

Heterogeneity in Inflation Forecasts

*Analysing the Performance of the Mean, Median and
Individual Forecasters*

NHH



Norwegian School of Economics
Master thesis - Major in Economic Analysis

Author:
Kaja Kierulf

Supervisor:
Krisztina Molnár

Bergen, Spring 2013

This thesis was written as a part of the Master of Science in Economics and Business Administration at NHH. Please note that neither the institution nor the examiners are responsible - through the approval of this thesis - for the theories and methods used, or results and conclusions drawn in this work.

Acknowledgements

I would like to thank my supervisor Krisztina Molnár for her guidance and support. In particular I would like to thank her for engaging discussions helping me in finding and developing an interesting subject for my thesis.

I would also like to thank Yngve Borgan for generous help and patience throughout the process of writing this thesis. I would also like to thank him for cooking dinner every day for the past six months.

Bergen, June 2013

Kaja Kierulf

Abstract

In this thesis heterogeneity in inflation forecasts is analysed. First, the Survey of Professional Forecasters and the Michigan survey are analysed with focus on the mean, median and individual forecasters. The forecasters are analysed with respect to forecast bias and forecast accuracy. The performance of the individual forecasters are analysed and compared to the performance of the mean and median forecasters. The mean and median forecasters are found to be less biased and more accurate than the individual forecasters. Second, a simple dynamic stochastic general equilibrium model describing an economy with heterogeneous consumers is analysed. The consumers in the model use two predictors and switch between the predictors based on their past performance. The predictors in the model are also analysed with focus on forecast bias and forecast accuracy. The mean predictor is found to be both less biased and more accurate than the two predictors used by the consumers in the model. However, the results are more mixed for the median predictor which is found to be less biased, but less accurate than the predictors used by the consumers.

Contents

1	Introduction	1
2	Evaluating Forecasts	5
2.1	Evaluating Forecasts	5
2.1.1	Measure of Forecast Bias	5
2.1.2	Measures of Forecast Accuracy	6
2.2	Tests of Hypotheses	7
2.2.1	Tests of Forecast Bias	7
2.2.2	Tests of Forecast Accuracy	10
2.2.3	Correlation of Forecast Errors	11
3	Surveys	13
3.1	Survey of Professional Forecasters	13
3.1.1	Calculations	13
3.1.2	Analysing Forecast Bias	17
3.1.3	Analysing Forecast Accuracy	19
3.2	Michigan Survey	23
3.2.1	Calculations	24
3.2.2	Analysing Forecast Accuracy	27
3.3	Comparing the Surveys	30
4	The Model	33
4.1	The model	33
4.1.1	Predictors	36

5	Simulation	39
5.1	Approach for Simulating the Model	39
5.2	The Standard Model	40
5.2.1	Analysing Forecast Bias	42
5.2.2	Analysing Forecast Accuracy	44
5.2.3	Comments on the Performance of the Predictors	48
5.3	Robustness Analysis	49
6	Conclusion	51
A	Analysing the Preprocessed Dataset from the Michigan Survey	57
B	Tables Concerning the Robustness Analysis of the Model	59

Chapter 1

Introduction

The impact of inflation on the economy is widely discussed in the literature. Expectations of inflation is widely used by a variety of decision-makers from consumers to corporations and central banks. They are concerned with how the inflation affects the outcome of their decisions. Consequently forecasts of inflation becomes important for the decision-makers and for understanding developments in the economy.

Ang et al. (2007) find that surveys are better at forecasting inflation than both macro variables and asset markets. The master theses of Eikill (2012) and Øyvind Steira (2012) study the performance of individual forecasters in the Survey of Professional Forecasters. They find diversity in the individual forecasts where some perform good while others perform less good. In addition Eikill (2012) finds that the consensus forecasts perform better than the performance of the majority of the individuals.

In this thesis we will analyse the Survey of Professional Forecasters and the Michigan survey. The focus of the analysis is on heterogeneity in the forecasts and on the performance of the mean and median forecasters. The mean and median forecasters are analysed and compared to the individual forecasters. Most studies of inflation surveys focus only on the mean and median forecasts, thus analysing and working with the data with focus on heterogeneity and individual forecasts might contribute to better understanding of the surveys.

Brock and Hommes (1997) study heterogeneity in expectations in a cobweb

model. Agents in their model choose to use one of several predictors based on their past performance. This results in dynamics in the predictors used by the agents in the model. Anufriev et al. (2012) use a simple frictionless dynamic stochastic general equilibrium model to analyse inflation dynamics. Their analysis focuses on alternative interest rate rules when agents have heterogeneous expectations and update their beliefs as in Brock and Hommes (1997).

We will use the same model as Anufriev et al. (2012) to analyse heterogeneity in inflation forecasts. We also analyse the performance of the mean and median predictors by comparing them to the other predictors in the economy. To our knowledge no papers focus on the performance of the mean and median predictor compared to the other predictors in a model. Thus the analysis is valuable and contributes to new information in the literature.

There are two main findings in this thesis. First, there is heterogeneity in the inflation forecasts. As expected there is heterogeneity in the forecasts studied in the surveys and there is also heterogeneity in the model we analyse. The main focus of the model is on heterogeneity in the inflation forecasts. There are two predictors in the model and the consumers switch between the predictors based on past performance. Both the predictors persist in the long run and thus there is heterogeneity in the model.

The second finding is that the mean is a good forecaster. Ang et al. (2007) find that the surveys provide good forecasts compared to macroeconomic variables and asset markets. We study the mean and median forecasters in the surveys and find that they are less biased and more accurate than the individual forecasters. Thus the mean and median forecasters perform well because all the information obtained from the individual forecasts is utilised. In the analysis of the theoretical model we also find that the mean predictor performs well. The predictor is less biased and more accurate than the two predictors used by the consumers in the economy. The median predictor in the model does not perform equally well because the model is simple and not well suited for producing median forecasts. In addition to the good performance of the mean predictor, we observe that it is sluggish. The mean predictor is observed to be sluggish both in the surveys and in the model.

In the following we start in chapter 2 by presenting the evaluation of the forecasts in the surveys and in the theoretical model. We also describe how the forecasts in the surveys are tested. Chapter 3 contains the analysis of the surveys. The surveys analysed in this thesis are the Survey of Professional Forecasters and the Michigan survey. Chapter 4 presents the model analysed in the thesis. Chapter 5 analyses the model by performing simulations. The conclusion is given in chapter 6.

Chapter 2

Evaluating Forecasts

In this chapter the theory behind the evaluation of the forecasts is presented. The forecasts will be evaluated based on the accuracy and the bias of the forecasts. The theory concerning the testing of the forecasts in the surveys is also included.

2.1 Evaluating Forecasts

The forecasts are evaluated by measuring the mean error (ME), the mean absolute error (MAE), the root mean squared error (RMSE) and the absolute error (AE) of the forecasts. The ME, MAE and RMSE are commonly used in the literature, see Batchelor (2001) and Croushore (1993). The performance of the mean, median and individual forecasters will be measured.

The error measures used in this thesis are all based on the forecast error. The forecast error at time t is the difference between the actual and forecasted value of the variable in question, e.g. $A_t - F_t$. The forecast is valid at time t , but the forecast is given at an earlier point of time.

2.1.1 Measure of Forecast Bias

The mean error (ME) is given by

$$ME = \frac{1}{T} \sum_{t=s}^T (A_t - F_t),$$

where T is the total number of observations. The ME measures the mean bias of a forecaster. For a biased forecaster the ME will usually be different from zero. For an unbiased forecaster the ME will be close to zero. However, an ME close to zero is not sufficient to deem a forecaster unbiased. A positive ME indicates that on average the forecaster underestimates the inflation while a negative ME indicates overestimation. The absolute value of the ME is also used to focus on the magnitude of the bias of the forecaster and not whether the forecaster overestimates or underestimates the inflation.

2.1.2 Measures of Forecast Accuracy

The mean absolute error (MAE) is given by

$$MAE = \frac{1}{T} \sum_{t=s}^T |A_t - F_t|.$$

A more accurate forecast will result in a smaller value of the MAE than a less accurate forecast.

The root mean squared error (RMSE) is given by

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=s}^T (A_t - F_t)^2}.$$

A more accurate forecast will result in a smaller value of the RMSE than a less accurate forecast. Compared to the MAE, large forecast errors are penalized more heavily than small forecast errors as the error is squared.

One of the surveys studied in this thesis is the The Surveys of Consumers (Michigan survey). The survey is designed to be a representative sample of the American households and is not concerned with following individuals over time. Each time the survey is conducted about 60% of the panellists are new respondents while the rest are interviewed for a second time, see Curtin (2013b). When there is only one observation in the time-series, the MAE and RMSE will break down to the absolute error (AE)

$$AE = |A_t - F_t|.$$

The individual forecasters in the Michigan survey cannot be studied with time and therefore the AE is the only measure used to analyse the data in this survey. The measure is also calculated for the SPF to be able to compare the accuracy of the mean and median forecasts in the two surveys. A similar situation arises when the model is analysed. Only the predictors in the model, and not the consumers, can be followed with time. Thus the predictors can be analysed with respect to all the error measures described while the consumers can only be analysed with respect to the AE measure.

2.2 Tests of Hypotheses

We want to compare the performance of the mean and median forecaster to the performance of the individual forecasters. In the SPF we can follow the forecasts of an individual with time and thus follow his performance. However, only single forecasts can be studied in the Michigan survey as we cannot follow individuals with time. Therefore we also compare the mean and median forecasts for each quarter to the single forecasts for each quarter by comparing their AE. Tests of hypotheses are used to analyse the significance of the observed differences. The tests are only conducted for the surveys and not for the model. The results in the model are based on large amounts of data and consequently they are only marginally susceptible to inaccuracies. The reference used for the tests is Walpole et al. (2007) unless otherwise stated.

2.2.1 Tests of Forecast Bias

A forecaster is biased if the ME is not close to zero. The ME of the mean, median and individual forecasters are tested to determine if the estimated ME differs significantly from zero. The hypothesis tested is

$$\begin{aligned} H_0 : \mu &= 0 \\ H_1 : \mu &\neq 0, \end{aligned} \tag{2.1}$$

where μ is the estimated ME. We assume the MEs to be normally distributed and the forecasts made by the same forecaster to be independent of each other, hence we can apply the central limit theorem. The assumptions might not hold. Some forecasters only participate in the survey for a short period of time and thus applying the central limit theorem is inaccurate. To help reduce this inaccuracy only forecasters participating in the survey a minimum number of times will be used in this analysis. In addition the forecasts made by the same forecaster are likely not to be completely independent of each other. Due to these limitations the test results will be interpreted with care. The variance is unknown and therefore the test statistic used is Student t-distributed. The reference for this test of bias is Stekler (2002).

The hypothesis (2.1) is tested for each individual forecaster in addition to the mean and median forecasters. Regarding the individual forecasters, we are mainly interested in the combined result, e.g. whether the null hypothesis is rejected for at least one individual. We assume the ME of different individuals to be independent of each other. This assumption is reasonable although individuals might use the same analysis and models to form their forecast. When all the null hypotheses are true and the test statistics are independent, the p-values of each test of bias can be combined to

$$\chi^2 = -2 \sum_{n=1}^N \ln(p_n).$$

The test statistic χ^2 is chi-squared distributed with $2N$ degrees of freedom. The total number of forecasters is N and p_n is the p-value of the hypothesis for the n 'th forecaster. The reference used for combining the tests is Bickel and Doksum (2007).

