Department of Economics and Finance Chair Econometric Theory

Prediction Markets and Macroeconomic Forecasting

CANDIDATE
Federico Angelini
ID 659911

THESIS SUPERVISOR
Prof. Giuseppe Ragusa

THESIS CO-SUPERVISOR
Prof. Pierpaolo Benigno

ACCADEMIC YEAR 2014/2015

Table of Contents

Introduction

Can economic prediction markets data be used by forecasters in their models? Do prediction market prices contain relevant information that can be exploited for increasing the accuracy of forecasting models? Do prediction market forecasts tend to be more accurate than the ones provided by professional forecasters? The following text has tried to address these questions using the available data about economic prediction markets and survey forecasts.

Prediction markets have been largely applied to forecast future events: from the next United States president to the next quarter earnings per share of Google. They are platforms where participants can buy or sell contracts tied to future events. All the participants that bet on the outcome that turns out to be the actual realization are remunerated while the others lose the amount of money they had betted. Since contracts exchanged on prediction markets are similar to binary options, it is possible to interpret the contract price as a probability that the underlying event will occur. For instance, as I am writing, on the Iowa Electronic Market, which is the most important political prediction market, the contract tied to the winning of the Republican Party in the next presidential election is exchanged at 0,377$. This means that prediction market participants believe that with a probability of 37,7% the next United States President will be republican. Conversely, a democratic candidate is believed to win the presidential election with a probability of 61,4%. As an example, let's assume that the Republican Party will win the election. This means that if an investor buys now a contract tied to the victory of a republican candidate, I will gain 0,623$ for each 1$ he has betted.

Although not as liquid as political prediction markets, a macroeconomic prediction market called *"Economic Derivative Market"* was run between October 2002 and September 2005 by Goldman Sachs and Deutsche Bank operating as counterparts. Four economic variables were the object of separate auctions: Non-farm payroll, initial unemployment claims, retail sales and ISM index[1].

Wolfers and Zitzewitz (2004) and Wolfers and Gurkaynak (2006) were two of the first papers to deal with this market. By collecting and gathering data about each of the auctions held in the Derivative Market, the authors have carried out several analyses. First, prediction markets forecasts have been compared with the ones of professional forecasters. Although the two were not published

[1]The market eventually moved to the Chicago Mercantile Exchange (CME) and three new contracts were introduced: U.S. GDP, Eurozone inflation index and U.S. International trade balance. Nowadays, the Nadex holds an economic prediction market on non-farm payrolls, Fed interest rate and unemployment claims.

on the same day[2], prediction market forecasts have been proved to be more accurate than consensus forecasts. Moreover, the two sets of predictions differ also in terms of efficiency. The authors have noticed how consensus forecast errors tend to be correlated over time, while prediction market forecasts do not show this feature. This might imply that consensus predictions can be improved and information can be better used.

Second, authors have shown how time series of the standard deviation of the auction is able to better mimic the economic uncertainty than the common measure of disagreement between professional forecasters. Performing a regression analysis, disagreement has been proved to be correlated even if loosely with the uncertainty measure computed from the time series of standard deviation. According to the authors, this difference between the two measures can be explained by the fact that disagreement among forecasters can be low even in the context of high uncertainty[3].

Finally, authors have investigated which predictions better anticipate the reactions of financial markets to the economic release. By regressing the change of the financial variable[4] simultaneously on the prediction market surprise and the consensus forecast surprise, prediction markets have been proved to better mimic the response of financial markets to macroeconomic news or surprises. Since then numerous papers have used the information embedded in the standard deviation of the economic derivative market as a proxy for the economic uncertainty.

Beber and Brandt (2006) have used the standard deviation computed from the prices of vanilla options of each macroeconomic derivative auction to construct a variable of macroeconomic uncertainty. This has then been used to study the effect of macroeconomic uncertainty on the bond and stock markets. This analysis has revealed that, after the macroeconomic release, financial market implied volatility tends to decrease, as the bigger is the ex-ante uncertainty surrounding the auction[5]. This suggests that for economic releases with low levels of uncertainty, financial volatility does not vary significantly once the actual value is released.

Second, the authors have found that the volume of transactions in the option market is correlated with the uncertainty measure derived from the prediction market. This tends to increase once the uncertainty is settled, namely when the actual data is released. This behaviour evidences that market

[2] Consensus forecasts are available one week before the release day of the economic variable they are predicting, while prediction markets forecasts are made one day before the announcement.

[3] The so-called stick to the consensus bias. That is, professional forecasters tend to adjust their predictions to be in line with consensus forecast during unstable economic environment.

[4] Namely, the change in both the S&P 500 and 10-years Treasury note yield registered 5 minutes before and 25 minutes after the release.

[5] Similar findings have been illustrated by Fornari (2004) and Huang (2008).

participants tend to wait until the actual macroeconomic value is released to take financial positions in the market.

In addition, stock and bond markets respond differently to macroeconomic uncertainty. Bonds show the tendency to be more sensitive to uncertainty while stocks are less sensitive.

Finally, non-farm payrolls economic derivative has been proven to be either the most liquid market and have the largest impact on financial market volatility.

The *Economic Derivative Market* has not been the unique economic prediction market recently. Florian et al. (2011) have analysed the performance of the *Economic Indicator Exchange* (from now on, EIX). This was an economic prediction market run online with play-money in 2009, which allowed participants to forecast the released value of 5 variables regarding the German economy: GDP, inflation, investments, export and unemployment. This market has a design that differs from the *Economic Derivative Market*.

To begin with, instead of holding auctions before the release day, the EIX worked as a continuous double auction. In this way, participants could constantly interact between one with the other. This has helped to solve the liquidity that is one of the main issues of prediction markets.

Secondly, instead of several binary contracts, a single stock with a linear pay-off was exchanged[6].

Finally, to maintain the participants active and motivated in the market, monthly and yearly prizes were hand out to the participants that provided the best forecasts. The authors have collected and scrutinized 11,708 transactions performed in the market. The results were intriguing. Forecasts made by EIX participants have been able to outperform in terms of accuracy the Bloomberg forecasts. Because the market was continuously open, it has been also possible to compare the two forecasts at an equal distance from the release day. This has allowed getting over the possible bias deriving from the gap of time between the prediction market and surveys forecasts. Even sharing the same information set, prediction markets have been proved to provide more accurate predictions than professional forecasters.

Starting from these previous findings, I have tried to analyse prediction markets data under a different lens. Three different but simple experiments have been carried out.

First, as it has been done in Wolfers and Gurkanyak (2006), prediction markets and surveys forecasts have been compared to study which of the two provides better predictions. However, instead of simultaneously regress the economic released value on the predictions made by prediction markets and professional forecasters, I have set up three different models: a naïve model

---

[6] The mechanism of the market will be explained more in depth in the second chapter. Here just the main differences with respect to other prediction markets have been exposed.

(benchmark), an AR(1) model integrated with prediction market forecasts and another AR(1) using, this time, consensus forecasts. Regression results have shown how, even by using past observation of the interest variable, prediction markets tend to slightly outperform the consensus model in terms of accuracy.

Second, I have focused on the non-farm payrolls data to study if prediction market model results more accurate also for one-step-ahead predictions. Indeed, since prediction markets forecasts are available only one day before the economic release it could be questionable their use in terms of forecasting models. There is no great return in having a prediction the day prior the announcement. Because of this, the time series of the prediction market forecasts made one month before the new release has been created. In this analysis, I have decided to focus only on the non-farm payrolls prediction market data, since it has been proved to be the most liquid economic derivative.

In addition to the three models introduced above, a fourth model using monthly observations of business confidence and industrial production has been set up. All the variables used in these models have been rendered stationary by taking the first difference. Conversely to the case where prediction markets data were used to predict the next day economic release[7], the naïve model has been the one with the higher level of accuracy for the 1-step-ahead forecasts.

Third, I have used data about non-farm payrolls prediction market to investigate if they can result useful in the prediction of the unemployment rate. Prediction market data has been proved to provide slightly better predictions than surveys data. Moreover, it has been set out a simple linear regression model employing as predictors the one-lag unemployment observation, the prediction market forecasts for the non-farm payrolls and the actual non-farm payrolls release. Interestingly, besides the lagged unemployment, the unique observation that has resulted being negatively and statistically correlated with the unemployment has been the prediction market forecast.

Finally, I have tried to examine the relationship between *macroeconomic uncertainty* and volatility in the financial market. Instead of using the standard deviation of the prediction markets as a measure of macroeconomic uncertainty, I have preferred to employ the *macroeconomic surprise or news*. Following the approach of Balduzzi et al. (2001) and Andersen et al. (2002)[8], I have defined the *surprise effect* as the difference between the realized value and the prediction market forecast. Then to assess the effect of this measure on the implied financial market volatility, the change

---

[7] This can be thought as a form of 0-step-ahead predictions.

[8] Both papers deal with impact of macroeconomic news or surprise on financial variables. The former focuses on the US treasury market, while the latter on the foreign exchange market. The same variable has been used by Gurkaynak (2005) to judge the effect on long-term interest rate. Wolfers and Gurkanyak (2006) and Beber (2008) have employed the *macroeconomic surprise* derived from prediction markets to investigate his impact on the S&P500 return.

between the VIX closing price on the day before the announcement and the VIX opening price on the day of the announcement it has been regressed on the *surprise effect*. Regression results have shown that positive *surprise effects* tend to decrease the implied volatility, while the opposite occurs for *negative surprises*.

The rest of the text is organized in the following way. Chapter 1 illustrates the magnitude of macroeconomic forecasting errors and tries to gather all the relevant literature in order to present an explanation of these errors. Chapter 2 introduces first what prediction markets are and how they work. Then, it reports all the main interesting features of prediction markets to evidence how and why information derived from prediction markets should be used. The final part of the chapter introduces the analyses and models that I have proposed. Finally, chapter 3 explains the available dataset and discusses the results obtained.

Chapter 1: Macroeconomic Forecasting Errors

"*Prediction is very difficult, especially if it's about the future*."

