecasts provided by CE with fixed horizon forecasts produced by WAN12. The comparison indicates that WAN12 is a good approximation for fixed horizon forecasts. Three points in appendix 3 illustrates this: (1) the WAN12 and CE fixed horizon forecasts follow each other closely in figures 24 to 27, (2) the forecast accuracy is similar for the two methods (table 5), and (3) correlation between the methods is high (table 5). We therefore continue to use WAN12 to calculate

[11] We could have used another approximation instead of 2 months. However, we chose 2 months as we have been given limited information from CE of whether or not another approximation would have been more accurate.

[12] 4.87 percent * 2/12 + 4.30 percent * 10/12 = 4.40 percent

mean, high and low forecasts for the next 12 months for all variables in the time period from October 1989 to January 2010. Also, this paper uses mean forecasts for the next 12 months in the analysis if not stated otherwise (high and low will only be used when looking at disagreement in section 8).

Generating 12-month fixed horizon actual values

WAN12 produced forecasts for the next 12 months, and we therefore need actual values for the next 12 months to compare against these forecasts. Below we explain how these values are calculated for real GDP growth, consumer price growth and unemployment rates.

Annual actual real GDP is calculated by taking a weighted average of quarterly real GDP (appendix 4). The actual real GDP growth is then computed as the rolling annual growth for the next 12 months. For example, real GDP for the next 12 months was JPY 423,757 billion in January 1989, and JPY 447,370 billion in January 1990. Real GDP growth for the next 12 months in January 1990 was therefore 5.6 percent[13].

WAN12 forecast consumer price growth for the next 12 months by computing the average for the next 12 months. These forecasts differ from actual consumer price growth that is presented in figure 2 as the actual values go from one point in time to another, and are not the average over the period. For example, consumer price growth in January 1990 represents the change in the consumer price index (CPI) from January 1989 to January 1990. The actual values then fluctuate more than forecasts, partly due to monthly seasonality. A moving average is therefore used to compute smoother time series (appendix 4) that are comparable with forecasts.

Unemployment rates are reported on a monthly basis in figure 2, e.g., the unemployment rate in January 2000 is the actual rate reported in this month. We need actual values for the next 12 months and we generate these values by a simple method weighting monthly unemployment rates for the next 12 months (appendix 4). For example, the average of monthly unemployment rates reported from January 1990 to December 1991 is used to compute the actual unemployment rate for the next 12 months in January 1990.

---

[13] (447,370/423,757) – 1 = 5.6 percent

Autocorrelation

When testing forecasts over a 12-month horizon, we face the issue of overlapping observations. This means that shocks will affect the forecast errors for several consecutive periods because the forecasts span a longer period of time than the sampling frequency of forecasts. For example, an oil-price shock in August 2008 affects any forecast errors that include August 2008, meaning that forecast errors from surveys taken from August 2007 to August 2008 are all correlated. Autocorrelation is therefore inevitable in our data.

Conventional computation of standard errors in OLS regressions requires that the errors are serially uncorrelated and that they inhibit constant variance. The variance of the residuals is not necessarily constant in our time series, since some periods can be more difficult to forecast than others. Our data therefore probably display heteroskedasticity in addition to autocorrelation.

When testing hypotheses about forecasting equations and allowing for these overlapping observations, one alternative is to define the sampling interval to be equal to the forecast interval. The sample is then cut into 12 pieces and every 12th observation is used. This, however, limits the number of observations, and clearly does not make use of all available data. Another alternative is to adjust the covariance matrix as suggested by Brown and Maital (1981) using the method of Hansen and Hodrick (1980), or perhaps as modified by Newey and West (1987) to guarantee a positive definite covariance matrix[14]. Croushore (2010) executes both alternatives, and reports that they are largely consistent. Newey-West (1987) is easily implemented in Stata, and we choose to use this method in our paper. This method will help us overcome autocorrelation and heteroskedasticity in the error terms in our data[15].

---

[14] The methods compute a weighted variance-covariance matrix, which effectively gives less weight to the errors made in the high-variance or highly-serial correlated observations. Practically, this will give higher standard errors and lower t-statistics than OLS.

[15] The Newey–West variance HAC (heteroskedasticity and autocorrelation consistent) estimator handles autocorrelation up to and including a lag of m. Thus, it assumes that any autocorrelation at lags greater than m can be ignored. A question is how many lags to. An overlap of 12 creates an MA(11) in the errors, and we have therefore set the maximum lag-length to 12.

Section 5: Methodology for evaluating forecasts

Sections 3 and 4 introduced the data that will be used in our analysis. This section outlines the methodology we will apply in order to analyze this data. The methodology can be divided into two parts: (1) measures used to analyze forecast accuracy and (2) tests of rationality.

Forecast accuracy

Forecast accuracy is assessed by comparing mean forecasts against actual values, and this paper uses the following four measures: (1) mean absolute error (MAE), (2) root-mean-squared error (RMSE), (3) mean prediction error (MPE), and (4) mean normalized squared error (MNSE). We have chosen these methods because they are commonly used among researchers to test forecast accuracy[16].

Mean absolute error (MAE)

The first measure we will use to analyze the forecast accuracy is mean absolute error (MAE):

$$MAE = \frac{\sum_{t=0}^{N} |(A_t - F_t)|}{N}$$

Where A_t is actual values, F_t is forecasts[17], N is the number of observations and t denotes time. MAE is the average of all differences between actual values and mean forecasts (forecast errors), disregarding the sign of the error. For MAE, a forecast that is 1% too low (an error of +1%) and another that is 1% too high (an error of -1%) both represent absolute errors of 1%. The closer the MAEs are to zero, the more accurate are the forecasts.

Root-mean-squared error (RMSE)

The second measure that is used to analyze the forecast accuracy is root-mean-squared error (RMSE):

[16] See for example Batchelor (2000), Blix et al. (2001), Mestre (2007) and Mankiw et al. (2003)

[17] A_t and F_t are mostly actual values and forecasts for the next 12 months, but in some occasions they span another time period, e.g. for 3-month CDs in three months.

$$RMSE = \sqrt{\frac{\sum_{t=0}^{N}(A_t - F_t)^2}{N}}$$

RMSE is computed by squaring all errors, thus disregarding their signs, as in the case of the MAE. These squared errors are then averaged to give the mean squared error (MSE). As its name suggests, the RMSE is the square root of the MSE. The forecast accuracy improves as the RMSE moves closer to zero. The difference between MAE and RMSE is that RMSE heavily penalizes forecasters who make a few large errors, relative to forecasters who make a larger number of small errors[18].

Mean prediction error (MPE)

The third measure we use to analyze forecast accuracy is mean prediction error (MPE):

$$MPE = \frac{\sum_{t=0}^{N}(A_t - F_t)}{N}$$

MPE is a simple average of forecasting errors and should be close to zero over time in order for a forecast to be unbiased. MPE differs from MAE and RMSE by taking the sign of the difference (+/-) into account. A positive ex-post forecast error indicates that forecasts have on average underestimated actuals, while a negative ex-post forecast error indicates the opposite, i.e. the forecasts have on average overestimated actuals. MPEs can therefore be used to analyze bias as well as forecast accuracy.

Mean normalized squared error (MNSE)

The last measure to analyze forecast accuracy is mean normalized squared error (MNSE):

$$MNSE = \sqrt{\frac{\sum_{t=0}^{N}\frac{(A_t - F_t)^2}{\sigma_p^2}}{N}}$$

[18] The MAE implicitly assumes that the seriousness of any forecasts error depends directly on the size of the error, so that an error of ±2% is treated as twice as serious as an error of ±1%. The RMSE implicitly assumes that the seriousness of any error increases sharply with the square of the size of the error, so that an error of ±2% is treated as four times as severe as an error of ±1%.

σ_p^2 is the standard deviation of the actual values of a variable in period p. MNSE divides the squared error by the standard deviation of a variable's actual values in a period, and thereby adjusts prediction errors for volatility in actual values.

Rationality tests

The hypothesis of rational expectations from section 2 does not necessarily imply that economic agents' forecasts are always correct, but that they do not exhibit systematic mistakes. The popularity of the rational expectations hypothesis in macroeconomic modeling has motivated numerous authors to test it on survey expectation data (Gerberding 2006). The rationality tests are conducted in two parts: (1) test of bias and (2) tests of forecast efficiency.

Test of bias

Under the hypothesis of rational expectations, forecast errors should have a mean value of zero. A test for bias examines whether expectations are centered on the right value or if they differ systematically from actual values. The most common test for bias is to regress actual values of a variable against a constant and the corresponding forecasts:

$$A_t = \alpha + ßF_t + \varepsilon_t$$

and then test the joint null hypothesis that α= 0 and *ß* = 1[19]. Inability to reject this hypothesis indicates that agents' forecasts are considered unbiased.

However, Holden and Peel (1990) have shown that while the condition α = 0 and *ß* = 1 is sufficient for unbiasedness, it is not necessary. This means that unbiasedness is still possible even if the null hypothesis is rejected by the data. Mankiw and Shapiro (1986) also highlighted the fact that because of the autoregressive nature of the right-hand side of the equation, there is a small-sample bias that tends to reject the null of rationality too often. Both articles propose to regress forecast errors on a constant instead and test whether the constant can be restricted to zero:

$$A_t - F_t = \alpha + \varepsilon_t$$

[19] This test is known as a Mincer-Zarnowitz (1969) test, although the condition that α= 0 and *ß* = 1 was first suggested by Theil (1966) (from Croushore (2010)).

As demonstrated by Holden and Peel (1990), the condition $\alpha = 0$ is both necessary and sufficient for unbiasedness. We will use this method to test for bias. All tests are carried out in Stata using the Newey-West method instead of OLS regression as there is autocorrelation and heteroskedasticity present in our time-series (see section 4 under Autocorrelation).

Tests of efficiency

To earn the declaration of full rationality, more than unbiasedness is required: the forecast error must be uncorrelated with the entire set of information that is available to the forecaster at the time the forecast is made; a criterion known as efficiency. To reveal if agents failed to employ important information in making their forecasts, efficiency tests determine whether forecast errors are correlated with such important information. In the following, we will present four tests of efficiency: (1) adding lagged actual values, (2) adding forecasts, (3) adding lagged forecast errors, and (4) adding full information set. The first three tests examine weak-form efficiency, while the last one tests strong-form efficiency.

Efficiency-test 1 – Adding lagged actual values

The rationality concept implies that prediction errors at least must be uncorrelated with historical information on prior realizations of the variable being forecasted. This is a weak-form efficiency condition and specifies that if lagged values of the actual variable, A, are added to the right-hand side of the regression model, they should appear with zero coefficients (Lovell 1986). The first test of efficiency tests this condition:

$$A_t = \alpha + ß_1 F_t + ß_2 A_{t-12} + \varepsilon_t$$

In the spirit of Holden and Peel (1990) and Mankiw and Shapiro (1986), we will adjust this to

$$A_t - F_t = \alpha + ß_1 A_{t-12} + \varepsilon_t$$

Now, the coefficients α and $ß_1$ should not differ significantly from zero. Usually, the most recent realized value of the forecasted variable known to the forecasters is employed. However, if forecasts are done for the next 12 months, the real corresponding values will be the realized values for the last 12 months, and not monthly annualized values which can contain seasonal noise. Actual values for the last 12 months are therefore used on the right hand side[20]. Tests are carried out for α = 0, β = 0 and the joint hypothesis of both being zero (α = β = 0).

Efficiency-test 2 – Adding forecasts

In the second test of efficiency, forecast errors are regressed on a constant and forecasts are added to the right-hand side of the equation. The test is carried out to see whether there is information in the forecasts themselves that can be used to predict forecast errors (Mankiw et al. 2003):

$$A_t - F_t = \alpha + ß_1 F_t + \varepsilon_t$$

These regressions should have no predictive power if the forecasters are efficient and rational. Tests are carried out for α = 0, β = 0 and the joint hypothesis of both being zero (α = β = 0).

Efficiency-test 3 – Adding lagged forecast errors

The next regression tests whether forecast errors are persistent. If they are, this year's error can be forecasted on last year's errors, and forecasts can be improved by knowing the last forecast error.

$$A_t - F_t = \alpha + ß_1(A_{t-12} - F_{t-12}) + \varepsilon_t$$

This is basically a test of autocorrelation, which we know by construction is present in the data (see part about autocorrelation). However, the specific test here is if the error made is still persistent 12 months later. When interpreting the coefficient, it states the extent to which errors made a year ago persist in today's forecast, as done in Mankiw et al. (2003).

Efficiency-test 4 – Adding relevant publicly available information

The test for bias and the three efficiency tests are tests of weak-form rationality. The full rationality assumption has a more demanding implication: it requires that any other variables known to the forecaster (for example publicly available information on the

[20] Real GDP is published quarterly, with a lag of two months. It is important that the values included in this regression was actually known at the time the forecasts were carried out, and values for the first quarter in a year is not known until May. We have taken this into account, and actual values of GDP added to the right hand side are values actually known at the time of the corresponding forecast (revisions not taken into consideration).

growth rate of money supply, unemployment rate, etc.) must also be uncorrelated with forecast errors as a condition of strong-form rationality. (Lovell 1986)

A test for this strong-form rationality is to regress the forecast error on the variables assumed to be included the information set. The information set presumably includes all the variables that would be contained in a sophisticated economic model of the variable being analyzed. One then tests individually or jointly for the significance of the items in the information set. If the variables in the information set are significantly correlated with the forecast error, this finding suggests that agents failed to sufficiently take account of such information in producing their inflation forecasts. (Thomas 1999)

$$A_t - F_t = \alpha + ß_1 F_t + ß_2 y_{t-12} + ß_3 \pi_{t-12} + ß_4 U_t + ß_5 i_t + \varepsilon_t$$

This test asks whether the forecasts done for each variable take adequate account of publicly available information at the time the forecasts are carried out. Forecast errors are regressed on recently available information on macroeconomic data. We include the forecast itself (F), and GDP growth (y) and inflation (π) over the last 12 months, as well as the current unemployment rate (U) and interest rate (i) at the time the forecast is carried out[21]. This list is not exhaustive for all relevant information in forecasting macroeconomic variables, but according to the literature, it should be sufficient to test for strong-form efficiency. Significance is tested for each variable and we test joint significance for the lagged information variables.

[21] As mentioned earlier in the paper, one might object that using real-time data would better reflect the information available when forecasts were made (Croushore 2010), but most of these variables are only subject to small revisions. The exception is GDP growth, which can be considerably revised. We have to take into consideration that this might induce errors in the regression.

Part 2: Analysis of forecasts throughout business cycles

The rest of this paper analyzes the Consensus Economics forecasts for Japan, and the analysis is divided into two segments. This first part of the analysis (Part 2) looks at forecasts throughout the business cycles that were outlined in section 1. The second part (Part 3) will look closer at real GDP growth, interest rates and inflation expectations in different monetary policy regimes.

This first part of the analysis is organized in the following way. First, in section 6, we take a preliminary look at forecast behavior by studying different figures that compare our time series of forecasts with actual values. Second, section 7 studies forecast accuracy and rationality over the whole time period. Third, we divide the full sample into contractions and expansions as defined by ESRI (2010), to see how forecasts perform in different parts of a business cycle (section 8). Finally, in section 9, we look in more detail at interesting observations in specific contractions and expansions.

Previous research on survey expectations has mainly focused on forecast accuracy and rationality of expectations in different countries (for example Blix et al. (2001)), and on expectations of different economic agents (see section 3). In addition, inflation expectations have been frequently studied. However, to our knowledge, no research has been done where the main focus has been on the behavior and performance of forecasts over business cycles.[22] Moreover, we include forecasts of several macroeconomic variables in addition to the usual inflation forecasts.

[22] Two exceptions are Thomas (1999) and Mehra (2002), who briefly touch upon different stages in business cycles explaining over- and underestimation of inflation.

Section 6: Preliminary look at the dataset in business cycles

The first section of the analysis takes a preliminary look at the dataset. Three different figures of the selected time series are presented to create a visual of how expectations performed from October 1989 to January 2010[23]. These figures compare mean, high and/or low forecasts for each variable with actual values for the next 12 months (the exception is the 3-month CDs that display graphs for forecasts and actual values in three and twelve months). Contractions in these four business cycles are illustrated by light grey shaded areas on all figures that divide the time period into contractions and expansions in this paper, and expansions are present outside the shaded areas.

The first figure of each variable shows the forecast band alongside actual values. The forecast band is assembled by taking low, mean and high forecasts, and it captures the span of all forecasts within our time period. This figure is useful to explore how actual values have moved compared to forecasts. For example, one may expect that actual values do not move outside forecast bands, but, as we will see, this happens on several occasions.

The second figure for each macroeconomic variable compares mean forecasts against actual values. This figure is useful as it takes a closer look on mean forecasts together with actual values, which are the two most interesting measures to compare. The figure tells us when forecasts over- and underestimate actual values, while still taking the level of the variable into account.

Finally, the third figure is similar to the second figure as it also compares mean forecasts against actual values. However, this figure displays the size of forecast errors over the time period and ignores the level of actual values. Appendix 5 contains a further description of these three types of figures.

Observable patterns in these three figures are now discussed for real GDP growth, consumer price growth and unemployment rates for the next 12 months. Finally, observations in figures of 3-month CDs in 3 and 12 months are presented.

[23] That is, the forecast in January 2010 is for January to December of 2010, and is compared with the actual outcome of the respective variables for 2010.

Real GDP growth for the next 12 months

Some interesting observations are made by studying figures 5 to 7, which compare forecasts and actual values for real GDP growth. First, actual real GDP growth moves outside the forecast band at several occasions, particularly in contractions when there is a steep fall in actual values. Second, we observe that while actual real GDP growth tends to stay in the upper part or above of the forecast band in times of expansions, the opposite is the case in contractions, when actual real GDP growth stays in the lower part or below the forecast band. This pattern suggests that forecasters overestimate real GDP growth in contractions and underestimate real GDP growth in expansions. Third, actual real GDP growth moves relatively quickly into the lower part of the forecast band when there is a negative trend in the actual real GDP growth and into the upper part when there is a positive trend. Also, these changes from under- (over-) to overestimation (underestimation) appear right before contractions (expansions) start. Fourth, AFEs tend to be higher in contractions than in expansions. This is particularly true for the Asian crisis and the subprime crisis. However, actual real GDP growth appears to have fallen quicker and deeper in these periods compared to other periods. Finally, forecasters do not seem able to predict turning points in the economy. Actual values almost invariably peak and turn or hit a trough and turn long before the forecasts catch up.

Consumer price growth for the next 12 months

Figures 8 to 10 compare forecasts and actual values for consumer price growth. The first important point to reflect on the actual consumer price growth seems to operate closer to the forecast band compared to what we saw for actual real GDP growth. This can indicate that forecasts for inflation are more accurate than forecasts for real GDP growth. Second, the pattern of overestimation in contractions and underestimation in expansions is not as clear for consumer price growth as it was for real GDP growth. For example, forecasts overestimated the actual consumer price growth in most of the expansion after the asset price bubble. Third, AFEs seem to be highest in the subprime crisis for consumer price growth. However, actual inflation also fell faster in this contraction than in other periods. These observations are in line with what we saw for real GDP growth. Fourth, AFEs in the Asian crisis look surprisingly low considering the rapid fall in actual inflation. Finally, turning points are not predicted for inflation either, but forecasters seem to pick up changes more quickly for inflation than they do for real GDP.

Unemployment rates for the next 12 months

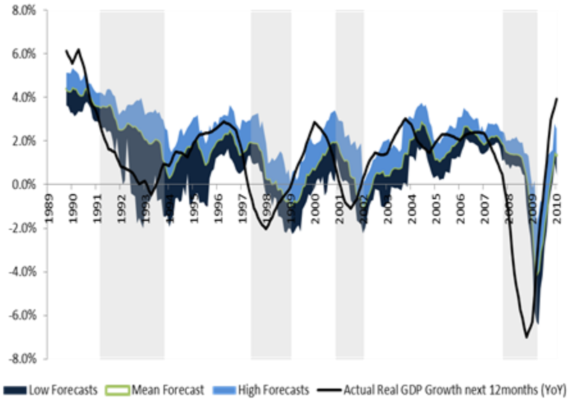
Various patterns can be found by studying figures 11 to 13, which compare forecasts and actual values for unemployment rates. First, the actual unemployment rate seldom moves outside the forecast band, and forecasts seem to be even more accurate for unemployment rates than forecasts for consumer price growth. Second, unemployment rate forecasts appear to overestimate the actual unemployment rate in expansions and underestimate it in contractions. Third, there is no clear pattern of forecast accuracy in different time periods, but the forecast accuracy during the asset price bubble seems particularly high. Finally, while there are no clear peaks or troughs in unemployment, the forecasts still lag behind changes in actual values.

3-month CDs in three and 12 months

Figures 14 to 16 and 17 to 19 display forecasts and actual values for 3-month CDs in 3 and 12 months respectively. These are some of the comments after looking at these figures: First, forecasts for 3-month CDs in three months are more accurate than forecasts for 3-month CDs in 12 months. Second, mean forecasts tend to overestimate actual values for 3-month CDs. An exception to this pattern is the underestimation present in the expansion before the asset price bubble. Finally, the forecast accuracy seems poorest in the period from the end of 1989 to the end of 1995.