In addition to test whether the forecasters are biased, we want to test if the individual forecasters are more biased than the mean and median forecasts. The test of the mean forecaster is constructed by finding the absolute value of the ME of the mean and the individual forecasters. We assume the ME of the mean forecaster to be constant and independent of the ME of the individual forecasters. In addition the ME of the individual forecasters are assumed to be independent of each other. These assumptions are not unreasonable as there are numerous

forecasters in the SPF. The hypothesis for comparing the bias of the individual forecasters to the bias of the mean forecaster is

$$\begin{aligned} H_0 : \mu &= \mu_0 \\ H_1 : \mu &> \mu_0. \end{aligned} \tag{2.2}$$

μ_0 is the mean forecaster's absolute value of ME and μ is the the mean of the individual forecasters' absolute value of ME. The test performed compares the bias of the mean forecaster to the mean bias of the individual forecasters. If the null hypothesis is rejected we know that, at the specified level of significance, the mean forecaster is less biased than the mean bias of the individual forecasters. However, it does not imply that the mean forecaster is less biased than all the individual forecasters. In the further discussion a rejection of the null hypothesis will for convenience be formulated as the mean forecaster is less biased than the individual forecasters. The number of periods an individual forecaster participates in the survey varies. Some individual forecasters only participate for one period while others stay in the survey for several decades. The mean forecaster gives a forecast in every period and therefore comparing the MEs directly is inaccurate. To help reduce the inaccuracy, only forecasters participating in the survey a minimum number of times will be used in the analysis and the results will be interpreted with care.

We also want to compare the bias of the median forecaster to the bias of the individual forecasters. We will use a sign test for this purpose, but the assumptions and the subsequent discussion in the previous paragraph is valid for this test as well. The hypothesis tested is

$$\begin{aligned} H_0 : \mu &= \mu_0 \\ H_1 : \mu &> \mu_0, \end{aligned} \tag{2.3}$$

where μ_0 is the absolute value of the median forecaster's ME and μ is the median of the individual forecasters' absolute value of ME. The number of individual forecasts with larger and smaller absolute values of ME are counted. Possible forecasts with equal absolute value of ME as the median forecaster are eliminated from the sample. The test statistic is a binomial random variable, representing

the number of observations with ME larger than μ_0 . The test is a nonparametric test and it does not utilise all the information provided by the data. The test is therefore less efficient and in general more data is needed for possibly rejecting the null hypothesis.

2.2.2 Tests of Forecast Accuracy

In this section we discuss the testing of the error measures MAE, RMSE and AE. These error measures concern the accuracy of the forecasts.

We start by comparing the MAE of the mean forecaster to the MAE of the individual forecasters. The assumptions and the subsequent discussion of hypothesis (2.2) is valid for this test as well. The following hypothesis is tested

$$\begin{aligned} H_0 : \mu &= \mu_0 \\ H_1 : \mu &> \mu_0, \end{aligned} \tag{2.4}$$

where μ_0 is the MAE of the mean forecaster and μ is the mean of the MAE of the individual forecasters. The central limit theorem is applied and the test statistic used is Student t-distributed. Only forecasters participating in the survey a minimum number of times is used in the analysis.

When comparing the MAE of the individual forecasters to the MAE of the median forecaster a sign test will be applied. The assumptions discussed concerning the hypothesis (2.2) is valid for this test as well. The hypothesis tested is

$$\begin{aligned} H_0 : \mu &= \mu_0 \\ H_1 : \mu &> \mu_0, \end{aligned} \tag{2.5}$$

where μ_0 is the MAE of the median forecaster and μ is the MAE of the median of the individual forecasters. The number of forecasts with larger and smaller values of MAE are counted and the test statistic is a binomial random variable as previously described concerning hypothesis (2.3).

Testing the RMSE is similar to testing the MAE. The RMSE of the individual forecasters are compared to the RMSE of the mean forecaster and the hypothesis tested is corresponding to hypothesis (2.4). The RMSE of the individual forecasters

are also compared to the RMSE of the median forecaster and the hypothesis tested is corresponding to hypothesis (2.5).

The AE of the forecasts are also tested, but the approach is different from testing the MAE and RMSE. The approach is different because when we analyse the AE we do not follow individual forecasters with time, but analyse each quarter independently. We assume the AE of the mean forecast to be a constant and independent of the single AEs. The number of forecasts each quarter varies from 9 to 53 in the SPF and consequently assuming the mean forecast to be independent might be inaccurate. The number of observations each quarter in the Michigan survey is about 500 and assuming the mean forecast to be independent of the single forecasts is not unreasonable. The hypothesis for each quarter is equal to the hypothesis (2.4) where μ_0 is the AE of the mean forecast each quarter and μ is the mean of the AE of the single forecasts.

A similar approach is used for the AE of the median forecaster. The AE of single forecasts are compared to the AE of the median forecast at each quarter. The hypothesis at each quarter is equal to (2.5) where μ_0 is the AE of the median forecast and μ is the median AE of the single forecasts.

2.2.3 Correlation of Forecast Errors

The one year ahead forecasts of inflation are obtained from the survey every quarter. The forecasts will overlap as all the forecasts are used when analysing the survey. A shock a specific quarter will affect all the yearly values the quarter is included. Four yearly error terms are dependent on the same quarter and will therefore be affected by the same shock. The error terms for an individual forecaster will be serially correlated for the different quarters. When we test if the forecasters are biased, we will adjust for the serial correlation when the standard errors are calculated. Newey and West (1987) introduce a heteroscedastic and autocorrelated consistent covariance matrix. We will use these standard errors, but in our dataset there are several missing variables and the standard errors applied when testing if the individuals are biased are calculated by the Amplitude Modulated estimator as suggested by Datta and Du (2012). Thus the standard

errors applied when testing if the individual forecasters are biased are Newey-West standard errors adjusted for missing values.¹

Adjusting for correlated standard errors is not necessary when testing and comparing an error measure of the mean or median forecaster to the individual forecasters. The standard error in these tests are concerned with the uncertainty across the individual forecasters, not across time. We assume that the forecasts made by the individual forecasters are independent. Therefore we assume that there is no autocorrelation between the error measures of the individual forecasters.

The figures displaying the forecasts, the realised inflation and the results of the surveys include observations for every quarter. The figures will be interpreted with care as the realised inflations in these figures will be serially correlated.

¹The forecast errors obtained for four consecutive quarters will be serially correlated. Consequently a forecast error a given quarter will be serially correlated with three other forecast errors and the lag used in the calculations of the Newey-West standard errors is three.

Chapter 3

Surveys

In this chapter the surveys are discussed and analysed. The Survey of Professional Forecasters (SPF) is analysed with respect to forecast bias and forecast accuracy. The Surveys of Consumers (Michigan survey) is only analysed with respect to forecast accuracy as the measure of forecast bias is not applicable. At the end of the chapter the analyses of the surveys are compared.

3.1 Survey of Professional Forecasters

The Survey of Professional Forecasters (SPF) is a quarterly survey conducted in the United States by the Federal Reserve Bank of Philadelphia. It contains forecasts of several macroeconomic variables including forecast of inflation. The participants in the survey are limited to those currently generating forecasts for their employers or clients or those who have done so in the past. These limitations are used to ensure the quality and accuracy of the survey. The information and data used in this thesis is found on the web page of the survey, SPF (2013a).

3.1.1 Calculations

The forecasts used in this thesis are the forecasts of the consumer price index. The forecasts are quarterly forecasts, but we are interested in the one year ahead forecasts. Following Stark (2013b), the one year ahead forecast is calculated by

finding the geometrical average given by

$$\pi_{year} = 100 * \left(\sqrt[4]{\left(1 + \frac{\pi_{q1}}{100}\right) \left(1 + \frac{\pi_{q2}}{100}\right) \left(1 + \frac{\pi_{q3}}{100}\right) \left(1 + \frac{\pi_{q4}}{100}\right)} - 1 \right). \quad (3.1)$$

π_{year} is the calculated forecast of the yearly inflation, given in percentages. π_{qn} is the forecast the n 'th quarter after the survey is conducted, also given in percentages. The data used for the realised inflation is obtained from the same data source, the SPF web page SPF (2013a). The realised inflation is also given in quarterly values while we are interested in the yearly values. The procedure described above is applied to obtain the yearly values of the realised inflation. The yearly mean and median forecasts are found by first finding the underlying mean and median forecasts for each quarter before the yearly values are calculated. For convenience, the subscript of the yearly inflation will be dropped.

Missing data is eliminated from the dataset and possible outliers are left untreated. Other methods could have been applied, but the forecasters are professionals and believed to have specific reasons for their forecasts. Using statistical methods to assign forecast values where they are left blank or to normalize extreme values seems unjust and will not reflect the opinions of the panellists. The yearly forecast by a specific panellist will not be calculated if the panellist does not give a forecast for all the quarters in the coming year. In total, 344 yearly forecasts are not calculated because forecasts for at least one quarter is missing. The number of forecasts each quarter varies from 9 to 53 which is a relatively small number considering that the accuracy of the mean and median forecasts might be dependent on the number of forecasts.

The forecasts given in the SPF are coded with an identification number for each panellist. Using this number it is possible to track an individual panellist's replies over time.¹Tracking the replies of an individual forecaster enables calculation of the forecaster's mean error (ME), mean absolute error (MAE) and root mean squared error (RMSE). The error measures are calculated only for individual forecasters responding to the survey a minimum number of times. This restriction is applied to ensure there is enough data to analyse the individual forecasters. Following Clements (2008) the individual forecasters have to participate in the

survey minimum 12 times. After applying this restriction we are left with 101 forecasters. However, when the absolute error (AE) is analysed, the performance of the individual forecasters are not followed with time. Thus the dataset includes all the forecasts and is not restricted to contain only the forecasters responding at least 12 times.

Figure 3.1 displays all the forecasts of the inflation together with the mean and median forecasts and the realised inflation. The overall trend of the mean and median forecasts follow the overall trend of the realised inflation. However, the mean and median forecasts seem to be lagging the realised inflation. Forecasters are not able to foresee a shock so one would expect some lag in the estimates, but a lag over several quarters cannot be explained by this argument. In addition the mean and median forecasters are less volatile than the realised inflation. We also observe that the mean and median forecasts continuously underestimate or overestimate the inflation for consecutive quarters and sometimes years. This suggests that the errors made by the forecasts are correlated. As discussed in chapter 2 some correlation in the errors is expected which can explain some of the features observed in the figure. However, when the forecasts underestimate or overestimate the inflation continuously for more than a year, this cannot be the sole cause of the correlated errors. We also note that there is heterogeneity in the single forecasts. The dispersion of the forecasts persists throughout the entire forecast period. However, there are few extreme outliers observed in the figure and the majority of the forecasts are close to the mean and median forecasts.

¹Two minor issues arise when using the identification numbers in the surveys, see Stark (2013a). First, for some of the older responses the identifier might have been used for different participants. Some of the identification numbers drop out of the survey for a large number of periods before re-entering which suggests that the same identifier might have been used for different participants. Second, when an individual changes employment but remains in the survey, there is a discussion whether the identification number should follow the individual or the firm. The guideline used in this situation is for the identifier to follow the individual if the forecast is mostly associated with the individual, but to follow the firm if the forecast is mostly associated with the firm. Consequently we cannot for certain say that the measures calculated are correct for the individuals, but they are at worst good approximations.

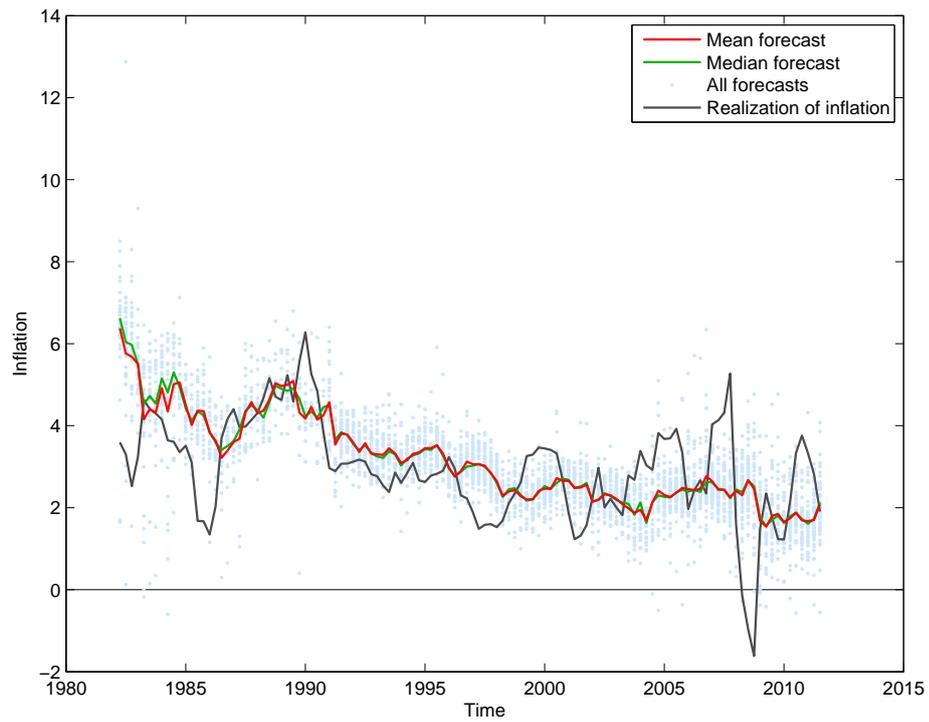


Figure 3.1: The forecasted values of the inflation in the SPF. The mean and median forecasts are also included together with the realised inflation. The values of the mean and median forecasts are similar and consequently overlapping for several periods which makes it difficult to observe both estimates in the plot.