--Nils Bohr, Nobel laureate in Physics

As for every science it is fundamental to go through closely the past to make advances. Looking back and analysing what has gone wrong can give valuable information. It can point out the main errors, so it is easier to do not make them again. But it also can shed the light on where improvements are needed and, therefore, illustrate the future path. With this view in mind, I have structured the following chapter. Before starting to advance claims about how prediction markets and the information embedded in them can help the forecasting science, I forced myself to investigate the main problems in this field. I believe, or at least I hope, that this has helped me to clarify my ideas. At the same time, it has enabled me to give structure to my entangled flows of thoughts. Obviously, my intent in this first chapter is not do deeply analyse all the issues regarding macroeconomic forecasting. Although very stimulating, this task is well beyond my abilities. It is also out of my research question. So I will simply report the main drawbacks the literature has highlighted over time. This will help to build a general framework in which I will introduce my ideas. Using a famous economics motto, in this first chapter I will stand on the shoulder of giants.

The structure of the chapter is the following. The first part will be devoted to the major flaws of macroeconomic forecasting. The second part will deal with the possible explanations of these failures. Several sources of errors have been identified by the literature, so I will try to present them in a concise way without sacrificing the importance of the findings. The last part will introduce the concept of prediction markets and how they can be used to improve the practice of forecasting.

1.1 Forecast errors: How much big they are?

When it comes to study the overall accuracy of macroeconomic forecasts, two variables are usually taken into consideration: the mean absolute error (MAE) and the root mean square error (RMSE). The MAE is the mean of the absolute value of forecast errors, namely:

$$\frac{1}{N}\sum_{i=1}^{N}|e_t|$$

where e_t is the forecast error made in period t. Namely: $e_t = y_t - \hat{y}_{t|t-1}$.

The RMSE, instead, is analytically defined as:

$$\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}$$

where $\hat{y}_i$ is the forecast made at time $i-1$ and y_i the actual realization. The RMSE differs from the MAE since it does not attach to each error an equal weight. Indeed, by taking the square of the errors before they are averaged, the RMSE gives higher weights to bigger error. As a result, RMSE allows penalizing more forecasting models producing large but infrequent errors.

The two measures can be compared in order to perform an accurate diagnosis of the nature of the forecasting errors made by a model. If the RMSE is larger than the MAE, than the greater is the variance of the forecasting errors in the sample. Although the two are commonly reported in forecasting literature, the RMSE is usually preferred. In order to pin down the more accurate model, a relative comparison, in terms of RMSE, with benchmark models is performed. The RMSE of the model to be tested is compared with the RMSE of the benchmark model. If the RMSE is smaller than the one of the benchmark, the forecasting model is more accurate.

As for the US, as a preliminary analysis of forecast performances, I choose to focus on the *Survey of Professional Forecasters* (*SPF)* and on their predictions for the GDP growth. The Survey of Professional Forecasters was established in the 1968 and is one of the oldest macroeconomic surveys in the United States. Originally, it was conducted jointly by the American Statistical Association and the National Bureau of Research. However, in 1990, the Federal Reserve Bank of Philadelphia took over the survey. The SPF produces forecasts for several economic variables: Gross Domestic Product, CPI and PCE Inflation and Unemployment. Although each survey participant has an identification number, names of the firms for which they work remain anonymous. For what concerns the nature of the models used to carry out predictions, panelists employ mathematical models[9] but they modify the forecasts adding personal or subjective adjustments.

The choice of the SPF has been motivated by different arguments. To begin with, most of the prevalent literature regarding forecasting accuracy deals with it. Second they are anonymous, and then there is no way to gain publicity from their predictions[10]. Moreover, they report the same forecast that they sell in the market. So these two arguments should ensure that the SPF forecasts

---

[9] The majority of the panelists uses a combination of structural models (IS/LM or AD/AS) and time series models (VAR/VEC).

[10] This argument will be analysed later in the chapter.

are likely to value accuracy over other goals. By analysing the performance of the GDP forecasts over the period 1985-2013, we can notice that forecasting accuracy is influenced mainly by two factors: forecasting horizon and data revision (Stark, 2010). Data evidence that forecast errors tend to be bigger for higher forecasting horizons. The difference between the RMSE of the 1-year-ahead forecast and the RMSE of the 1-quarter-ahead forecast using the initial release is equal to 0.54. The continuous revision of economic data operated by statistical agencies is another element that affects negatively the forecasting accuracy. By comparing the RMSE of the 1-year-ahead GDP forecast using the initial release and the RMSE of the forecast employing the latest vintage release, it is possible to notice an increase in the RMSE of the 13%. Using the available data is also possible to study how the SPF consensus forecasts perform with respect to two benchmark models. The first one is a *no-change model*, which uses the latest quarter release value as a forecast for the next period forecast.

$$\hat{y}_t = y_{t-1}$$

The second one is an *autoregressive model*, which uses past observations of the variable to be forecasted weighted by coefficients estimated with past data.

$$\hat{y}_t = \hat{\phi}_0 + \sum_{i=1}^{p} \hat{\phi}_i y_{t-i} + u_t$$

Both the benchmark models are outperformed by the SPF forecasts, with the *no-change model* displaying the lowest relative accuracy. The edge in accuracy tends to be bigger for closer forecasting horizons. Interestingly, the RMSE ratio between the consensus model and the autoregressive model tends to be very close to 1 for the two, three and four-quarter ahead forecasts.

Reifschneider and Tulip (2007) have made a more complete analysis about macroeconomic forecasts in the US. They have evaluated the forecasting performance of six different institutions[11] over the period 1986 and 2006. For what concerns GDP forecasts, the authors have found that precision tends to decrease as the forecasting horizon increases. The average RMSE for the one-year-ahead annual GDP growth is slightly below 1.4. Interpreting the real meaning of these values is not always straightforward. Fortunately, the authors have reported also the standard deviation of the GDP growth over the period studied, to quantify how much the forecasts were accurate. Indeed,

[11] Federal Reserve Board, Congressional Budget Office, Blue Chip, SPF, Administration and FOMC members.

by dividing the RMSE of the forecast by the standard deviation of the variable to be forecasted, it is possible to assess how much a forecast is informative. The lower is the ratio the higher is the predictive power of a forecast. Results for the GDP growth forecasts were not shining. For the lowest horizon, namely the one-quarter-ahead GDP forecast, the ratio is close to 0.9. As one might expect, the RMSE-standard deviation ratio becomes bigger as the forecast horizon increases. This implies that predictions too far from now should be given less importance.

As for the Europe, Oller and Barot (1999) have investigated the accuracy of OECD and national institute forecasts of 13 European countries. They have found that, during the 1971-1995 period, the mean absolute error of one-year ahead GDP annual growth forecast was equal to 1.4%.

Similar results are reported in Batchelor (2000). International Monetary Found (IMF), OECD and Consensus forecasts for G7[12] countries from 1990 to 2000 are compared in the paper. This paper is interesting for two reasons. The first one concerns the overall accuracy of GDP forecasts during the 1990s. For each target year, the previsions for the annual growth of GDP made in May and October of the previous year and in May of the target year are studied. This approach allows gaining some useful insights on how the accuracy of the previsions varies over time. From the data, it can be noticed how the MAE decreases as the target year is approached. As one might expect, the MAE is below 1 only for the prevision made during the target year. Although there is an increase in efficiency, the average MAE over the three periods is close to 1.4. The comparison between the accuracy of the Consensus with respectively the OECD and the IMF is the other point to be noticed. The author investigates which of these three institutions have been more accurate over the years. The Consensus, namely the pool of private sector forecasters, has been proved to provide forecasts more accurate and informative than the IMF and OECD. Although it is true that IMF and OECD have qualitatively better information, the decrease in the error variance achieved by pooling different private forecasts outperforms this information advantage. Same findings are found in other works (Loungani, 2001).

1.1.2 Systematic Errors

As it will be discussed later, the economy is a complicated system. None is exactly sure of the position of the economy in one year. Because of this, economic forecasts are valuable. They can transmit important information about the future. As it has been reported earlier, these predictions have a great component of uncertainty, which implies forecast errors. However, when predictions for a longer period are made, accuracy, in the statistical sense, is not of primary importance. What

[12] US, Japan, Germany, France, UK, Italy and Canada.

really matters is to get the direction of the economy, not the exact point. Whether it will run into a recession or in expansion. Being able to know in advance where the economy is going is extremely important. Countries can carry out economic policies to mitigate downturns and limit potential social and economic damages. Financial companies can decide where to allocate their money more efficiently. However as it will be immediately described, detecting in advance economic turns is one of the biggest challenges of the forecasting science. To begin with, there are methodological issues. It is necessary to transform quarterly forecasts into monthly data signalling probability recession. Secondly, the model could predict a turning point that does not occur in reality. So higher level of confidence is needed to identify one (McKnees, 1992). Fildes and Stekler (2002) analysed the overall performance of macroeconomic forecasting both in the US and UK. They show that the biggest issue affecting both US and UK macroeconomic predictions is the inability to foresee recessions. Loungani (2001) have confirmed these findings. Focusing on the period going from 1989 to 1998, he has found that forecasters have been able just to anticipate 2 of the 60 recessions that took place. Two-thirds of the recessions were still not reported in April of the year in which the recession occurred. Moreover, in eighty per cent of the case studied, forecasts made in October of the year of the recession underestimated the magnitude of the recession.

Tendency to make systematic errors both in estimating GDP and inflation is another major drawbacks affecting forecasters (Zarnowitz and Braun, 1992). They are more likely to overestimate GDP growth during slowdowns and underestimate it during booms. To make bad things worst, Flides and Stekler (2002) notice that systematic errors occurred during period of great perturbations. Although this could be used as an excuse of poor performances, the authors were concerned about this aspect. Indeed, according to them, forecasts are more valuable during periods of uncertainty. They are instruments that policy makers might use to plan future economic policies. Because of this, relying on inaccurate predictions during turbulent economic moments could be costly.

1.2 Explanation of Forecast errors

Common problems in economic forecasting have been briefly reported. But the main point has not been touched yet. Why do forecast errors occur? Of course is not simple to identify a single reason for which the forecast failed. Because there are several reasons from which forecast errors can derive, it is not possible to exactly disentangle one from the other. In the following pages are reported the main causes the literature has identified as explanation of forecast errors.