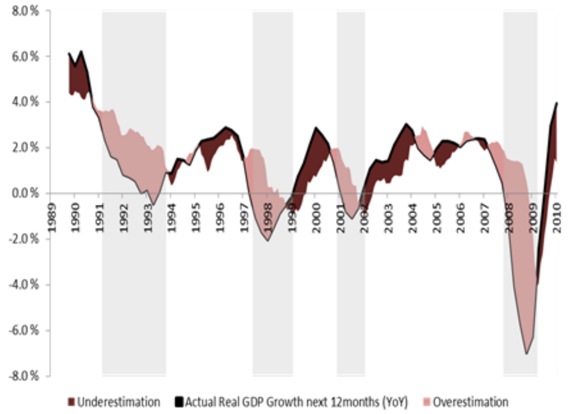
Figure 5-7: Real GDP growth for the next 12 months

Figure 5: Forecast band and actual values



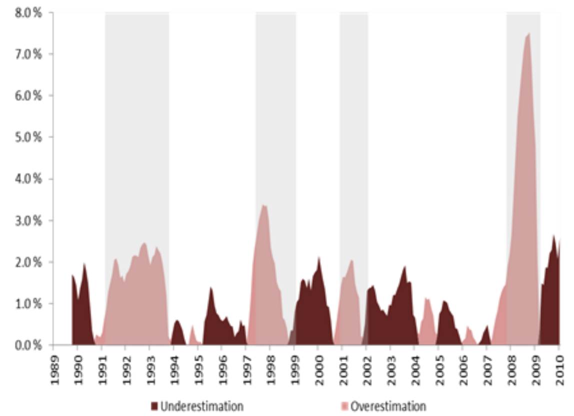
Source: Consensus Economics and Thomson Reuters Datastream

Figure 6: Over- and underestimation - Mean forecasts vs actual values



Source: Consensus Economics and Thomson Reuters Datastream

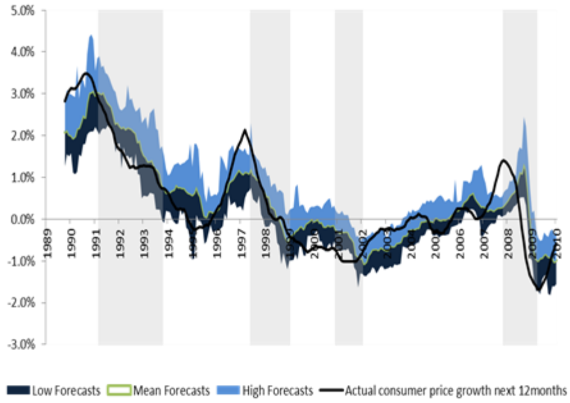
Figure 7: Absolute Forecast Error - Difference between mean forecasts and actual values



Source: Consensus Economics and Thomson Reuters Datastream

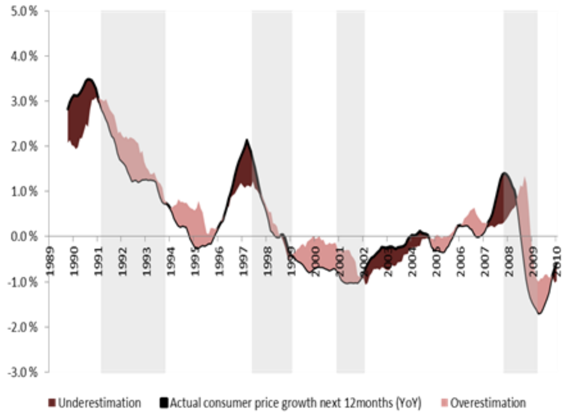
Figure 8-10: Consumer price growth for the next 12 months

Figure 8: Forecast band and actual values



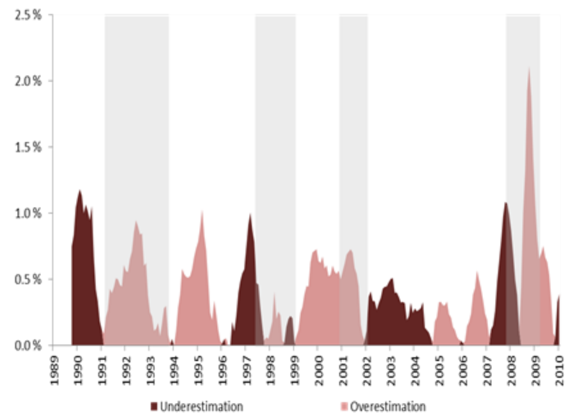
Source: Consensus Economics and Thomson Reuters Datastream

Figure 9: Over- and underestimation - Mean forecasts vs actual values



Source: Consensus Economics and Thomson Reuters Datastream

Figure 10: Absolute Forecast Error - Difference between mean forecasts and actual values

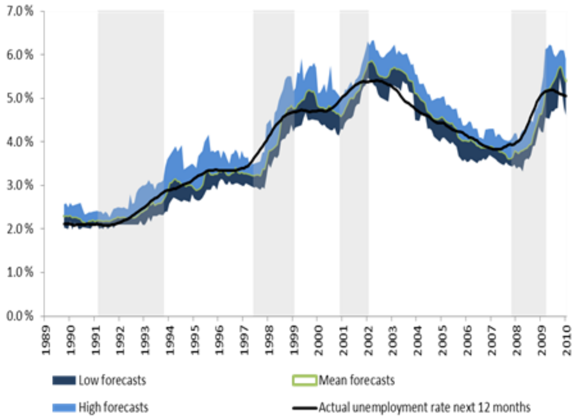


Source: Consensus Economics and Thomson Reuters Datastream

Figure 11-13: Unemployment rate for the next 12 months

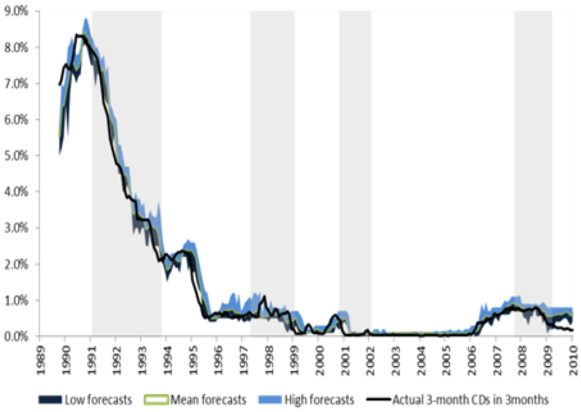
Figure 14-16: 3-month CDs in 3 months

Figure 11: Forecast band and actual values



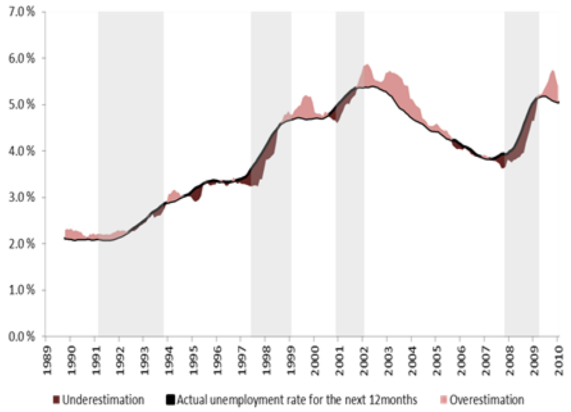
Source: Consensus Economics and Thomson Reuters Datastream

Figure 14: Forecast band and actual values



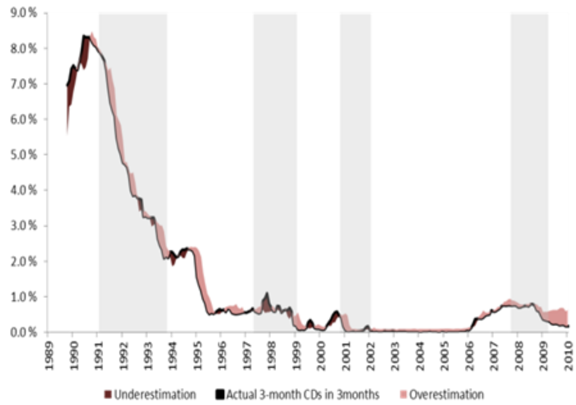
Source: Consensus Economics and Thomson Reuters Datastream

Figure 12: Over- and underestimation - Mean forecasts vs actual values



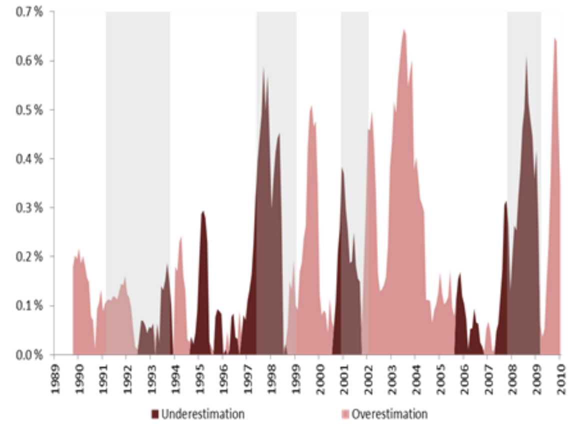
Source: Consensus Economics and Thomson Reuters Datastream

Figure 15: Over- and underestimation - Mean forecasts vs actual values



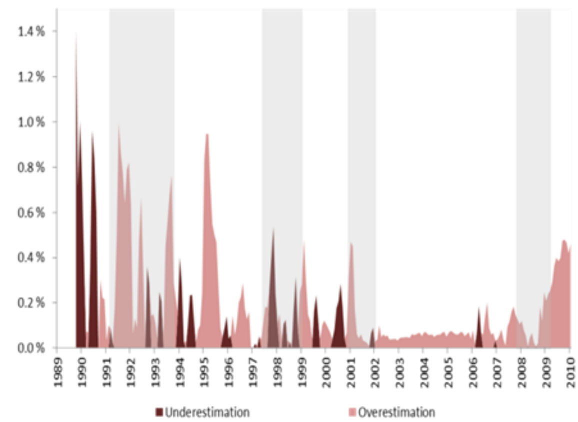
Source: Consensus Economics and Thomson Reuters Datastream

Figure 13: Absolute Forecast Error - Difference between mean forecasts and actual values



Source: Consensus Economics and Thomson Reuters Datastream

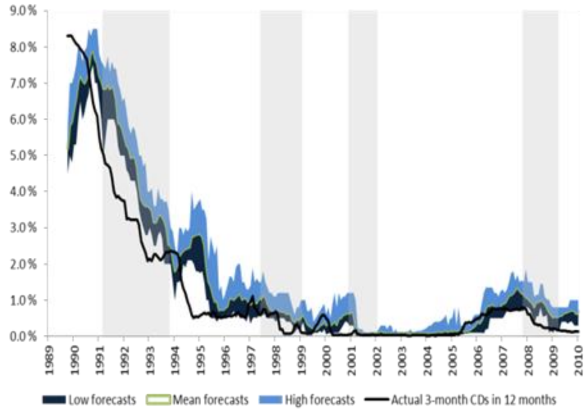
Figure 16: Absolute Forecast Error - Difference between mean forecasts and actual values



Source: Consensus Economics and Thomson Reuters Datastream

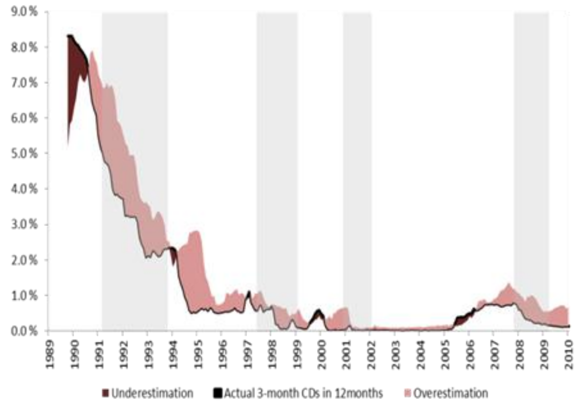
Figure 17-19: 3-month CDs in 12 months

Figure 17: Forecast band and actual values



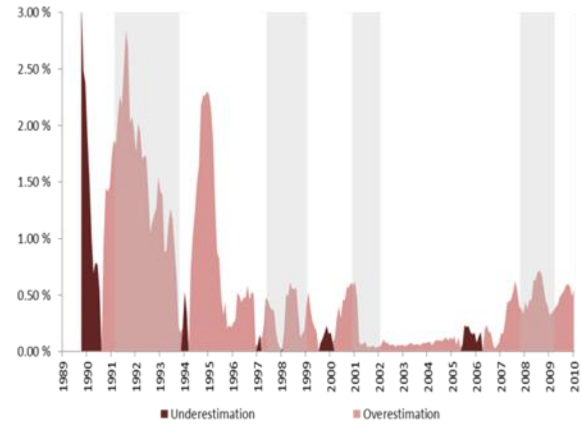
Source: Consensus Economics and Thomson Reuters Datastream

Figure 18: Over- and underestimation - Mean forecasts vs actual values



Source: Consensus Economics and Thomson Reuters Datastream

Figure 19: Absolute Forecast Error - Difference between mean forecasts and actual values



Source: Consensus Economics and Thomson Reuters Datastream

Section 7: Analysis of CE forecasts from 1989 to 2010

The time series in this paper spans two decades and can be divided into different sub-periods. However, we begin with analyzing forecasts in the full sample before separating the sample into expansions and contractions. The last section mentioned that forecast accuracy seemed to vary among variables, and this section will begin by analyzing this observation more closely. After this analysis, we investigate whether CE forecasts are superior to theoretical forecast models in Japan. We then look at forecast bias for each variable, before we evaluate whether the expectations are rational in the sense implied by Muth (1961), both in testing bias and efficiency. Some concluding remarks are given in the end of this section.

Forecast accuracy differs among variables

In our preliminary look at the data we found indications of variation of forecast accuracy between different variables. For example, actual consumer price growth seemed to operate closer to the forecast band compared to actual real GDP growth. MAEs and RMSEs for the different variables over the whole sample support these findings (table 1). Unemployment rate forecasts are most accurate according to these two measures (MAE is 0.20 percent and RMSE is 0.26 percent), while forecasts of real GDP growth are least accurate (MAE is 1.40 percent and RMSE is 1.94 percent). Also, forecasts for 3-month CDs in 3 months are more accurate than forecasts for 3-month CDs in 12 months.

Why does forecast accuracy differ between variables? One possibility is that volatility in actual values cause differences in forecast accuracy. For example, the actual unemployment rate is known to be a slow reacting indicator and this might explain the seemingly high forecast accuracy of this variable. However, forecasts of unemployment rates are still relatively accurate after adjusting for fluctuations in the actual values (MNSE of 0.25 (table 1)), and real GDP growth forecasts remain least accurate (MNSE of 0.86 (table 1)). So even though volatility in actual values might explain some of the variation in forecast accuracy, it does not explain it all.

Forecast accuracy probably also depends on the information flow of different variables. For example, actual values for real GDP are published every quarter, while consumer price growth, unemployment rates and 3-month CDs are published at least every month. A higher release frequency will improve forecasters' information flow and, likely, their forecasts. Blix et al. (2001) also find better forecasts for inflation than for real GDP growth. In addition, real GDP growth is more heavily revised than the other variables, and these revisions might be hard to predict. Another possible explanation of why forecasts of real GDP growth are the least accurate is that GDP is simply harder to forecast, as it comprises more input than price indices like the CPI or unemployment and interest rates.

Finally, forecasts of 3-month CDs in 3 months are probably more accurate than forecasts of 3-month CDs in 12 months because of the obvious fact that it becomes more difficult to make accurate forecasts as the time span of forecasts increases.

Table 1: Forecast accuracy of CE forecasts in the full sample

Measure	Full sample *Real GDP growth*	*Consumer price growth*	*Unemployment rate*	*3-month CDs in 3m*	*3-month CDs in 12m*
MAE: Mean absolute error	1.40 %	0.45 %	0.20 %	0.20 %	0.60 %
RMSE: Root mean squarred error	1.94 %	0.57 %	0.26 %	0.31 %	0.91 %
MPE: Mean prediction error	-0.41 %	-0.11 %	-0.03 %	-0.08 %	-0.45 %
STDEV: Standard deviation	2.25 %	1.19 %	1.06 %	2.31 %	1.99 %
MNSE: Mean normalized squarred error	0.86	0.48	0.25	0.13	0.46

Forecasts are more accurate than theoretical models

So far, this paper has only analyzed forecasts from experts gathered by Consensus Economics. However, theoretical models can also be used to create forecasts, and an interesting question is: can market experts beat theoretical forecast models? This paper addresses this question by comparing the forecast accuracy of mean forecasts from CE with forecasts from theoretical models (which are outlined in appendix 6).

The same measures that were used to assess the forecast accuracy of forecasts from Consensus Economics (MAE, RMSE and MPE) are used to assess the forecast accuracy of theoretical models. In addition, Theil's U-statistics and Diebold-Mariano statistics are used to compare CE forecasts with theoretical models. The methodology and results of this comparison are outlined in appendix 6, and we will only present our main findings here.

Overall, our results from this analysis indicate that CE forecasts are more accurate than theoretical models. CE forecasts have lower MPEs and RMSEs than all theoretical models for real GDP growth, consumer price growth and unemployment rate. Also, the MPE is closest to zero for the CE forecasts, and Theil's U-statistic is below 1 for all theoretical models. Diebold-Mariano statistics show that CE forecasts are significantly better than all theoretical models for these variables[24].

CE forecasts are more accurate than most theoretical models for 3-month CDs, but the random walk model that takes the current 3-month CD as a forecast of the 3-month CD in 3 and 12 months is almost as accurate as CE forecasts (Theil's U-statistic of 0.97 and 0.94). Also, the CE forecasts are not significantly better than this model (DM statistics of 0.60 and 0.54). In other words, experts could not beat the simplest theoretical model when forecasting 3-month CDs. This result may be due to monetary policy regime changes from 1989 to 2010, which we will investigate further in part 3.

Croushore (2010) and Ang et al. (2007) also investigate whether simple time-series models can do better than survey forecasts for inflation and find that they do not. We add to this literature by studying additional variables and focusing on forecasts in Japan. In section 3 we concluded that consensus expert survey forecasts were the best forecasts to use, and since these forecasts also beat the theoretical models, we are further reassured that the CE forecasts are among the best forecasts for macroeconomic variables in Japan.

Forecasts overestimate actual values on average, but not significantly

From the figures in section 6, it is not easy to see whether forecasters systematically over- or underestimate actual values over the full sample. The forecasts seem to interchange between over- and underestimation. The exception is forecasts for 3-month CDs in 12 months, where forecasts almost invariably stay too high (see figures 17 to 19). MPEs in table 1 give us more specific information, and reveal that in fact all variables show negative forecast errors, implying overestimated expectations. On average, actual values for all macroeconomic variables in question turn out to be lower than expected.

[24] One exception is the Diebold-Mariano statistic for the WMA model for real GDP growth and unemployment rate, which forecasts are not significantly different from CE forecasts. However, this model is "post-based" (appendix 6). We have also conducted tests following Stock and Watson (1999). The results from these tests are similar to our findings from Diebold-Mariano statistics. Forecasts from theoretical models seem to add little or nothing to forecasts from CE.

To dig deeper, we carry out formal tests of bias (see table 2). Are the forecasts biased on average? The results show that neither expectations for real GDP growth, inflation nor unemployment rates are biased; the mean forecast error is not significantly different from zero for any of these variables. For interest rates, however, the null hypothesis of unbiasedness is rejected for both 3-month and 12-month forecasts. The errors are quite small, especially for the 3-month horizon, but they are consistent. A negative development in actual values in the period might explain the overestimation[25], but this will be investigated later in the paper.

Table 2: Rationality tests for CE forecasts in the full sample

	Rationality	Full sample				
		Real GDP growth	*Consumer Prices*	*Unemployment Rate*	*3 month CDs in 3m*	*3 month CDs in 12m*
Bias[1]	α[5]	-0.0041	-0.0011	-0.0003	-0.0008*	-0.0045**
	$\alpha = 0$	0.2542	0.2979	0.5792	0.0496	0.0029
Eff. test 1[2]	$\alpha = 0$	0.9855	0.3597	0.1001	0.0078	0.0024
	$\beta = 0$	0.0232	0.7957	0.0723	0.9199	0.3019
	Ho $(\alpha=\beta=0)$	0.0692	0.5813	0.1921	0.0066	0.0000
Eff. test 2[3]	$\alpha = 0$	0.7146	0.2390	0.0500	0.0110	0.0289
	$\beta = 0$	0.4206	0.8402	0.0298	0.9602	0.2454
	Ho $(\alpha=\beta=0)$	0.2975	0.4982	0.0856	0.0070	0.0000
Eff. test 3[4]	$\alpha = 0$	0.1437	0.0872	0.6994	0.0022	0.0054
	$\beta = 0$	0.5640	0.5521	0.5904	0.0006	0.6828
	Ho $(\alpha=\beta=0)$	0.2912	0.2309	0.7944	0.0004	0.0003

[1] $A_t - F_t = \alpha + \varepsilon$ [2] $A_t - F_t = \alpha + \beta * A_t + \varepsilon$ [3] $A_t - F_t = \alpha + \beta * F_t + \varepsilon$

[4] $A_t - F_t = \alpha + \beta * (A_{t-12} - F_{t-12}) + \varepsilon$

[5] *The first row displays coefficients for bias, while the other values in the table are p-values from rationality tests*

Forecasts are weak-form rational, but strong-form irrational

We will now investigate whether expectations are rational[26]. We have already done the first test of rationality, and seen that forecasts are unbiased for the full sample, forecasts are unbiased, except for interest rate forecasts, which have been significantly overestimated in the survey period. We will now continue to test for weak-form rationality with the first three efficiency tests. The fourth and final efficiency test examines strong-form rationality.

[25] A point mentioned by Thomas (1999) and Mehra (2002) for inflation forecasts in the US.

[26] See section 5 for methodology.

The three tests of weak-form efficiency display a similar pattern (see table 2). Actual values for real GDP growth (efficiency test 1) and forecasts for unemployment rate (efficiency test 2) seem to contain information about forecast errors (since tests of $\beta = 0$ are rejected for these two variables). Nevertheless, this is not sufficient to reject weak-form rationality, since the null hypothesis of rationality (the joint hypothesis of $\alpha = \beta = 0$) is not rejected for any of the tests for real GDP growth, inflation and unemployment rate. We conclude that forecasts are rational for these variables for the full sample. As for bias, the story is different for interest rate expectations when testing for weak-form rationality. All tests are rejected, but they are mainly rejected because of bias, not inefficiency.

When testing for strong-form efficiency, we add additional information by including variables to the right-hand side of the equation. These variables are forecasts and actual values for the variable in question, as well as other actual values of relevant variables known at the time of the forecast. Following Ball and Croushore (1995) and Mankiw et al. (2003), we interpret the coefficients in the tests according to economic theory and practice[27]. Failure to take into account the interaction of macroeconomic variables may lead to wrong conclusions and forecasts.

The null hypothesis of strong-form rationality is rejected for real GDP growth and unemployment rate, but not for inflation (see table 18, appendix 7). In the following, we will discuss findings for each variable respectively.

Forecasters seem to overreact to news in other macroeconomic variables when forecasting real GDP growth. All signs of the coefficients indicate that the variables' effect on real GDP growth is less than forecasters believe, but only coefficients for inflation and interest rates are significantly different from zero. In addition, the joint significance test of all coefficients equaling zero is rejected, leading us to reject strong-

[27] For example, high real GDP growth implies higher activity in the economy, and the unemployment rate is likely to fall. Failure to take higher real GDP growth into account when forecasting unemployment rates would therefore likely result in overestimation of unemployment and a negative forecast error. This will be reflected in a negative sign on the coefficient of real GDP growth. If the coefficient is positive, however, forecasters have on average overestimated the effect higher real GDP growth will have on unemployment.

form efficiency for real GDP forecasts, which indicates that these forecasts are not strong-form rational.

As for GDP growth, the forecasters seem to overreact to information from other macroeconomic variables when forecasting unemployment rate. Only the coefficient for interest rates is significantly different from zero, but the joint test of all coefficients being zero is rejected, and we conclude that forecasts of unemployment are not strong-form efficient and therefore irrational.

All in all we find mixed evidence on under- and overreaction to information in other macroeconomic variables when forecasting inflation. It seems that inflation forecasts are efficient in this stronger-form test, and therefore rational[28]. This is quite surprising as most strong-form efficiency tests reject rationality of inflation forecasts (see for example Gerberding (2006) and Thomas (1999)). However, this result changes if we include money supply in the regression (see table 19, appendix 7). This variable is strongly significant in predicting forecast errors, and the joint hypothesis of strong-form rationality is now rejected for inflation as well. Forecasters do not seem to make sufficient use of information contained in the level of money supply when forecasting inflation. Strong-form rationality is also rejected for interest rate forecasts when money supply is included in the test.

Concluding remarks on forecasts for the full sample

All in all, CE forecasts for the main macroeconomic variables seem to be quite accurate when looking at the full sample, even though strong-form rationality is rejected. They forecast better than theoretical models, and stand up to tests of weak-form rationality. This is in accordance with several papers, including Croushore (2010) and Gerberding (2006). The exception is forecasts of interest rates, which are both biased and inefficient, and are therefore considered irrational in our sample.

When looking at figures 5 to 13 in section 6, forecasts for real GDP growth, consumer prices and unemployment rates appeared to have systematic patterns over the business

[28] None of the coefficients are significantly different from zero, and signs on the coefficients are spread evenly between indication of under- and overreaction. This conclusion is further strengthened as the joint test of all coefficients being zero is not rejected.

cycle. Even though weak-form rationality is not rejected for the full sample, we are motivated by this observation to examine how expectations behaved in periods of contractions and expansions. For example, graphs 6 to 8 seem to indicate rather consistent overestimation of GDP growth in contractions, and underestimation in expansions. Are expectations irrational in both upturns and downturns in the economy, but in opposite directions, so that on average rationality is not rejected for the full sample? The next two sections will look closer at this.