3.1.2 Analysing Forecast Bias

In this section the bias of the forecasters is analysed. The mean error (ME) and the absolute value of this measure is used to analyse the bias.

Analysing ME

The bias of the individual, mean and median forecasters are analysed by analysing the ME of the forecasters. For an unbiased forecaster the ME will be close to zero. The ME of the individual, mean and median forecasters are displayed in figure 3.2(a). Some individual forecasters generally overestimate the inflation and the ME of these individual forecasters are negative. Other individual forecasters generally underestimate the inflation and the ME for these individual forecasters are positive. The ME of some individual forecasters are close to zero and thus these individual forecasters might be unbiased. However, an ME close to zero is not sufficient to deem a forecaster unbiased and therefore we are only deem forecasters biased. The ME of the mean and median forecasters are both slightly negative and thus on average these forecasters slightly overestimate the inflation.

The individual, mean and median forecasters are tested separately for bias. The combined result of the individual forecasters gives a p-value of 0, e.g. at least one individual forecaster is biased. Analysing the tests further reveals that 38% of the individual forecasters are deemed biased at the 5% level of significance. The mean and median forecasters are also tested for bias. The p-value for the mean and median forecasters when testing for bias is 6.4% and 5.3% respectively. Thus strictly applying a significance level of 5% these forecasters are not deemed biased. However, as earlier discussed some of the assumptions might not hold and applying the level of significance strictly seems unwise. Thus we will not deem the predictors biased, but we will keep in mind that using a slightly larger level of significance the predictors would be deemed biased.

The absolute value of the ME of the individual, mean and median forecasters are displayed in figure 3.2(b). The larger the absolute value of the ME, the more biased the forecaster. This measure ignores whether the forecaster overestimates or underestimates the inflation and thus the only focus is the magnitude of the

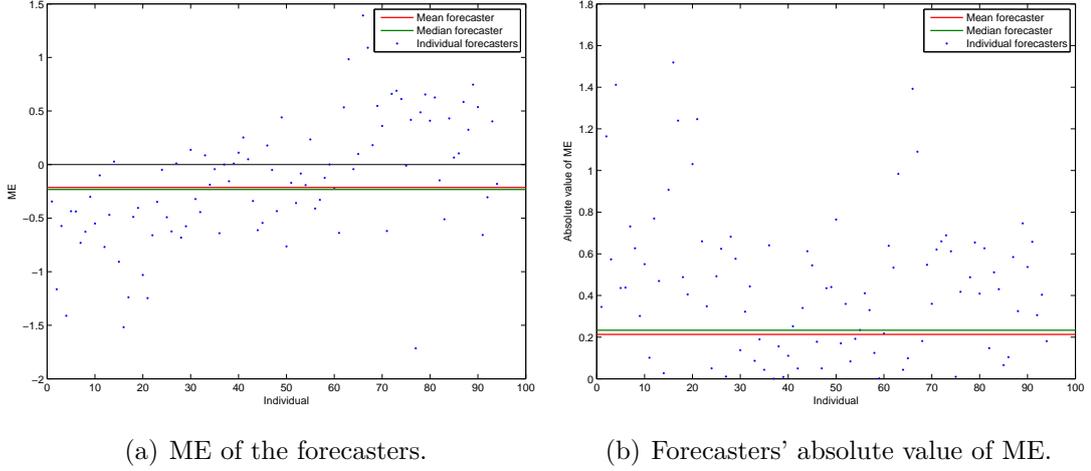


Figure 3.2: ME and absolute value of the ME of the forecasters in the SPF. The values of the mean and median forecasters are similar and therefore overlapping in the figures. The only reason for numbering the individual forecasters is to be able to separate them in the figures.

Compared to	Smaller absolute value of ME	Larger absolute value of ME
Mean	29.8%	70.2%
Median	30.9%	69.1%

Table 3.1: Percentages of forecasters that are more or less biased than the mean and median forecasts according to the absolute value of the ME.

bias. The bias of the individual forecasters are compared to the bias of the mean and median forecasters. Some individuals are less biased while others are more biased than the mean and median forecasters. Table 3.1 presents the share of individual forecasters with smaller and larger absolute values of ME than the mean and median forecasters. 30% of the individual forecasters perform better than the mean and median forecasters while 70% perform worse.

The bias of the individual forecasters are tested and compared to the bias of the mean and median forecasters. The absolute value of the MEs are used in these tests. The p-value obtained when testing the mean forecaster is 0. Thus the mean forecaster is less biased than the individual forecasters. The p-value obtained when

Compared to	Smaller MAE	Larger MAE
Mean	29.8%	70.2%
Median	31.9%	68.1%

Table 3.2: Percentages of forecasters that are more and less accurate than the mean and median forecasts according to the MAE.

testing the median forecaster is $6 \cdot 10^{-5}$, e.g. the p-value is close to 0. Consequently the median forecaster is less also biased than the individual forecasters.

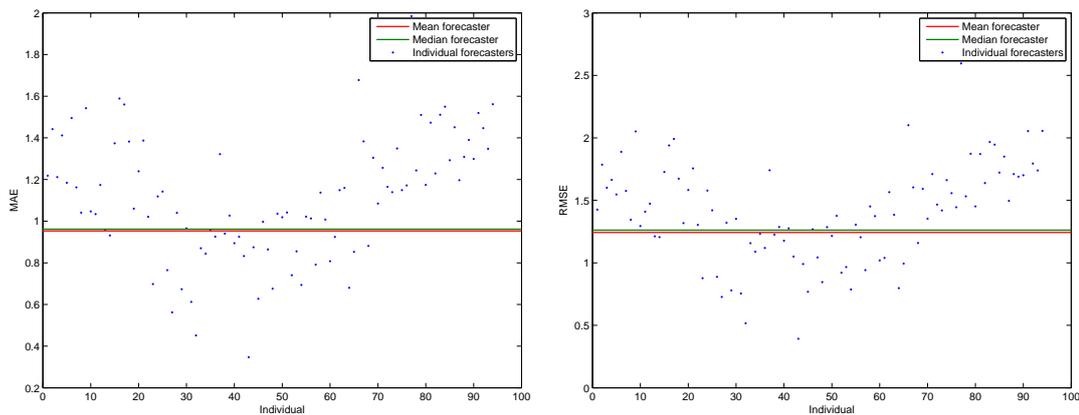
3.1.3 Analysing Forecast Accuracy

In this section the forecast accuracy is analysed. The measures used for this analysis are the mean absolute error (MAE), root mean squared error (RMSE) and absolute error (AE).

Analysing MAE

The MAE of the individual, mean and median forecasters are calculated and displayed in figure 3.3(a). The more accurate the forecaster, the smaller the MAE and thus the best forecasters have the smallest value of MAE. Some forecasters perform better than the mean and median while others perform worse. Table 3.2 shows the percentages of the individual forecasters performing better and worse than the mean and median forecasters by comparing the MAE. According to this measure 30% of the forecasters are more accurate than the mean and median forecasts while 70% are less accurate.

The MAE of the individual forecasters are tested and compared to the MAE of the mean and median forecasters. The p-value obtained when testing the mean forecaster is 0. Thus the mean forecaster is a more accurate forecaster than the individual forecasters according to the MAE measure. The p-value obtained when testing the median forecaster is $1 \cdot 10^{-4}$, e.g. the p-value is close to 0. Thus the median forecaster is a more accurate forecaster than the individual forecasters when testing the MAE of the forecasters.



(a) MAE of the forecasters.

(b) RMSE of the forecasters.

Figure 3.3: MAE and RMSE of the individual forecasters in the SPF. The values of the mean and median forecasts are similar and therefore overlapping in the figure. The individual forecasters are numbered only to be able to separate them in the figures.

Analysing RMSE

The RMSE of the individual, mean and median forecasters are calculated and displayed in figure 3.3(b). The more accurate the forecasters, the smaller the RMSE and thus the best forecasters have the smallest values of the error measure. Some forecasters are more accurate than the mean and median forecasters while others are less accurate. Table 3.3 shows the percentages of forecasters that are more and less accurate than the mean and median forecasters by comparing their RMSE. 33% of the forecasters are more accurate than the mean and median forecasters while 67% are less accurate.

Comparing the MAE results to the RMSE results we notice that a larger share of the forecasters are more accurate than the mean and median forecasters when the RMSE measure is used. The RMSE penalizes larger errors more heavily than the MAE. The results might lead to the conclusion that the mean and median forecasters make larger errors when they make an error compared to the individual forecasters. However, this conclusion seems unreasonable when observing the behaviour of the mean and median forecasters. A possible explanation for this

Compared to	Smaller RMSE	Larger RMSE
Mean	33.0%	67.0%
Median	33.0%	67.0%

Table 3.3: Percentages of forecasters that are more and less accurate than the mean and median forecasts according to the RMSE.

curious result is that the individual forecasters do not have to participate in every quarter. Some individual forecasters are likely to participate in the survey only for a period in which inflation is relatively easy to forecast. On the other hand the values of the mean and median forecasters are calculated for every quarter and thus periods when inflation is more difficult to forecast are included.

The RMSE of the mean and median forecasters are tested and compared to the RMSE of the individual forecasters. The p-value when testing the mean forecaster is 0. Thus the mean forecaster is determined to be significantly more accurate than the individual forecasters according to the RMSE. The p-value obtained when testing the median forecaster is $3 \cdot 10^{-4}$, e.g. the p-value is close to 0. Thus the median forecaster is deemed to be significantly more accurate than the individual forecasters when analysing the RMSE.

Analysing AE

The forecasts are analysed separately for each quarter when using the AE measure, e.g. we do not follow individual forecasters with time. The main reason for performing this analysis for the SPF is to be able to compare the results to the Michigan survey and to the theoretical model. Individual forecasters cannot be followed in the Michigan survey and thus the AE is the only measure used to analyse the Michigan survey. We will discuss the Michigan survey in section 3.2. When analysing the theoretical model we are only able to follow the predictors with time. We are not able to follow individual forecasters in with time and thus analysing the AE is valuable for comparing the surveys to the model. The model will be presented in chapter 4 and analysed in chapter 5.