1.2.1 Model errors

If it were possible to have a perfect model, probably no errors would occur. Predictions would be very accurate and we would be able to better manage the economy. Developments in computer science and other technologies have fuelled this dream. But, essentially, it is just a dream. Reality is different. Economists spend time working on their models in the attempt to get the best representation of the world. They have to identify relevant variables, then express their relationships and, finally, estimate parameters. Assumptions about variables are made to simplify this process. Because of the underlying complexity of macro-econometric models, some mistakes can be done. Variables can be omitted, parameters can be wrongly estimated, relationship between variables can be mistaken and model can be misspecificated. Consequently, final prediction can be negatively influenced. In their detailed papers, Hendy and Clements (1998,1999,2001) build up a forecast-error taxonomy.

This includes the following sources of deviation between forecasted value and actual realization:

i. Structural change affecting deterministic (ia) and stochastic components (ib)
ii. Model misspecification of deterministic (iia) and stochastic elements (iib)
iii. Forecast-origin inaccuracy
iv. Estimation uncertainty
v. Inherent stochastic nature of economies

The authors display the effects of each of these errors by using a simple first-order vector autoregressive process. After introducing a structural change in the parameters of the model, they study how this affects the final prediction. Shifts in the coefficients of deterministic trends are highlighted as the most dangerous. According to the authors, they are responsible of forecasting failure. If forecasters are not able to timely detect changes, the model they are working with is not anymore a good representation of the economy. Consequently, forecast errors can occur. However, structural breaks are difficult to anticipate because of the complexity of the economy. They can regard underlying changes in the structure of the economy for which forecasters do not have data or information. Structural breaks can be different one from the other, so even if forecasters possess information about past breaks these could not be useful.

1.2.2 The environment: an unstable and uncertain context

Forecasting errors are used to blame forecasters for the poor performance of macroeconomic forecasting. This is a myopic way of addressing the problem. Although they are not faultless, forecasters face one of the most difficult tasks in the economic science: trying to set out relationship between economic variables. The economy is an extremely complex dynamic system populated by one of the most complex creature in the world, the human being. A myriad of different economic agents interact constantly. From their interactions, macroeconomic aggregates arise. On top of that, the underlying structures, from political regimes to new technology, regulating the interactions of the economic agent evolve continuously. In this unstable context, it is quite obvious that forecasts can be wrong and forecast errors can occur. Secondly, future is full of uncertainties. Hendry and Clements (2001) distinguish two forms of uncertainty threatening forecasts. For the first one, forecasters understand the probabilities involved and, therefore, they can to a certain extent factor them in their predictions. In this first set belong events such as political changes, fluctuations in harvests or variations in oil supplies. Rephrasing the Rumsfeld words, we might refer to these circumstances as *known unknowns*. Although difficult to timely predict them, these are events of which forecasters have some form of past knowledge. Beliefs about these events can be formed by weighting current and past information. For instance, by enquiring why and when fluctuations in the oil price have occurred over time, forecasters might assess the probability of a future shock in the oil price.

The second form of uncertainty is, by his nature, unpredictable. Following the earlier jargon, these are the *unknown unknowns*. Unfortunately for the forecasting science, this second form of uncertainty is the predominant. Several mathematical techniques[13] are used to transform time series into stationary time series, namely process whose main moments do not change over time.

As it has been reported earlier, anticipate turning points or structural breaks (so, *unknown unknowns*) is problematic for forecasters. Predicting the behaviour of a process that keeps changing every time is extremely difficult. However, by assuming stationarity, is possible to get over these problems. Once the series is rendered stationary, predictions can be carried out. Then is sufficient to use the inverse of the mathematical technique used to transform the series, to have prediction of the original series. Indeed, the "optimal theory" of forecasting (Klein, 1971), is built on two main assumptions: (i) the model is a good representation of the economy and (ii) the underlining structure of the economy is stable. The second point is the more controversial. Common sense teaches us that

[13] If the series displays a long-run trend, it can be rendered stationary by subtracting this component. The new series is trend-stationary. Another method is to take the changes of the series over a period of time (i.e, monthly, quarterly, annual). In this case the series is difference-stationary. Being able to distinguish between a trend-stationary and difference-stationary is not always possible. Unit-root test can be performed to unravel this argument.

future is very unlikely to be similar to the past[14]. Economics too is not invulnerable to this principle. So assuming that variables are stationary, namely that future values are similar to past ones might be problematic. Not only variables have different statistical properties over time, but also relationships between variable change over time (D'Agostino, Gambetti, Giannone 2010). Stock and Watson (1993) offer some useful insight regarding this point. In their paper, they investigate why the experimental recession index (XRI)[15] has not been able to predict the economic downturn occurred in the US in the October 1990. The performance of the index until the summer of 1990 was good. The RMSE of the index, from the end of 1988 to mid1990, was slightly above 1.3. Moreover, the slowdown of 1989 was effectively predicted. However, the recession of the summer of 1990 was completely missed. According to the authors, the failure was not due to misspecification in the recession index or miss-estimation of parameters. The change in the behaviour of some variables was the explanation of this failure. Indeed, several economic variables that had explained well the GDP dynamics at the end of the 1980s were not able to do so in the summer of 1990. Although these indicators were able to predict crisis in the past, they failed to anticipate the one occurred during the 1990. The authors then add that there were alternative indicators not used in the XRI that could have increased the range of recession probabilities. If an index with these alternative indicators had been built, the recession could have been predicted. These findings prove how the relationships between economic variables are not always constant. Variables that are leading indicators of the health of the economy during a business cycle can become lagging indicators in another one. Only two[16] of the seven leading indicators used to predict recessions from 1990 to 2001, had anticipated the 2007 recession (Nate Silver, 2014). Okun Law is another example of how interaction between economic variables can change over time. This one describes a relationship between GDP growth and unemployment. The main idea is that an increase in the GDP should be followed by a decrease in the unemployment rate. Although this relationship between these two variables still exists, the impact of the GDP on the unemployment rate has constantly varied over time (Knotek, 2007). The author uses a rolling regression to perform such study. From the paper, it can be concluded that the coefficient has not been constant. High period of fluctuations have been registered from the first years of the 1970 until the mid-1990s. From the 1997, Okun's coefficient has shown a tendency to constantly increase, being close to -0.040 in the

[14] This approach is the most common in all the main methodologies used to describe time series process. A simple example is given by an autoregressive process of order (p), where is assumed that the future value of a variable is a combination of the (p) past variables.

[15] It is an index predicting the probability of a recession six month ahead. It is built from the econometric model developed by Stock and Watson (1989).

[16] Namely, housing prices and temporary hiring.

last period of the regression. Underestimating the complex nature of the economy and its constant changes can result in big macroeconomic forecasting failures.

1.2.3 Biases

Although the economy is already an extreme complex environment, there also human reasons that can increase the probability of forecasting failure. One of the most known is the overconfidence. People tend to show beliefs about future events that are not in line with the actual probabilities. They are prone to stick to their vision even if new information shows the opposite. Clements (2003) have run several regression-based tests to judge whether the inflation forecast probabilities of the survey of professional forecasters (SPF) differ from the actual probabilities. He has showed that the forecast distributions are flawed by a tendency to over-estimate the uncertainty and under-estimate the level of inflation.

No incentive to do better than other forecasters is another problem that affect forecast accuracy. Because of this, it is usually more convenient to simply adjust the prediction derived by forecasting model to be in line with the forecast consensus. This is especially true in the case of the SPF. Since every forecaster is anonymous, there is no great motivation to outperform the others forecasters. By using data on the level of disagreement in the SPF forecasts (McKnees 1992), it can be noticed that previsions have become less dispersed. Hence, the level of disagreement, namely how much forecasts are different one from the other, has moved toward 0.5 over the past 50 years. GDP, inflation and unemployment forecasts show the same pattern. One could be tempted to explain this homogeneity in the SPF previsions by assuming that all the forecasters have models that are somewhat similar. At the same time, the fact that every forecaster has the same amount of information could be another argument to explain why this homogeneity occurs (Zarnowitz, 1992). Although they are not wrong, these two explanations are not entirely true. Subjective or personal adjustments are another aspect to bear in mind when forecasting accuracy is analysed.

Although studies (McKness, 1990) have proven that judgemental adjustments can account for an increase of roughly 15% in the accuracy of the forecast, they introduce as well some form of bias. According to Laster and Bennett (1997) each forecaster, at the moment he provides the prevision, weights in two different goals. The first one and, most obvious, is to be as accurate as possible. None wants to provide forecasts that are completely useless because of their poor accuracy. Being inefficient could be costly in term of future working possibilities both for the company and the forecaster producing the forecast. The second goal is more "*egoistic*". Forecasts are not only the products provided by forecasting companies, but they are also instruments that can be used to increase the publicity of the company itself. Although irrational, it could be more efficient for a forecasting company eager to be noticed to come up with outlying forecasts. The risk to be wrong is considered less important than the possibility to gain publicity. Of course, the less is the company reputation the higher will be the incentive to act in this way. However, this is not the unique motivation that can explicate this behaviour. As it has been documented by Laster and Bennet (1997), the industry in which the forecasting company works has a huge impact on which of the two goals - accuracy and publicity - is deemed more important. On one hand, non-financial corporations using forecasts mainly for internal proposal are more concerned with accuracy. On the other hand, publicity is regarded more valuable for consulting firms. Media coverage is what they are after, so it is rational to produce forecasts that draw attention. Between these two extremes there are other categories, for which both accuracy and publicity are equally important. Econometric forecasting companies, for instance, aimed at producing accurate forecasts but they also want to consolidate their position in the market in which they operate. The same rationale holds for banks and brokerages. Publicity is valuable, but it cannot ruin the credibility of the company. Nature of the company is another explanation of the divergence from the consensus. Independent forecasters show more aggressiveness[17] in their predictions, while large corporations tend to follow more the consensus. As it has been showed by the literature (Lamont, 2002), the higher the company has to lose from their predictions the more conservative will be.