We do not see any particular pattern of interest rate forecasts in business cycles, as overestimation seems to occur in most expansions and contractions (figures 14 to 19). As we pointed out in section 1, experts did not see the severity of the asset price bubble, and were overly optimistic for too long a time. This optimism may have caused overestimation in a long period after the bubble burst and is perhaps the explanation of the significance bias for 3-month CDs. This is consistent with the overestimation of inflation in the same period. Lower inflation than expected can have led to lower interest rates than expected. In addition, from 1999 onwards, the interest rates were basically zero, making underestimation literally impossible until the end of the zero interest rate policy.

The noteworthy patterns of interest rate expectations prompt us to delve further into the seemingly sluggish interest rate forecasts. However, analyzing forecasts for 3-month CDs according to different monetary policy regimes seems more suitable and makes more sense in the light of Japan's economic history. Expectations of interest rates will therefore not be included in the next two sections that examine forecasts in the business cycle. Instead, we will examine interest rate forecasts separately in the third part of this paper, in which we study the monetary regimes in Japan and how they connect to the irrational expectations of interest rates.

Section 8: Forecast performance in expansions versus contractions

In this section, we will analyze how forecasts perform in expansions and contractions, as defined by ESRI (see section 1). This is done in order to more closely examine how expectations behave over the business cycle – are there systematic differences between contractions and expansions? In section 6, we saw indications of such differences, and we begin with analyzing the observation of forecasts being more accurate in expansions than in contractions. Then we study differences in disagreement among forecasters in these sub-periods, before looking into patterns of over- and underestimation. Finally, we show that forecasters are irrational in contractions and expansions. Possible explanations for this section's findings are provided along with the different observations.

Forecasts are more accurate in expansions than in contractions

Forecast accuracy seems to be higher in expansions than in contractions. This result can be observed by comparing MAEs and RMSEs in expansions and contractions for the different variables (table 10, appendix 7). MAEs and RMSEs are lower in expansions than contractions for all variables. For example, the MAE for real GDP growth is 0.90 percent in expansions and 2.37 percent in contractions.

Higher volatility in macroeconomic variables in contractions compared to expansions may explain higher forecast accuracy in expansions. In our sample, the volatility of actual real GDP growth and unemployment rate has been higher in contractions than in expansions from 1989 to 2010. The forecast accuracy of unemployment rates is slightly higher in contractions compared to expansions after adjusting for differences in volatility, but real GDP growth is still more accurate in expansions when comparing MNSEs. MNSEs are also higher in contractions for consumer prices and 3-month CDs. These results indicate that differences in volatility only partly explain why forecast accuracy is higher in expansions. (Table 10, appendix 7)

Forecasters seldom predict turning points in the economy. There are turning points before both contractions and expansions, but the sudden shock leading to contractions often seems more surprising than the turnaround before an expansion. Unexpected events can have sudden and large impacts on the real economy. For example, there was a severe decline in global economic activity after Lehman Brothers went bankrupt on

September 15th, 2008 (Quintanilla and Sprinzen 2008), when liquidity dried up and fear replaced optimism. The massive measures taken by authorities around the globe in response to this crisis made the expansion that followed more foreseeable than the downfall. If the shocks leading the economy into contraction are more unexpected than the take-offs initiating expansions, forecast accuracy will be lower in contractions than in expansions, a point also made by Blix et al. (2001).

Disagreement is higher in different parts of the business cycle

All in all, it seems logical that higher uncertainty of future economic development makes it harder to produce accurate forecasts in contractions than in expansions. If this is the case, it is interesting to see whether forecasters disagree more in contractions than expansions. In most textbook macroeconomic models, people form expectations based on a mutual information set shared by everyone (Mankiw et al. 2003). In addition to having the same information available, people also have the same information processing capabilities. The models therefore assume that everyone has the same expectations, and we simply accept this since the models tells us to. But this is an assumption easily rejected by the data. Anyone who has looked into survey data on expectations can see that disagreement among forecasters is substantial, and the forecast bands in figures in section 6 suggest that disagreement also occurs in our data.

We have conducted an analysis of disagreement of forecasts inspired by Mankiw et al. (2003), who analyze 50 years of inflation expectations data from several sources. Their research documents substantial disagreement among both consumers and professionals about expected future inflation. We carry out the same tests, but use different measures of disagreement (standard deviation and high-minus-low instead of the interquartile range). Appendix 8 contains a more thorough discussion of these measures and a description of the disagreement tests. We test forecasts for real GDP growth and unemployment rate in addition to inflation forecasts. Disagreement is studied in expansions versus contractions, and the analysis is done for the two raw time series, current year and next year, and the constructed forecasts for the next 12 months[29].

---

[29] We do not have standard deviation for forecasts for the next 12 months, and hence only high-minus-low is analyzed for this time series.

We find that disagreement is significantly higher in contractions than in expansions for expectations of real GDP growth for current year and next 12 months' forecasts (see table 20, appendix 8). If contractions are times of higher uncertainty, this is as expected, since the dispersion in how things will develop is big, perhaps especially in the beginning of a crisis. However, it is somewhat surprising to see that the picture has changed when looking at tests done for next year forecasts. Here, disagreement is higher in expansions than in contractions for real GDP growth. This observation might be explained with the following reasoning: in expansions, forecasters will either expect the uptrend to continue next year, or predict a turning point. Contractions, one the other hand, are shorter and faster, leading most forecasters to expect normalization in the economy within a couple of years.

For current year unemployment forecasts, the disagreement level is not significantly different in expansions compared to contractions. However, disagreement is higher in expansions than contractions for next year forecasts and forecasts for the next 12 months. The explanation for this, as far as we see it, is related to the one we gave for real GDP growth for next year forecasts. In a contraction, forecasters almost invariably expect unemployment to rise over the medium term, especially since unemployment rate is a lagging variable. In expansions, on the other hand, it is more difficult to see decide if the economy will continue to grow or if it will reach a peak and turn.

For inflation, disagreement is not significantly different in the two sub-periods. Interestingly, there seems to be less overall dispersion in inflation forecasts than for real GDP growth forecasts, perhaps indicating that inflation is easier to predict, and thus giving rise to less disagreement. Both levels for SD and HML are roughly half the levels of GDP growth. This is in accordance with our findings of better forecast accuracy for inflation than for real GDP. Unemployment is the variable least disagreed upon, consistent with the findings for forecast accuracy (section 7).

What makes disagreement rise and fall?

The second disagreement test explores further what makes disagreement rise and fall (see table 21, appendix 8). In order to do this analysis we need to study fixed horizon forecasts and the test is therefore only carried out for HML for the next 12 months. High actual values and high forecasts for real GDP growth result in lower disagreement of real GDP growth compared to low actual values and low forecasts. The opposite is true for inflation and unemployment rate, where high actual values and high forecasts correlate with high disagreement. If actual real GDP growth increases, disagreement falls, and if actual real GDP growth decreases, disagreement rises. The opposite is true for unemployment, where an increase in unemployment rate corresponds to higher disagreement, and a decrease gives less disagreement. Changes in actual inflation are not significant and therefore inconclusive.

Real GDP growth is usually higher in expansions than contractions, and the results from the second disagreement test therefore support the finding that disagreement of real GDP growth is lower in expansions than in contractions. Unemployment rate usually increases in contractions and decreases in expansions, and disagreement about unemployment rate should therefore be higher in contractions compared to expansions. However, in the first disagreement test, we found the opposite. A possible explanation is that the level of the unemployment rate has been higher in expansions than in contractions[30].

Our findings add to the literature on disagreement in forecasts by linking the level of disagreement to different stages of the business cycle. Not only do forecasters seem more or less irrational throughout the business cycle, they also seem to disagree more with one another during certain parts of the cycle.

Overestimation and underestimation in contractions and expansions

For the full sample, we found that forecasters slightly overestimate actual values on average, but the results were not significant for real GDP growth, inflation and unemployment rates. If we separate the full sample into periods of contractions and expansions, the findings become more graded and interesting.

There is systematic overestimation of real GDP growth in contractions (MPE -2.27 percent) and underestimation in expansions (MPE 0.56 percent), see table 10 (appendix 7). Furthermore, a formal test of bias where we include a dummy for times of contractions, shows that the bias is significant in both contractions and expansions (table 12, appendix 7). Real GDP growth is expected to fall in contractions and rise in expansions, but the rate of change and how fast it occurs seems to take forecasters by

[30] The average actual unemployment rate was 3.99 percent in expansions and 3.30 percent in contractions.

surprise in both contractions and expansions. The pattern is economically similar for unemployment expectations. Actual unemployment rates usually increase in contractions and decrease to some extent in expansions, and the tests of bias show a significant under-prediction of unemployment in contractions and over-prediction in expansions. Inflation expectations display significant bias of overestimation in contractions, but the underestimation in expansions is small and not significant.

The biases indicate that forecasters systematically underestimate the magnitude of downturns in the economy in contractions and the pace of recovery in expansions.

Forecasts are not rational in contractions and expansions

We have shown that bias is systematic in contractions and expansions, an indication of irrational expectations. We will follow the same procedure for the efficiency tests. Splitting the full sample in two by using dummy variables for times of contractions clearly displays inefficiency in both expansions and contractions, though the results differ somewhat between the variables (table 13 to 17, appendix 7).

For expectations about real GDP growth, all tests of rationality are rejected in both sub-periods. Bias is bigger in contractions for this variable, but forecasts are more inefficient in expansions. Forecasts of unemployment rate are judged to be irrational by two out of three efficiency tests in expansions, but none of the tests are rejected for contractions[31]. Both lagged values and forecasts seem to contain information about forecast errors in expansions for this variable. For inflation expectations, rationality is only rejected for contractions, and none of the added variables can explain forecast errors. These results make inflation expectations the most efficient of the three variables.

The efficiency tests thus display results that are more indecisive than bias tests, but they still paint a clear picture that forecasters are irrational in both contractions and expansions, enforcing the pattern found in figures 5 to 13. Even though forecasts exhibit weak-form rationality in the full sample of 21 years, more thorough testing of the upturns and downturns in the economy tells a story of consistent bias and irrationality.

[31] However, the p-value for the null of rationality is low in all tests for unemployment rate forecasts in contractions.

Possible reasons for bias and inefficiency

Our results from rationality tests are consistent with most previous empirical studies, which almost invariably find that survey expectations are not fully rational in the sense implied by Muth (1961) (e.g. Gerberding (2006), Mankiw et al. (2003) and Roberts (1998)). Several papers have found that forecasts are weak-form rational, particularly when testing long periods and several surveys. Croushore (2010) shows this, and also points to sub-periods during which expectations have been irrational in the US. We have taken this one step further by linking the irrationality to specific sub-periods, namely periods of contractions and expansions. Thomas (1999), De Long (1995) and Dotsey and DeVaro (1995) find that inflation expectations underestimate inflation when inflation is rising, and overestimate inflation when inflation is declining. Again, we relate this to periods of contractions and expansions, and expand the research by also including expectations of unemployment rate and real GDP growth.

Even though the forecasts pass the tests for weak-for efficiency and unbiasedness for the full sample, we have shown that this is not the case for sub-periods when dividing the sample into contractions and expansions. Should we proceed to conclude that forecasters are irrational? It seems plausible, but we will not jump to conclusions. There are several reasons why signs of bias and inefficiency might not necessarily mean irrational expectations. Some papers are very blunt in concluding that agents are not rational, but we wish to contemplate this further and provide possible explanations for our findings – especially alternative reasons for our test results. Some of the most interesting points in the literature on why failing tests of rationality do not automatically mean that forecasters are irrational are: 1) individuals may have different information and processing capabilities, 2) forecasters may have incomplete information and problems distinguishing between temporary and permanent shocks, 3) informational frictions may exist, 4) forecasters can behave strategically, 5) using revised data may induce bias, and 6) studying only certain sub-periods in the data can yield irrational outcomes. These points are elaborated below and commented on in relation to our data and findings. The list is not exhaustive, however, and it is beyond the scope of this paper to systematically work through all the issues stated.

Different information

If different individuals have different information, tests based on the mean forecast can be biased (Keane and Runkle 1990). Some tests of rationality therefore require micro data. However, our mean is collected from expert economists, who should have access to the same information and have roughly the same capabilities. The tests carried out with the mean forecasts should therefore be valid.

Incomplete information and problems distinguishing between permanent and temporary shocks

With incomplete information about the nature of the shocks hitting the economy, it may be rational for forecasters to adjust their forecasts only gradually in response to new information. This may lead to continuous bias over a period. However, even if data and model uncertainty give rise to such persistence in forecast errors, they should not persist indefinitely, though the adjustment process may take considerable time (Gerberding 2006).

Early studies in the literature of inflation expectations deemed agents irrational by the mere existence of serial correlation in the error terms. As we have pointed out earlier on, there will inevitably be autocorrelation in our time series since they inhibit overlapping observations. As mentioned, we correct for autocorrelation (as far as possible) in our tests by using Newey-West HAC estimators. But as we could see from efficiency test number 3, adding lagged variables to the right-hand side of the regression still added explanatory power for real GDP forecasts in expansions. Is this a sure sign of irrationality? Not necessarily. Cukierman and Meltzer (1982) and Cukierman (1986) argued that even completely rational agents might be incapable distinguishing correctly between temporary and permanent shocks to aggregate supply or demand. If permanent shocks are mistakenly viewed as being temporary (which can be rational beliefs), agents can repeatedly make one-sided forecast errors. This is plausible during turning points in the economy – it is hard to say whether a turnaround is only temporary, or if the tides have changed, until the development seems persistent.

Informational friction and sticky expectations

There are strong assumptions behind the rational expectations hypothesis (see section 2). Therefore, researchers have tried to explain the findings of irrational expectations by proposing alternative models to the rational expectation hypothesis. Mankiw et al.

(2003) suggests a "sticky-information" model to answer their findings of disagreement and irrational expectations. In this model, economic agents update their expectations only periodically because of costs of collecting and processing information. They find that the model broadly matches many of the facts of their data, including several departures from full rationality, such as autocorrelation in forecast errors and insufficient sensitivity to recent macroeconomic news[32]. The sticky-information model seems sensible for households, but for professionals it should be different, since their job is to collect and process this information.

The model of Marcet and Nicolini (2005) is in the same spirit. They use a "sticky-expectations" model to examine the observed relationship between money growth and inflation. Instead of shocks and information being immediately understood by all agents, they propose a model in which agents learn how to formulate their expectations by observing the behavior of the economy. This model is also more appropriate for explaining the behavior of households than the behavior of experts' expectations.

Furthermore, Andolfatto et al. (2005) show that rational expectations may, in small samples, fail conventional tests of unbiasedness and efficiency[33]. This occurs if there is informational friction or the economy is not totally transparent, which is in accordance with the sticky-information and –expectation models above. This is relevant for our data, especially since the Japanese economy has been a difficult case for the last 20 years.

Strategic forecasting

Some professional forecasters may behave strategically and fail to reveal their true forecasts (Ehrbeck and Waldmann 1996). They may have other objectives than minimizing forecast errors, for example avoiding risks to their reputations as forecasters

[32] We find that forecasters are over-sensitive to macroeconomic news when making forecasts, which is opposite of Mankiw et al.'s results. However, they only test inflation, while our findings are significant only for real GDP and the unemployment rate.

[33] Andolfatti et al. (2005) construct a DSGE model and embed an assumption of rational expectations. Still, standard regressions run on equilibrium realizations of inflation and inflation expectations revealed an apparent bias. The null hypothesis of rational expectations was incorrectly rejected in several cases, which they interpret as casting doubt on conventional interpretations of evidence on irrationality.

(Lamont 1995). A large forecast error that appears as an outlier compared to the mean may be stigmatizing and embarrassing for the forecaster. In an attempt to avoid this situation, forecasters may try to minimize the deviation of their forecast from the consensus mean instead of stating their true expectations. This is termed strategic forecasting. Concerns about reputation should not apply to consumers and households, but it does seem relevant for professionals, and may therefore apply to our data.

Using revised data

In efficiency test 1 we include lagged actual values of the variable being forecasted. If the beta is significantly different from zero, informational efficiency is rejected. However, it must be ensured that the variables included in the regression were actually available to the survey respondents when they answered the survey. This implies that we should use "real-time data", as pointed out by Croushore (2010) and mentioned earlier in the paper. Following this line of thought, one should not include variables whose time series are subject to major revisions after initial publication, such as real GDP growth.

We have chosen to include real GDP growth in efficiency tests, but we should be aware of the variable's shortcomings when interpreting its coefficients. Statistics that are revised drastically could potentially give a wrong short-term view, and efficiency tests can be wrongfully rejected. This calls for using real-time GDP, but Webb (1987) argued that even if real-time data is used, and rationality still is rejected, data revisions and changing coefficients over time would prevent using the results to make better forecasts. In testing for informational efficiency, money market rates are good, since they are never revised. Inflation rates are only subject to minor revisions, as is the unemployment rate. The inclusion of real GDP growth calls for a watchful eye, but most variables in our efficiency tests are still robust.

Some specific sub-periods can be irrational

Early studies found expectations to be irrational (see also part about survey expectations), but this finding has later been shown to be specific to the data and period tested, for instance by Croushore (2010). He finds no evidence of bias in testing his full data sample, but when looking at sub-samples he finds evidence of bias, especially in the first part of the sample period. He concludes that the apparent bias and inefficiency found in the early literature seems to have dissipated. He calls for deeper investigation into possible bias in particular sub-samples of the data, which is what we have done in this paper. We find that these sub-periods are more specific than Croushore indicated.

As mentioned, Thomas (1999), Mehra (2002) and others have suggested that the bias in the survey forecasts may vary across accelerating versus decelerating inflation environments, or across the business cycles. They indicate that this might be true for other variables as well, a belief consistent with our findings.

Even though Croushore (2010) finds episodes in which forecasters made persistent forecast errors, he concludes that the periods are so short, that by the time they can be identified, they have nearly disappeared. He tries to improve the survey forecast by using information in the forecast errors, but doing so in real-time proves very difficult and only leads to increased forecast errors. A relevant topic for further research is whether one can improve forecasts knowing that there is systematic bias in contractions and expansions[34].

Concluding remarks on irrationality

We are not able to make a strong conclusion about irrational forecasts due to the reasons stated above. As mentioned, it is beyond the scope of this paper to systematically work through all of the catches, but we encourage future research to look especially into the matter of permanent versus temporary shocks and the use of real-time data. Nevertheless, the evidence of bias and inefficiency seems strong and consistent in sub-periods, leaving us with a conclusion of systematic irrationality during contractions and expansions. This is as far as we know a new finding, and we welcome further research on business cycle sub-periods done with data from other countries.

[34] An obvious dilemma arises if this is to be done in real-time, since contractions and expansions are often dated and decided on with a considerable lag.

Section 9: Forecast behavior in different expansions and contractions

The previous section studied forecasts in expansions and contractions, and this section further divides the time series into each specific contraction and expansion. This separation enhances our understanding of forecast behavior as some expansions and contractions differ from what we usually observe in these periods. This section takes a closer look at the following three observations we found in section 6: (1) the subprime crisis seemed to have the least accurate forecasts, (2) the forecast accuracy looked unusually high for consumer price growth in the Asian crisis and for unemployment rate in the asset price bubble, and (3) consumer price growth forecasts did not have a particular pattern of over- or underestimation in expansions and contractions. In addition, we study forecast rationality in the expansion after the asset price bubble.

Subprime crisis has the least accurate forecasts

Overall, forecast accuracy seems to be poorest in the subprime crisis as the MAEs are highest for all variables in this contraction. High MAEs in the subprime crisis might arise from the subprime crisis' severe effect on the real economy (section 1). The movement in the real economy probably took the market by surprise as few could foresee that this crisis would hit as quickly and powerfully as it actually did. This is supported by strong bias for both real GDP growth and unemployment rate expectations in this period[35]. The subprime crisis' MNSEs are in the middle range, and the volatility in actual macroeconomic variables is therefore the main reason for why this period has the lowest forecast accuracy of all contractions and expansions. (Table 11, appendix 7)

Unusually high forecast accuracy in two episodes

In section 8, we found that forecast accuracy is higher in expansions than contractions. However, there are some exceptions: forecasts for consumer price growth are most accurate in the Asian crisis and forecasts for unemployment rate are most accurate in the asset price bubble when we compare MAE, RMSE and MNSE for different contractions and expansions. These measures are closest to zero and there is no sign of bias in these two episodes and variables. This is surprising as there is no obvious reason of why forecasts should perform better in these episodes compared to others; for

[35] Though forecasts of unemployment rate are deemed bias for contractions as a whole, the sub-prime crisis is actually the only contraction where the bias is significant for unemployment expectations.

example low MNSEs imply that volatility cannot explain these episodes. (Table 11, appendix 7)

Consumer price growth is both over- and underestimated in expansions

The previous section stated that there was no significant bias in expansions for consumer price growth, but bias is present in two of the expansions when we break our sample into different expansions. Forecasts underestimate consumer price growth in the expansion before the asset price bubble and overestimate it in the expansion before the dot-com bubble. Both of these forecast errors are statistically significantly different from zero (but in opposite directions which explains why inflation expectations in expansions as a whole are not deemed biased). (Table 12, appendix 7)

Underestimation in the expansion before the asset price bubble is probably due to booming asset prices, such as land and stock prices, which increased the wealth of the Japanese people and led to higher consumer price growth through the wealth effect. It seems that the high inflation took market experts by surprise.

The overestimation in the expansion before the dot-com bubble might be explained by an unexpectedly strong deflation after the Asian crisis. Forecasters seem to have predicted that the consumer price growth would increase after Japan entered a deflationary environment at the end of Asian crisis. These expectations might have their origins in Japan's latest experience with a deflationary environment in 1995 and 1996, when consumer price growth started to increase not long after it became negative. However, consumer price growth actually decreased further in the expansion before the dot-com bubble and Japan experienced a long period of deflation.

Forecasts in the expansion before the Asian crisis are surprisingly rational

The expectations in the expansion before the Asian crisis are more rational compared to other sub-periods. None of the variables are judged to be biased, and only a few efficiency tests reject rationality. In addition, forecast accuracy is among the best for all variables in this period. Looking closer at this expansion, the recovery in the economy was slower and less sudden than during the other expansions. If forecasters usually underestimate the rate and level of change, their expectations should follow the actual development more closely in this period, which they apparently did. (Table 15 to 17, appendix 7)

Part 3: Analysis of forecasts in periods of different monetary policies

The third and final part of this paper assesses the criticism from Bernanke, the OECD and the IMF of the BOJ's monetary policies that were outlined in section 1. We aim to answer the following questions by analyzing forecasts in periods of different monetary policies: Was the BOJ's monetary policy from 1991 to 1994 too defensive according to private sector's forecasts in this period? Did the BOJ lose credibility when they broke out of the zero interest rate policy?

Part 3 mainly analyzes interest rate expectations, but also includes analyses of real GDP growth and inflation forecasts when appropriate. The analysis is motivated by the following finding in section 7: interest rates have not followed any particular pattern in business cycles. Also, it makes sense to analyze interest rates in periods of different monetary policy regimes considering Japan's history over the last 21 years.

The advantages of the dataset is that it enables us to analyze what the private sector expected the BOJ would do while at the same time assessing the private sector's perceptions of the future state of the Japanese economy. The analysis is mainly descriptive; the possible explanations of the results represent our thoughts and intuition and should not be considered as hard evidence. Still, we are able to make robust conclusions in this section because the empirical analysis of forecast accuracy and bias support our findings.

It would have been interesting to analyze central bank announcements and changes in monetary policies' immediate effect on forecasts. However, data limitations have prevented us from doing this analysis (daily observations of forecasts would have been more appropriate to use than monthly observations). For example, an announcement or interest rate change from the BOJ a month from now may affect CE forecasts. However, monthly forecasts may also be affected by other events in the same month, and it is then impossible to say how much of the forecast change is due to the central bank's action. Galati et al. (2010) also point out this problem. We encourage other researchers to look into this type of event studies, for example by using market data of interest rates.