The AE is calculated for single forecasts as well as the mean and median

Compared to	Smaller AE	Larger AE
Mean	43.0%	57.0%
Median	43.4%	56.5%

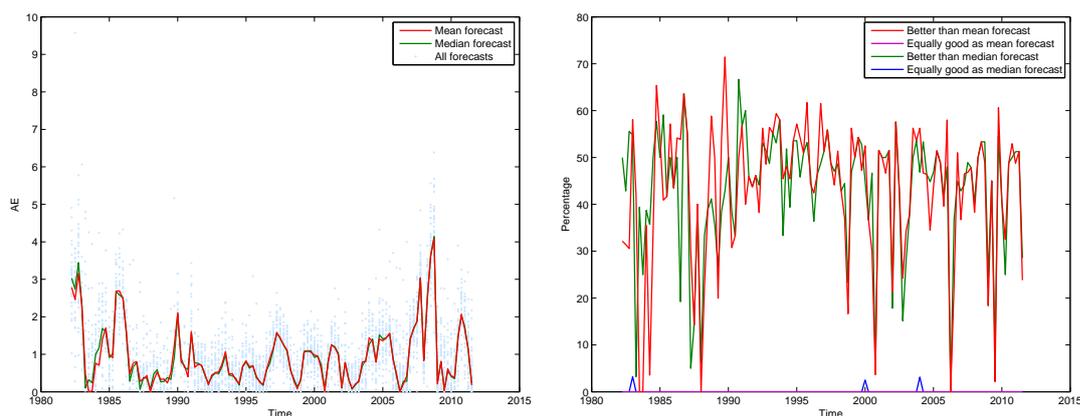
Table 3.4: Percentages of forecasts performing better, equal or worse than the mean and median forecasts according to the AE.

forecasts. A small value of the AE indicates that the forecast is accurate while a large value indicates that the forecast is inaccurate. Comparing the AE of the single forecasts to the AE of the mean and median forecasts for the same quarter gives the relative performance of the mean and median forecasts. By repeating the procedure for the whole survey we find how well the mean and median forecasts perform. The overall results are summarized in table 3.4. 43% of the forecasts performed better than the mean and median forecasts and 57% performed worse.²

To further analyse the performance of the mean and median forecasts, the AE of the single, mean and median forecasts are plotted as a function of time in figure 3.4(a). The AE of the mean and median forecasts fluctuates which means that the accuracy of the estimates varies with time. Comparing the AE of the single forecasts to the AE of the mean and median forecasts for each quarter results in figure 3.4(b). There is little stability in the relative performance of the mean and median forecasts. Sometimes they are more accurate than most of the single forecasts while other times they are less accurate.

The AE of the single forecasts are tested and compared to the AE of the mean and median forecasts. The tests are performed each quarter. Analysing

²There were not expected to be any forecasts equal to the mean and thus performing equally good as the mean. However, approximately nobody performing equally good as the median is not an expected result. The median forecast is calculated using the median forecasts for each quarter before calculating the yearly value as shown in equation (3.1). If a panellist forecasts the median each quarter he will also forecast the yearly median forecast. However, he might for example forecast the median value in one quarter and some other values for the other quarters. These forecasts are highly unlikely to produce the yearly median value. Consequently approximately no panellists forecast the median yearly values and approximately no panellists end up with the same AE as the median forecast.



(a) AE of the forecasts.

(b) Comparing the AE of the single forecasts to the mean and median forecast.

Figure 3.4: AE of the forecasts in the SPF are calculated and plotted as a function of time. The AE of the single forecasts are compared to the AE of the mean and median forecasts. Percentages of the single forecasts that are more and equally accurate as the mean and median forecasts are displayed.

the tests we find how many percentages of the quarters the mean and median forecasts perform better than the single forecasts. The level of significance used in this analysis is 5%. In 83% of the quarters the mean forecasts are deemed more accurate than the single forecasts. The corresponding percentage for the median forecasts is 20%. The test applied when the median forecasts are analysed is a sign test. A sign test is a less efficient test and rejecting the null hypothesis based on this test will in general require a larger sample size. For comparison the sign test is also applied for the mean forecasts. The result is that 27% of the quarters the mean forecasts are more accurate than the single forecasts. Thus it is likely that some of the seemingly large difference in the performance between the mean and median forecasts can be explained by the test used.

3.2 Michigan Survey

The Surveys of Consumers (Michigan survey) is conducted by the Michigan University each month. The survey concerns a variety of economical variables includ-

ing the forecast of the one year ahead inflation. The respondents are statistically designed to be representative for all American households, excluding Alaska and Hawaii. Minimum 500 interviews are conducted each month, see MichiganSurvey (2013b). The information and data used in this analysis is found on the web page of the Michigan survey, MichiganSurvey (2013a).

3.2.1 Calculations

In this section we will discuss the dataset used in the analysis of the Michigan survey focusing mainly on missing data and outliers. Different coding is used for different types of missing data allowing it to be treated differently depending on the type of missing data. The procedures used are the same as described in Curtin (2013a). The data used in this thesis is the one year ahead inflation expectation.

Some of the responses in the survey are coded as "not available". Mostly this reflects the respondent's refusal to answer the question. The interview is voluntary and the respondents are told so at the beginning of the interview. In addition the respondents are told that they can skip any question if they prefer to do so. The dataset used in this thesis contains less than 0.13% responses listed as not available. These responses are eliminated from the dataset.

Some respondents reply that they do not know whether prices will go up or down during the next 12 months. Less than 0.91% of the respondents give this reply and these observations are eliminated from the dataset.

Some respondents reply that they think prices will go up, but that they do not know by how much. Similarly some respondents reply that they think prices will go down, but that they do not know by how much. These observations are coded as either "don't know by how much, but prices will go up" (DK UP) or "don't know by how much, but prices will go down" (DK DW). The share of the respondents replying DK DW in the dataset is 0.48%. The share of the respondents replying DK UP is 7.0% and the share varies between 1.6% and 23.0% for different quarters. This information is not missing, but incomplete, and we should use this partial information when calculating the mean and median forecasts. If we eliminate these

observations the estimates of the mean and median will be biased.³

The responses DK UP and DK DW are replaced by a distribution of responses instead of using a point estimate which is commonly used to treat missing and incomplete data. The distribution of the respondents who reply that prices will increase and by how much is found for each quarter. The respondents who reply DK UP are distributed similarly as the ones with complete information. To increase the accuracy the respondents are split into fractions whenever appropriate. E.g. let 10% of the cases with complete information for a particular quarter forecast prices to increase by 2%. If there are 14 respondents who reply DK UP, 1.4% of these respondents are estimated to forecast prices to increase by 2%. The same approach is applied to find estimates for the DK DW replies.

The estimated mean is unaffected by whether a point estimate or a distribution is used when handling the partial information DK UP and DK DW. The estimate of the median is however improved by using a distribution instead of a point estimate. The point estimate used if the median were to be calculated would have been the median of the forecasts with increasing prices for the DK UP replies. The median of the forecasts of increasing prices is an upper bound for the median of the whole sample at that time and consequently it is biased upwards. Whenever this point estimate is a part of the calculated median of the whole dataset at that time, the estimate might be biased. Using the point estimate would also imply that all DK UP cases contributes to increased median. The median of the known increasing prices is always larger than the median of the sample as a whole. Whereas when using a distribution, some of the values imputed for the DK UP cases have a smaller value than the overall median. In addition these problems are increased because the survey only records integer values. The corresponding

³To easier explain why the mean and median estimates will be biased we introduce two variables, KNOW UP and KNOW DW. Respondents replying they expect prices to increase and are able to quantify the price increase are given the code KNOW UP. Similarly the respondents replying they expect prices to decrease and are able to quantify the price decline are given the code KNOW DW. The ratio of those responding DK UP to those responding DK DW does not equal the ratio of those responding KNOW UP to those responding KNOW DW. Thus the estimates will be biased if the data is eliminated.

arguments can be made for the DK DW replies, but using a distribution instead of a point estimate will not have the same effect because the median forecaster never predicts prices to decrease. Nevertheless, for consistency, the same approach is used for the DK DW replies.

Following Curtin (2013a), replies smaller than -10% or greater than $+50\%$ are assumed to be outliers. 0.36% of the recordings were extreme negative values and were assumed to be outliers. There were no extreme positive values which might suggest that the dataset was already processed for positive outliers. The outliers are assumed to contain valid information and are therefore not eliminated, but truncated. This treatment ensures the valid information to be contained, but limits the effect extreme values might have. Truncation of the outliers has no effect on the median estimate, but will most likely improve the mean estimate.

Minimum 500 interviews are conducted each month, but due to missing data a smaller number might be the basis for the calculations in this thesis. About 650 replies are used in most of the 1980's, but since the end of that decade the number has decreased and stayed at about 450 replies. The results will be compared to the results for the SPF and consequently only quarterly data is used. The responses obtained in the last month of each quarter is used in the analysis.

Figure 3.5 displays all the forecasts of the inflation together with the mean and median forecasts and the realised inflation. The forecasts displayed in the figure are those obtained after processing the dataset as described in this section. The dispersion of the single forecasts is large even though the extreme values have been truncated. The heterogeneity in the forecasts persists throughout the forecast period. Even though the dispersion is large, the mean and median forecasts follow the overall trend of the realised inflation. The mean and median forecasts rarely take the same value. This is partly because the median forecasts only take integer values while the mean forecasts can take any decimal number. Whenever the difference between the mean and median forecasts is larger than one it is not only because the median forecasts take an integer value, but because of the different methods of estimation. In addition the mean and median forecasts seem to be lagging the realised inflation. The mean and median forecasts also seem to

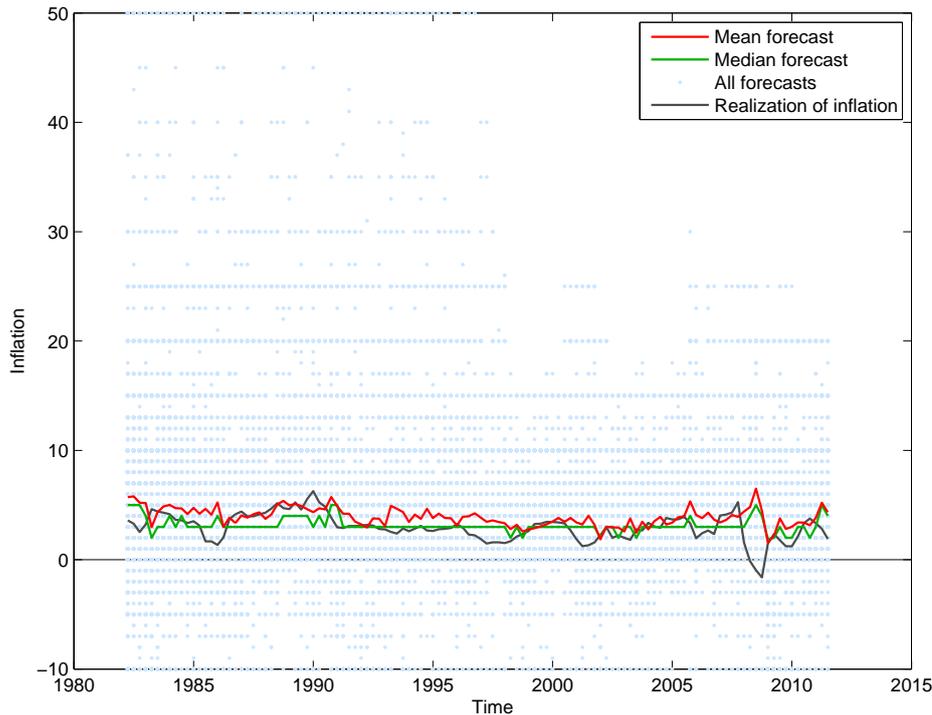


Figure 3.5: The forecasted values of the inflation in the Michigan survey. The mean and median forecasts are also included together with the realised inflation.

overestimate and underestimate the inflation for several consecutive quarters and sometimes years. This effect might partly be caused by the correlated forecast errors discussed earlier, but when the effect lasts for more than a year this cannot be the only explanation.

3.2.2 Analysing Forecast Accuracy

The forecast accuracy is analysed in this section. We cannot follow individual forecasters with time in the Michigan survey. Consequently the only error measure calculated for this survey is the absolute error (AE).