1.2.4 What does the data say?

An immense literature about macroeconomic forecasting accuracy exists. However, it has not been found any models or forecasting techniques able to outperform the others consistently yet. Moreover, forecast errors have been a constant over time. Data can provide some useful insights regarding this issue. Forecast errors cannot be explained only by model errors or biases in the process of forecasting. It is equally important to consider the role of the economic data. They are extremely important in the forecasting process. By using data with low quality, forecasters[18] risk to incur in parameter estimation errors. Wrongly estimates of parameters, increases the possibility to come up with predictions that are not accurate. Poor data can be the result of several factors. In this section I will introduce some of them. To begin with, the impact of data revision on forecast will be

[17] The mean absolute deviation (MAD) from consensus for individual forecasters is 0.82, while for industrial corporations is 0.37.

[18] Here I am not assuming that forecasters deliberately use data with poor quality. What I am saying is that they have no choice. For example, in order to estimate the next quarter GDP growth, they have to use economic data that are very likely to be revisited by statistical agencies.

analysed. Then, a more tricky issue will be introduced: the predictive power of economic data used to make forecasts.

1.2.4.1 The Data Revision Issue

One common problem with economic data is their continuous revisions[19]. Statistical agencies normally release data one or two moths later the economy activity has occurred. For instance, in the U.S., the Bureau of Economic Analysis publishes the *"advance estimate"* of the GDP, namely the first estimation of GDP over the quarter, roughly one month later the end of the quarter. Then, a second estimate and a third estimate are announced, respectively, two and three months later the end of the quarter. This implies that reliable data for economic activity of the first quarter is available only at the end of June. Similar patterns are for other macroeconomic variables. The magnitude of these revisions must be quantified in order to study how they affect economic forecasts. Croushore (2002) notices that the process of revision varies over countries[20] and across variables. How data are gathered, analysed and constructed influences their final revisions. Poor methods of analysing data, obviously, lead to inaccurate estimations, which are sensible to several revisions. For the U.S. GDP growth rate, the difference between the first and the final revision is around 1.2 percentage points[21]. Arouba (2008) estimates how large revisions are for several US economic variables. He compares the first estimation of a given economic variables with his revision three years later. By computing for each variable his signal to noise ratio[22], he notes that revisions affect variables in different fashions. The lower is the signal noise ratio, the less accurate are the estimates and so the bigger is the expected revision. Inflation displays a low signal to noise ratio (<0.2). Nominal output, real output, industrial production, real final sales and non-farm payroll

---

[19] Following the Croushore (2002) classification, data revisions can be either *information-based* (new information occurs and past data must be uptadet) or *structural* (changes in the methodology of managing the data).

[20] McKenzie (2007), found that the RMAR (defined as the ratio between mean absolute revision of variable and mean absolute growth rate of variable over time) is quite different among countries. Canada and Spain have the lowest RMAR (<0.2), while Belgium, Denmark, Italy, Japan, Norway, Netherlands, Czech Republic and Portugal have the higher RMAR (>0.4). UK and US have a RMAR between 0.2 and 0.4. As the author notices, these numbers should not be interpreted as a measure of poor quality in the data. They might, for instance, show an increase in the quality after the first release.

[21] http://www.bea.gov/national/pdf/revision_information/relia.pdf

[22] It is the ratio between the standard deviation of revision and the standard deviation of final growth.

employment have a modest signal to noise ratio (between 0.2 and 0.4). Labour productivity is the variable with the higher level of signal to noise ratio (>0.4).

Literature has proved that data revisions are systematic, but how much they alter economic forecasts? Stark and Croushore (2002) address this question. They notice that revisions of economic data affect forecasts in three separated ways. Firstly, they modify the inputs forecasters use to carry out predictions[23]. Secondly, they alter parameters estimations. Finally, they influence the choice of the lag structure of the model[24]. The authors show, by employing a simple autoregressive process of order 1 (AR(1)), the effects of data changes on forecasts. For an AR(1) model:

$$y_t = \mu + \phi y_{t-1} + e_t$$

The 1-step-ahead forecast with available data at vintage v – suppose this the first release - is the following:

$$y_{t|t-1,v} = \hat{\mu}_v + \hat{\phi} y_{t-1,v}$$

For the same horizon, but using a new set of data published at vintage w – this is the second release - the forecast is:

$$y_{t|t-1,w} = \tilde{\mu}_w + \tilde{\phi} y_{t-1,w}$$

Because the set of data changed, the parameters estimated are different. By subtracting the first equation to the second one, it is possible to assess the effect of the data revision on the forecast:

$$y_{t|t-1,w} - y_{t|t-1,v} = \tilde{\mu}_w + \tilde{\phi} y_{t-1,w} - \hat{\mu}_v - \hat{\phi} y_{t-1,v} \tag{1}$$

$$= (\tilde{\mu}_w - \hat{\mu}_v) + (\tilde{\phi} y_{t-1,w} - \hat{\phi} y_{t-1,v})$$

The first bracket contains the change in the mean of the process due to the data revision. The second bracket express the combined effect of the change in the parameter estimation and the different observations used to make the forecast.

---

[23] For short-term model assigning more importance to recent observations, this effect is the most perilous.

[24] Information-criteria used by forecasters to pin down the optimal lag is influenced by data revisions, because they give rise to new form of forecast errors.

As they authors notice, the final impact on forecasts is strictly related with the magnitude of the data revision. If it is modest, the two forecast will differ, *ceteris paribus*, by a small amount. Indeed, both the first and second bracket will be slightly different, so the difference between the two forecasts will be close to zero. Conversely, the bigger is the revision, the wider will be the discrepancy between the two forecasts. Hence, the lower will be the accuracy of the first forecast. Beside direct effects of data revisions on forecast accuracy, the literature has investigated if such revisions introduce news in the data or reduce the overall noise. Statistical agencies could either report its sample information or use other variables to produce optimal estimation of the true value of the data to be published. The actual methodology employed helps to answer to the question posed earlier. In the first case, the initial estimate of the data can be thought as an observation of the final data but with measurement errors. As time goes by and more samples are analysed, data revisions reduce these errors. Consequently, data revisions are correlated with earlier data. Moreover, they are predictable and cannot be thought as optimal forecasts of subsequent releases. In the second case, the statistical agency combines all the information available at that time to estimate the true value of the data. This implies that later data revisions can incorporate only new information. As a consequence, data revisions do not reduce the noise in data but add news. Literature has provided mixed findings.

In their seminal paper, Mankiw and Shapiro (1986) use U.S. GNP data from 1975 to 1982 to test the nature of data revisions. The authors find that provisional estimates of the final value are efficient forecasts[25]. Because of this, the authors have concluded that data revisions add new information instead of reducing the noise. Mork (1987) has shown that the second revision of US GDP increases the amount of information in the data. However, the initial release is a mixture of both news and noise. A solution to the issue of data revisions is offered by the combination of large factor models and principal component analysis. The main idea is to extrapolate the underlining, but unobservable, factors that move the variables. By assuming that data revisions are uncorrelated across variable, this technique is able to depurate forecasts from the negative impacts of data revisions.

1.2.4.2 Noise versus Signal in the Data

A second interesting point to study, concerning economic data, is how much information they embedded. Forecasting, at his core, aims to extrapolate information from data in order to produce

[25] Since the forecast is built using all the relevant and available information at that specific time, future revisions are uncorrelated with the revision itself.

future predictions. It can use both past observations of the variables of interest or other variables believed to have a predictive power. However, without meaningful information, every forecasting technique is inherently and intrinsically useless. If the data we employ are not expressive, they cannot convey any information. So, the question is: Do they macroeconomic data contain enough predictive information? As always, the answer is not so straightforward. To have a picture of the degree of information contained in data, is possible to use the signal to noise ratio. Before going through this measure, it is important to clarify what information mean in forecasting. Every data incorporates some form of information. These can be either valuable or not. This is why several methods have been created to extrapolate useful information in the data. Employing the statistical jargon, the part of valuable information conveyed by data is defined signal. The rest is labelled as noise. If the data we are dealing with are full of noise, it is an arduous task to grasp the underlying level of information. The consequences of such event are perilous, as will be reported later. A synthetic measure of the quality of the data can be computed by simply dividing the signal variance by the noise variance. This is exactly the signal to noise ratio. The higher it is the most informative are the data. The following example can help to understand how a low signal to noise ratio can influence a forecast. Suppose we are interested in estimating the parameter β of a simple linear regression. However, we believe the data we are using are influenced by some noise. For example, the data used are first estimation of economic variables that are very likely to be revised in the future. This means that:

$$\ddot{x_1} = x_1 + v$$ [26]

Here x_t can be thought as the *signal*, while v is the noise. Unfortunately, we can observe only $\ddot{x_t}$. So the linear regression model becomes:

$$y = \beta_1(\ddot{x_1} - v) + e$$
$$y = \beta_1\ddot{x_1} + (e - \beta_1 v)$$

Our goal is to estimate the parameter β using the information available at t as accurately as possible. The usual OLS estimation with measurement error yields:

$$\widehat{\beta_1} = \frac{cov(y, \ddot{x_1})}{var(\ddot{x_1})} = \frac{cov(\beta_1\ddot{x_1} + (e - \beta_1 v), \ddot{x_1})}{var(\ddot{x_1})}$$

[26] Where v is assumed to be normally distributed with mean 0 and variance σ_v^2.

$$plim\,\hat{\beta}_1 = \beta_1\left(1 - \frac{\sigma_x^2}{\sigma_x^2 + \sigma_v^2}\right)^{27}$$

$$\frac{plim\widehat{\beta_1}}{\beta_1} = \frac{1}{\left(\frac{\sigma_x^2 + \sigma_v^2}{\sigma_x^2}\right)} = \frac{1}{1 + \frac{1}{\sigma_x^2 / \sigma_v^2}} \qquad (2)$$

From the equation (2), the effect of signal to noise ratio on the estimation of the parameter is visible. The higher is the signal to noise ratio (σ_x^2 / σ_v^2), the closer to 1 the right hand side of the equation will be. However, if the ratio is low, the term on the right will be different from zero and the estimated parameter will be inconsistent with the optimal one. Inconsistent estimated parameters are more likely to yield flawed forecasting performances.