This part of the analysis has two sub-parts. Section 10 begins with a preliminary look at forecast behavior of real GDP growth, consumer price growth and, most importantly, 3-month CDs in the three most distinct monetary policy periods. These periods are the "wait-and-see" monetary policy regime from 1991 to 1994, the "the zero interest rate policy" in Japan from February 1999 to June 2006 and the monetary policy after the zero interest rate policy. Section 11 analyzes observations made in section 10 by looking at forecast accuracy and bias (see section 5), and explains possible reasons behind these findings.

Section 10: Preliminary look at forecasts in periods of different monetary policies

This section takes a preliminary look at forecast behavior in the three distinctive periods of different monetary policies in Japan. 3-month CD is the main variable to be discussed, but the section also includes interesting observations of forecasts for real GDP growth and consumer price growth. Discussions of the findings follow in the next section.

Monetary policy regime from 1991 to 1994

In section 1 we mentioned that Bernanke (1999) criticized the monetary policy in Japan for being too defensive from 1991 to 1994, and that Nakaso (2001) called this a "wait-and-see" policy. We will also use this term to describe the monetary policy regime in this period. The actual 3-month CDs declined in this period and the private sector seem to have underestimated this decline (figure 28 to 33, appendix 9). Also, forecasters seem to have overestimated real GDP growth and inflation in this period (figure 34 to 39, appendix 9).

Monetary policy regime from February 1999 to June 2006

We mentioned in section 1 that the BOJ implemented a zero-interest rate policy in February 1999, and that this policy temporarily ended in August 2000 when the BOJ raised the overnight call rate. However, the zero-interest rate policy was soon reinforced in March 2001 in the dot-com bubble, and the policy lasted until July 2006 (section 1). In this part of the paper we label the monetary policy from February 1999 to July 2006 as the zero interest rate policy because the BOJ kept the interest rates close to zero in this period.

Some interesting observations are made in the zero interest rate monetary policy period when studying figures 28 to 39 (appendix 9). First, forecasts for 3-month CDs seem relatively accurate. Second, forecasts seem to overestimate 3-months CDs except for the underestimation in the period when the BOJ temporary broke out of its policy. Finally, there is a drop in the magnitude of the overestimation of 3-month CDs around the period where the BOJ reinforced its zero interest rate policy (March 2001).

Monetary policy regime after June 2006

The BOJ ended its zero interest rate policy in July 2006, a few months after the deflationary environment temporarily ended (section 1). However, private sector experts did not predict the breakout from the zero interest rate policy and began to underestimate interest rates. The underestimation only occurred for less than a year as forecasters rapidly adjusted their forecasts, and began to overestimate the interest rate. Also, the BOJ stopped raising interest rates and started to decrease them in the subprime crisis. The overestimation seems more severe in this period than under the zero interest rate monetary regime, and the interest rate expectations did not fall when the BOJ started to decrease interest rates. These forecast patterns after June 2006 can be found by looking at figures 28 to 33 in appendix 9.

Section 11: Forecast behavior in distinctive monetary policy periods

In this section we look more closely at observations made in the previous section by using measures for forecast accuracy and bias as outlined in section 5. Findings in distinctive monetary policy periods are discussed chronologically. After this discussion, we explain how regime changes may explain test results of forecast accuracy and bias. Lessons from forecasts in monetary policy regimes are discussed in the end.

Forecast behavior in wait-and-see monetary policy period

The previous section mentioned that the private sector seems to have overestimated real GDP growth, inflation and interest rate in the wait-and-see period. This observation is supported by studying MPEs for this period in table 22. For example, MPEs equal -0.24 percent and -1.24 percent for forecasts of 3-month CDs in 3 and 12 months. This overestimation is statistically significant for interest rate and inflation, and we will now discuss possible reasons for these findings. (Table 22 and 23, appendix 10)

The overestimation of CE forecasts for 3-month CDs suggests that the BOJ reacted more proactively to the crisis than what the private sector expected. The BOJ lowered the uncollateralized overnight call rate and the official interest rate throughout the wait-and-see policy. The official discount rate went from 6 percent to 1.75 percent, while the uncollateralized overnight call rate declined from 8.3 percent to 2.3 percent (figure 40, appendix 9). The magnitude of the decline in interest rates was higher than expected by experts in the private sector.

Why did the BOJ lower interest rates in the wait-and-see period? It may be that the BOJ was surprised by the development in the real economy in this period. As mentioned above, we identified that real GDP growth and inflation fell faster than the private sector expected from 1991 to 1994. The optimism of the Japanese economy in this period, as outlined in section 1, may explain why forecasts overestimated inflation and real GDP growth. Experts in the private sector did not seem to understand how overvalued the asset prices were and how much this would affect the real economy. The BOJ may have shared the same optimism, and perhaps they also overestimated real GDP growth and inflation in this period. This view is supported by the Japanese authorities' expectation of a collateral pick-up and a turnaround in the development of asset prices after the asset price bubble (section 1). Perhaps the BOJ had been repeatedly surprised by the fall in economic activity and the development in prices, and as a result, lowered the interest rates to counteract this development.

Overall, the criticism of BOJ's monetary policy from 1991 to 1994 from Bernanke and others seem harsh because the decline in interest rates proves that measures were in fact taken. Also, the BOJ's monetary policy was too defensive if Bernanke's criticism is right, but the private sector's overestimation of interest rate indicates that the BOJ did more than expected[36]. Finally, if monetary policies were too weak due to the BOJ's optimism of the Japanese economy in this period, experts' overestimation of real GDP growth and inflation suggest that the private sector was equally mistaken.

The measures taken by the BOJ were perhaps not large enough, but that is easy to say 10 years later. However, one might criticize the Japanese authorities for the weak disclosure requirements of NPLs. The size and sincerity of non-performing loans would have been discovered faster if proper disclosure requirements had been in place. Proper disclosure requirements could have encouraged stronger measurements from the BOJ and probably a more painful contraction, but most likely a faster recovery would follow. Instead, Japan entered a deflationary environment and the vicious circle dubbed the lost decade. The disclosure would also have given the private sector more information, and it is likely that forecasts would have been adjusted.

Forecast behavior in zero interest rate policy period

Observations made in the previous section are confirmed by analyzing CE forecasts in the zero interest rate period. Forecasts of 3-month CDs are relatively accurate measured by MAE, RMSE and MNSE. The MPEs of 3-month CD forecasts indicate that forecasters overestimate interest rate, and this overestimation is significant in tests of bias. (Table 22 and 23, appendix 10)

The high forecast accuracy of interest rates suggests that forecasters believed the BOJ in this period. When the BOJ entered the zero interest rate policy, they announced that they would keep its policy "until deflationary concern has been dispelled" (section 1). The strong deflationary environment lasted until May 2006 and signaled to forecasters

[36] However, there is a difference of what the private sector *expects* the BOJ will do, and what private agents believe the bank *should* do. For example, the private sector may have adjusted interest rate forecasts to reflect a defensive BOJ, even though private agents preferred more aggressive measures.

that the interest rate would remain at zero, and we then expect experts to adjust forecasts accordingly, which our results indicate that they did.

The drop in the private sectors' interest rate expectations around March 2001 enhances the view of a high confidence in the BOJ's zero interest rate policy among the private sector. When the interest rate is basically zero, one cannot underestimate it, implicating that overestimation is the only possible bias in this period. However, such a consistent overestimation throughout the period may indicate that some of the experts in the private sector did not believe that the BOJ would stay true to its zero interest rate policy, perhaps marked by the unexpected interest rate increase in August 2000[37]. An announcement to follow zero interest rate policy until concern about deflation has been dismissed is not very trustworthy when the bank increases interest rates while deflation is still highly present. Still, interest rates forecast fell drastically after the BOJ reintroduced their zero interest rate policy in March 2001.

The lack of higher inflation expectations may suggest that the private sector had low faith in the BOJ's capabilities of ending the deflation, and this belief among private experts may have its origins in the BOJ's temporary brake from the zero interest rate policy. To raise inflation expectations was one of the objectives of the BOJ's zero interest rate policy, as higher inflation expectations transfer into higher actual inflation. It is therefore interesting to notice that expectations underestimated inflation from 2002 to 2005. The consumer price growth was still negative, but turned out to be higher than expected in this period. The private sector could have increased inflation forecasts when they realized that they underestimated inflation, but they did not.

To sum up, it seems that the private sector believed the BOJ's zero monetary policy despite the temporary break from this policy, but the distortion may have lowered the private sector's confidence in the BOJ's ability to end deflation.

[37] Also, we are looking at mean forecasts and it is unlikely that *all* respondents will forecast a zero interest rate. In other words, the mean will be dragged up by those who predict a higher interest rate. It is possible that the median forecasts would have been more accurate under the zero interest rate regime, but we have not investigated this due to data limitation. However, we encourage other researchers to look into median forecasts.

Forecast behavior in the period after zero interest rate policy

In the previous section we observed that forecasts overestimated 3-month CDs in most of the period after the zero interest rate policy ended. The negative MPEs in table 22 support this observation, and the bias tests find this overestimation to be statistically significant. Another observation is that forecast errors were higher in this period compared to errors in the zero interest rate policy regime. This finding is reflected in relatively high MAEs, RMSEs and MNSEs.

The BOJ only increased the official interest rate once more after the increase in July 2006[38]. The stagnation of the inflation increase and following decrease until the second half of 2007 probably dissuaded the BOJ from raising the interest rate further as they may have been afraid of a new period of deflation. Also, the fall in inflation probably proves that the BOJ acted too early when they ended their zero interest rate policy; and that OECD's warnings (stated in section 1) were accurate.

The overestimation of interest rates from 2007 by the private sector may have its origins in an unexpected monetary policy in this period. The BOJ began to decrease the interest rates on October 31st, 2008 (Bank of Japan 2008). The low interest rates surprised the private sector that may have expected higher interest rates when the inflation increased rapidly in the year before October 2008. This expectation seems sensible considering the BOJ's statement when the zero interest rate policy was introduced (section 1), and the BOJ's action in July 2006.

The forecast errors are surprisingly high after October 2008 considering that Japan then practically returned to a deflationary environment and a zero interest rate policy. However, two factors may explain the private sector's high interest rates forecasts in this period: (1) lower credibility in the BOJ's monetary policies and (2) higher risk premiums.

First, a possible loss of confidence in the BOJ from the private sector may have raised interest rate forecasts. The higher errors may indicate that the private sector regards the BOJ's policy as unstable because interest rates seem to increase as soon as there is inflation in the Japanese economy. Perhaps forecasters believed that inflation could replace deflation within a few months considering the high volatility of inflation at this

[38] The official interest rate was increased to 0.75 percent on February 21st, 2007.

time, and they therefore increased their interest rate forecasts. The possible loss of confidence in the BOJ keeping interest rates low in this period is contradictory to what we find in the zero interest rate policy period.

Second, one may claim that interest rates forecasts are affected by private sector's expectations of future interest rate premiums. Deposits are insured in Japan (Nakaso 2001) and 3-month CDs are therefore practically risk-free. However, Shoven et al. (1991) state that certificates of deposit are not completely risk-free because of two types of risks. First, the authorities may not fulfill their promise of insuring deposits. Second, a bank failure may be costly for depositors even though the authorities honor the insurance policy because the depositor may face reinvestment risk when payments are delayed. The private sector may have expected higher risk premiums for 3-month CDs when economic uncertainty increased in the subprime crisis, and this may partly explain the high forecasts of interest rate in this period. However, the low actual values for 3-month CDs indicate that the risk premium did not increase as 3-month CDs were virtually identical to interest rates set by the central bank.

Above, we argued that the following three factors may explain the overestimation of interest rates after the zero interest rate policy: (1) unexpected monetary policy, (2) lower credibility in the BOJ's monetary policies, and (3) higher risk premiums. It is impossible to state how much of the overestimation is due to each factor without further analysis, and we encourage future research in this area. However, it seems likely that lower credibility in the BOJ's monetary policy is part of the reason why interest rates forecasts were so high in this period. This loss in credibility may weaken the BOJ's capability to extract information from markets (Aoki and Kimura 2007), which again makes it more challenging to conduct an optimal monetary policy.

Regime changes as an explanation of forecast inaccuracy and bias

When studying interest rate expectations, we have found inaccuracy and consistent bias over the full sample and in sub-periods. Does this make interest rate expectations irrational? Most papers test their entire data sample without looking at sub-periods. For inflation and interest rate expectations, this implicitly assumes the entire period to be one monetary regime with a stable relationship between output, inflation and interest rates. However, both the economy and monetary policy evolve, and regime switches may occur. Moreover, such changes can affect the behavior of expectations as well, which, as we have seen in this part of the paper, has been the case for interest rate expectations in Japan.

Furthermore, unforeseeable regime changes can result in forecasts with systematic errors in certain periods, even when the agents are fully rational (Thomas 1999). Suppose an agent rationally assigns a positive probability of the adoption by the central bank of a zero-inflation target. If the central bank fails to implement such a target, the agent's inflation forecast can appear biased. Nevertheless, the forecast may have been rational: the forecaster found it probable that the central bank would be able to implement the target, and even though the result proved otherwise ex post, the estimate might have been correct ex ante. It is not unlikely that this is true for interest rate expectations in Japan.

Regime shifts could also produce rejections of rationality in our tests, which assume a stationary environment (Evans and Wachtel 1993). We have analyzed three distinctive periods in Japan's monetary policy. It is unlikely that our data is stationary considering the change from a wait-and-see policy to a zero interest rate policy. The lack of stationary data can explain the rejections of rationality in the full sample. In addition, an announcement of a zero-policy is merely a signal from the central bank. Building the trust and confidence needed for expectations to follow takes time, and during this period, forecast errors may be consistent, but still rational.

During the last crisis, there were discussions on how inflation expectations were affected by the crisis, and in particular on whether the anchoring properties of long-run inflation expectations changed as the crisis unfolded (Svensson 2009). Galati et al. (2010) tested inflation expectations in the US, the Euro area and in the UK and found evidence of long-run inflation expectations becoming less firmly anchored during the financial crisis. It would be interesting to do similar tests for the crises in Japan, but we only have survey data. The horizon and low frequency of the survey data makes it difficult to formally test changes in the anchoring properties of long-term expectations (Ibid.). Nevertheless, this is a field for future research.

Reflections from forecasts in periods of different monetary policy regimes

We identified in the previous section that interest rates were held lower than the private sector expected from mid-2008 and onwards, in accordance with the economy heading into a new period of stronger deflation than expected. This is also consistent with the strong measures taken by central banks around the globe in the subprime crisis. However, the price development was pointing in the right direction as the recovery gained a foothold, and we have yet to see how long the central bank will keep interest rates low. If history has taught them anything, and IMF's, OECD's and Bernanke's criticisms have been heard, perhaps the Bank of Japan will keep the interest rates low until deflation is overcome and the Japanese economy has returned to a sustainable growth path with price stability. The arrival of such a positive economic environment was further set back by the devastating earthquake and tsunami that struck Japan on March 11th (ABC News 2011). The severe situation at the Fukushima Daiichi nuclear power plant is being stabilized, but the mishandling of the nuclear crisis has damaged confidence in politics and authorities. Whether the effect on the economy will be long lasting is yet to see. As we have seen, forecasting is no simple exercise, and though tests have shown signs of irrationality, it seems unreasonable to expect that forecasters can predict everything. To illustrate, who saw the earthquake coming?

Conclusion

We conduct a comprehensive study of private sector forecasts in Japan from 1989 to 2010. Japan's economy has had a distinctive development over the two last decades compared to other countries. This paper has thoroughly analyzed the private sector's expert forecasts for real GDP growth, inflation, unemployment rate and interest rate in Japan in this period. Results for the full sample have mainly been in accordance with earlier research on expectations. First, forecasters are more capable of making accurate forecasts for some macroeconomic indicators than others, and volatility and information flow in actual values can partly explain why. Second, private sector forecasts significantly outperform theoretical forecast models. Finally, forecasts are rational for real GDP growth, inflation and unemployment rate when looking at the full sample.

Most research does not analyze forecasts in sub-periods. This paper divides the full sample into sub-periods of contractions and expansions, and this separation provides contradictory results to findings for the full sample. Forecasts are not rational throughout the business cycle as macroeconomic variables are systematically over- and underestimated in contractions and expansions. Higher forecast accuracy in expansions indicates that shocks leading to contractions are more unexpected by experts in both timing and magnitude compared to turnarounds in expansions. Especially, the downturn in the subprime crisis is an example of this finding. Moreover, disagreement varies throughout the business cycle; disagreement between forecasters is usually low when there is high economic growth, low inflation and low unemployment, or expectations of this.

Periods of different monetary policy regimes have been used to analyze interest rate forecasts. The BOJ's monetary policy from 1991 to 1994 has received criticism, but the private sector's overestimation of interest rates suggests that the central bank did more than expected. The private sector seemed to believe the zero interest rate policy from 1999 to 2006 despite the temporary end of this policy in 2000 to 2001, but this distortion appears to have lowered their faith in the BOJ's ability to end deflation. The private sector's forecast accuracy of interest rate has been poor since the zero interest rate policy unexpectedly ended in July 2006, indicating a loss of confidence in the BOJ's monetary policy.

While consensus forecasts contain weaknesses, the findings in this paper and the literature indicate that these forecasts are the best in predicting macroeconomic variables. In addition, they are greatly listened to, as best exemplified by following daily news. Perhaps we can conclude that economic forecasters are like cross-eyed javelin throwers: They don't win many accuracy contests, but they keep the crowd's attention.

References

ABC News (2011): "*Key facts: Earthquakes in Japan*", ABC News online, last updated March 12 2011.

Akhtar, M. A., Cornelis A. Los, and Robert B. Stoddard, (1983): *"Surveys of Inflation Expectations: Forward or Backward Looking?"* Federal Reserve Bank of New York, Quarterly Review, 1983 volume 8.

Andolfatto, David, Scott Hendry and Kevin Moran (2005): *"Are Inflation Expectations Rational?".* Journal of Monetary Economics, Volume 55, Issue 2, March 2008, Pages 406-422.

Ang, Andrew, Geert Bekaert and Min Wei (2007): *"Do Macro Variables, Asset Markets or Surveys Forecast Inflation Better?"*, Journal of Monetary Economics, 54, 1163 – 1212.

Aoki, Kosuke and Takeshi Kimura (2007): "*Uncertainty about Perceived Inflation Target and Monetary Policy"*, Deutsche Bundesbank, Discussion Paper Series 1: Economic Studies, No. 10/2007.

Bank of Japan (2008): "*On Monetary Policy Decisions*", October 31 2008. Retrieved from www.boj.or.jp/en/ on May 29 2011.

Batchelor, R. A. (1990): *"All Forecasters are Equal"*, Journal of Business Economics and Statistics, 8, 143-144. In Batchelor (2000).

Batchelor, R. and P. Dua (1995): *"Forecaster Diversity and the Benefits of Combining Forecasts"*, Management Science, Vol. 41, No. 1, 68-75.

Batchelor, Ray (2000): "*The IMF and OECD versus Consensus Forecasts"*, City University Business School, London (August 2000).

Bates, J. M. and C. W. J. Granger (1969): *"The Combination of Forecats",* Operational Research Quarterly, Vol. 20, No. 4 (Dec 1969), 451 – 468.

Berkmen, S. Pelin (2010): "*Selected Issues: The Impact of Fiscal Consolidation and Structural Reforms on Growth in Japan*", IMF Country Report No. 10/212.

Bernanke, Ben S. (1999): ”*Japanese Monetary Policy: A Case of Self-Induced Paralysis?*”, Princeton University.

Blix, M., J. Wadefjord, U. Wienecke and M. Ådahl (March 2001): “*How good is the forecasting performance of major institutions?”*, Monetary Policy Departement, Economic Review 3/2001.

Brown, Bryan W. and Schlomo Maital (1981): *“What do Economists Know? An Empirical Study of Experts's Expectations”,* Econometrica, Vol. 49, No. 2 (Mar 1981), 491 – 504.

Bullard, James and Kaushik Mitra (2001): “*Learning About Monetary Policy Rules*”, Federal Reserve Bank of St. Louis, Working Paper 2000-001E, January 2002.

Caballero, R. J., T. Hoshi and Kashyap A. K. (2006): “*Zombie Lending and Depressed Restructuring in* Japan”, NBER Working Paper 12129.

Capistrán, Carlos and Allan Timmermann (2006): *“Disagreement and Biases in Inflation Expectations”*, Working Paper Banko de México, June 2006.

Clemen, Robert T. (1989): *“Combining Forecasts: A Review and Annotated Bibliofigurey”* International Journal of Forecasting, 5, 559-583.

Consensus Economics (2011), webpage www.consensuseconomics.com (20 March 2011)

Croushore, Dean (2010): *“An Evaluation of Inflation Forecasts from Surveys Using Real-Time Data”*, The B.E. Journal of Macroeconomics, Vol 10, No. 1.

Croushore, Dean and Laurance Ball (1995): *“Expectations and the Effects of Monetary Policy”*, NBER Working Paper 5344.

Cukierman, Alex (1986): “*Measuring Inflationary Expectations*”, Monetary Economics, 17, 315 – 324.

Cukierman, Alex and Allan H. Meltzer (1982): *“What do Tests of Market Efficiency Show in the Presence of Permanent-Transitory-Confusion”*. In Cukierman (1986)

DeLong, J. Bradford (1995): *“America's Peacetime Inflation: The 1970s”,* University of Berkely and NBER.

Diebold, F. X. and R. S. Mariano (1995) *"Comparing Predictive Accuracy"*, Journal of Business and Economic Statistics, 13, 253-263.

Dotsey Michael and Jed L. DeVaro (1995): *"Was the Disinflation of the Early 1980s Anticipated?"*, Federal Reserve Bank of Richmond Economic Quarterly, Vol. 81, No. 4: 41—59.

Economic and Social Research Institute (2010): "*The provisional determination of Business Cycle Peak and Trough*", The Cabinet Office.

Ehrbeck, Tilman and Robert Waldmann (1996): *"Why are Professional Forecasters Biased? Agency Versus Behavioral Explanations"*, Quarterly Journal of Economics, 111, p. 21-40.

Evans, Martin and Paul Wachtel (1993): *"Inflation Regimes and the Sources of Inflation Uncertainty"*, Journal of Money, Credit and Banking, Vol. 25 (Aug 1993), 475 – 511.

Fujii, Mariko and Kawai, Masahiro (2010): "*Lessons from Japan's Banking Crisis, 1991–2005*", Asian Development Bank Institute, ADBI Working Paper 222.

Galati, Gabriele, Steven Poelhekke, and Chen Zhou (2010): *"Did the Crisis Affect Inflation Expectations"*, Netherlands Central Bank, International Journal of Central Banking, Vol. 7, No. 1.

Garrett, Thomas A. and John C. Leatherman (2000): *"An Introduction to State and Local Public Finance",* Scott Loveridge ed., The Web-Book of Regional Science, The Regional Research Institute, West Virginia University.

Gerberding, Christina (2006): "*Household versus expert forecasts of inflation: New Evidence from European survey data*", Deutsche Bundesbank.

Gaspar, Vitor, Frank Smets and Dacid Vestin (2006): "*Optimal Monetary Policy under Adaptive Learning*", available at http://repec.org/sce2006/up.17883.1140458209.pdf

Hansen, Lars P. and Robert J. Hodrick (1980): *"Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis"*, The Journal of Political Economy, Vol. 88, No. 5 (Oct 1980), 829 – 853

Hayakawa, Hideo and Eiji Maeda, (2000): "*Understanding Japan's Financial and Economic Developments Since Autumn 1997*", Research and Statistics Department, Bank of Japan, Working paper 00-1.