Compared to	Smaller AE	Equal AE	Larger AE
Mean	26.5%	0%	73.5%
Median	13.0%	16.4%	70.6%

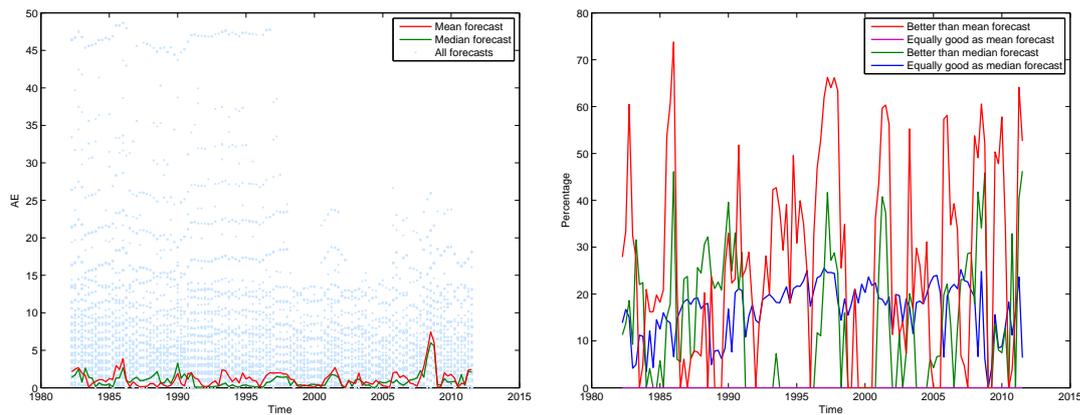
Table 3.5: Percentages of the single forecasts that are more, equally and less accurate than the mean and median forecasts according to the AE.

Analysing AE

The AE of the single forecasts and the AE of the mean and median forecasts are calculated for each quarter. An accurate forecast results in a smaller value of AE than a less accurate forecast. The AE of each single forecast is compared to the AE of the mean and median forecasts for the same quarter to measure the relative accuracy of the mean and median forecasts. The overall shares of the single forecasts that are more, equally and less accurate than the mean and median forecast can be found in table 3.5. The mean and median forecasts perform better than 70 – 75% of the single forecasts. The estimates of the mean and median forecasts are based on the processed dataset. However, when comparing the performance of the mean and median forecasts to the single forecasts, the DK UP and DK DW replies are not included. These panellists were not able to quantify their forecast and thus to compare these replies to the mean and median forecast would be imprecise. We also note that the AE of several forecasts are equal to the AE of the median forecasts. This will occur whenever a single forecast is equal to the median forecast.⁴

To further analyse the performance of the mean and median estimates the AE of the forecasts are plotted as a function of time in figure 3.6(a). The accuracy of the mean and median forecasts varies with time and is particularly poor in the beginning of the financial crisis that started in 2008. The comparison of the mean and median forecasts to the single forecasts is also plotted as a function of time. The results can be found in figure 3.6(b) and it is clear that the relative performance of the mean and median forecasts vary with time. E.g. sometimes

⁴For comparison the estimates of the mean and median are also calculated based on the preprocessed dataset, see appendix A.



(a) AE of the forecasts.

(b) Comparing the AE of the single forecasts to the mean and median forecast.

Figure 3.6: AE of the forecasts in the Michigan survey are plotted as a function of time. The AE of the single forecasts are also compared to the AE of the mean and median forecasts. Percentages of the single forecasts that are more and equally accurate as the mean and median forecasts are displayed.

the mean and median forecasts are more accurate than most of the single forecasts while other times they are less accurate than most of the single forecast.

The accuracy of the single forecasts are tested and compared to the accuracy of the mean and median forecasts. The tests are performed for each quarter. We analyse the tests by finding how many quarters the mean and median forecasts are deemed more accurate than the single forecasts. The result is given in percentages of the total number of quarters analysed. The level of significance used is 5%. The mean forecasts are found to be more accurate than the single forecasts in 78% of the quarters. When analysing the median forecasts a sign test is used. A sign test is less efficient and generally more data is needed to reject the null hypothesis. However, the median forecasts are found to be more accurate than the single forecasts in 94% of the quarters. This might indicate that in the Michigan survey the median forecasts are more accurate when forecasting the inflation than the mean forecasts. Using a sign test when analysing the mean forecasts, the corresponding number is 76%. The difference in performance might be caused by the extreme forecasts made by some panellists in the Michigan survey. These

responses affect the mean forecasts more than the median forecasts and might contribute to the mean forecasts performing poorer than the median forecasts.

3.3 Comparing the Surveys

The two economical surveys analysed in this thesis both include the one year ahead forecast of inflation. However, there are significant differences between the surveys. The SPF is conducted among professional forecasters while the Michigan survey is designed to represent the American households. In addition the Michigan survey includes about a tenfold number of respondents compared to the SPF.

The surveys are evaluated over a time span where both surveys have adequate forecasts and where the inflation has been realised. The time period used is the 1st quarter 1982 to the 2nd quarter 2012. Both the mean and the median forecasts are included in the analysis as neither is significantly superior to the other, see Curtin (2013a). The SPF is conducted quarterly while the Michigan survey is conducted monthly. For easy comparison of the surveys only quarterly data from the Michigan survey is used. The data obtained in the last month of each quarter in the Michigan survey is used in order to match the forecasting periods for the two surveys. Records show that since the 2nd quarter of 1990 the panellists in the SPF report their forecasts in the middle of each quarter, see SPF (2013b). The panellists in the Michigan survey gain an advantage by the late reporting because they have more information when forecasting the inflation.

There is heterogeneity in the forecasts in the surveys and the heterogeneity persists throughout the forecast period. There is less diversification in the SPF forecasts than the Michigan forecasts. This is to be expected as professionals have more knowledge and understanding of the economical situation, enabling them to make better forecasts. However, the diversification in the Michigan forecasts is surprisingly large as several forecasts are outside the interval -5% to 20% . Comparing the mean and median forecasts in the two surveys, it is apparent that their values are more similar to each other in the SPF than in the Michigan survey. Only integer values are recorded in the Michigan survey which contributes to the

difference between the mean and median forecasts. In addition the mean forecasts are more affected by the large positive outliers and consequently the forecasts in general seem to predict a higher level of inflation than the median forecasts.

Many assumptions are made when analysing the error measures and performing the tests. These assumptions are described in section 2.2. As previously discussed some of these assumptions are likely to be inaccurate and thus the results discussed in this chapter might be inaccurate. On the other hand the assumptions enables us to perform statistical inference with reasonable accuracy, but we will keep in mind the discussion concerning the assumptions.

The bias of the forecasters is analysed in the SPF. However, the bias of the forecasters in the Michigan survey is not analysed as we cannot follow individual forecasters with time. Some of the individual forecasters in the SPF are deemed biased while some are not deemed biased. Neither the mean or the median forecasters are deemed biased, but the p-values of the tests are small and adjusting the level of significance only a couple of percentages will alter the results. The ME of the mean and median forecasters in the SPF are -0.21 and -0.23 respectively. The values show that the ME of a forecaster is not necessarily zero even though the forecaster is not deemed biased. Consequently the forecaster might be slightly biased, but not enough to be deemed biased by the statistical test. The ME of the mean and median forecasters are negative and consequently they overestimate the inflation.

The accuracy of the forecasters is also analysed. As the individual forecasters can be followed with time in the SPF both the MAE and RMSE measures are analysed in addition to the AE. The mean and median forecasters are more accurate than the individual forecasters according to the MAE and RMSE measures. The error measures are not calculated for the Michigan survey as we cannot follow the individual forecasters with time in this survey.

The AE of the forecasts are analysed for both the SPF and the Michigan survey. The relative accuracy of the mean and median forecasts can be compared to each by using this measure. The AE of the forecasts in the Michigan survey are generally larger than the AE of the forecasts in the SPF because the diversification in the

Michigan survey is larger. This is to be expected as the forecasters in the SPF are professionals. The mean and median forecasts in the Michigan survey perform well compared to the single forecasts in the survey. The similar measures for the SPF do not perform equally well. This might be caused by the professionals in the SPF performing well and hence the mean and median forecasts are less likely to perform better than the single forecasts. Further adding to this the single forecasts in the Michigan survey are expected to be less accurate as they are forecasted by consumers. In addition the number of forecasts might influence the accuracy of the mean and median forecasts. The mean and median forecasts are believed to increase in accuracy with increased number of panellists. There are more panellists in the Michigan survey than in the SPF. Consequently the large number of panellists might contribute to the good performance of the mean and median forecasts in the Michigan survey relative to the performance of the single forecasts.

To sum up, the mean and median forecasts in the surveys perform well. Specifically the mean and median forecasters are less biased and more accurate than the individual forecasters. The mean and median forecasts in the Michigan survey performs particularly well compared to the single forecasts in the same survey. In addition we observe that the mean and median forecasters appear to be sluggish compared to the realised inflation.

Chapter 4

The Model

In this chapter the theoretical model is presented. The model is a simple frictionless dynamic stochastic general equilibrium model. The main focus of the model is on heterogeneous predictors applied by the consumer. The model will not be solved analytically, but will be analysed by performing simulations in the next chapter.

4.1 The model

The model used in this thesis is among others described in Woodford (2003) and Anufriev et al. (2012). There are no frictions in the economy, neither in the goods nor in the financial markets. The markets in the economy are perfectly competitive and the prices adjust continuously so the markets clear. There is no cash in the economy because there is no need for cash as there is no friction in the financial markets. Only one period risk free bonds can be traded. There are no bonds issued by the fiscal government, but there are privately issued bonds. The monetary authority issues base money which is the monetary liabilities of the central bank. There are no arbitrage opportunities in the economy and thus riskless nominal assets with similar short maturity are perfect substitutes whether they are issued by private parties or by the central bank. The central bank can influence the prices in the economy by adjusting the interest rates paid on the liabilities of the central bank. However, as this is an endowment economy, the

central bank cannot influence the real economic activity.

Consumers in the economy maximize their expected present value of total utility

$$\max \hat{E}_s \left[\sum_{t=s}^{\infty} \beta^t u(C_t) \right]. \quad (4.1)$$

\hat{E}_s is the subjective expectation operator of consumers in period s . β is the discount factor and takes a value in the interval $(0, 1)$. C_t is the consumption of the economy's single good in period t . u is the utility of the consumer and dependent only on the consumption C . u is assumed to be concave and strictly increasing in C .

Consumers face the budget constraint

$$B_t + P_t C_t \leq (1 + i_{t-1})B_{t-1} + P_t Y_t$$

each period. B_t is the sum of the household's bonds and base money, measured at the end of period t . P_t is the price of the economy's good in period t and i_t is the interest rate. C_t and Y_t is the consumption and exogenous endowment of the single good.

The transversality condition is introduced to ensure maximization of utility. If consumers accumulate wealth in the limit, they cannot be maximising consumption as they can buy more goods and increase their consumption. The transversality condition is

$$\lim_{t \rightarrow \infty} \hat{E}_s \beta^t B_t = 0.$$

The maximisation problem (4.1) is solved and the first order conditions of the optimisation problem gives

$$u'(C_t) = \beta(1 + i_t) \hat{E}_t \left[u'(C_{t+1}) \frac{P_t}{P_{t+1}} \right],$$

which is the Euler equation describing the optimal consumption. Let the endowment each period be constant, e.g. $Y_t = Y$. The markets in the economy clears and therefore $C_t = Y$. Combining this with the Euler equation, we are left with

$$\frac{1}{(1 + i_t)} = \beta \hat{E}_t \left[\frac{P_t}{P_{t+1}} \right],$$

which is known as a Fisher equation. Assuming the steady-state inflation to be zero, log-linearisation gives

$$i_t = r_t + \hat{E}_t \pi_{t+1}. \quad (4.2)$$

r_t is the real interest rate and π_t is the inflation at time t .

The central bank in the economy is assumed to follow the interest rate rule

$$i_t = r_t + \varphi \pi_t + u_t. \quad (4.3)$$

The central bank follows the Taylor principle by letting $\varphi = 1.5$ as proposed in Taylor (1993). By following this interest rate rule, the central bank is responding more than one-to-one to changes in inflation and avoids accumulation. The target inflation is implicitly assumed to be 0 when the central bank follows this interest rate rule.

u_t is a shock caused by the central bank behaving differently than expected. It might be caused by exogenous shifts in the central banks reaction to the economic environment or by problems regarding the measurement of inflation. We will consider two types of shocks in the model. One of the shocks we will study is a white noise drawn from the normal distribution, e.g. $u_t = \varepsilon_t$ where $\varepsilon \sim \text{i.i.d } N(0, \sigma_\varepsilon^2)$. The main focus of the analysis will be on the model when the shock is a white noise. To introduce persistence in the shock an autoregressive model of order 1 (AR1) is used, e.g. $u_t = \rho u_{t-1} + \varepsilon_t$ where $\varepsilon \sim \text{i.i.d } N(0, \sigma_\varepsilon^2)$. The extent of persistence in the model is dependent on the value of the parameter ρ . A high parameter value results in a highly persistent process while a low parameter value results in a less persistent process. Setting $\rho = 0$ reduces the AR1 process to a white noise.