An interesting approach to study the information quality of the economic data is proposed by Ormerod (2000). By using findings in the Random Matrix Theory, the author explores whether economic data are statistically different from a random process. The idea behind the study is the following. If the data[28] available to forecasters are generated by a random process, past information are not functional for predicting future values. The test is structured in the following way. For 17 advanced economies, the annual rate of growth of the GDP from 1870 to 1994 is computed. Then a delay matrix[29] with lags up to 12 years is built for each of the economies. The lag has been chosen to be consistent with the literature estimation of the length of the business cycle. From the delay matrix, $\mathbf{Z}$, the correlation matrix, $\mathbf{C}$, is computed.[30] This is one is supposed to contain all the relevant information inside the time series. The eigenvalues of the correlation matrix are calculated several times in order to create an interval. The range of the eigenvalues is then compared to the range of eigenvalues of the other correlation matrix derived by a random matrix. For a random matrix of order *TxN*, the range of the eigenvalues of its correlation matrix is given by the following formula:

---

[27] Since: $Cov(\ddot{x}_1 v) = Cov(x_1 + v, v) = \sigma_v^2$ and $Cov(\ddot{x}_1, e) = 0$.

[28] Past observations of GDP.

[29] For a given vector Tx1 vector, x, the delay matrix, Z, is a Tx(M+1) matrix containing m lags of the vector x. In this paper, the delay matrix has an order of 112x13.

[30] $\boldsymbol{C} = \frac{1}{(T-M)}\boldsymbol{Z}'\boldsymbol{Z}$

$$\lambda_{(max,\min)} = \sigma^2 \left(1 \pm \frac{1}{\sqrt{Q}}\right)^2 \tag{3}$$

Where σ^2 is the variance of the elements of the correlation matrix and Q is the ratio of T and N.
For a correlation matrix derived by a random matrix with the same order of the delay matrix $\mathbf{Z}$, the equation (3) implies that the eigenvalues should have a value between 0.329 and 2.005[31]. Surprisingly, only for the UK (2.181 and 2.055) and US (2.174 and 2.051)[32] the eigenvalues lie outside the calculated range of those of a random matrix. The dominant economic position firstly of the UK and then of the US, trough the period analysed for the study, is advanced as a possible explanation by the author. Because of their economic power, the two countries have been able to operate as an isolated and, hence, less random system.
The author extends the analysis further. Focusing only on the US and UK, quarterly data[33] of the GDP growth has been used to carry out the same analysis described earlier. The results are mixed. As for the US, the eigenvalues of the correlation matrix fell well outside the range of the ones implied by the random correlation matrix[34]. Conversely, the eigenvalues of the UK correlation matrix are not different from the ones of the random matrix. This suggests that just the US data contain some level of information that can be exploited to predict future value. As I mentioned earlier, having data largely corrupted by noise is perilous for researchers.
So, what are the dire consequences of dealing with noisy data? In his speech at the University of London, Ben Broadbent (2013) discusses some of these consequences. To begin with, noisy data make difficult to distinguish between good models and bad models. Assuming a signal to noise ratio of 0.5, a model that is 100% wrong[35] has a 40% chance to outperform the real model. As the SNR increases, as expected, the probability that the wrong model does better than the good one decreases more rapidly. The author continues by illustrating that, in order to confidently reject the wrong model, we need vast sample of data. However, when it comes to macroeconomic data, we can count only on a small subset of reliable data. This makes the work of forecasting even more difficult. Second, noisy data makes the task to predict rare events or shifts (*turning points)* problematic. Models, owing to different estimated parameters, produce diverse probability or range of probabilities for some future events. How it is possible to sort the best one out? In an environment

[31] These values are corrected for small-sample bias, since the formula (3) is valid in the limit case.

[32] In parenthesis are reported the maximum eigenvalues of both UK and US.

[33] For the US, the sample of data goes from 1947Q1 to 1999Q3, while for the UK from 1955Q1 to 1999Q1.

[34] Using the formula (3) with, as always, m equals to 12, the range of the eigenvectors is .209 to 2.308.

[35] For example, in the univariate case, a model whose estimated **B'** is two times the real **B (B'=2B).**

of noisy data and where frequencies of rare events are rare, thousand of years of data are needed to exactly pin down the model showing the more accurate probability.

Finally, mistaking correlation with causation is very likely when noisy data are used. This has become truer with the boom of Big Data. Nowadays, thousands of economic data are available just one-click away. So the chance to find apparent correlation between variables has become more frequent. If noise is mistaken for signal, the model estimated will be corrupted. All these issues make difficult to properly differentiate a good model from a bad model. This is why sometimes the great accuracy of a model could be explained in term of good luck and not in term of superiority.

Chapter 2: Prediction Markets

In the following chapter prediction markets will be introduced. There is a vast literature about this argument. Indeed prediction markets have been used for several reasons and in several fields. In order to be consistent with my research question, features of prediction markets about macroeconomic events have been investigated. The entire chapter has been built on three questions: What, Why and How. First, what prediction markets are? In this section, the structure of these markets will be studied. Few words will be spent also on how negotiations occur and what are the types of contracts exchanged. Second, the Why question will be tackled: Why do prediction markets could help macroeconomic forecasting? In this section, will be highlighted prediction markets features that could be used to improve the forecasting science. Finally, it will be analysed how these features can be implemented in simple forecasting macro-econometric models. Since several points of contact exist between the second and third question – Why and How-, I have decided to do not discuss them independently. In this way, I have tried to do not repeat the same concepts twice and give more flow to the text.

2.1 What prediction markets are?

Prediction markets are markets in which people can bet on future events. Economic agents can bet virtually on everything. From the next US president, to how much dollars a new Hollywood film will gain at the box office or to the next level of US unemployment. In prediction markets, people interact between them by exchanging contracts. These are related, as I said above, to future events. If the event underlying the contract occurs, the "*investor"* will be paid one dollar for each contract he bought. Conversely, the contract will pay out zero yielding a loss for the investor equal to the sum of money spent on the contracts. For instance, at the moment of writing, the contract linked to the victory of Hilary Clinton at the next democratic primary is 0.85 $. So, if Hilary Clinton wins the primary – and the underlying event occurs-, the contract would yield a profit of 0.15$ for each contract bought. This type of contract is defined as *winner-takes all* in the literature and it can be thought as a binary option (Wolfers, 2004).
Although this one is the most traded, two other contracts exchanged on prediction markets exist. The first on is an *index* contract, for which the payoff depends on the final value of the underlying event. For example, an *index* contract pays 1$ for every 1000 dollar cashed by a film at the box office. The second form, is a *spread* contract, which costs 1$ and pays 2$ if the value of the underlying event is above a given threshold. For what concerns economic events, the *winner-takes*

all contract is the most employed due to simplicity and clarity reasons. So, I will discuss the other main features of this contract only (See Wolfers 2004, for other contracts uses).

Before going through the information that can be derived from these contracts, it is useful to spend some words on the way auctions are run. Although the other contracts are traded mainly with the use of continuous double auction, *winner-takes all* macro-economic contracts are traded following a pari-mutuel system. The payout is unknown until the end of the underlying event of the contract. All the money betted in the auction are summed and then distributed to the winner. To clarify this point, suppose that a prediction market about the next US GDP quarterly growth is run. It is not required to exactly predict the value of the GDP, but the interval in which it will lie. Let's say that intervals are 0.5 wide. Once the market is open, economic agents bet on the future level of GDP growth according to their beliefs. So, if an agent believes that next-quarter GDP growth will be 0.6, he will buy a given number of the contract whose growth interval is 0.51-1. All the other agents act following this procedure. The market is closed 1 minute before the GDP growth announcement and all the money betted are summed in a pool. The final sum is 100$. After the announcement, GDP growth is revealed to be 0.7. As a consequence, the agents that have bought the contract whose range is 0.51-1, will split the entire sum. If, for example, 25$ of 0.51-1 range contracts have been bought, investors will receive 4$[36] for each dollar wagered. As I said earlier, the final equilibrium-winning price is unspecified until the end of the auction. Of course indicative prices are quoted during the auction, but they can continuously vary over time in response to new exchange of contracts.

As Wolfers and Gürkaynak (2006) have evidenced, the pari-mutual system helps to increase the liquidity of the market. Indeed, buyers and sellers must not be matched as in a normal market. If economic agents bet on different outcomes of a future event, the functioning of the market will be ensured and clearing will be feasible. Liquidity is a fundamental issue in prediction market, as it will be discussed in a couple of pages. The more liquid is a market the more it will attract new participants. This in turn will increase the amount of information in the market by averaging different beliefs about the outcome of a future event. More information might not result in more prediction accuracy, but it will surely increase the level of informativeness of the market.

The structure of economic prediction markets and the type of contracts have been introduced, but no words have been spent on the economic events underlying these contracts. From 2002 to 2007, the largest economic prediction market was the one jointly run by Goldman Sachs and Deutsche

---

[36] The 4$ gain is computed in the following way: $\frac{\sum_{i=1}^{N} x_i}{\sum_{j=1}^{M} y_j}$, where N is total number of bets, M is amount of wagers in the winning contract. In our example, the final winning price is equal to: $\frac{100\$}{25\$} = 4\$$.

Bank. In this market, it was possible to trade contracts tied to several US macroeconomic variables: non-farm payrolls, ISM purchasing manager index[37], initial jobless claims and retail sales. The number of *auctions* for each of these events was different. Non-farm payrolls auction was run two different times: the first on the day before the event and the second on the same day of the announcement. Initial jobless claims, retail sales and ISM purchasing manager index had just one auction on the same day of the announcement.

The possibility to trade on the same day of the event is an important thing to be noticed. To begin with, it allows factoring in the prediction all the relevant information regarding the event that will be traded. This, in turn, might lead to an increase in the overall forecasting accuracy and efficiency. Secondly, it reduces the premium risk required from investors wanting to hedge again the underlying event risk. This second point is important, as it will be reported in the next section, in the analysis of contract prices as beliefs about the outcome of an event.