Hirose, Yasuo and Shinsuke Ohyama (2009): "*Identifying the Effect of Bank of Japan's Liquidity Facilities: The Case of CP Operations During Financial Turmoil*", Bank of Japan Working Paper Series, No. 09-E-7.

Holden, K. and D. A. Peel (1990): *"On Testing for Unbiasedness and Efficiency of Forecasts"*, The Manchester School Vol. LVIII No. 2 (Jun 1990), 120 – 127.

Hoshi, Takeo and Anil K. Kashyap (2004): "*Japan's Financial Crisis and Economic Stagnation*", Journal of Economic Perspectives, Vol. 18, No. 1, Winter 2004, Pages 3-26.

Hoshi, Takeo and Anil K Kashyap (2008): "*Will the U.S. Bank Recapitalization Succeed? Eight Lessons from Japan",* NBER Working Paper No. 14401.

IMF (2000): "*Japan: Staff Report for the 2000 Article IV Consultation*", IMF Staff Country Report No. 00/98, August 2000.

Keane, Michael P. and David E. Runkle (1990): *"Testing the Rationality of Price Forecasts: New Evidence from Panel Data"*, American Economic Review, Vol. 80, No. 4 (Sep 1990), 714 – 735.

Kershoff, G.J. and B.W. Smit (2002): *"Conducting Inflation Expectation Surveys in South Africa"*, South African Journal of Economics, Vol. 70, No. 3 (Mar 2002), 205 – 212.

Lamont, Owen (1995): *"Macroeconomic Forecasts and Microeconomic Forecasters"*, NBER Working Paper No. 5284 (Oct 1995).

Lipsky, John (2009): "*Building a Post-Crisis Global Economy—An Address to the Japan Society"*, IMF, speeches 2009 (December 10).

Llambros, Louis and Victor Zarnowitz (1987): *"Consensus and Uncertainty in Economic Prediction"*, Jounal of Political Economy, Vol. 95, No. 3, 591 – 662.

Lovell, Michael C. (1986) "*Tests of the Rational Expectations Hypothesis*", The American Economic Review, Vol. 76, No. 1 (March 1986) page 110 – 124.

Mankiw, Gregory N. and Matthew D. Shapiro (1986): *"Do we Reject too Often? Small Sample Properties of Tests of Rational Expectations Models",* Economics Letters 20 (1986), North Holland.

Mankiw, G.N., R. Reis and J. Wolfers (2003): "*Disagreement about Inflation Expectations*". NBER Working Paper 9796.

Marcet, Albert and Juan Pablo Nicolini (2005): *"Money and Prices in Models of Bounded Rationality in High Inflation Economies"*, European Central Bank Working Paper No. 469 (April 2005).

McNees, S. K. (1987): *"Consensus Forecasts: Tyranny of the Majority?"*, New England Economic Review, Nov/Dec, 15 – 21. In Batchelor (2000).

Mehra, Yash P. (2002): *"Survey Measures of Expected Inflation: Revisiting the Issues of Predictive Content and Rationality"*, Federal Reserve Bank of Richmond Economic Quarterly Volume 88/3 Summer 2002.

Mestre, Ricardo (February 2007): "*Are survey-based inflation expectations in the Euro area* informative?", European Central Bank Working Paper Series no. 721.

Mincer, Jacob A. and Zarnowitz, Victor (1969): *"The Evaluation of Economic Forecasts"*, NBER.

Molnár, Krisztina and Sergio Santoro (2006): "*Optimal Monetary Policy When Agents Are Learning"*, Institute of Economics, Hungarian Academy of Science, MT-DP-2006/1.

Muth, John F. (1961): *"Rational Expectations and the Theory of Price Movements"*, Econometrica, Vol. 29, No. 3 (Jul 1961), 315 – 335.

Nakaso, Hiroshi (2001): "*The financial crisis in Japan during the 1990s: how the Bank of Japan responded and the lessons learnt*", Bank for International Settlements (BIS), BIS Papers no.6.

Naszodi, Anna (2010): "*Testing the asset pricing model of exchange rates with survey data*", European Central Bank, Working Paper No. 1200, May 2010.

Newey, W. K. and K. D. West (1987): *"A Simple Positive Semi-Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix",* Econometrica, Vol. 55, No. 3, 703-8.

Noda, Yoshihiko. (2010): *"The Global Economy and the Japanese Economy: Combating Deflation"*, IMF, 2010 Annual Meetings, Press Release No. 13.

OECD (2006): "*Economic Survey of Japan, 2006*", OECD, Policy Brief, July 2006.

Oda, Nobuyuki and Kazuo Ueda (2005): "*The Effects of the Bank of Japan's Zero Interest Rate Commitment and Quantitative Monetary Easing on the Yield Curve: A Macro-Finance Approach*", Bank of Japan, Working Paper No. 05-E-06, April 2005.

Okina, K., M. Shirakawa and S. Shiratsuka (2001): "*The Asset Price Bubble and Monetary Policy: Japan's Experience in the Late 1980s and the Lessons*", IMES, Bank of Japan, Monetary and Economic Studies (Special Edition).

Pearce, Douglas K. (1979): *"Comparing Survey and Rational Measures of Expected Inflation: Forecast Performance and Interest Rate Effects"*, Journal of Money, Credit and Banking, Vol. 11, No. 4 (Nov 1979), 447 – 456.

PFP Wealth Management (2009): *"Chronicle of a death foretold",* Novermber 30 2009, Retrieved from thepriceofeverything.typepad.com at May 16 2011.

Quintanilla, Rodrigo and Scott Sprinzen (2008): "*Broader Lessons From Lehman Brothers' Bankruptcy*", Standard & Poor's, Ratings Direct, September 17 2008.

Roberts, John M. (1998): *"Inflation Expectations and the Transmission of Monetary Policy"*, Board of Governors of the Federal Reserve System, Paper No. 98, Oct 1998.

Schrooten, Mechthild (2000): "*Japan: End of Zero Interest Rate Policy - Beginning of Long-term Upswing?*", Springer Link, Economic Bulletin, Volume 37 No. 9, September 2000.

Shinotsuka, Eiko (2000): "*Monetary Policy and Structural Reform of the Japanese Economy* ", Bank of Japan, speech at the Meeting on Economic and Financial Matters in Nagano on February 4, 2000.

Shiratsuka, Shigenori (2003): "*Asset Price Bubble in Japan in the 1980s: Lessons for Financial and Macroeconomic Stability*", The Institute for Monetary and Economic Studies (IMES), Bank of Japan, Discussion Paper No. 2003-E-15.

Sommer, Martin (2009): "*Why Has Japan Been Hit So Hard by the Global Recession?*", IMF, Staff Position Note, SPN/09/05.

Stock, James H. and Mark W. Watson (1999): "*Forcasting Inflation*", NBER, Working Paper 7023, March 1999.

Stone, Douglas and William T. Ziemba (1993): *"Land and Stock Prices in Japan"*, Journal of Economic Perspectives, Volume 7, Number 3, Pages 149 – 165.

Sturm, J-E and J. de Haan (2009): "*Is Central Bank Communication Really Informative when Forecasting Interest Rate Decisions? New Evidence Based on a Taylor Rule Model for the ECB*", CESifo Area Conference, Munich, 20-21 February.

Svensson, L. (2009): *"Monetary Policy with a Zero Interest Rate"*, Speech at SNS, Stockholm, February 17, 2009.

Syed, M., K. Kang and K. Tokuoka (2009): "*"Lost Decade" in Translation: What Japan's Crisis could Portend about Recovery from the Great Recession"*. IMF Working Paper, WP/09/282, Asia and Pacific Department.

The Bank of Tokyo-Mitsubishi (2006): "*BOJ ends Zero Interest Rate Policy*", The Bank of Tokyo-Mitsubishi UFJ,Ltd., Economic Review, No. 12 July 2006.

Theil, Henri (1966): *"Applied Economic Forecasting",* Amsterdam: North Holland, 1966. In Croushore (2010).

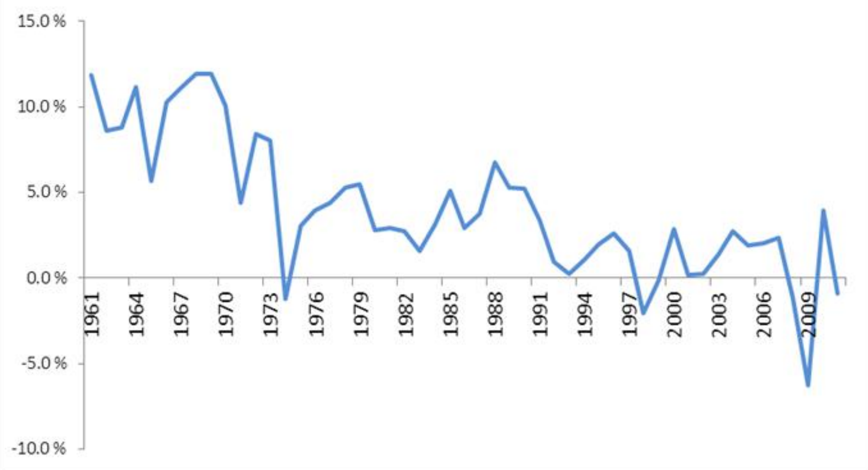
Thomas, Jr. Lloyd B. (1999): *"Survey Measures of Expected U.S. Inflation"*. Journal of Economic Perspectives; Volume 12, Number 4; Fall 1999; Pages 125-144.

Webb, Roy H. (1987): *"The Irrelevance of Tests for Bias in Series of Macroeconomic Forecasts"*, Federal Reserve Bank of Richmond.

Zarnowitz, Victor (1982): *"The Accuracy of Individual and Group Forecasts From Business Outlook Surveys"*, NBER Working Paper No. 1053, Dec 1982.

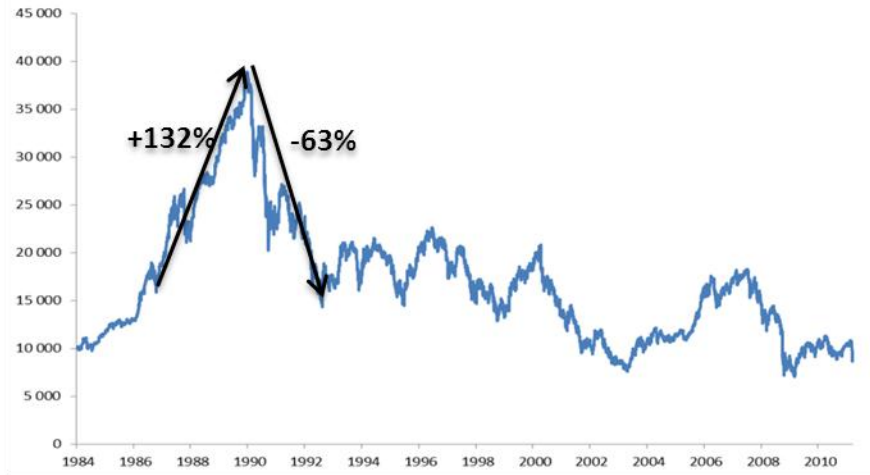
Appendix 1: Historic development in real GDP, Nikkei 225 and land prices

Figure 20: Real GDP growth for Japan (YOY%)



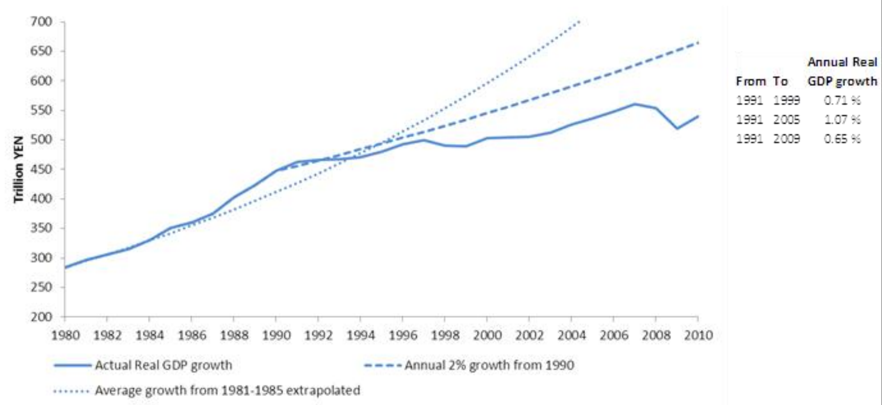
Source: Thomson Reuters Datastream

Figure 22: Nikkei 225



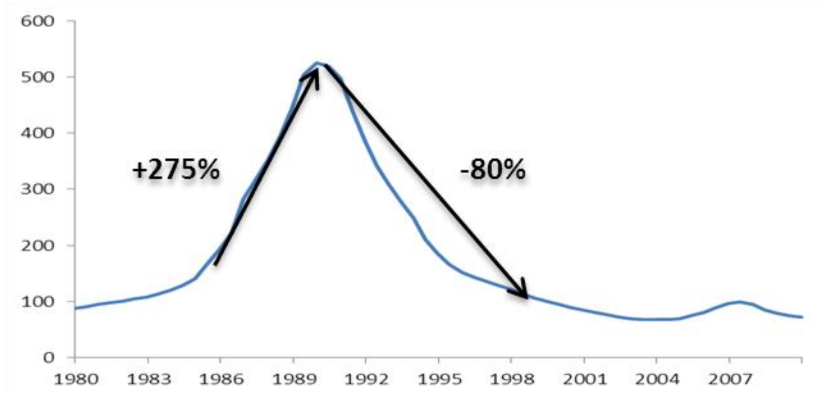
Source: Thomson Reuters Datastream

Figure 21: Real GDP for Japan - Actual values compared to two growth models



Source: Thomson Reuters Datastream

Figure 23: Land Price Index for the Commercial Area in Japan's 6 big cities (2000=100)



Source: Thomson Reuters Datastream

Appendix 2: Historic development of macroeconomic variables in business cycles

Section 1 ended with a description of common trends in macroeconomic variables in expansions and contractions. This appendix contains a more detailed discussion of the historic development of macroeconomic variables in different expansions and contractions for the five business cycles from figure 1. Figures 2 to 4 in section 1 and table 3 and 4 summaries are the main sources for the following discussion.

Table 3: Historic development of macroeconomic variables in sub-periods

Period	*Real GDP growth**	*Consumer prices growth**	*Unemploy-ment rate***	*3-month CDs***	*Private consumption growth**	*Industrial production growth**
Expansion before asset price bubble	5.5 %	2.0 %	-0.4 %	3.4 %	4.8 %	6.5 %
Asset price bubble	1.5 %	1.7 %	0.2 %	-5.6 %	1.5 %	-4.6 %
Expansion after asset price bubble	1.9 %	0.8 %	0.9 %	-1.8 %	2.2 %	3.8 %
Asian crisis	-1.2 %	-0.1 %	1.2 %	0.1 %	-0.6 %	-2.3 %
Expansion after Asian crisis	1.4 %	-0.4 %	-0.2 %	0.0 %	0.9 %	5.0 %
Dot-com bubble	0.0 %	-1.3 %	0.8 %	-0.5 %	1.5 %	-13.2 %
Expansion after dot-com bubble	1.9 %	0.1 %	-1.4 %	0.6 %	1.3 %	3.9 %
Subprime crisis	-2.7 %	0.0 %	1.4 %	-0.3 %	-1.4 %	-25.7 %
Expansion after subprime crisis	0.4 %	-0.7 %	-0.6 %	-0.2 %	0.7 %	14.0 %

**Annual growth*

***Change over a period = Actual value in the end of the period - Actual value in the beginning of the period*

Table 4: A closer look at the historic development of unemployment rate and 3-month CDs in sub-periods

Period	Unemployment rate			3-month CDs		
	Start of period	*End of period*	*Change*	*Start of period*	*End of period*	*Change*
Expansion before asset price bubble	2.6 %	2.2 %	-0.4 %	4.7 %	8.1 %	3.4 %
Asset price bubble	2.4 %	2.6 %	0.2 %	8.1 %	2.5 %	-5.6 %
Expansion after asset price bubble	2.6 %	3.5 %	0.9 %	2.3 %	0.6 %	-1.8 %
Asian crisis	3.3 %	4.5 %	1.2 %	0.6 %	0.7 %	0.1 %
Expansion after Asian crisis	4.7 %	4.5 %	-0.2 %	0.6 %	0.6 %	0.0 %
Dot-com bubble	4.4 %	5.2 %	0.8 %	0.6 %	0.1 %	-0.5 %
Expansion after dot-com bubble	5.4 %	4.0 %	-1.4 %	0.1 %	0.8 %	0.6 %
Subprime crisis	3.7 %	5.1 %	1.4 %	0.8 %	0.5 %	-0.3 %
Expansion after subprime crisis	5.2 %	4.6 %	-0.6 %	0.4 %	0.1 %	-0.2 %

Historic development in the 1st business cycle

The economic activity was high in Japan in the expansion period before the asset price bubble. This period had high real GDP growth (5.5 percent annually), high consumer price growth (2 percent annually), low unemployment rate (2.4 percent), high real private consumption growth (4.79 percent annually) and strong industrial production growth (6.53 percent annually) compared to other periods. A high increase in interest rate (3-month CDs increased from 4.7 to 8.1 percent from the beginning to the end of this expansion) reflects tightening in monetary policy in this period.

A slowdown in the Japanese economy followed the high growth period when the asset price bubble burst. The real GDP growth remained positive, but was only 1.5 percent in the asset price bubble. Consumer price growth (inflation) declined after its peak in 1991, and the unemployment rate started to increase at the end of the contraction. The interest rate and real private consumption growth dropped, and there was a negative industrial production growth (-4.6 percent).

Historic development in the 2nd business cycle

Real GDP growth improved to 1.90 percent in the expansion after the asset price bubble, but was far below the high growth in the previous expansion. The deflationary environment Japan entered in this period is recognized by a decline towards zero for both inflation and interest rates. Furthermore, the increase in unemployment rates also supports our view that Japan still had some underlying problems from the asset price bubble (section 1) even though real consumption growth and industrial production growth recovered to respectively 4.79 and 6.53 percent.

The problems in the Japanese financial system that were revealed in the Asian crisis led Japan into a recession as real GDP growth rapidly declined and turned negative. The deflationary environment continued with an interest rate and inflation close to zero. Also, the downturn led to higher unemployment rate and negative growth in both private consumption and industrial production.

Historic development in the 3rd business cycle

The real GDP growth, private consumption growth and industrial production growth recovered after the Asian crisis. This development proves that the stimulus policy from the Japanese authorities was effective. However, lack of improvement in unemployment rate and a stronger deflationary environment might indicate that stronger measures should have been taken.

The turmoil and re-emerged fear of a financial break-down in the dot-com bubble led to a severe fall in industrial production, which experienced a negative annual growth of 13.2 percent in this period. Also, the annual real GDP growth fell and was approximately zero over the period (0.02 percent). Unemployment continued its upward trend and deflation sustained. Real private consumption moved against other indicators as it improved (was 1.54 percent over the period).

Historic development in the 4th business cycle

Effective measures from Japanese authorities and favorable global conditions led to an annual real GDP growth of 1.9 percent in the expansion period before the subprime crisis. Also, the unemployment rate broke out of its upward trend that had lasted since the asset price bubble, and declined 1.4 percent. Real private consumption growth and industrial production growth were positive, respectively 1.29 percent and 3.91 percent. However, Japan did not break out of the deflationary environment as inflation and interest rate approximately zero (0.14 and 0.16 percent).

The global credit crunch and the following slowdown in the global economy caused large movements in Japanese macroeconomic variables in the subprime crisis. This contraction had lowest real GDP growth, real private consumption growth and industrial production growth of all expansions and contractions from 1986 to 2010. Also, the unemployment rate experienced the highest growth in this contraction. Deflation was still present.

Historic development in the 5th business cycle

Japan's rapid recovery after the subprime crisis came as global and domestic measures proved to be effective. Real GDP growth, real private consumption growth and industrial production all saw a rapid bounce-back in this period, and the unemployment rate came down 0.6 percent. The deflationary environment worsened in the beginning of the expansion as inflation declined. However, the inflation now has an upward trend and this movement might indicate an end to the long lasting deflation period in Japan.

Appendix 3: Fixed horizon forecasts - CE versus WAN12

Every quarter CE asks the panelists to do forecasts for the next 6-7 quarters. Hence, we get quarterly fixed horizon data that can be compared to our fixed horizon forecasts created by using the WAN12-formula. These fixed horizon forecasts from CE are only provided for real GDP growth, personal consumption, industrial production, consumer prices and 3-month CDs. We will compare CE fixed horizon forecasts for real GDP growth and consumer prices with fixed horizon forecasts from WAN12 to check how accurate our method is. CE fixed horizon forecasts go from the first quarter in 2003 to the second quarter in 2012. However, we will only use fixed horizon forecasts up to the last quarter in 2010.

Fixed horizon forecasts from WAN12 seem to be good approximations for fixed horizon forecasts from CE by the look of figures 24 and 25. Fixed horizon forecasts from WAN12 seem follow fixed horizon forecasts from CE closely in the relevant time period. Figures 26 and 27 show the absolute errors (AEs) for the forecasts errors (FEs) for fixed horizon forecasts for WAN12 and CE. These figures give a more disturbing picture of the quality of fixed horizon forecasts from WAN12 as a few large AEs are present. However, these errors appear to be averaged out over time according to table 5, which shows the forecast accuracy of fixed horizon forecasts from WAN12 and CE. The forecast accuracy is actually slightly better for WAN12's forecasts as the MAEs, RMSEs and MPEs are closer to zero, with the exception being the MPE for consumer price growth.

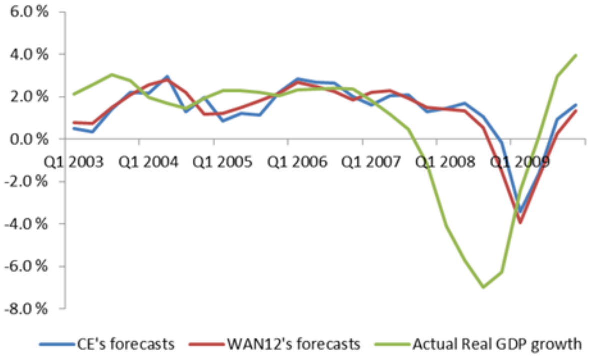
A limitation of the results is that this paper only compares WAN12 with CE fixed horizon forecasts in the latter end of the time period, and not before 2003. This is due to data availability of CE fixed horizon forecasts before 2003.