Inserting the central bank's interest rate rule, equation (4.3), into equation (4.2) gives

$$\pi_t = \frac{1}{\varphi} \tilde{E}_t \pi_{t+1} + u_t \quad (4.4)$$

where \tilde{E}_t is the aggregate expectation in the economy. This expectation is the weighted sum of the subjective expectations used in the economy, see Branch and McGough (2009) and Anufriev et al. (2012).

4.1.1 Predictors

Consumers in the economy have heterogeneous expectations and they choose between using the biased and the rational predictor. The biased predictor forecasts a constant level of inflation, c , and is not affected by the evolution of the inflation. The rational predictor is based on the assumption that it is the only predictor in the economy. The predictor forecasts the rational inflation based on this assumption. It is not rational within the model because it fails to recognise the existence of the biased predictor in the economy. When the shock is a white noise, the rational predictor at time t is $E_t^R(\pi_{t+1}) = 0$. If the shock is an AR1 process, the rational predictor at time t is $E_t^R(\pi_{t+1}) = \frac{\varphi\rho^2}{\varphi-\rho}u_{t-1}$.

For each time step the share of the consumers using the two predictors are updated based on the past performance as described among others in Brock and Hommes (1997) and Lines and Weserhoff (2010). The shares of the biased and rational predictors at time t are

$$n_t^B = \frac{\exp(\lambda a_t^B)}{\exp(\lambda a_t^B) + \exp(\lambda a_t^R)} = \frac{1}{1 + \exp[\lambda(a_t^R - a_t^B)]} \quad (4.5)$$

$$n_t^R = 1 - n_t^B.$$

The parameter $\lambda \geq 0$ measures how sensitive consumers are to using the most attractive predictor. If $\lambda = 0$, the fractions are $n_t^B = n_t^R = 0.5$ as the consumers do not differentiate between the predictors. In the limit where $\lambda \rightarrow +\infty$ all consumers in the economy will use the most attractive predictor. Following Lines and Weserhoff (2010) we will give an example of the effect of the sensitivity parameter. We will analyse the effect of the parameter based on a model with two predictors, A and B . Let the prediction error of predictor A be 1 percentage point. The prediction error of predictor B is 1% less than the error of A , e.g. the error of predictor B is 0.99 percentage points. Given a parameter value of 1, 49.50% of the population will choose predictor A while 50.50% of the population will choose predictor B . For comparison, let the value of the parameter be 10 and let the predictors make the same mistakes. Using these assumptions, 45.04% of the population will choose predictor A while 54.96% of the population will choose predictor

B. The share of the population choosing the predictor with the smallest error has increased because the propensity to use the best predictor has increased.

The attractiveness of the predictors in equation (4.5) are given by

$$a_t^B = - \left[\pi_{t-1} - E_{t-2}^B(\pi_{t-1}) \right]^2 = - [\pi_{t-1} - c]^2$$

$$a_t^R = - \left[\pi_{t-1} - E_{t-2}^R(\pi_{t-1}) \right]^2 .$$

a_t^B is the attractiveness of the biased predictor at time t and E_{t-2}^B is the biased predictor at time $t - 2$ which is always equal to the constant c . The corresponding variables are used for the rational predictor. The most attractive predictor will be used by the biggest share of the consumers. The maximum attractiveness of a predictor is 0 which occurs when the predictor has made no forecasting error the previous time step. All other predictions will cause the attractiveness to be negative.

It is important to be aware of the time lag between the evaluation and the use of the predictors. The attractiveness of the predictors in period t are based on the performance of the predictors in period $t - 1$. The attractiveness determines the fractions using the two predictors in period t which again determines the expected inflation in the economy in period $t + 1$. The expected inflation in period $t + 1$ influences the actual inflation in period t .

Chapter 5

Simulation

In this chapter the model is simulated and analysed. The main focus of the analysis is on the performance of the predictors in the model. The predictors are analysed both with respect to forecast bias and forecast accuracy.

5.1 Approach for Simulating the Model

The values of the parameters used for simulation purposes are based on Anufriev et al. (2012). The monetary policy in the model is expressed by the parameter φ in equation (4.4). Using a value less than 1 will lead to a cumulation process and will contradict the Taylor rule. To avoid an unstable situation, the value of the monetary parameter is set to 1.5. The shock used in the standard model is a white noise and the standard deviation (σ_ε) of the white noise is 0.5. The sensitivity parameter (β) measures how fast consumers switch between predictors. The sensitivity parameter in the model is set to 1. The biased forecaster will constantly predict the inflation to be 1%. The rational predictor will as discussed in section 4.1.1 constantly predict the inflation to be 0. There is no cost using either predictor. The inflation is initially set to 0 and the forecasts are initially used by equal shares of the population e.g. 50% of the consumers use each forecast.

The main focus of the analysis is on the mean predictor. The mean predictor is the weighted average of the predictors used by the consumers in the model. E.g.

the mean predictor is the weighted average of the rational and biased predictor in the model. There will be less focus on the median predictor in the analysis of the model. The median predictor in the model will be equal to either the biased or the rational predictor. This is not a good predictor when comparing it to the real world where there are several different forecasts and the median often is quite similar to the mean predictor. There will be little emphasis on the median predictor in this chapter even though it was analysed in chapter 3 concerning the surveys. The predictor will not be displayed in the figures to easier visualise the results for the mean predictor. However, the error measures of the median predictor are calculated to evaluate the performance of the predictor and for the completeness of the analysis.

Several measures are estimated in this chapter to evaluate the model. To ensure accurate information 10000 series with 10000 time steps are simulated. A computer is not able to produce completely random numbers, but calculates the pseudorandom numbers based on an algorithm. The algorithm is dependent on an initial value, called the seed. Different seeds will produce different series of random variables. To avoid differences in the results due to the seed and not the parameter values, the same seed is used at the very beginning of all simulations. The same seed is also used for the figures and consequently they can be compared to each other with a meaningful result. The differences observed are only caused by the different parameter values and not by different random numbers.

5.2 The Standard Model

The model is simulated using the standard parameter settings. The results of a particular simulation can be found in figure 5.1(a). The mean predictor is lagging the realised inflation. This is caused by the lagging behaviour of the predictors as discussed in section 4.1.1. Let the forecasts be valid in time period $t+1$. The shares in period t determines the forecasts in period $t+1$ and they are again determined by the performance of the predictors in time $t-1$. Thus there is a time lag of two steps between the valid period of the forecast and the performance of the predictors

which the forecast is based up on. The lagging behaviour of the mean predictor sometimes leads to large prediction errors. In particular this error becomes large when the inflation changes from one extreme value to another extreme value with opposite sign. In these periods the mean predictor is increasing or decreasing while the realisation of the inflation is moving in the opposite direction. The lagging behaviour of the mean predictor is also observed when analysing the SPF and Michigan survey in chapter 3.

The mean predictor always takes a value in the interval made between the two predictors, with 0 and 1 as limits. There is however no upper or lower limit for the realised inflation. Consequently the mean predictor is not able to forecast extreme values of inflation when the inflation exceeds the interval and extreme values occur. Given the standard parameters, the variance of the mean forecast is estimated to 0.04 while the variance of the realised inflation is estimated to 0.27. The smaller variance of the mean predictor is caused by the limits of the predictor while there are no limits for the realised inflation. This feature is also observed and commented when the SPF is analysed in section 3.1. The variance of the mean and median forecasts is smaller than the variance of the realised inflation.

The model is analysed with respect to the heterogeneity in the forecasts. We are interested in finding if both the biased and the rational predictor persist in the long run. The model is simulated with time-series consistent of 10000 time steps. Neither predictor dies away thus the heterogeneity in the model is persistent in the long run.

The percentages of the population using the two predictors corresponding to figure 5.1(a), are displayed in figure 5.1(b). The share using a specific predictor varies considerably as can be seen in the figure. However, for the particular sequence shown in the figures, the consumers are never using just one of the predictors. The consumers will use only one of the predictors whenever the forecast error made by one of the predictors is large enough for the software to interpret the share of the population using the predictors as 0 and 1. By comparing the two figures, we can observe that the best forecast is used by the biggest share of the population. This is in line with the earlier discussion in chapter 4 when we

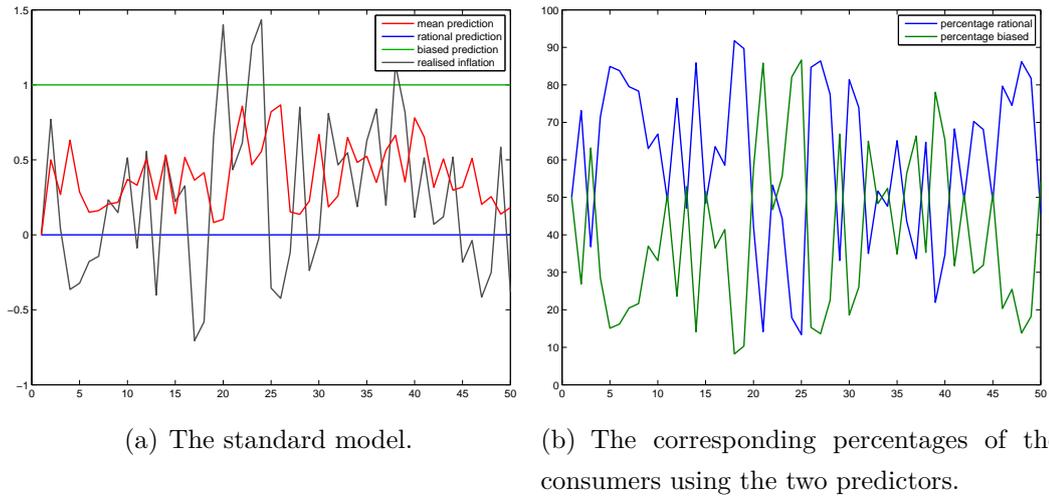


Figure 5.1: The model is simulated with the standard parameter settings. The predictors and the realised inflation are displayed as a function of time. The corresponding percentages of the consumers using the two predictors are also displayed as a function of time.

presented the theoretical model. However, there is a time lag between the evaluation and the use of the predictors as earlier discussed. The variation in the shares also affects the realised inflation because the expected inflation in the economy changes whenever the shares change. The mean predictor is a weighted average of the two predictors in the economy. The lagging behaviour of the shares in the model results in the mean predictor lagging the realised inflation. The mean forecasters in the survey analysed in chapter 3 were also found to be lagging the realised inflation. In addition we note that the most attractive predictor has the heaviest weight and consequently the value of the mean predictor is closest to the best predictor. This is an advantageous feature of the mean predictor contributing to good performance.

5.2.1 Analysing Forecast Bias

In this section the forecast bias of the predictors is analysed. The mean error (ME) and the absolute value of this error measure is used to analyse the bias.

Analysing ME

The results of the ME analysis are displayed in table 5.1. The median predictor is the least biased predictor as the absolute value of the ME is the smallest of all the predictors. The predictor slightly overestimates the inflation as the ME is slightly negative. As earlier described this predictor is not a good model for the median forecasters analysed in the surveys. However, the median predictor in the model is able to switch between the rational and biased predictor in a manner that makes it only slightly biased. The mean predictor is the second least biased predictor according to this measure. The predictor is somewhat biased and overestimates the inflation as the ME is negative. The rational predictor is biased and the predictor underestimates the inflation as the ME is positive. The predictor does not realise the existence of the biased predictor and thus it fails to recognise the upward pressure on the inflation from the biased predictor. The predictor called the biased predictor is the most biased predictor in the model. The predictor consequently overestimates the inflation as the ME of the predictor is negative.