2.2 Why and How prediction markets can help macroeconomic forecasting?

In the last section, the structure of macroeconomic prediction markets and how contracts are exchanged have been examined. However, the question regarding why prediction markets should be helpful for macroeconomic forecasting is still pending. The answer to this question can be found in some peculiar features of the prediction markets.

In the first part of this section will be explained why prediction markets are able to incorporate efficiently information and provide accurate prevision of future events.

The second section will deal with another question: how relevant information embedded in the prediction markets can be used in a macroeconomic forecasting context.

2.2.1 Why prediction markets data should be used?

2.2.1.1 Efficiency in the aggregation of information

Economic markets have been established to answer to different needs. The most obvious one is to allow trading between two or more people, buyers and sellers. The reasons behind the trade can be different: exchange of goods or capitals, allocation of resources, necessity to hedge from some risk or possibility to make a profit. By the interaction of different economic agents, markets incorporate information. For instance, a drop in the price of oil could express the discovery of a new fuel that can be used for alimenting automobiles.

[37] It is a monthly index reflecting the sentiment of companies. It takes into account orders, inventories, employment and production variations.

In predicting the next future, information has a huge role. In an uncertain system as the economy, the most accurate and reliable is the level of information in possess, the highest should be the accuracy of predictions. However, extracting information from markets is not always simple[38]. Prediction markets offer a solution. They are very likely one of the best form of market for aggregating information (Tetlock 2004, Wolfers and Zitzewitz 2004 and 2012). Because of their structure, participants in these markets have monetary incentives[39] to seek out as more information as possible about the event they are betting on. It must be noticed that is not always easy to get by valuable information. This could be sound odd in a world where by simply pushing down letters on a computer is virtually possible to discover any news. However, two problems arise. First, discovering information could be costly. Not all-relevant information is free. Some forms of payment are required to get access to news or database. Second going through all the information is time consuming. Because of the vast flow of news, it is not possible to examine carefully all of them. So without incentives, information might be overlooked. However, in prediction markets, bettors are motivated to find and use new information, because they can get a profit out of it.

By having more accurate information than other participants, they can provide better predictions and, in turn, increase the possibility to gain money. Although information seeking is performed also in other markets, prediction markets differ from these by offering incentives to reveal information. To elaborate this point, it is useful to treat it analytically. Wolfers and Zitzewitz (2006) have shown that by maximizing the following expected utility:

$$Max\ E(U_j)_{x_j} = q_j\ Log\left(y + x_j(1-\pi)\right) + (1-q_j)Log(y - x_j\pi)$$

The optimal quantity of contracts a bettor j should buy is:

$$x_j^* = \frac{q_j - \pi}{\pi(1-\pi)}y \tag{4}$$

where $q_j, \pi\ and\ y$ are respectively the prior belief of the participant j about the event, the price of the contract and the bettor income. From the equation (4) – that will be analysed more accurately

[38] From markets full of noise traders, which do not trade on any particular information, the task of getting information out of the market is arduous to say at least.

[39] Wolfers (2004,2007) has tested whether using play-money instead of real money can affect the accuracy of prediction markets. By comparing two prediction markets, one with real money and the other one with play money, the author has not found any discrepancy between the two.

later on -, it can be notice that if the bettor hides his belief q_j[40] he will not buy the optimal quantity of contracts, so it will under or over exposed himself to the uncertainty of the outcome. For example, by buying more contracts, that is over-predicting the probability of the outcome with respect to his belief, he will not optimally weight the expected return and the risk of the bet. He will run some risk not backed up by his belief. Essentially, he will act irrationally.

So, the only way to avoid this problem is to reveal his belief. This point should not be overlooked. Indeed it is probably one of the most important implications of prediction markets. By being well-informed, by acting coherently with their beliefs and by having economic incentives to do so, participants are lead to provide predictions as accurate as possible. It is possible to argue that their loss function and their accuracy function are the same thing. Although it could be straightforward, this is not always the case in forecasting. As it has been reported earlier in the text (Laster and Bennet, 1997) forecasters could not have the incentive to be as accurate as possible. The sticking to the consensus *bias* or the need to stand out from the crowd to increase their publicity, are two factors that can affect the accuracy of forecasters predictions. Before moving to the consequences stemming from the power of aggregate different information, it is worth to give another look to the equation (4). As it has been said earlier, this identifies the optimal amount of contracts that a risk-neutral investor with log utility is supposed to buy. By equating the supply to the demand of contracts (see Wolfers and Zitzewitz, 2006), is possible to show that:

$$\pi = \int_{-\infty}^{\infty} q f(q) dq = \overline{q} \tag{5}$$

So, the market-equilibrium price reflects the participants average beliefs. This gives a theoretical foundation for interpreting the price of a contract as the participants market probability of the underlying event. However such finding should be taken with some grain of salt, as the authors have noticed.

This equivalence is driven from the assumption of log utility preference. Some form of divergence - even if small[41] - between price and average beliefs can arise for different reasons. First, strongly dispersed beliefs[42] among individuals can lead the equilibrium price away from the mean of beliefs.

[40] Either by pretending to be more ($\tilde{q}_j > q_j$) or less confident ($\tilde{q}_j < q_j$) than how he actually is.

[41] It must be noticed that the magnitude of this divergence is very unlikely to be higher than .05.

[42] In Wolfers (2007), beliefs are drawn from a uniform distribution. For strongly dispersed beliefs, the author refers to a distribution with a range of 0.2 or more.

Second, aversion to risk is another important component to be taken into account. On one hand, risk-neutral investors tend to overestimate the probability of event with low likelihood and underestimate the probability of likely event. On the other, risk-averse investors behave in the opposite way. They have been proved to overestimate events with high probability and underestimate events with low probability. This could explain why differences between beliefs and equilibrium price occur in correspondence of events with extremely high or low probability. Gjerstad (2007) has found similar results regarding the equivalence between equilibrium price and mean belief. Moreover he has proved how with high level of risk aversion, bettors mean beliefs is equal to equilibrium price.

Manski (2004) has found out an apparent controversial finding. He has showed that, in an environment with heterogeneous beliefs and risk neutral investors, the equilibrium price should not be interpreted as a mean belief but as a bound on the mean belief. Wolfers and Zitzewitz (2006) have settled the issue by demonstrating that Manski's results stemmed from strong assumptions[43].

Wolfers and Gürkaynak (2006) have implemented these findings by investigating the presence of risk premium in prediction markets. In a context where equilibrium prices reflect the underlying probability of events, it is interesting to study if risk premium can influence this relationship. In order to perform this analysis, it has been computed the equilibrium price that a representative agent dealing with a risk is supposed to face:

$$\pi = \frac{p}{p + (1-p)\dfrac{U'(w)}{U'(\beta w)}} \tag{6}$$

where p is the probability that the risk takes place, β is the effect of the risk on the wealth of the agent and w is the wealth. From the analysis of the equation (6), it is possible to identify four different cases:

a. $\boldsymbol{\beta = 1}$. If the wealth remains constant in the two states no risk is present. So, no risk premium vanishes. As a consequence, contract price is equal to the probability of the event.
b. $\boldsymbol{U'(\beta w) = U'(w)}$. Under risk neutrality assumption, it is guaranteed the equivalence between price and probability.

[43] By assuming risk neutrality, all the investors bet their wealth whenever the price is different from their beliefs. The breakdown in the linearity between beliefs and number of contracts bought is responsible for the wedge between equilibrium price and market beliefs.

c. $\boldsymbol{\beta > 1.}$ If a positive shock to the wealth occurs, a risk-averse investor will under-price the value of the contract with respect to the probability of the event: $\boldsymbol{\pi < p}$.

d. $\boldsymbol{\beta < 1.}$ In response to a negative wealth shock, a risk-averse investor will hedge himself. As a consequence, the price of the contract will be higher than the actual probability of the event: $\boldsymbol{\pi > p}$.

To summarize, the price distribution of a risk adverse investor will diverge from the price distribution of risk neutral investor as a consequence of two factors: magnitude of shock (β) and degree of risk aversion ($\frac{U\prime(w)}{U\prime(\beta w)}$). This difference could be explained as premium risk. However, what do data say about this difference? Do prediction markets equilibrium price contains some form of risk premium? In order to answer to these questions, the authors regressed[44] the variation of the S&P 500 occurring immediately after the announcement of the data on the forecast error for each economic derivative contract. The idea behind this approach is the following. The difference between the predicted value and actual value of the macroeconomic variable the investor desires to hedge against is used as shock, while the change of the S&P 500 is thought as the effect produced by the shock on the wealth[45]. This approach allows having an estimate of the β[46] in the equation (6) that can be, in turn, used to assess how much equilibrium prices diverge from event probabilities. The higher is the shock on the wealth of the investor and, consequently, the higher is the risk premium demanded, the larger is the divergence between price and probability. Authors have found that the risk premium is very low for all the macroeconomic derivatives but the non-farm payrolls. This exception could be explained because the large impact of non-farm payrolls on the investors wealth. However, the authors notice that for normal level [47]of risk aversion, the discrepancy between price and probability should be around 4%.

2.2.1.2 The implications of efficient aggregation of information

Prediction markets have been proved to be able to efficiently aggregate information. From this, numerous consequences occur. First, there is a rapid incorporation of information in the prices of the contracts (Wolfers and Zitzewitz, 2004). Because the contracts exchanged in prediction markets are designed to be clear and not vague, only relevant-to-the event information is included in the

[44] $\Delta S\&P500_t = \alpha + \beta x_t$, where, $x_t = Actual_t - Forecast_t$.

[45] This approach implies that the wealth of the agent is invested in the S&P portfolio.

[46] According to the authors, the relationship between the regression beta, $\hat{\beta}$, and the equation beta, β, is the following: $\beta = 1 + \hat{\beta}$.

[47] For higher level (<20) of risk aversion, the distance between prices and probabilities is more consistent.

prices. This is due to the fact that only informed take part to the market. Although desirable, this feature has also a drawback. Uninformed traders, what Fischer Black (1987) defines as *noise traders*, are mainly excluded from the markets. This reduces, in some cases, the liquidity of the market and, consequently, the interest in it could decrease.