Table 5: Comparison of WAN12 forecasts and CE's fixed horizon forecasts

	MAE	*RMSE*	*MPE*	*Correlation*
Real GDP Growth				
- WAN12	1.83 %	2.69 %	-0.57 %	0.95
- Consensus Economics	1.86 %	2.85 %	-0.62 %	
Real Consumer Prices Growth				
- WAN12	0.46 %	0.64 %	-0.09 %	0.97
- Consensus Economics	0.48 %	0.70 %	-0.05 %	

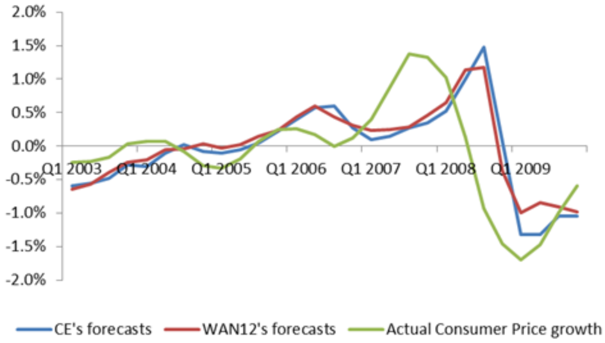
Figure 24 to 27: Forecasts from WAN12 versus CE’s fixed horizon forecasts

Figure 24: Comparison of WAN12 forecasts and CE's fixed horizon forecasts: Real GDP growth



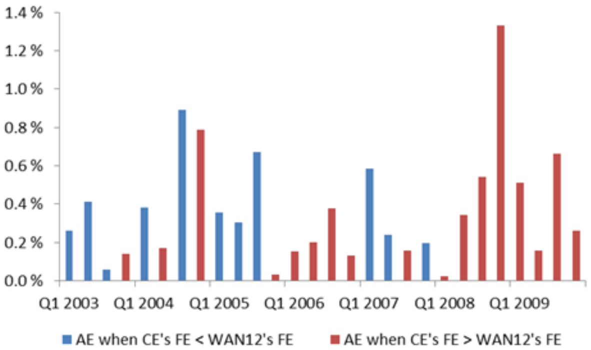
Source: Consensus Economics and Thomson Reuters Datastream

Figure 25: Comparison of WAN12 forecasts and CE's fixed horizon forecasts: Consumer price growth



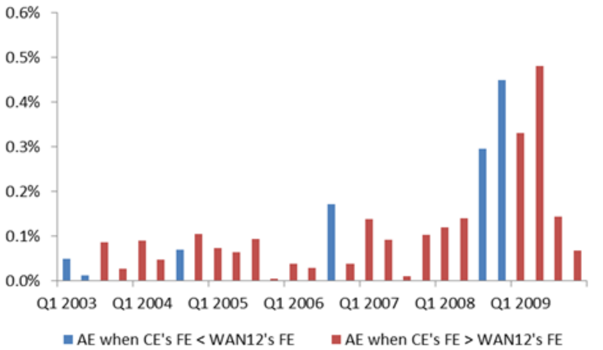
Source: Consensus Economics and Thomson Reuters Datastream

Figure 26: Absolute errors for WAN12 and CE's FEs for fixed horizon forecasts: Real GDP growth



Source: Consensus Economics and Thomson Reuters Datastream

Figure 27: Absolute errors for WAN12 and CE's FEs for fixed horizon forecasts: Consumer price growth



Source: Consensus Economics and Thomson Reuters Datastream

Appendix 4: Construction of actual values for the next 12 months

Section 4 transformed actual raw data for macroeconomic variables into comparable data, but there were three technicalities that this section left out. This appendix constructs actual real GDP, consumer price growth and unemployment rate for the next 12 months

Real GDP

Actual real GDP is given each quarter, and this paper uses the following weighted average to produce actual real GDP for the next 12 months:

$$Real\ GDP_{n12m} = Quarterly\ real\ GDP_1 * \frac{Remaining\ months\ of\ first\ quarter}{3}$$

$$+ \sum_{i=2}^{4} (Quaterly\ real\ GDP_i) + Quarterly\ real\ GDP_5 * \frac{3 - Remaining\ months\ of\ first\ quarter}{3}$$

$Real\ GDP_{n12m}$ is actual real GDP for the next 12 months, $Quarterly\ real\ GDP_i$ is quarterly real GDP in quarter i, i=1 is the current quarter and i=2 is the next quarter. This formula is best explained with an example: Real GDP is JPY 105,183 billion in the first quarter of 1990, JPY 106,766 billion in the second quarter of 1990, JPY 113,679 billion in the third quarter of 1990, JPY 121,742 billion in fourth quarter of 1990 and JPY 111,417 billion in first quarter of 1991. These numbers are used when we want to calculate actual real GDP for the next 12 months at February 1990. The remaining months of the first quarter in 1990 equals 2 months, and real GDP for the next 12 months in February 1990 is therefore JPY 449,448 billion[39].

Consumer price growth

Actual consumer price growth for the next 12 months is constructed by a moving average:

$$CPI_{n12m} = \frac{\sum_{t=0}^{12} CPI_{n12m,t-12}}{12}$$

CPI_{n12m} is actual consumer price growth for the next 12 months, t is the month at time t and $CPI_{n12m,t-12}$ is actual consumer price growth for the next 12 months for t-12 months ago. For example, if t equals December 2000 will this formula calculate the

[39] 449,448 = 105,183 * (2/3) + 106,766 + 113,679 + 121,742 + 111,417 * (1/3)

actual consumer price growth from December 2000 to December 2001. This is done by averaging the actual consumer price growths from January 2000 to January 2001, from February 2000 to February 2001, … , and from December 2000 to December 2001[40].

Unemployment rate

A simple weighted average formula is used to compute actual unemployment rate for the next 12 months:

$$Unemployment\ Rate_{n12m} = \frac{\sum_{t=1}^{12} Unemployment\ Rate_t}{12}$$

$Unemployment\ Rate_{n12m}$ is the unemployment rate for the next 12 months, $Unemployment\ Rate_t$ is the unemployment rate at time t, *t=1* is the current month and t=2 is the next month.

[40] This will smooth inflation substantially compared to monthly measures, but since forecasts originally are done for yearly inflation, we feel that using raw monthly figures would induce too much noise. However, we should be aware that this construction may make forecast accuracy better and more consistent.

Appendix 5: Description of figures in section 6

First figure for each variable

The dark blue shaded area in this graph is the difference between the mean forecasts and low forecasts and the light blue area is the difference between high forecasts and mean forecasts. The green line represents mean forecasts and the black thick line is actual values.

Second figure for each variable

The dark red shaded areas represent the forecast errors when the forecast means are below the actuals, signifying an underestimation of actual values made by forecasters. The light red shaded areas represent the forecast errors when the forecast means are above the actuals, signifying an overestimation of actual values made by forecasters. The black line is the actual values for the next 12 months.

Third figure for each variable

The figures have the same color definitions the second figure for each variable. However, the actual values are now the x-line and the heights of shaded areas represent absolute forecast errors (AFEs) between mean forecasts and actual values for the next 12 months.

Appendix 6: The theoretical forecast models versus CE forecasts

This appendix goes more into detail of the analysis of comparing CE's and theoretical forecast models' forecasts. First, we outline the theoretical forecast models, and then we move over to explaining the methodology behind the comparison of forecasts from these models with CE forecast. Finally, we compare the forecast accuracy of theoretical forecast models with CE forecasts.

The theoretical forecast models

The theoretical forecasts models are random walks (RW), moving averages (MA), a weighted moving average model (WMA), an exponential moving average model (EMA), an exponential smoothing average model (ESA) and ARIMA models.

Random Walks (RWs)

A random walk model is a backward looking model as it only uses historic values to predict the future. The assumption in these models is that it is impossible to say whether a variable will increase or decrease in the future, and that the best prediction is historic values. A RW model has the following annotation:

$$F_t = A_{t-1}$$

A_{t-1} is the actual value at time t-1 and F_t is the forecast value at time t. For example, "The random walk approach – Last 12 months" is a random walk model that forecast the real GDP growth for the next 12 months by looking at the real GDP growth for the last 12 months. "The random walk approach – Last 24 months" forecasts the real GDP growth for the next 12 months by taking the average of the actual real GDP growth for the last 12 months and the actual real GDP growth for the last 12 months for 12 months ago.

This approach is also used for consumer prices and unemployment rate. However, the approach is modified for the 3-month CDs as the current and last month's actual 3-month CDs rates are used as the best forecasts for the 3-month CDs in 3 and 12 months. These approaches are present in "RW – Current" and "RW – Last month".

Moving averages (MAs)

Moving averages are also backward looking models that use historic values to predict the future. The assumption behind these models is that the best forecasts are those that smooth out historic values. The logic behind smoothing historic values is that one historic value may be influenced by a shock that makes the value unusually high or low, and to adjust for seasonality. However, when several actual values are smoothed against each other will shocks neutralize each other and the computed number will be less influenced by certain shocks.

Forecasts from MA models depend on the average of historic values in N periods. Older data points are dropped off as new ones are added. The moving average model:

$$F_t = \frac{\sum_{i=1}^{N} A_{t-i}}{N}$$

The annotation for F_t and A_{t-i} is as before, and N is the number of time periods averaged.

This paper uses 11 different moving averages. For example, the "MA – Last 2m" for real GDP growth takes the average of actual real GDP growth for the last 12 months in the current and the last month to forecast real GDP growth for the next 12 months, "MA – Last 3m" takes the average of the last 3 months real GDP growth for the last 12 months to forecast real GDP growth for the next 12 months, and so on.

The same approach is used for consumer prices and unemployment rate. However, the "MA – Last 2m" for 3-month CDs takes the average of the current actual 3-month CDs and last month's actual 3-month CDs to forecast 3-month CDs in 3 and 12 months, and so on.

Weighted moving average model (WMA)

The weighted moving average model is a backward looking model as well, but the model is relatively sophisticated compared to the previous models. The model is a regression model that put different weights on certain variables that are used to produce forecasts, and can be outlined like this for real GDP growth:

Forecast for real GDP growth the next 12 months = w_1 * real GDP growth last 12 months_t + w_2 * real GDP growth last 12 months_{t-1} + … + w_{12} * real GDP growth last 12 months_{t-11} + w_{13} * Change in real GDP growth last 12 months_t + w_{14} * Change in real GDP growth last 12 months_{t-1} + … + w_{24} * Change in real GDP growth last 12 months_{t-11}

t applies to the current month in the model, t-1 refer to the last month, and so on. w_i represent the weights on the different variables in the model. The model is identic for consumer prices and unemployment rate, but the model is not used for 3-month CDs.

Solver in Excel is used to optimize the forecast accuracy of the model by adjusting the weights. However, the weights that optimize the forecast accuracy do not make economic sense (table 6). One might expect that recent observations would have higher weights than later observations, but this is not the case by the look of the optimized models. Another weakness with these models is that the model produces forecasts after all actual values are known and it is therefore appropriate to say that the model is "post-based".

Table 6: Weights in WMA model

# months ago	Real GDP growth		Consumer price growth		Unemployment rate	
	Actual value last 12m	*Change in actual value last 12m*	*Actual value last 12m*	*Change in actual value last 12m*	*Actual value last 12m*	*Change in actual value last 12m*
1 month	0.09	-0.64	-0.04	2.75	-0.02	2.75
2 months	0.60	-0.49	0.08	-0.87	0.13	-0.87
3 months	0.06	0.00	-0.01	-0.75	-0.01	-0.75
4 months	-0.06	0.37	-0.15	-0.96	-0.11	-0.96
5 months	0.01	-0.04	0.03	0.42	0.03	0.42
6 months	-0.21	0.07	0.55	-0.58	-0.03	-0.58
7 months	-0.25	0.09	-0.32	-0.39	-0.64	-0.39
8 months	-0.06	-0.07	0.13	0.41	0.07	0.41
9 months	0.01	0.01	0.01	-0.14	0.01	-0.14
10 months	-1.09	0.90	-0.50	-0.55	-1.04	-0.55
11 months	-0.73	-0.09	-2.71	-1.10	-1.37	-1.10
12 months	2.05	-0.73	3.65	2.84	4.00	2.84

Exponential moving average model (EMA)

The exponential moving average model uses both actual values and historic forecasts to make new forecasts. The model is outlined like this for real GDP growth:

Forecast for real GDP growth the next 12 months = w_1 * real GDP growth last 12 months$_t$ + $(1 - w_1)$ * Forecast for real GDP growth the next 12 months$_{t-1}$

t applies to the current month in the model, t-1 refer to the last month, w_i represent the weight on the real GDP growth for the last 12 months and $(1 - w_i)$ represent the weight on the forecast for real GDP growth for the next 12 months a month ago. The model is identic for consumer prices and unemployment rate. However, it is slightly different for 3-month CDs as the model then uses the current actual 3-month CD instead of looking at the actual 3-month CD for the last 12 months.

Solver in Excel is used to optimize the forecast accuracy of the model by adjusting the weights, and this means that this model is also "post-based".

Exponential smoothing average model (ESA)

The exponential smoothing average model uses both actual values and historic forecast errors to produce new forecasts. The model for real GDP growth:

Forecast for real GDP growth the next 12 months = Real GDP growth last 12 months + w_1 * Forecast error for real GDP growth

w_1 represent the weight on forecast errors. The forecast error is the difference between the forecast for real GDP growth for the next 12 months made 12 months ago and the actual real GDP growth for the last 12 months. The model is identic for consumer prices and unemployment rate. However, it is slightly different for 3-month CDs as the model then uses the current actual 3-month CD instead of looking at the actual 3-month CD for the last 12 months.

ARIMA

Autoregressive Integrated Moving Average (ARIMA) is more a complicated model. Autoregressive and moving average are two of the components of the model, while the process of translating the calculations into a metric that can be interpreted is referred to as integrated. Three assumptions need to be met to achieve best results of using the ARIMA approach: (1) the time series contain at least 50 data points, (2) the data series is stationary, e.g. the data series varies around a constant mean and variance, and (3) the series has a constant variance (homoscedastic). (Garrett and Leatherman 2000)

This paper uses Minitab to create forecasts for ARIMA (2,0,0), ARIMA (1,1,0), ARIMA (1,0,0) and ARIMA (0,1,1).

Methodology for comparing CE forecasts with theoretical forecasts models

Section 7 compares CE forecasts with forecasts from theoretical forecasts models by looking at MAE, RMSE, MPE and MNSE. In addition, Theil's U-statistics and Diebold-

Mariano test are included in this part of the analysis, and the explanation of these two measures is outlined below.

Theil's U-statistics

Theil's U-statistics is defined as:

$$Theil's\ U = \frac{RMSE\ of\ CE}{RMSE\ of\ naïve\ method}$$

Theil's U-statistics is a simple measure to compare the forecast accuracy of two different forecast methods. Theil's U-statistic compares the RMSEs of two forecast series. RMSE for CE forecasts is used as the benchmark because the objective is to find out whether market experts produce better forecasts than theoretical models. The forecast accuracy is better for expert forecasts than forecasts from theoretical forecast models if Theil's U-statistic is below 1, and the opposite is true if Theil's U-statistics is above 1. When Theil's U-statistic equals 1 is the forecast accuracy equivalent for the two methods in comparison.

Diebold-Mariano test

We test whether the forecasting performance of theoretical models is statistically different from CE forecasts by using a Diebold-Marino (DM) test. The null hypothesis to be tested is that the square realized forecast errors are the same:

$$H_0: E[(A_t - F_t^{CE})^2 - (A_t - F_t^{TM})^2] = 0$$

Where F_t^{CE} is CE forecasts at time t and F_t^{TM} is the forecast of the theoretical model in comparison at time t. The Diebold-Mariano statistic is

$$\frac{\sqrt{N}\bar{g}}{\sqrt{\hat{V}}} \sim N(0,1)$$

where $g_t = (A_t - F_t^{CE})^2 - (A_t - F_t^{TM})^2$ is the difference between the forecast errors at time t, and $\bar{g}$ is the time average of these differences. Finally, $\hat{V}$ is the estimated long-run variance of g_t. CE forecasts are significantly better than forecasts from theoretical models if the DM statistic is above 1.69 on a 5 percent critical level, and theoretical models are significantly better than CE forecasts if DM statistic is below -1.69. This test is in the spirit of Naszodi (2010)[41].

Results from comparing forecasts from CE with theoretical models

The following outlines our results from comparing forecasts accuracy of CE with theoretical forecast models for real GDP growth, consumer price growth, unemployment rate and 3-month CDs. Table 7 and 8 outlines MAEs, RMSEs, MPEs and Theil's U-statistics, and the results indicate that expert forecasts are better than forecasts from theoretical forecast models. However, the forecast accuracy of theoretical forecast models is closer to the forecast accuracy of CE forecasts for 3-month CDs compared to other variables.

Forecasts from CE have the lowest MAE and RMSE for real GDP growth (1.40 percent and 1.94 percent), for consumer price growth (0.45 percent and 0.57 percent) and unemployment rate (0.20 percent and 0.26 percent). This result is reflected in Theil's U-statistics that are below 1 for all theoretical forecast. The best theoretical forecast model is the weighted moving average model for these variables, and CE forecasts are not significantly better than this model for real GDP growth and unemployment rate (DM statistics of 1.34 and -0.47). However, this model is "post-based" and is therefore expected to provide relatively good forecasts. The number of variables in this model enhances this view. Also, remember that this model has weak economic interpretation (see above). CE forecasts' MPE is closer to zero than all theoretical forecast models for real GDP growth and unemployment rate, while the weighted moving average method is the only theoretical forecast model that has a MPE closer to zero than CE forecasts' MPE. Another observation is that almost every theoretical forecast model for unemployment rate has positive MPE, and therefore tend to underestimate actual unemployment rate.

For 3-month CDs in 3 months, CE forecasts have the fourth lowest MAE and the third lowest RMSE (0.20 percent and 0.31 percent), and the third closest to zero MPE (-0.08 percent). However, the simple random walk approach that makes forecasts equal to current 3-month CDs has almost the same forecast accuracy as CE forecasts (Theil's U equals 0.97), and the exponential moving average and exponential smoothing average

[41] Naszodi also suggests that we adjust for autocorrelation. However, the test results without this adjustment are so consistent that this adjustment will most likely not distort our results.

have slightly better forecasts (Theil's U statistics of 1.03 and 1.06). None of these models are significantly better or worse to forecast 3-month CDs in 3 months than CE forecasts (see DM statistics). For 3-month CDs in 12 months, CE forecasts have higher MAE than several theoretical models, and the MPE of CE forecasts are further from zero than most theoretical models. The simple random walk approach that makes forecasts equal to current 3-month CDs has almost the same forecast accuracy as CE forecasts (Theil's U equals 0.94), and the exponential smoothing average model has better forecasts (Theil's U statistics of 1.46). The ESA is significantly better than CE forecasts (DM statistics of -6.48), but this model is "post-based".

Table 7: Comparison of CE forecasts and forecasts from theoretical models

	Real GDP growth					Consumer price growth					Unemployment rate				
Forecast method	*MAE*	*RMSE*	*MPE*	*TU**	*DM***	*MAE*	*RMSE*	*MPE*	*TU**	*DM***	*MAE*	*RMSE*	*MPE*	*TU**	*DM***
CE Mean	1.40 %	1.94 %	-0.41 %			0.45 %	0.57 %	-0.11 %			0.20 %	0.26 %	-0.03 %		
RW - Last 12m	2.08 %	2.86 %	-0.40 %	0.68	5.69	0.70 %	0.91 %	-0.19 %	0.62	6.17	0.31 %	0.39 %	0.14 %	0.67	5.49
RW - Last 24m	2.06 %	2.76 %	-0.66 %	0.70	8.81	0.84 %	1.00 %	-0.22 %	0.57	9.93	0.43 %	0.51 %	0.19 %	0.52	9.44
MA - Last 2m	2.09 %	2.87 %	-0.47 %	0.68	6.00	0.71 %	0.93 %	-0.20 %	0.61	6.34	0.32 %	0.40 %	0.15 %	0.65	5.90
MA - Last 3m	2.10 %	2.87 %	-0.49 %	0.68	6.37	0.73 %	0.95 %	-0.21 %	0.60	6.50	0.33 %	0.42 %	0.15 %	0.63	6.30
MA - Last 4m	2.11 %	2.88 %	-0.52 %	0.67	6.68	0.74 %	0.97 %	-0.21 %	0.59	6.65	0.34 %	0.43 %	0.16 %	0.61	6.67
MA - Last 5m	2.13 %	2.88 %	-0.54 %	0.67	7.00	0.76 %	0.98 %	-0.22 %	0.58	6.81	0.35 %	0.44 %	0.16 %	0.60	7.03
MA - Last 6m	2.14 %	2.88 %	-0.57 %	0.67	7.37	0.77 %	0.99 %	-0.22 %	0.57	6.99	0.37 %	0.45 %	0.17 %	0.58	7.37
MA - Last 7m	2.14 %	2.89 %	-0.59 %	0.67	7.71	0.79 %	1.01 %	-0.23 %	0.57	7.19	0.38 %	0.47 %	0.17 %	0.57	7.70
MA - Last 8m	2.15 %	2.88 %	-0.61 %	0.67	8.04	0.80 %	1.02 %	-0.23 %	0.56	7.42	0.39 %	0.48 %	0.18 %	0.55	8.02
MA - Last 9m	2.15 %	2.88 %	-0.64 %	0.67	8.40	0.82 %	1.03 %	-0.23 %	0.55	7.68	0.40 %	0.49 %	0.18 %	0.54	8.32
MA - Last 10m	2.16 %	2.88 %	-0.66 %	0.67	8.67	0.83 %	1.04 %	-0.23 %	0.55	7.95	0.41 %	0.50 %	0.19 %	0.53	8.62
MA - Last 11m	2.16 %	2.88 %	-0.68 %	0.67	8.89	0.84 %	1.05 %	-0.23 %	0.54	8.25	0.42 %	0.51 %	0.19 %	0.52	8.90
MA - Last 12m	2.16 %	2.88 %	-0.70 %	0.67	9.07	0.86 %	1.06 %	-0.23 %	0.54	8.56	0.43 %	0.52 %	0.19 %	0.51	9.17
WMA	1.66 %	2.17 %	0.40 %	0.89	1.34	0.53 %	0.71 %	0.02 %	0.80	3.72	0.22 %	0.28 %	0.03 %	0.93	-0.47
EMA	2.06 %	2.76 %	-0.88 %	0.70	3.19	0.68 %	0.89 %	-0.19 %	0.64	5.56	0.29 %	0.38 %	0.14 %	0.70	4.18
ESA	2.10 %	2.84 %	0.83 %	0.68	3.26	0.71 %	0.91 %	0.19 %	0.63	6.02	0.31 %	0.39 %	-0.14 %	0.67	4.64
ARIMA (2.0.0)	2.12 %	2.89 %	-0.42 %	0.67	2.97	0.71 %	0.93 %	-0.19 %	N/A	5.93	N/A	N/A	N/A	N/A	
ARIMA (1.1.0)	2.12 %	2.88 %	-0.41 %	0.67	2.93	0.71 %	0.93 %	-0.19 %	0.67	5.96	0.31 %	0.39 %	0.14 %	0.67	5.57
ARIMA (1.0.0)	2.12 %	2.89 %	-0.43 %	0.67	2.96	0.73 %	0.95 %	-0.20 %	N/A	6.41	N/A	N/A	N/A	N/A	
ARIMA (0.1.1)	2.11 %	2.63 %	-1.50 %	0.74	4.60	0.97 %	1.35 %	-0.59 %	0.65	7.49	0.32 %	0.40 %	0.14 %	0.65	5.91