The mean and median forecasters in the SPF are not deemed biased by the tests performed. However, adjusting the level of significance only a couple of percentages will lead the mean and median forecasters in the survey to be deemed biased. Both the forecasters in the SPF slightly overestimate the inflation as the ME of the forecasters are slightly negative. These results are reproduced in the model. The mean and median predictor overestimate the inflation. In addition the mean and median forecasters are slightly biased.

The bias of the mean and median forecasters in the SPF are compared to the bias of the individual forecasters. The mean and median forecasters are found to be less biased than the individual forecasters. These results are reproduced in the model in the sense that the mean and median forecasters are the two least biased predictors. However, we are not able to follow individual forecasters in the model with time as they switch between the rational and biased predictor. Thus the results are reproduced in the sense that the mean and median forecasters are less biased than individual forecasters following either the rational or the biased predictor closely.

ME of mean	ME of median	ME of rational	ME of biased
-0.14	-0.06	0.27	-0.73

Table 5.1: Estimated ME of the predictors in the model.

5.2.2 Analysing Forecast Accuracy

In this section the forecast accuracy of the predictors are analysed. The mean absolute error (MAE), root mean squared error (RMSE) and absolute error (AE) of the predictors are used for this purpose.

Analysing MAE

The result of the MAE calculations are displayed in table 5.2. The MAE of the mean predictor is smallest and thus according to this measure the mean predictor is the most accurate predictor. The MAE of the rational predictor is slightly larger than the MAE of the mean predictor. Thus the rational predictor is slightly less accurate and the second most accurate predictor according to this measure. The median predictor is the third best predictor according to the MAE. The median predictor in the model is not a good comparison to the median forecasters in the surveys which is apparent by using this measure. The biased predictor is the least accurate predictor according to this measure.

The mean and median forecasters analysed in the SPF are found to be more accurate than the individual forecasters when testing the MAE. The mean predictor in the model is the most accurate predictor and thus we seem to be able to reproduce some of the findings in the SPF survey. However, the mean predictor in the model is only found to be more accurate than the other predictors in the model and not more accurate than individual forecasters as we cannot follow individual forecasters with time. The results will however be valid for individual forecasters following either the rational or the biased predictor closely.

Analysing RMSE

The results of the RMSE calculations are displayed in table 5.3. The mean predictor is the most accurate predictor according to the RMSE. The rational predictor is the second most accurate predictor only slightly less accurate than the mean predictor. The median predictor is the third most accurate predictor in the model. The median predictor is not a good comparison to the median forecaster analysed in the surveys. The biased predictor is the least accurate predictor in the model.

The mean and median forecasters analysed in the SPF are found to be more accurate than the individual forecasters when testing the RMSE of the forecasters. We are able to reproduce some of these findings as the mean predictor in the model is the most accurate predictor. However, we cannot follow individual forecasters with time in the model and thus we do not know how well the mean predictor performs compared to the individual forecasters. Nevertheless, individuals following either the rational or the biased predictor closely are likely to be less accurate than the mean predictor.

Analysing AE

Table 5.4 displays the results of the AE calculations. The table displays the percentages of the consumers that are more and less accurate than the mean and median predictors. A third of the consumers are forecasting more accurate than the mean predictor while the rest forecast less accurate. Thus we find that the mean forecasts perform better than most of the single forecasts. Only 14% of the consumers forecast more accurate than the median predictor and only 17% forecast less accurate. 69% of the consumers forecasts equally accurate as the median predictor and it is apparent that the median predictor in the model is not a good comparison to the median forecasts in the surveys.

The results of the AE analysis of the model are comparable with the results of the AE analyses of the surveys. When analysing the AE measure we do not follow individual forecasters with time and since we study the performance at each time step the analyses can be compared. The SPF and the Michigan surveys are both analysed using the AE measure. The mean and the median forecasts in both the

MAE of mean	MAE of median	MAE of rational	MAE of biased
0.45	0.55	0.47	0.77

Table 5.2: Estimated MAE of the predictors in the model.

RMSE of mean	RMSE of median	RMSE of rational	RMSE of biased
0.56	0.68	0.59	0.89

Table 5.3: Estimated RMSE of the predictors in the model.

surveys are found more accurate than the single forecasts according to the AE. The mean and median forecasts in the Michigan survey perform particularly well compared to the single forecasts in the same survey. The mean forecasts in the Michigan survey perform more accurate than 74% of the single forecasts. The corresponding number for the SPF is 57%. In comparison, the mean predictor is more accurate than 70% percentage of the consumers in the model. Thus we are able to reproduce some of the findings from the surveys.

The results for the model are more similar to the results for the Michigan survey than the SPF. This might suggest that the model is better suited for the Michigan survey than the SPF. In both the model and the Michigan survey the mean predictor is more accurate than the single forecasts. The mean forecasts perform well because they exploit the numerous forecasts in the survey. This is not to say that all forecasts in the Michigan survey and the model are poor, but that they can be combined to perform better. The single forecasts in the SPF are, however, thought to perform well as they are forecasted by professionals. Consequently combining the single forecasts to the mean forecast results in a smaller gain for the SPF.

Compared to	Smaller AE	Equal AE	Larger AE
Mean	33.1%	0%	66.9%
Median	13.5%	69.3%	17.2%

Table 5.4: Percentages of the single forecasts that are more, equally and less accurate than the mean and median forecasts according to the AE.

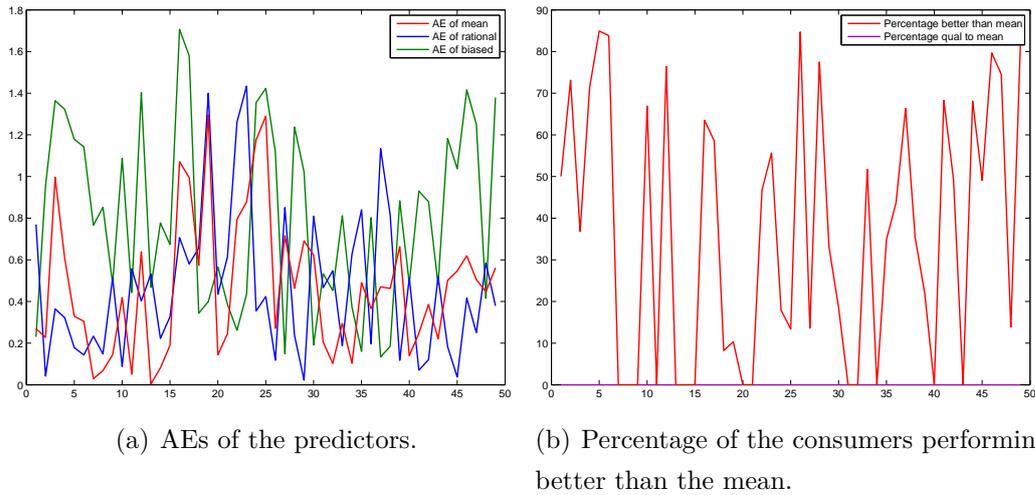


Figure 5.2: AE of the mean, rational and biased predictor are displayed as a function of time. The accuracy of the mean predictor is also compared to the accuracy of the predictors used by the consumers. The result is displayed as a function of time.

Further insight in the errors is gained by displaying the AE of the predictors with time. The results can be found in figure 5.2(a) which corresponds to figure 5.1. It is clear that the AE varies considerably with time for all the predictors. The biased predictor appears to perform considerably worse than the other predictors which also were the results of the MAE and RMSE analyses.

The relative performance of the mean predictor is displayed with time in figure 5.2(b). The AE of the mean predictor is compared to the AE of the forecasts used by the consumers for each time step. The share of the consumers performing more accurate than the mean predictor varies considerably with time. Some time steps the mean predictor is the most accurate predictor while other time steps it is not the most accurate predictor. The mean predictor is never the least accurate predictor as there is always some portion of the consumers performing less accurate than the mean.

5.2.3 Comments on the Performance of the Predictors

The mean predictor is the best predictor when the realised inflation is in between the biased and rational predictor and there are no large changes in the inflation so the effect of the time lag is minimized. Whenever the realised inflation is close to or exceeding the biased or rational predictor, one of these predictors is the best predictor. For the mean predictor to approach the rational predictor, the error made by the biased predictor must be large. If the error made by the biased predictor is large, the error of the mean predictor is larger than the error of the rational predictor as long as there is a small part of the population still using the biased predictor. Similar reasoning can be used for the reversed situation when the biased predictor performs well and the rational performs poorly. In addition, the time lag causes the mean predictor to be less accurate. Isolating this effect, the rational and biased predictors perform relatively better because they are not lagging the realised inflation as they are constant. The mean predictor is never the worst predictor, but it might be equally bad as one of the predictors. This will occur in the limit where only one predictor is used by all the consumers in the model. Whenever this situation arises the mean predictor will simultaneously be equally bad and good as the other predictor because there is only one predictor used by the consumers.

The main finding of the analysis of the model is that the mean predictor performs well. The predictor is compared to the other predictors in the model when analysing forecast bias and forecast accuracy. The median predictor is less biased than the mean predictor, but the median predictor is not an accurate predictor. Thus overall the mean predictor appears to be the best predictor. In addition we are able to analyse the accuracy of the mean predictor compared to the accuracy of the single forecasts. The mean predictor is more accurate than most of the single forecasts. This result is in line with the results of the analyses of the surveys since the mean forecasters in the surveys were found to be more accurate than the single forecast.

5.3 Robustness Analysis

In this section the parameter values are changed and the results analysed and compared to the standard model to see if they still hold.

The parameter value of the variance of the shock (σ_ε) is set to 0.25 and 1 in two separate simulations. The value of the parameter in the standard model is 0.5. The results hold for the absolute value of the ME when the parameter value of the variance changes. E.g. for both parameter values the median predictor is the least biased, the mean predictor is the second least biased, the rational predictor is the third least biased and the predictor called the biased predictor is the most biased predictor. The results of the MAE analysis hold except that the mean and rational predictor become equally accurate at the highest level of variance. Similar results are obtained for the RMSE measure. The results from the standard model holds for all the parameter values except for the highest level of variance where the rational predictor is slightly more accurate than the mean predictor. The results of the AE analysis also hold. The mean and median predictors are better than most of the consumers in the model. More detailed information about the performance of the predictors when the parameter value of the variance changes can be found in appendix B.

The value of the sensitivity parameter (β) is set to 0.2 and 5 in two separate simulations. The value of the parameter in the standard model is 1. The results of the standard model holds for the ME measure when the value of the sensitivity parameter changes. E.g. the median is the least biased predictor, the mean the second least biased predictor, the rational the third least biased predictor and the predictor called the biased predictor is the most biased predictor. The results for the MAE measure holds for all the values of the parameter except for the mean and rational predictor at the highest level of sensitivity. At this level of sensitivity the rational predictor is more accurate than the mean predictor according to the MAE measure. The results when using the RMSE measure hold for all the values used for the parameter except at the highest level of sensitivity. At this level of sensitivity the rational predictor is more accurate than the mean predictor according to the RMSE measure. The results of the AE analysis also hold when the sensitivity of the

consumers changes. The mean and median predictors are better than most of the consumers in the model. More detailed information concerning the performance of the predictors when the value of β changes can be found in appendix B.

The value of ρ is set to 0.5 and 0.9 in two separate simulations. The value of the parameter in the standard model is 0. Changing the value of this parameter, greatly changes the dynamics of the model. The rational predictor no longer constantly predicts the inflation to be 0, but the predicted inflation is dependent on the previous shock. The results of the standard model hold for the ME measure when the value of the parameter changes. E.g. the median predictor is the least biased, the mean predictor is the second least biased, the rational predictor is the third least biased and the predictor called the biased predictor is the most biased predictor. The results of the MAE analysis holds for all the values of ρ except the largest value. Using this value the rational predictor is more accurate than the mean predictor according to the MAE measure. Similar results are obtained for the RMSE measure. The results of the standard model hold for all the values of ρ except when the value of the parameter is equal to 0.9. At this level the rational predictor is more accurate than the mean predictor. The AE analysis still holds when the value of ρ changes. The mean and median predictors are better than most of the consumers in the model. More detailed information concerning the performance of the predictors when the value of ρ changes can be found in appendix B.