Second, prediction markets are difficult to manipulate (Wolfers and Leigh 2002, Rhode and Strumpf 2008). It could be legitimate to think that by betting large amount of money would be possible to influence the behaviour of the other bettors[48]. However, this does not happen in prediction markets. If bettors realize that price movements are not driven by new information, they start to trade between them to exploit differences between their beliefs and the prices. This would lead to a new equilibrium price that is in line with the one before the attempted manipulation.

Third, in addition to absence of manipulation, contracts do not offer arbitrage options. If the market probability of a given outcome increases, the opposite outcome will be deemed more unlikely. Moreover, contracts prices tied to the same event but exchanged in different prediction markets tend to move together.

Finally, prediction markets are proved to comply with the weak market efficiency condition. As well known, this implies that past returns cannot be used to predict future ones (Leigh, Wolfers, Zitzewitz 2003).

To summarize, it has been reported why prediction markets efficiently aggregate information. A synthetic measure, namely the equilibrium price of the auction, can be used to infer the market belief about the outcome of a future event. This one, due to the structure and the functioning of prediction markets, tends to be more accurate than surveys and pools in predicting future events. Wolfers and Gurkaynak (2006) have scrutinized data[49] from the Economic Derivatives Market run by Goldman Sachs and Deutsche Bank. The authors have compared these data with survey predictions, in order to assess which one was more accurate in predicting the outcome of macroeconomic variables. Prediction market forecast have been proved to be slightly more accurate than survey predictions. Some interesting findings have been explored.

First, prediction market forecasts showed lower mean absolute error and root mean square error. Second, they were more correlated with the realized value. Third, the authors have regressed both

[48] I believe that trying to manipulate the market could increase the liquidity of it by introducing noise trading. Although counterintuitive at a first sight, attempts to manipulate could be beneficial for prediction markets.

[49] Because only primary data for this market were available, the sample of the study consisted of 152 observations. Unfortunately, due to shutdown of the Economic Derivatives Market, it has not been possible to carry out other studies.

the economic derivatives and survey predictions on the actual macroeconomic value to test which predictor was more informative. Economic derivatives outperformed surveys[50].

Finally, the authors have investigated whether frequent biases in surveys were also present in prediction markets. It has been show that almost all prediction market forecasts are efficient in the sense that errors were not predictable[51]. One possible explanation for this slight superiority in accuracy could be explained by the proximity of prediction market auctions to the event to be forecasted. Indeed, as it has been written above, auctions are usually run the day before or on the same day of the announcement. However, this explanation does not seem sufficient. Efficient information aggregation and the fact that each participant has to make a bet on his prediction are elements that should not be overlooked in explaining this edge in accuracy.

2.2.1.3 Prediction Market Drawbacks

Before introducing how forecasters could use this information in forecasting macroeconomic variables, it must be analysed also what can hinder prediction market success.

First of all, contract questions should be posed to promote the interest of as vast as possible public, including *noise traders*. Ideally, the object of the market should be something for which everyone has an idea or belief about it. Heterogeneity of beliefs among traders is pivotal in prediction markets. Having different ideas is what drives people to trade[52] between them. It is what keeps prediction markets alive. If everyone shared the same belief, it is very unlikely that any form of trade would occur. People with identic beliefs will never enter into a trade, because none will take the opposite position in the exchange. They will be either all buyers or all sellers. This is why is fundamental to attract heterogeneous individuals to the market. This leads to an active market, where participants are more prone to look out for new information and they react more rapidly to them. As a consequence, prediction accuracy is boosted.

As noted by Wolfers and Zitzewitz (2006), it is quite a challenge to structure a prediction market that effectively weights interest and contractibility. Another element that can reduce the interest in a prediction market is insider information. If participants believe that someone might have private valuable information, they will likely not enter in the market because the possibility to make money

---

[50] Interestingly, survey coefficients were negative for all macroeconomic variables but business confidence.

[51] Only in the initial unemployment claims market, forecast errors were in some form predictable. According to the author, this flaw can be explained by the illiquidity of the market.

[52] Of course there are other reasons can be found to explain trade: entraintement, risk-loving or desire to become richer.

is very low. Heterogeneity of beliefs is vital for prediction markets, but private information can kill them.

Absence of information is also a source of failure for prediction markets. Wolfers and Zitzewitz (2006) have shown this point by using a prediction market on the likelihood of mass destruction weapons in Iraq. Although the question was of public interest – so the market was liquid -, participants wrongly overestimated the possibility to find mass destruction weapons. The inaccuracy of the prediction can be explained by the fact that no relevant information at all was available. As it has been said earlier, prediction markets are proved to provide accurate predictions. However, a form of bias it has been found in them (Wolfers and Zitzewitz, 2004, 2007 and Snowberg ,Wolfers and Zitzewitz 2012). It has been shown how prediction markets participants have issue in understanding low probability events. Under some circumstances, they tend to under-bet high probability events and over-bet low probability outcomes[53]. As a result, the equilibrium prices of these contracts are not in line with the actual probability of these events.

2.2.2 How prediction markets data should be used?

From all the evidences reported above, it is possible to evidence some interesting features of the prediction markets. First, prediction markets force participants to *put money where their mouths are*. As it has been evidenced, this increases the accuracy of the prediction. Second, because participants are allowed to bet on different outcomes tied to the same event, a probability distribution function can be derived. Indeed, economic questions in prediction markets cannot be posed with single *winner-takes all* contracts. For every auction, several contracts with different ranges linked to the same economic event are issued. By looking at the equilibrium prices of these contracts is possible to have an idea of how much the market values the probability of each outcomes. Depending on the contract ranges and the nature of economic variable, the number of contracts can vary from 12 to 19. In an uncertain system as the economy, probability distribution functions are valuable. Sometimes is difficult to interpret point forecasts. They are simple numbers, so they could be misleading. To get around this problem, predictions should be put in relation with other meaningful variables. One of these is economic uncertainty. To support this point, I will

[53] It is the so-called: *favorite-longshot* bias. This tendency has been vastly discussed in the literature. Possible explanations have been found in the love for risk for typical of some individuals or in the difficulty for economic agents to distinguish between small and tiny probabilities.

present an example. Suppose that I am given a forecast about the level of GDP growth in the next quarter.

The first thing that I could do is to compare this number with predictions of other forecasters. It could be either in line with the other predictions or different. But what it is really important to understand predictions is to identify how much uncertainty there is in the economy (Ericson, 1999). The higher it is, the most likely predictions could be wrong. One way to do that, could be to measure the magnitude of difference between forecasts. Here, precaution is needed. The difference between the number I am given and the other forecasts should not be interpreted as uncertainty. It is disagreement[54]. Literature has largely shown how disagreement is not a good proxy for uncertainty (D'Amico, 2008). So by comparing different forecasters, I cannot derive a measure of economic uncertainty that I can use to assess different predictions. What I really need is how much economic agents, in terms of probability, value outcomes different from the mean. That is, how much probability events are dispersed. This is where prediction markets step in. They provide this information through the implied probability distribution function of the auction. If the dispersion is limited, *ceteris paribus*, it is legitimate to be more confident of the goodness of the forecast, because the expected error implied from the probability distribution is low.

2.2.2.1 A Macroeconomic Prediction Market: The case of the EIX

Prediction markets have been proved to provide successful forecasts in several fields (O'Leary, 2011). However, applications to macroeconomic forecasting have been rare. An interesting approach has been proposed by Teschner et al. (2011). They have set up a prediction market in which economic agents interacted between them to predict the future value of some macroeconomic variables[55]. Although the final aim was to aggregate different beliefs and have market forecasts, the market created by the authors was different from usual prediction markets under some aspects. First of all, auctions were not held on the same day of the announcement of the economic variable, but for each event there was a continuous double auction. This lasts from the day after the prior announcement to the day before the new announcement. This feature is extremely important. It

[54] As it is presented in Wolfers and Gurkanyak (2006), disagreement could be low even in an uncertain economic context. Although it could be counterintuitive at a first sight, forecasters tend to fluctuate around the consensus when the level of uncertainty is high. As a result, predictions are homogenous one with the other and the level of disagreement is low. This is also why disagreement is not able to mimic uncertainty.

[55] GDP, Exports, Inflation, Investments and Unemployment.

allows investigating how the accuracy of market prediction forecasts evolves over time[56]. Second, instead of binary options – *winner-takes all options* -, contracts with linear payouts[57] were exchanged. According to the authors, the introduction of these contracts aimed at solving two common issues in prediction markets. First, by using one single contract instead of several contracts representing different outcomes tied to an economic event, liquidity in the market was enhanced. This, in turn, promoted a dynamic market in which traders actively interacted and information flow was more fluid. Second, because a single contract was exchanged in the market, the potential bias stemming from arbitrary range event intervals was solved (Sonneman et al., 2008).

Although transactions were carried out with play-money, monetary prizes were handed out on monthly and yearly basis. This feature has helped to address a twofold goal. First, monthly prizes kept participants constantly active on the market. They did not simply place an offer in the market and wait until the announcement, but they challenged one with the other in order to outperform all the other participants. Prizes were given to those that both increased the value of their portfolio and placed at least five orders on the market over the past month. The possibility to cash-in money over a short horizon has been a massive determinant in strengthening the motivation of participants. Second, because of monetary incentives investors were driven to reveal truthfully their information. Once the market was closed, the authors have scrutinized all the data and compared the market forecasts with the ones made by Bloomberg. In order to investigate the overall precision of predictions, several accuracy indicators[58] have been computed in three different points of time[59] prior to the first release of the macroeconomic data. From this analysis, two elements stand out. Firstly, as it is evident from the error measures computed by the authors, prediction markets outperform largely Bloomberg predictions both at 10 days and 1 day before the outcome. Secondly, the accuracy of prediction market does not constantly decreases over time as one might expect. The mean absolute errors increases during the first interval of time, namely between 10 and 5 days before the release of the data, going from 1.22 to 1.32. However, in the 5 days interval between the announcement, it decreases notably reaching a minimum of 1.08. Another interesting point about

[56] Prediction markets about election or economic questions – i.e, "*Will the Greece leave the Euro?*- offer the possibility to carry out such analysis. However, because the auction was run on the same day of the announcement, it was impossible to study the accuracy of macroeconomic derivatives predictions over time.