**TU = Theil's U-statistics*

***DM = Diebold-Mariano statistic*

Table 8: Comparison of CE forecasts and forecasts from theoretical models

	3-month CDs in 3 months					3-month CDs in 12 months				
Forecast method	*MAE*	*RMSE*	*MPE*	*TU**	*DM***	*MAE*	*RMSE*	*MPE*	*TU**	*DM***
CE Mean	0.20 %	0.31 %	-0.08 %			0.60 %	0.91 %	-0.45 %		
RW - Current month	0.19 %	0.32 %	-0.08 %	0.97	0.60	0.54 %	0.93 %	-0.35 %	0.94	0.54
RW - Last month	0.24 %	0.41 %	-0.10 %	0.76	4.64	0.59 %	1.00 %	-0.37 %	0.88	2.14
MA - Last 2m	0.22 %	0.36 %	-0.09 %	0.86	3.16	0.56 %	0.96 %	-0.36 %	0.91	1.35
MA - Last 3m	0.24 %	0.40 %	-0.10 %	0.78	4.44	0.58 %	0.99 %	-0.37 %	0.88	2.05
MA - Last 4m	0.26 %	0.44 %	-0.11 %	0.71	5.00	0.60 %	1.03 %	-0.38 %	0.86	2.63
MA - Last 5m	0.29 %	0.48 %	-0.12 %	0.65	5.36	0.62 %	1.05 %	-0.39 %	0.83	3.13
MA - Last 6m	0.31 %	0.52 %	-0.13 %	0.60	5.64	0.65 %	1.08 %	-0.40 %	0.81	3.57
MA - Last 7m	0.33 %	0.55 %	-0.14 %	0.56	5.86	0.67 %	1.11 %	-0.41 %	0.79	3.96
MA - Last 8m	0.35 %	0.59 %	-0.15 %	0.52	6.05	0.69 %	1.14 %	-0.42 %	0.77	4.30
MA - Last 9m	0.38 %	0.63 %	-0.16 %	0.49	6.22	0.71 %	1.17 %	-0.43 %	0.75	4.61
MA - Last 10m	0.40 %	0.67 %	-0.17 %	0.47	6.37	0.73 %	1.20 %	-0.44 %	0.73	4.87
MA - Last 11m	0.42 %	0.70 %	-0.18 %	0.44	6.51	0.75 %	1.22 %	-0.45 %	0.72	5.10
MA - Last 12m	0.44 %	0.74 %	-0.19 %	0.42	6.63	0.77 %	1.25 %	-0.46 %	0.70	5.30
EMA	0.18 %	0.30 %	-0.07 %	1.03	-0.70	0.53 %	0.91 %	-0.34 %	0.97	0.02
ESA	0.18 %	0.29 %	-0.06 %	1.06	-1.22	0.35 %	0.60 %	-0.19 %	1.46	-6.48
ARIMA (2.0.0)	0.23 %	0.38 %	-0.09 %	0.82	3.95	0.57 %	0.97 %	-0.36 %	0.91	1.50
ARIMA (1.1.0)	0.23 %	0.38 %	-0.08 %	0.82	3.91	0.57 %	0.97 %	-0.35 %	0.90	1.53
ARIMA (1.0.0)	0.24 %	0.41 %	-0.09 %	0.76	4.64	0.58 %	1.00 %	-0.37 %	0.88	2.10
ARIMA (0.1.1)	0.23 %	0.39 %	-0.08 %	0.81	4.13	0.57 %	0.98 %	-0.36 %	0.90	1.68

**TU = Theil's U-statistics*

***DM = Diebold-Mariano statistic*

Appendix 7: Results for forecast accuracy and rationality tests in business cycles

Forecast accuracy

Table 9: Forecast accuracy in the full sample

	Full sample				
Measure	*Real GDP growth*	*Consumer price growth*	*Unemployment rate*	*3-month CDs in 3m*	*3-month CDs in 12m*
MAE	1.40 %	0.45 %	0.20 %	0.20 %	0.60 %
RMSE	1.94 %	0.57 %	0.26 %	0.31 %	0.91 %
MPE	-0.41 %	-0.11 %	-0.03 %	-0.08 %	-0.45 %
STDEV	2.25 %	1.19 %	1.06 %	2.31 %	1.99 %
MNSE	0.86	0.48	0.25	0.13	0.46

Table 10: Forecast accuracy in contractions and expansions

	Expansions					Contractions				
Measure	*Real GDP growth*	*Consumer price growth*	*Unemployment rate*	*3-month CDs in 3m*	*3-month CDs in 12m*	*Real GDP growth*	*Consumer price growth*	*Unemployment rate*	*3-month CDs in 3m*	*3-month CDs in 12m*
MAE	0.90 %	0.43 %	0.19 %	0.18 %	0.49 %	2.37 %	0.51 %	0.23 %	0.24 %	0.82 %
RMSE	1.09 %	0.52 %	0.26 %	0.29 %	0.79 %	2.95 %	0.66 %	0.28 %	0.34 %	1.10 %
MPE	0.56 %	0.01 %	-0.11 %	-0.05 %	-0.25 %	-2.27 %	-0.34 %	0.14 %	-0.15 %	-0.82 %
STDEV	1.33 %	1.20 %	0.97 %	2.35 %	2.20 %	2.24 %	1.16 %	1.19 %	2.17 %	1.53 %
MNSE	0.82	0.43	0.26	0.13	0.36	1.32	0.57	0.23	0.16	0.72

Table 11: Forecast accuracy in different contractions and expansions

		Real GDP growth					Consumer price growth					Unemployment rate				
Period		*MAE*	*RMSE*	*MPE*	*STDEV*	*MNSE*	*MAE*	*RMSE*	*MPE*	*STDEV*	*MNSE*	*MAE*	*RMSE*	*MPE*	*STDEV*	*MNSE*
Contractions	**Asset Price Bubble**	1.88 %	1.94 %	-1.88 %	0.81 %	2.41	0.46 %	0.52 %	-0.46 %	0.57 %	0.91	0.09 %	0.11 %	-0.01 %	0.25 %	0.42
	Asian Crisis	1.80 %	2.13 %	-1.63 %	0.59 %	3.59	0.19 %	0.23 %	0.03 %	0.62 %	0.37	0.32 %	0.37 %	0.25 %	0.36 %	1.02
	Dot-Com Bubble	1.33 %	1.46 %	-1.04 %	0.53 %	2.77	0.46 %	0.52 %	-0.44 %	0.09 %	5.97	0.24 %	0.27 %	0.10 %	0.14 %	1.91
	Subprime Crisis	4.82 %	5.31 %	-4.77 %	2.26 %	2.35	1.01 %	1.15 %	-0.47 %	1.17 %	0.99	0.35 %	0.38 %	0.34 %	0.45 %	0.83
Expansions	**Pre Asset Price Bubble**	1.01 %	1.20 %	0.82 %	1.03 %	1.16	0.76 %	0.85 %	0.75 %	0.18 %	4.78	0.14 %	0.15 %	-0.14 %	0.01 %	24.36
	Pre Asian Crisis	0.57 %	0.75 %	0.23 %	0.80 %	0.93	0.43 %	0.53 %	-0.10 %	0.74 %	0.72	0.11 %	0.14 %	0.05 %	0.19 %	0.73
	Pre Dot-Com Bubble	1.27 %	1.37 %	1.09 %	0.80 %	1.71	0.51 %	0.54 %	-0.51 %	0.09 %	5.94	0.22 %	0.27 %	-0.17 %	0.06 %	4.45
	Pre Subprime Crisis	0.81 %	0.94 %	0.34 %	0.58 %	1.63	0.31 %	0.37 %	0.12 %	0.39 %	0.94	0.22 %	0.30 %	-0.16 %	0.54 %	0.55
	Post Subprime Crisis	2.04 %	2.08 %	2.04 %	2.21 %	0.94	0.46 %	0.52 %	-0.37 %	0.37 %	1.40	0.34 %	0.41 %	-0.34 %	0.05 %	8.30

Rationality tests in business cycles

Results from all tests of rationality follow below. First a table for tests of bias, then tables for weak-form efficiency and finally tables containing results from strong-form efficiency tests.

Results for tests of bias[42]

Table 12: Results for bias tests for the full sample and all sub-periods in the business cycles

	Real GDP growth	*Consumer Prices*	*Unemployment Rate*	*3 month CDs in 3m*	*3 month CDs in 12m*
Full sample					
Bias (α)	-0.0041	-0.0011	-0.0003	-0.0008*	-0.0045**
p-value α = 0	0.2542	0.2979	0.5792	0.0496	0.0029
Contractions					
Bias (α)	-0.0227***	-0.0034*	0.0014*	-0.0015*	-0.0082**
p-value α = 0	0.0000	0.0101	0.0367	0.0327	0.0011
Expansions					
Bias (α)	0.0056**	0.0001	-0.0012*	-0.0005	-0.0025
p-value α = 0	0.0020	0.9232	0.0239	0.359	0.0936
Different conractions					
Contraction - Asset price bubble	-0.0188***	-0.0046***	-0.0001	-0.0031**	-0.0159***
Contraction - Asian crisis	-0.0163*	0.0003	0.0025	0.0006	-0.0033***
Contraction - Dotcom bubble	-0.0104*	-0.0044**	0.001	-0.0011	-0.0016
Contraction - Subprime crisis	-0.0477***	-0.0047	0.0034***	-0.0011**	-0.0049***
Different expansions					
Expansion before the asset price bubble	0.0082	0.0075***	-0.0014***	0.0035	0.0037
Expansion before the asian crisis	0.0023	-0.001	0.0005	-0.0013	-0.0073*
Expansion before the dotcom bubble	0.0109**	-0.0051***	-0.0017	-0.0001	-0.002
Expansion before the subprime crisis	0.0034	0.0012	-0.0016	-0.0006***	-0.001
Expansion after the subprime crisis	NA	NA	NA	NA	NA

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

$A_t - F_t = \alpha + \varepsilon$

Results for tests of weak-form efficiency[43]

The first set of tables show efficiency tests 1, 2 and 3 for all the macroeconomic variables we have studied. They show results for the full sample and for the split sample of contractions and expansions. The second set of tables show the same tests for all specific contractions and expansions, for real GDP growth, consumer price growth and unemployment rate.

[42] Table for bias displays the constant (corresponding to MPE) for all tests in addition to p-values for the full sample and contractions/expansions.

[43] All efficiency tables display p-values. See section 5 for more info on the tests and how they are interpreted.

Table 13: Results for weak-form efficiency tests for the full sample and in expansions/contractions

Weak-form efficiency tests		Real GDP growth			Consumer price growth			Unemployment rate		
		Full sample	Contractions	Expansions	Full sample	Contractions	Expansions	Full sample	Contractions	Expansions
Eff.test 1[1]	α = 0	0.9855	0.0168	0.0000	0.3597	0.2833	0.8378	0.1001	0.5798	0.0906
	β = 0	0.0232	0.6284	0.0436	0.7957	0.2271	0.1914	0.0723	0.2000	0.0181
	Ho (α=β=0)	0.0692	0.0000	0.0000	0.5813	0.0028	0.4250	0.1921	0.0514	0.0027
Eff.test 2[2]	α = 0	0.7146	0.0087	0.0000	0.2390	0.1295	0.6706	0.0500	0.9782	0.0290
	β = 0	0.4206	0.4553	0.0548	0.8402	0.1805	0.0867	0.0298	0.4878	0.0026
	Ho (α=β=0)	0.2975	0.0000	0.0000	0.4982	0.0005	0.2285	0.0856	0.0686	0.0000
Eff.test 3[3]	α = 0	0.1437	0.0000	0.0132	0.0872	0.0145	0.6193	0.6994	0.0435	0.0377
	β = 0	0.5640	0.5000	0.0000	0.5521	0.3776	0.3873	0.5904	0.9895	0.9945
	Ho (α=β=0)	0.2912	0.0000	0.0000	0.2309	0.0441	0.5208	0.7944	0.1140	0.1138

[1] $A_t - F_t = \alpha + \beta * A_{t-12} + \varepsilon$ [2] $At - Ft = \alpha + \beta * Ft + \varepsilon$ [3] $At - Ft = \alpha + \beta * (At-12 - Ft-12) + \varepsilon$

Table 14: Results for weak-form efficiency tests for the full sample and in expansions/contractions

Weak-form efficiency tests		3-month CDs in 3 months			3-month CDs in 12 months		
		Full sample	Contractions	Expansions	Full sample	Contractions	Expansions
Eff.test 1[1]	α = 0	0.0078	0.8353	0.0049	0.0024	0.0003	0.0114
	β = 0	0.9199	0.0010	0.0605	0.3019	0.0000	0.8169
	Ho (α=β=0)	0.0066	0.0001	0.0100	0.0000	0.0000	0.0372
Eff.test 2[2]	α = 0	0.0110	0.9113	0.0051	0.0289	0.0130	0.0170
	β = 0	0.9602	0.0005	0.0722	0.2454	0.0000	0.9457
	Ho (α=β=0)	0.0070	0.0000	0.0116	0.0000	0.0000	0.0306
Eff.test 3[3]	α = 0	0.0022	0.0226	0.0447	0.0054	0.0120	0.0127
	β = 0	0.0006	0.1817	0.0015	0.6828	0.1334	0.4538
	Ho (α=β=0)	0.0004	0.0443	0.0061	0.0003	0.0000	0.0418

[1] $A_t - F_t = \alpha + \beta * A_{t-12} + \varepsilon$ [2] $At - Ft = \alpha + \beta * Ft + \varepsilon$ [3] $At - Ft = \alpha + \beta * (At-12 - Ft-12) + \varepsilon$

Table 15: Results for weak-form efficiency tests for different expansions and contractions – Real GDP growth

Real GDP growth		Contraction - Asset price bubble	Contraction - Asian crisis	Contraction - Dotcom bubble	Contraction - Subprime crisis	Expansion before the asset price bubble	Expansion before the asian crisis	Expansion before the dotcom bubble	Expansion before the subprime crisis	Expansion after the subprime crisis
Eff.test 1[1]	α = 0	0.0000	0.0000	0.0033	0.0052	0.1630	0.1172	0.0012	0.0000	NA
	β = 0	0.3468	0.0000	0.0000	0.5190	0.3778	0.5877	0.0105	0.0016	NA
	Ho (α=β=0)	0.0000	0.0000	0.0000	0.0000	0.0002	0.2348	0.0000	0.0000	NA
Eff.test 2[2]	α = 0	0.0094	0.0001	0.0153	0.0000	0.0000	0.0457	0.0000	0.0000	NA
	β = 0	0.9957	0.0000	0.0012	0.1184	0.0000	0.2046	0.0557	0.0000	NA
	Ho (α=β=0)	0.0000	0.0000	0.0000	0.0000	0.0000	0.1105	0.0001	0.0000	NA
Eff.test 3[3]	α = 0	0.0000	0.0000	0.0001	0.0018	NA	0.3559	0.0001	0.0990	NA
	β = 0	0.1655	0.0000	0.0004	0.7361	NA	0.4450	0.1153	0.1855	NA
	Ho (α=β=0)	0.0000	0.0000	0.0000	0.0000	NA	0.0422	0.0000	0.0444	NA

[1] $A_t - F_t = \alpha + \beta * A_{t-12} + \varepsilon$ [2] $At - Ft = \alpha + \beta * Ft + \varepsilon$ [3] $At - Ft = \alpha + \beta * (At-12 - Ft-12) + \varepsilon$

NA - not available due to Insufficient observations to do Newey-West regressions

Table 16: Results for weak-form efficiency tests for different expansions and contractions – Consumer price growth

Consumer price growth		Contraction - Asset price bubble	Contraction - Asian crisis	Contraction - Dotcom bubble	Contraction - Subprime crisis	Expansion before the asset price bubble	Expansion before the asian crisis	Expansion before the dotcom bubble	Expansion before the subprime crisis	Expansion after the subprime crisis
Eff.test 1[1]	α = 0	0.4633	0.0072	0.0000	0.0000	0.0561	0.8794	0.0000	0.9064	NA
	β = 0	0.0989	0.0012	0.0000	0.0000	0.2157	0.0220	0.0344	0.0282	NA
	Ho (α=β=0)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0508	0.0000	0.0000	NA
Eff.test 2[2]	α = 0	0.6585	0.8865	0.0000	0.5792	0.0000	0.0506	0.0000	0.4750	NA
	β = 0	0.2181	0.7630	0.0000	0.3644	0.0000	0.1113	0.0000	0.0050	NA
	Ho (α=β=0)	0.0001	0.8984	0.0000	0.3075	0.0000	0.1439	0.0000	0.0000	NA
Eff.test 3[3]	α = 0	0.0000	0.0304	0.3451	0.0000	NA	0.9326	0.0000	0.3222	NA
	β = 0	0.7802	0.0004	0.1085	0.0000	NA	0.7222	0.3003	0.4424	NA
	Ho (α=β=0)	0.0000	0.0000	0.0000	0.0000	NA	0.3392	0.0000	0.4540	NA

[1] $A_t - F_t = \alpha + \beta * A_{t-12} + \varepsilon$ [2] $At - Ft = \alpha + \beta * Ft + \varepsilon$ [3] $At - Ft = \alpha + \beta * (At-12 - Ft-12) + \varepsilon$

NA - not available due to Insufficient observations to do Newey-West regressions

Table 17: Results for weak-form efficiency tests for different expansions and contractions – Unemployment rate

Unemployment rate		*Contraction - Asset price bubble*	*Contraction - Asian crisis*	*Contraction - Dotcom bubble*	*Contraction - Subprime crisis*	*Expansion before the asset price bubble*	*Expansion before the asian crisis*	*Expansion before the dotcom bubble*	*Expansion before the subprime crisis*	*Expansion after the subprime crisis*
Eff.test 1[1]	$\alpha = 0$	0.0001	0.0000	0.0000	0.5249	0.0000	0.0686	0.2783	0.0000	NA
	$\beta = 0$	0.0000	0.0000	0.0000	0.5983	0.0000	0.0441	0.3442	0.0000	NA
	Ho (α=β=0)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0292	0.0057	0.0000	NA
Eff.test 2[2]	$\alpha = 0$	0.0000	0.0000	0.0000	0.1985	0.0000	0.4547	0.0000	0.0000	NA
	$\beta = 0$	0.0000	0.0000	0.0000	0.4861	0.0000	0.4807	0.0000	0.0000	NA
	Ho (α=β=0)	0.0000	0.0000	0.0000	0.0000	0.0000	0.5691	0.0000	0.0000	NA
Eff.test 3[3]	$\alpha = 0$	0.1759	0.0000	0.0000	0.0000	NA	0.0541	0.0027	0.1921	NA
	$\beta = 0$	0.0024	0.0000	0.0000	0.1853	NA	0.0054	0.1440	0.2307	NA
	Ho (α=β=0)	0.0093	0.0000	0.0000	0.0000	NA	0.0079	0.0017	0.0669	NA

[1] $A_t - F_t = \alpha + \beta * A_{t-12} + \varepsilon$ [2] $At - Ft = \alpha + \beta * Ft + \varepsilon$ [3] $At - Ft = \alpha + \beta * (At\text{-}12 - Ft\text{-}12) + \varepsilon$

NA - not available due to Insufficient observations to do Newey-West regressions

Results for tests of strong-form efficiency

First table shows strong-form efficiency test for forecasts of real GDP growth, consumer price growth and unemployment rate. The second table shows results where money supply is included in the information set, for consumer price growth and interest rates expectations.

Table 18: Results for strong-form efficiency tests– Real GDP growth, consumer price growth and unemployment rate

Efficiency test 4 - Strong-form	Forecast errors real GDP growth	Forecast errors consumer price growth	Forecast errors unemployment rate
GDP mean forecast next 12 months	-0.1400		
	(0.1777)		
p-value $\beta_{Forecast\ GDP} = 0$	0.4316		
Inflation mean forecast next 12 months		-0.2186	
		(0.2911)	
p-value $\beta_{Forecast\ CPI} = 0$		0.4535	
Unemployment rate mean forecast next 12 months			-0.1116
			(0.0973)
p-value $\beta_{Forecast\ UR} = 0$			0.2524
Actual GDP last 12 months	-0.2698	0.0986	0.0433*
	(0.1452)	(0.0602)	(0.0212)
p-value $\beta_{GDP} = 0$	0.0645	0.1029	0.0422
Actual inflation last 12 months	-1.5168*	-0.2680	0.0963
	(0.6561)	(0.1991)	(0.0764)
p-value $\beta_{Inflation} = 0$	0.0216	0.1795	0.2088
Actual unemployment rate last 12 months	0.2374	-0.0648	-0.0769
	(0.4131)	(0.1812)	(0.1281)
p-value $\beta_{UR} = 0$	0.5659	0.7208	0.5489
Actual 3-month CDs at date	0.8923**	0.1512	-0.1405***
	(0.2680)	(0.1158)	(0.0266)
p-value $\beta_{CDs} = 0$	0.0010	0.1928	0.0000
Constant	-0.0119	0.0002	0.0080***
	(0.0177)	(0.0075)	(0.0023)
p-value $\alpha = 0$	0.5030	0.9828	0.0007
H_0 ($\beta_{GDP} = \beta_{Inflation} = \beta_{UR} = \beta_{CDs} = 0$)	0.0001	0.3065	0.0000
Observations	243	243	243

Newey-West standard errors in parantheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 19: Results for strong-form efficiency tests– Consumer price growth, 3 month CDs in 3 months and 12 months - including money supply in information set

Efficiency test 4 - Strong-form	Forecast errors consumer price growth	Forecast errors 3m CDs in 3 months	Forecast errors 3m CDs in 12 months
Inflation mean forecast next 12 months	-0.3201		
	(0.2829)		
p-value $\beta_{Forecast\ CPI} = 0$	0.2591		
3-month CDs in 3 months mean forecast		-0.6879**	
		(0.2342)	
p-value $\beta_{Forecast\ CD3m} = 0$		0.0036	
3-month CDs in 12 months mean forecast			-0.7666
			(0.4390)
p-value $\beta_{Forecast\ CD12m} = 0$			0.0821
Money supply (M2) seasonally adj. Y-o-Y growth	0.1213***	0.0721**	0.2078**
	(0.0320)	(0.0259)	(0.0637)
p-value $\beta_{Forecast\ M2} = 0$	0.0002	0.0058	0.0013
Actual GDP last 12 months	0.0855	0.0188	0.0697
	(0.0572)	(0.0161)	(0.0449)
p-value $\beta_{GDP} = 0$	0.1363	0.2417	0.1219
Actual inflation last 12 months	0.0407	0.0037	-0.0624
	(0.1134)	(0.0398)	(0.0818)
p-value $\beta_{Inflation} = 0$	0.7199	0.9269	0.4465
Actual unemployment rate last 12 months	-0.1631	-0.0318	-0.1387
	(0.1747)	(0.0609)	(0.1647)
p-value $\beta_{UR} = 0$	0.3515	0.6023	0.4008
Actual 3-month CDs at date	-0.0547	0.5933**	0.3756
	(0.0916)	(0.2269)	(0.3965)
p-value $\beta_{CDs} = 0$	0.5509	0.0095	0.3444
Constant	0.0026	-0.0005	0.0008
	(0.0071)	(0.0026)	(0.0075)
p-value $\alpha = 0$	0.7182	0.8351	0.9164
H_0 ($\beta_{M2} = \beta_{GDP} = \beta_{Inflation} = \beta_{UR} = \beta_{CDs} = 0$)	0.0000	0.0000	0.0289
Observations	242	242	242

Newey-West standard errors in parantheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix 8: Tests of disagreement

In this appendix we outline the measures we use in disagreement tests and disagreement test 1 and 2. Finally, the results from the disagreement tests are presented.

Measures to use in disagreement tests

Several papers have employed data on the dispersion in inflation expectations as a rough proxy for "inflation uncertainty". The Mankiw et al. (2003) approach is to interpret different inflation expectations as reflecting disagreement in the population. That is, different forecasts reflect different expectations. Llambros and Zarnowitz (1987) make a distinction between disagreement and uncertainty. They take intrapersonal variation as uncertainty and interpersonal variation as disagreement, and find that while there are pronounced changes through time in disagreement, uncertainty varies very little. We do not have variables that show intrapersonal variation, but we do have variables showing interpersonal variation, for example a high-minus-low-variable[44] (constructed by taking the difference between the high and low forecasts) and standard deviation. We will use these two measures and carry out tests on disagreement in forecasting GDP growth, inflation and unemployment rate.

Disagreement test 1 - T-tests of different disagreement level

The first test on disagreement is simply to look at the mean disagreement in different periods. We use simple t-tests to see whether the mean of standard deviation (SD) and high minus low (HML) differs significantly in different sub-periods. The null hypothesis is that disagreement is equal. The test is carried out for the time series of current year and next year, as well as for the constructed next 12-months data. For the latter, we do not have data for standard deviation, since it does not seem reasonable to create them in the same manner, i.e. using WAN12.