The results discussed in this section only concern the performance of each predictor relative to the performance of the other predictors when the parameter values change. Whether each predictor independently performs better or worse when the values of the parameters change is not discussed. Studying and explaining the effects of the changes in parameter values further would be interesting but is beyond the scope of this thesis.

Chapter 6

Conclusion

The focus of this thesis is on heterogeneity in inflation forecasts. We analyse the performance of the mean, median and individual forecasters. We start by analysing the Survey of Professional Forecasters and the Michigan survey. Next, a model describing an economy with heterogeneous consumers is studied. There are two predictors used by the consumers in the model and both forecast a constant level of the inflation. However, the consumers in the model switch between the predictors depending on their past performance.

There are two main findings in the thesis. First, there is heterogeneity in the forecasts. Studying the forecasts in the surveys it is apparent that there is heterogeneity in the individual forecasts. There is also heterogeneity in the model as there are two predictors used by the consumers in the model. The heterogeneity in the model is persistent, e.g. in the long run both predictors are used by the consumers in the economy.

The second finding is that the mean forecaster is a good forecaster. The focus of the analysis is on the performance of the mean and median forecasters compared to the individual forecasters. The analysis of the surveys reveals that the mean and median forecasters are both less biased and more accurate than the individual forecasters. The mean predictor in the model also performs well. It is both less biased and more accurate than the two predictors used by the consumers in the model. The mean predictor is a good predictor in the model because it is a self-

referential model. However, it is not trivial that the mean predictor performs well. Both monetary policy and shocks affect the inflation, thus the mean predictor at one time step could have been a poor predictor for the inflation the next time step. Although the mean predictor is a good predictor, we observe that it is sluggish compared to the realised inflation. This result is observed both for the surveys and for the model. The median predictor in the model does not perform equally well as the mean predictor. This is because the model is simple and not well suited for producing median forecasts.

Further work

A more thorough analysis of the model when the parameter values are changed would be an interesting topic for further studies. In particular, a more thorough analysis when the shock is an autoregressive process would be interesting. Changing the type of shock would greatly change the dynamics of the model. Next, we find that the mean predictor in the model is less biased and more accurate than the two predictors in the economy. We also find that the mean predictor is more accurate than the consumers in the model. However, we are not able to use the most common measures of accuracy and no measure of bias for the analysis of individual consumers. Developing and analysing a model where we are able to follow individual agents' choices and performance would be interesting. This would enable us to analyse both the forecast accuracy and the forecast bias more thoroughly.

Bibliography

Andrew Ang, Geert Bekaert, and Min Wei. Do macro variables, asset markets, or surveys forecast inflation better? *Journal of Monetary Economics*, 54:1163–1212, 2007.

Mikhail Anufriev, Tiziana Assenza, Cars Hommes, and Domenico Massaro. Interest rate rules and macroeconomic stability under heterogeneous expectations. *Macroeconomic Dynamics*, pages 1–31, 2012.

Roy Batchelor. How useful are the forecasts of intergovernmental agencies? the imf and oecd versus the consensus. *Applied Economics*, 33:225–235, 2001.

Peter J. Bickel and Kjell A. Doksum. *Mathematical Statistics*. Person Prentice Hall, 2007.

William A. Branch and Bruce McGough. A new keynesian model with heterogeneous expectations. *Journal of Economic Dynamics & Control*, 33:1036–1051, 2009.

William A. Brock and Cars H. Hommes. A rational route to randomness. *Econometrica*, 65:1059–1095, 1997.

Michael P. Clements. Consensus and uncertainty: Using forecast probabilities of output declines. *International Journal of Forecasting*, 24, 2008.

Dean Croushore. Introducing: The survey of professional forecasters. *Business Review*, November/December 1993.

Richard T. Curtin. Surveys of consumers, procedure to estimate price expectations. <http://new.sca.isr.umich.edu/fetchdoc.php?docid=24772>, Feb 2013a.

Richard T. Curtin. Surveys of consumers, sample design. <http://new.sca.isr.umich.edu/fetchdoc.php?docid=24773>, Mar 2013b.

Deepa Dhume Datta and Wenxin Du. Nonparametric hac estimation for time series data with missing observations. Unpublished, International Finance Discussion PaperS, 2012.

Karen Oftedal Eikill. Are individual forecasters rational?, 2012.

Marji Lines and Frank Weserhoff. Inflation expectations and macroeconomic dynamics: The case of rational versus extrapolative expectations. *Journal of Economic Dynamics & Control*, 34:246–257, 2010.

MichiganSurvey. Surveys of consumers. <http://new.sca.isr.umich.edu/>, Mar 2013a.

MichiganSurvey. Surveys of consumers, description. <http://new.sca.isr.umich.edu/fetchdoc.php?docid=24774>, Mar 2013b.

Whitney K. Newey and Kenneth D. West. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55, May 1987.

SPF. Survey of professional forecasters. <http://www.phil.frb.org/research-and-data/real-time-center/survey-of-professional-forecasters/>, Feb 2013a.

SPF. Dates of previous surveys. <http://www.phil.frb.org/research-and-data/real-time-center/survey-of-professional-forecasters/spf-release-dates.txt>, Mar 2013b.

- Tom Stark. Survey of professional forecasters, caveats. <http://www.phil.frb.org/research-and-data/real-time-center/survey-of-professional-forecasters/spf-caveats.pdf>, Feb 2013a.
- Tom Stark. Survey of professional forecasters, documentation. <http://www.phil.frb.org/research-and-data/real-time-center/survey-of-professional-forecasters/spf-documentation.pdf>, Feb 2013b.
- Herman O. Stekler. The rationality and efficiency of individuals' forecasts. *A Companion to Economic Forecasting*, pages 222–240, 2002.
- John B. Taylor. Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series on Public Policy*, 39:195–214, 1993.
- Ronald E. Walpole, Raymond H. Myers, Sharon L. Myers, and Keying Ye. *Probability & Statistics for Engineers & Scientists*. Person Prentice Hall, 2007.
- Michael Woodford. *Interest and Prices*. Princeton University Press, 2003.
- Øyvind Steira. How accurate are individual forecasters?, 2012.

Appendix A

Analysing the Preprocessed Dataset from the Michigan Survey

The analysis in section 3.2 of the Michigan survey is based on the processed dataset. For comparison the estimates of the mean and median are also calculated based on the preprocessed dataset. In this dataset only forecasts with complete information are used and the outliers are not truncated. The performance of these estimates can be found in table A.1. There is approximately no difference in the relative performance of the mean and median forecasts regardless of whether they are based on the preprocessed or processed dataset.

Compared to	Smaller AE	Equal AE	Larger AE
Mean	25.8%	0%	74.2%
Median	13.7%	16.3%	70.1%

Table A.1: Percentage of the single forecasts performing better, equal or worse than the mean and median forecasts. The mean and median forecasts are based on the preprocessed dataset.

Appendix B

Tables Concerning the Robustness Analysis of the Model

Tables displaying the results of the robustness analysis of the model.

	ME of mean	ME of median	ME of rational	ME of biased
$\sigma_\varepsilon = 0.25$	-0.13	0.08	0.26	-0.74
$\sigma_\varepsilon = 0.5$	-0.14	-0.06	0.27	-0.73
$\sigma_\varepsilon = 1$	-0.15	-0.13	0.29	-0.71
	MAE of mean	MAE of median	MAE of rational	MAE of biased
$\sigma_\varepsilon = 0.25$	0.24	0.37	0.30	0.74
$\sigma_\varepsilon = 0.5$	0.45	0.55	0.47	0.77
$\sigma_\varepsilon = 1$	0.85	0.90	0.85	1.00
	RMSE of mean	RMSE of median	RMSE of rational	RMSE of biased
$\sigma_\varepsilon = 0.25$	0.31	0.46	0.37	0.79
$\sigma_\varepsilon = 0.5$	0.56	0.68	0.59	0.89
$\sigma_\varepsilon = 1$	1.07	1.13	1.06	1.24

Table B.1: The estimated ME, MAE and RMSE of the predictors in the model for different values of the standard deviation (σ_ε). The standard value of the parameter is 0.5.

Compared to mean	Smaller AE	Equal AE	Larger AE
$\sigma_\varepsilon = 0.25$	25.8%	0%	74.2%
$\sigma_\varepsilon = 0.5$	33.1%	0%	66.9%
$\sigma_\varepsilon = 1$	38.7%	0%	61.3%
Compared to median	Smaller AE	Equal AE	Larger AE
$\sigma_\varepsilon = 0.25$	11.1%	63.6%	25.3%
$\sigma_\varepsilon = 0.5$	13.5%	69.3%	17.2%
$\sigma_\varepsilon = 1$	10.4%	78.4%	11.1%

Table B.2: Percentages of forecasts performing better, equal or worse than the mean and median predictors according to the AE for different values of the standard deviation (σ_ε). The standard value of the parameter is 0.5.

	ME of mean	ME of median	ME of rational	ME of biased
$\beta = 0.2$	-0.16	-0.04	0.32	-0.68
$\beta = 1$	-0.14	-0.06	0.27	-0.73
$\beta = 5$	-0.10	-0.09	0.20	-0.79
	MAE of mean	MAE of median	MAE of rational	MAE of biased
$\beta = 0.2$	0.42	0.56	0.48	0.72
$\beta = 1$	0.45	0.55	0.47	0.77
$\beta = 5$	0.50	0.52	0.48	0.84
	RMSE of mean	RMSE of median	RMSE of rational	RMSE of biased
$\beta = 0.2$	0.53	0.69	0.60	0.84
$\beta = 1$	0.56	0.68	0.59	0.89
$\beta = 5$	0.63	0.66	0.60	0.98

Table B.3: The estimated ME, MAE and RMSE of the predictors in the model for different values of the sensitivity (β). The standard level is $\beta = 1$.

Compared to mean	Smaller AE	Equal AE	Larger AE
$\beta = 0.2$	32.2%	0%	67.8%
$\beta = 1$	33.1%	0%	66.9%
$\beta = 5$	35.4%	0%	64.6%
Compared to median	Smaller AE	Equal AE	Larger AE
$\beta = 0.2$	21.1%	54.2%	24.6%
$\beta = 1$	13.5%	69.3%	17.2%
$\beta = 5$	3.5%	91.9%	4.6%

Table B.4: Percentages of forecasts performing better, equal or worse than the mean and median predictors according to the AE for different values of the standard deviation (β). The standard value of the parameter is 1.

	ME of mean	ME of median	ME of rational	ME of biased
$\rho = 0$	-0.14	-0.06	0.27	-0.73
$\rho = 0.5$	-0.12	-0.04	0.23	-0.77
$\rho = 0.9$	-0.02	-0.01	0.05	-0.95
	MAE of mean	MAE of median	MAE of rational	MAE of biased
$\rho = 0$	0.45	0.55	0.47	0.77
$\rho = 0.5$	0.54	0.59	0.57	0.87
$\rho = 0.9$	0.93	0.95	0.91	2.09
	RMSE of mean	RMSE of median	RMSE of rational	RMSE of biased
$\rho = 0$	0.56	0.68	0.59	0.89
$\rho = 0.5$	0.68	0.74	0.71	1.06
$\rho = 0.9$	1.18	1.21	1.14	2.73

Table B.5: The estimated ME, MAE and RMSE of the predictors in the model for different values of ρ . The value of the ρ in the standard model is 0.

Compared to mean	Smaller AE	Equal AE	Larger AE
$\rho = 0$	33.1%	0%	66.9%
$\rho = 0.5$	39.5%	0%	60.5%
$\rho = 0.9$	43.2%	2.5%	54.3%
Compared to median	Smaller AE	Equal AE	Larger AE
$\rho = 0$	13.5%	69.3%	17.2%
$\rho = 0.5$	11.5%	73.0%	15.5%
$\rho = 0.9$	7.4%	83.9%	8.7%

Table B.6: Percentages of forecasts performing better, equal or worse than the mean and median predictors according to the AE when the parameter value of ρ changes. The value of ρ in the standard model is 0.