Disagreement test 2 – What makes disagreement rise and fall?

The second set of tests carried out for disagreement is bivariate and multivariate regressions of disagreement (high minus low next 12 months).

[44] Mankiw et al. (2003) use interquartile (removing upper and lower quartile) based on the judgement that extreme observations are not particularly informative. We will use high minus low since that is what we have available.

$$HML_t = \alpha + ß_1 A_{t-12} + ß_2 F_t + ß_3 \Delta A_{t-12} + \varepsilon_t$$

On the right-hand-side of the regression we include actual values of the variable over the last 12 months, the forecast for the next 12 months, and the change in the actual value of the variable over the last 12 months. This test describes how behaviors in underlying variables affect disagreement, e.g. in testing disagreement in forecasts for real GDP growth, a negative sign on the first beta-coefficient shows that high real GDP growth leads to low disagreement[45].

Results from disagreement tests

Table 20: Results from t-test for different disagreement level for forecasts in the full sample and in expansions/contractions

Disagreement test 1	Current year forecasts		Next year forecasts		Next 12m forecasts
	SD	*HML*	*SD*	*HML*	*HML*
Real GDP growth					
Expansions	0.41%	1.59%	0.71%	2.65%	2.12%
Contractions	0.51%	2.08%	0.55%	2.07%	2.71%
Full sample	*0.44%*	*1.75%*	*0.65%*	*2.45%*	*2.32%*
Difference	*-0.0011*	*-0.0050*	*0.0016*	*0.0058*	*-0.0059*
H_0 (Difference = 0)	0.0003	0.0001	0.0000	0.0001	0.0000
Inflation					
Expansions	0.22%	0.89%	0.36%	1.30%	1.13%
Contractions	0.23%	0.88%	0.35%	1.28%	1.25%
Full sample	*0.22%*	*0.88%*	*0.36%*	*1.29%*	*1.17%*
Difference	*-0.0001*	*0.0001*	*0.0002*	*0.0003*	*-0.0012*
H_0 (Difference = 0)	0.4304	0.9223	0.3149	0.7119	0.0574
Unemployment rate					
Expansions	0.14%	0.53%	0.32%	1.15%	0.83%
Contractions	0.13%	0.49%	0.22%	0.74%	0.74%
Full sample	*0.13%*	*0.52%*	*0.29%*	*1.01%*	*0.80%*
Difference	*0.0000*	*0.0004*	*0.0011*	*0.0041*	*0.0009*
H_0 (Difference = 0)	0.7603	0.2588	0.0000	0.0000	0.0196

[45] The fourth regression in the tables is a multivariate regression, while the first three regressions are bivariate. The multivariate regression will not necessarily correspond to the bivariate in terms of significance. This is probably because of multicollinearity in the variables (e.g. high real GDP growth last period leads to high real GDP growth forecasts for the next period). Still, the joint hypothesis of all coefficients being zero is rejected for all three variables tested.

Table 21: Results from disagreement test 2 for the full sample – What makes disagreement rise and fall?

Disagreement test 2	*Actual values*	*Forecasts*	*Change*	*Multivariate*
Real GDP growth				
Actual real GDP growth last (β_1)	-0.1495**			-0.1223
	(0.0485)			(0.1553)
Mean real GDP growth forecast next 12 months (β_2)		-0.2235**		-0.0151
		(0.0807)		(0.2550)
Changes in actual real GDP growth last 12m (β_3)			-1.0528*	-0.7444
			(0.4992)	(0.4833)
Constant (α)	0.0252***	0.0265***	0.0227***	0.0247***
	(0.0018)	(0.0015)	(0.0017)	(0.0015)
H_0 ($\beta_1 = \beta_2 = \beta_3 = 0$)				0.0002
Inflation				
Actual inflation last 12 months (β_1)	0.1783***			0.1306
	(0.0458)			(0.0699)
Mean inflation forecast next 12 months (β_2)		0.1920***		0.0675
		(0.0497)		(0.0770)
Changes in actual inflation last 12m (β_3)			-0.1448	-0.4153
			(0.3936)	(0.3365)
Constant (α)	0.0106***	0.0107***	0.0117***	0.0105***
	(0.0005)	(0.0006)	(0.0007)	(0.0005)
H_0 ($\beta_1 = \beta_2 = \beta_3 = 0$)				0.0030
Unemployment rate				
Actual unemployment rate last 12 months (β_1)	0.0983*			-0.1816
	(0.0399)			(0.1300)
Mean unemployment rate forecast next 12 months (β_2)		0.1271***		0.2784*
		(0.0373)		(0.1285)
Changes in actual unemployment rate last 12m (β_3)			3.8611***	1.6731
			(0.9597)	(1.0147)
Constant (α)	0.0043*	0.0030	0.0076***	0.0037**
	(0.0017)	(0.0015)	(0.0004)	(0.0014)
H_0 ($\beta_1 = \beta_2 = \beta_3 = 0$)				0.0000

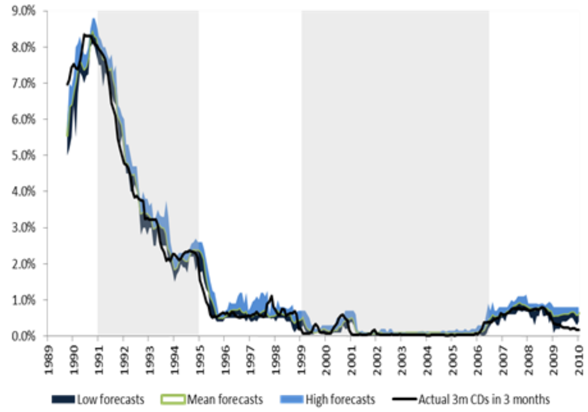
Newey-West standard errors in parantheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix 9: Figures for different monetary policy regimes

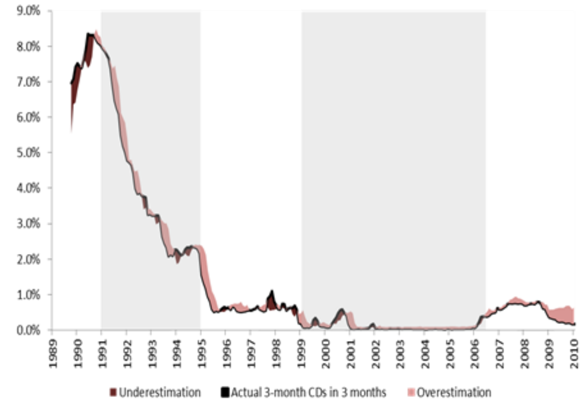
Figure 28-30: 3-month CDs in 3 months

Figure 28: Forecast band and actual values



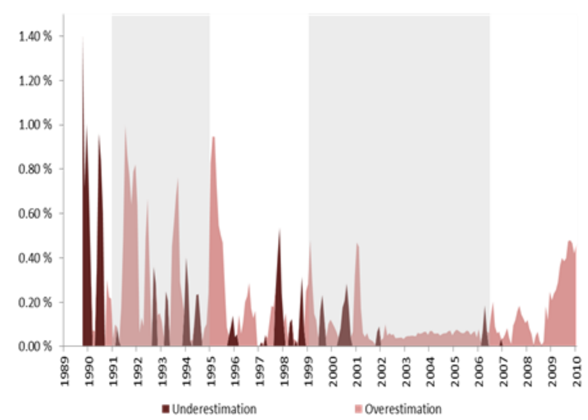
Source: Consensus Economics and Thomson Reuters Datastream

Figure 29: Over- and underestimation - Mean forecasts vs actual values



Source: Consensus Economics and Thomson Reuters Datastream

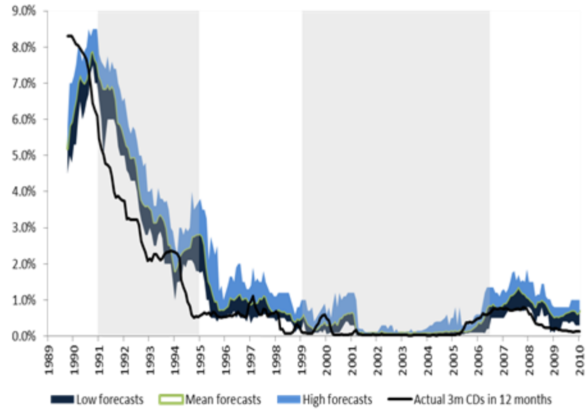
Figure 30: Absolute Forecast Error - Difference between mean forecasts and actual values



Source: Consensus Economics and Thomson Reuters Datastream

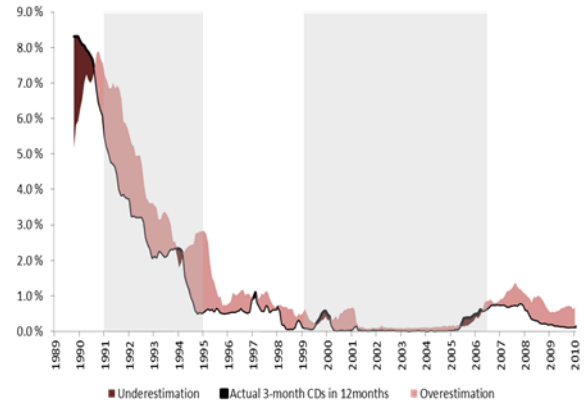
Figure 31-33: 3-month CDs in 12 months

Figure 31: Forecast band and actual values



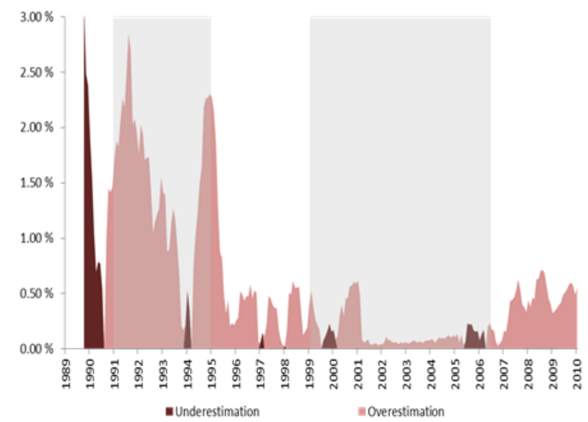
Source: Consensus Economics and Thomson Reuters Datastream

Figure 32: Over- and underestimation - Mean forecasts vs actual values



Source: Consensus Economics and Thomson Reuters Datastream

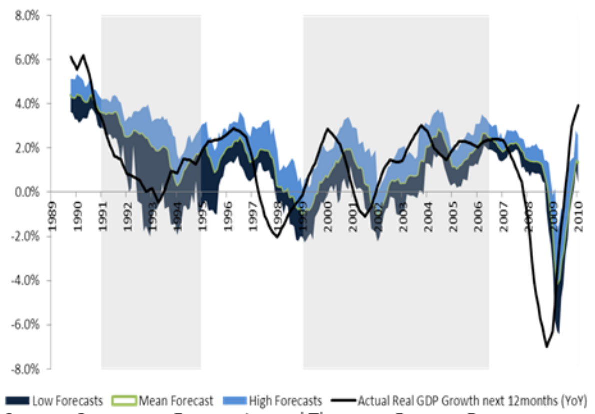
Figure 33: Absolute Forecast Error - Difference between mean forecasts and actual values



Source: Consensus Economics and Thomson Reuters Datastream

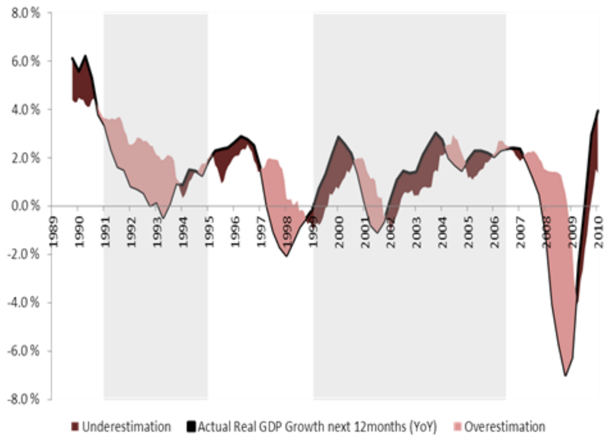
Figure 34-36: Real GDP growth for the next 12 months

Figure 34: Forecast band and actual values



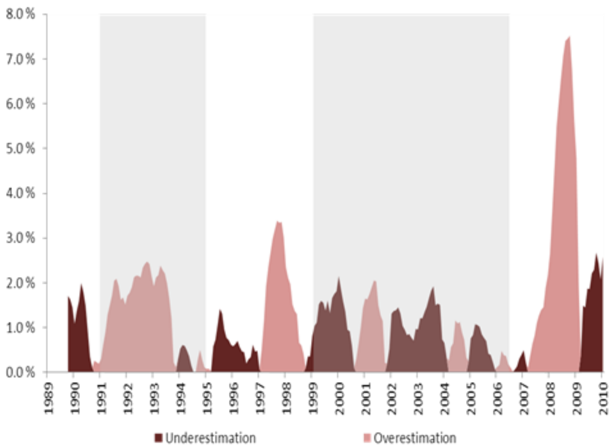
Source: Consensus Economics and Thomson Reuters Datastream

Figure 35: Over- and underestimation - Mean forecasts vs actual values



Source: Consensus Economics and Thomson Reuters Datastream

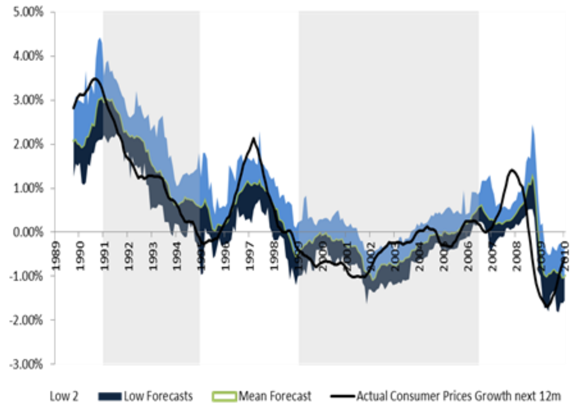
Figure 36: Absolute Forecast Error - Difference between mean forecasts and actual values



Source: Consensus Economics and Thomson Reuters Datastream

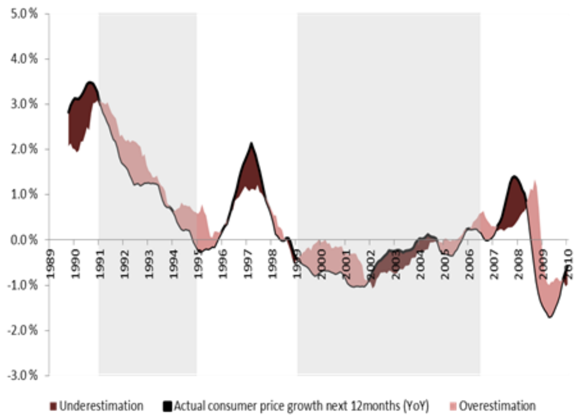
Figure 37-39: Consumer price growth for the next 12 months

Figure 37: Forecast band and actual values



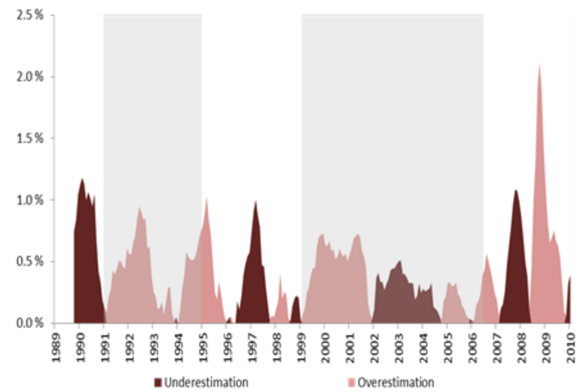
Source: Consensus Economics and Thomson Reuters Datastream

Figure 38: Over- and underestimation - Mean forecasts vs actual values



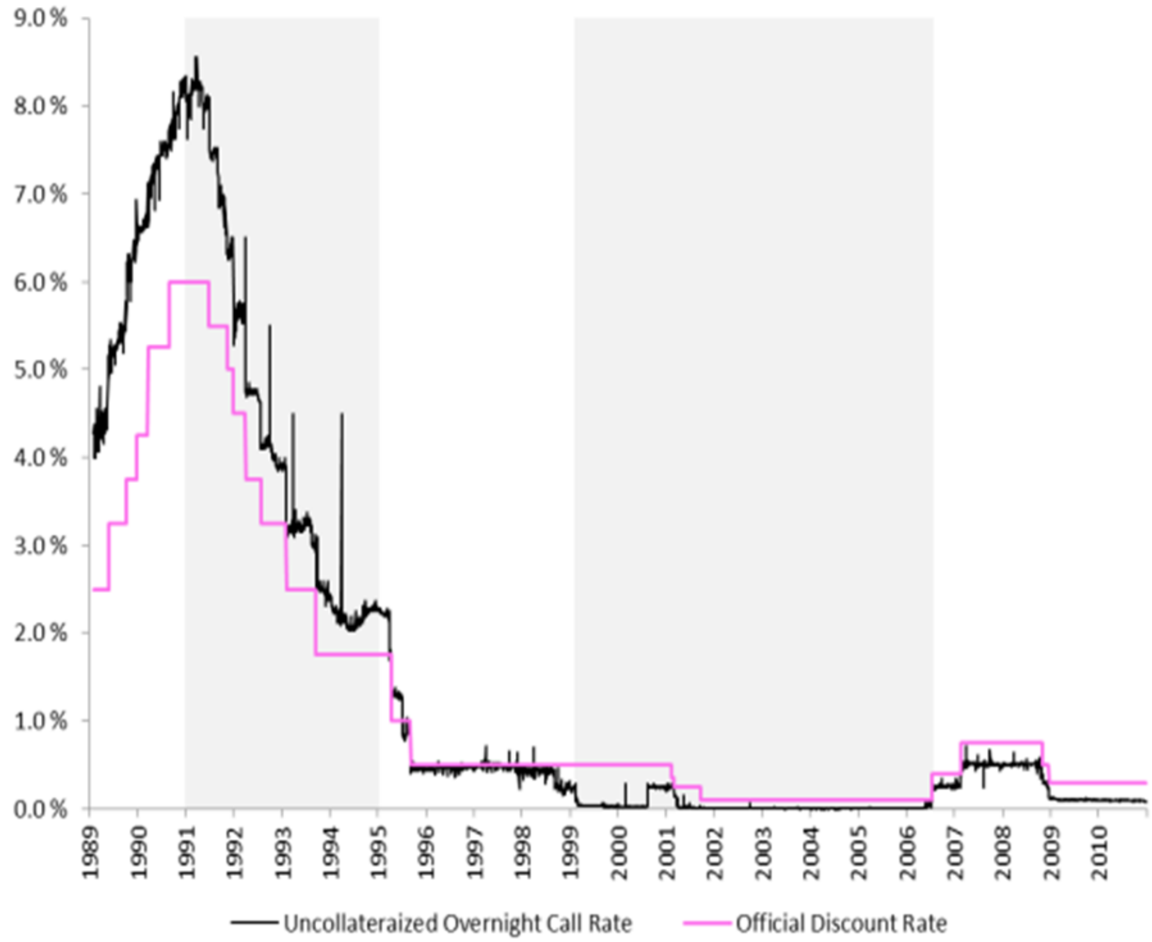
Source: Consensus Economics and Thomson Reuters Datastream

Figure 39: Absolute Forecast Error - Difference between mean forecasts and actual values



Source: Consensus Economics and Thomson Reuters Datastream

Figure 40: Uncollateralized Overnight Call Rate and Official Discount Rate



Source: Thomson Reuters Datastream

Appendix 10: Forecast accuracy and bias in different monetary policy regimes

Figure 41: Timeline of periods of different monetary policy regimes

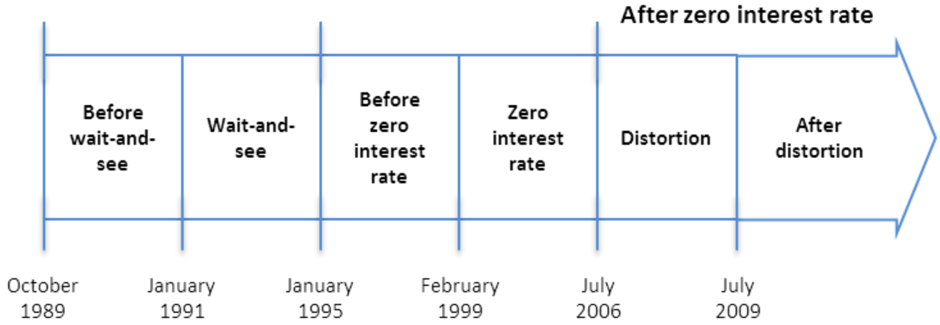


Table 22: Forecast accuracy of CE forecasts in periods of different monetary regimes

	Consumer price growth					3-month CDs in 3 months					3-month CDs in 12 months				
Monetary policy period	*MAE*	*RMSE*	*MPE*	*STDEV*	*MNSE*	*MAE*	*RMSE*	*MPE*	*STDEV*	*MNSE*	*MAE*	*RMSE*	*MPE*	*STDEV*	*MNSE*
All periods	0.45 %	0.57 %	-0.11 %	1.19 %	0.48	0.20 %	0.31 %	-0.08 %	2.30 %	0.13	0.60 %	0.91 %	-0.45 %	1.99 %	0.46
Before wait-and-see	0.66 %	0.75 %	0.42 %	0.39 %	1.91	0.48 %	0.62 %	0.12 %	0.66 %	0.94	1.69 %	1.86 %	-0.42 %	1.58 %	1.18
Wait-and-see	0.48 %	0.55 %	-0.47 %	0.68 %	0.81	0.35 %	0.44 %	-0.24 %	1.29 %	0.34	1.28 %	1.44 %	-1.24 %	1.08 %	1.33
Before zero interest rate policy	0.30 %	0.40 %	0.08 %	0.86 %	0.46	0.15 %	0.19 %	-0.01 %	0.23 %	0.82	0.31 %	0.36 %	-0.29 %	0.27 %	1.36
Zero interest rate policy	0.35 %	0.40 %	-0.12 %	0.40 %	1.01	0.08 %	0.11 %	-0.05 %	0.18 %	0.62	0.15 %	0.21 %	-0.10 %	0.26 %	0.81
Distortion	0.88 %	1.01 %	-0.21 %	1.18 %	0.86	0.14 %	0.17 %	-0.14 %	0.19 %	0.92	0.48 %	0.50 %	-0.48 %	0.25 %	2.02
After distortion	0.34 %	0.39 %	-0.11 %	0.34 %	1.15	0.44 %	0.44 %	-0.44 %	0.03 %	16.40	0.55 %	0.55 %	-0.55 %	0.01 %	58.06

Table 23: Results for bias tests for CE forecasts in periods of different monetary regimes

Monetary policy period	*Real GDP growth*	*Consumer price growth*	*3-month CDs in 3 months*	*3-month CDs in 12 months*
All periods	-0.0041	-0.0011	-0.0008*	-0.0045**
Before wait-and-see	0.001	0.0042	0.0012	-0.0042
Wait-and-see	-0.0086	-0.0047***	-0.0024*	-0.0124***
Before zero interest rate policy	-0.0036	0.0008	-0.0001	-0.0029***
Zero interest rate policy	0.0031	-0.0012	-0.0005***	-0.0010*
Distortion	-0.0309**	-0.0021	-0.0014**	-0.0048***
After distortion	NA	NA	NA	NA

$^{*}p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$

$A_t - F_t = \alpha + \varepsilon$