[57] $p = 100 + \alpha g_{x,t}$, where $g_{x,t}$ is the growth over the period *t* of the variable *x*. That is : $g_{x,t} = \frac{x_t - x_{t-1}}{x_{t-1}}$. α was arbitrarly set equal to 10. It must be noticed that, under this prediction market structure, equilibrium prices cannot be identified as events probability. The price, p, can be thought as the most probable outcome, given all the available information at a given time, according to the market participants.

[58] Both the mean absolute error and the root mean square have been used.

[59] Predictions accuracy was studied at 10,3 and 1 days before the release.

the accuracy of prediction market can be found in the discrepancy between the MAE and RMSE. On the day before the release, the mean absolute error of the prediction market is 1.08, while the root mean square error is 2.4. This difference could imply the presence of large but infrequent errors made by the participants.

2.2.2.2 Prediction Market Forecasts: Predicting macroeconomic release

This study above presented is another strong evidence of the forecasting accuracy of prediction markets. A vast number of economic agents with different beliefs driven by economic incentives are able to provide predictions that are more accurate than the ones of professional forecasters. So, the next big question is: How forecasters should handle information arising from prediction markets? How they should use these data? Due to the complexity of the issue, there is not a definite answer.

The starting point is to circumscribe and identify the potential advantages, in terms of forecasting, of prediction markets. Then, see how to factor in these advantages in simple macroeconomic forecasting models. So, what are the strong points of macroeconomic prediction markets?

First of all, prediction markets allow deriving a complete probability distribution function of all the possible outcomes. This measure could provide useful information about the overall level of uncertainty in the economy. Something for which is not easy to find meaningful data about.

Second, for what concerns macroeconomic variables, prediction markets provide more accurate predictions[60] than surveys. Once these elements have been identified, the next step is to conceive how to integrate them in macro-econometric models. In the following lines, I will propose some ideas on how the get the most out of this information and implement it in simple forecasting models. In the next chapter, I will use the available data on economic prediction markets to test these proposals and assess how much they can be helpful for macroeconomic forecasting.

The first group of implementations stems from the superiority of prediction markets in providing predictions with respect to surveys. Knowing that the prediction market about the economic release of interest efficiently aggregate different beliefs, it is legitimate to exploit this information. Hence, I can set up a simple linear regression like the following:

[60] In this context, I am referring to macroeconomic variables object of contracts in prediction markets. Namely, Macroeconomic Derivatives.

$$y_t = \alpha + \varphi y_{t-1} + \beta x_t + e_t \text{[61]} \tag{7}$$

In this way, all the relevant pieces of information embedded in the prediction market[62] can be extrapolated and fit them in a simple autoregressive process of order 1.

Of course other models could have been used to perform the same analysis. However, the choice of an AR(1) has been motivated by two arguments. First, literature (Nelson and Plosser, 1982) has evidenced how several macroeconomic variables can be described as random walk process.

Second, autoregressive processes are easy to deal with and can be easily implemented. Once all the parameters of the model have been estimated, it is possible to forecast the next value.

Moreover, it can be interesting to study if prediction market forecasts can be used to make 1-step-ahead forecast of the economic variable of interest. This translates in investigating the accuracy performance of the following model:

$$y_t = \alpha + \beta x_{t-1} + e_t \tag{8}$$

Here, since the goal is to produce a forecast for the next period, variables have been differenced in order to deal with stationary variables. So, y_t represents the change in the release macroeconomic variable between t and $t-1$, while x_{t-1} is equal to difference in the prediction market forecasts between $t-1$ and $t-2$.

2.2.2.3 Prediction Market Forecasts: Predicting the unemployment

It is also intriguing to study if the information expressed in economic prediction markets can be used to forecast other macroeconomic variables. Unfortunately, the choice of which macroeconomic variables to forecast has been hindered by the available data about prediction market. Indeed, it does not seem reasonable to use, for example, initial unemployment claims to forecast the next level of inflation[63]. The risk of such analysis could have been to mistake

[61] Going on with the example of the text : $\boldsymbol{y_t}$ is the level of initial claims at time t, $\boldsymbol{x_t}$ is the equilibrium price of the auction and $\boldsymbol{e_t}$ is the error term.

[62] In this example, equilibrium price is what we are looking for. Since it is computed from the probability distribution of the auction, it summarizes all the information we need.

[63] The idea was to find prediction market variables theoretically tied to the macroeconomic variable to forecast.

correlation with causation. With this in mind, the most straightforward choice would have been to use non-farm payrolls and initial unemployment claims to forecast unemployment. However, since the initial unemployment claims number of monthly observation was limited, this variable has not been used. So, it has been set up only one model with prediction market data about non-farm payrolls.

$$u_t = \alpha + \varphi u_{t-1} + \beta x_t + e_t \quad (9)$$

[64]

where x_t is the prediction market forecast of non-farm payrolls at time t.
Starting from this model, two analysis have been carried out. To begin with, it has been compared the prediction accuracy of the *prediction market* model with the accuracy of a simple AR(1) model. Second, it has been plugged both the actual release of the non-farm payrolls and the survey forecast in the above model to study what happens to the estimated prediction market parameter. The idea behind this second analysis derives from the following question: Does the prediction market forecast lose all his information in a forecasting model when the actual release of the data and the survey prediction are used? This question will be answered in the next chapter.

2.2.2.4 Prediction Market Forecasts: Studying the financial volatility

The second application of prediction markets data to macroeconomic forecasting aims at exploiting valuable information about the sentiment of prediction markets participants to see if it can reflects economic uncertainty. In order to do that, past literature has focused on the standard deviation derived from the probability distribution of the underlying event of the prediction market. As it has been mentioned earlier, just a handful[65] of forecasters offer probability distribution of their forecasts. This measure is crucial in evaluating the degree of underlying uncertainty related with the forecast. If the probability distribution is largely dispersed around the point forecast, the lower will be the level of confidence that a forecaster has in his prediction. By not having a single forecast point but an entire distribution with measure of probability attached to every event, public can better assesses the value of the forecast. Wolfers and (2006) have tested whether this variable,

---

[64] Going on with the example of the text : $\boldsymbol{y_t}$ is the level of initial claims at time t, $\boldsymbol{x_t}$ is the equilibrium price of the auction and $\boldsymbol{e_t}$ is the error term.

[65] Bank of England and the Survey of Professional forecasters publish along with their point forecasts, density forecasts for macroeconomic variables as inflation and GDP.

namely the time series of the standard deviation of a given economic auction, reflects the underlying uncertainty in the economy[66]. In order to do that, the authors have regressed, for each macroeconomic derivative, the time series of the standard deviation on the values that the VIX index had three months, two months and one day before the announcement. They have discovered that the level of uncertainty in single prediction markets is not able to explain the past and current volatility[67] of the economy, but it seems more related to the specific data release. However, I believe that could be interesting to expand this analysis a little bit further.

Following Andersen et al. (2003), I have computed the forecast error for each auction and for each variable. This measure has been used to proxy the surprise effect of the release[68]. Second, in order to have a form of comparison between the auctions, the standard deviation of all the forecast errors – surprise effects - has been computed. Third, the forecast error has been divided by the standard deviation of the surprise effects.

$$VIX_t - VIX_{t-1} = \alpha + \beta \underbrace{\left(\frac{\left(x_t - \hat{x}_{t|t-1}\right)}{\bar{\sigma}}\right)}_{\xi_t} + \epsilon_t$$

$$\Delta VIX_t = \alpha + \beta \xi_t + \epsilon_t \tag{10}$$

The surprise effect is supposed to incorporate all the main relevant pieces of information about uncertainty that can be taken out of prediction markets. Finally, this variable has been regressed on the change in the VIX closing price occurred between the opening price on the day of the announcement and the closing price on the day before the announcement.

2.3 What it has not been done?

Unfortunately, due to the closure of the macroeconomic derivatives market, other interesting analysis could have not been performed for the absence of data. For instance, if it had been possible to obtain data about the GDP auction, then it would have been easier to test the following studies. First, since GDP auctions were held respectively two months before, one month before and on the

[66] The authors have used the VIX as a proxy of the economic uncertainty.

[67] In this case is used the term volatility instead of uncertainty, because it is referred to the VIX index. However, the volatility of the VIX can be thought as an index of the uncertainty in the economy.

[68] If it is true that prediction market price reflects all the available information, market movements should be explained only by the difference between the forecast and the release.

same day of the announcement, it could have been possible to assess how the accuracy of prediction markets evolve over time. Indeed for the aforementioned macroeconomic prediction markets this was impossible because, due to liquidity issues, only two auctions close to the release day were run. As reported earlier, a similar study about the German GDP has been performed and some findings have been discovered. However, it would have been more fascinating to run the same the analysis for the US GDP since it would have been a more deep market.
Second, prediction markets could have offered monthly flows of information about the GDP growth that could have been exploited by now-casting models. Since GDP data about the current quarter are available with a lag of at least one month after the end of the quarter, forecasters virtually do not have any contemporaneous information about the GDP. However, prediction markets about GDP quarter growth could have solved this problem. Forecasters could interpret the prediction market forecast for each auction as a preliminary estimation of the first estimate of the GDP published by the statistical agency. Then they could insert this value in their model and update their forecasts for the next quarter GDP growth. Let’s make an example to explain this sequence. Suppose that we are on the first of January. The first estimate – *the advance estimate* - of first quarter GDP growth is going to be published by the Bureau of Economic Analysis in the last week of April. This implies that the first prediction market about the GDP growth will be held on the last week of January. In absence of other data, forecasters can feed their now-casting models with the prediction arising from the first auction and update their GDP forecasts. Then, they will do the same thing both for the second prediction market, two months before the release, and the last prediction market, on the same day of the announcement. Moreover, they can also assess how much this variable is able to reduce forecast errors. A quick look at the RMSE of the model before and after the use of prediction market forecast as input would reveal the impact on the overall accuracy of this variable. Combining prediction markets forecasts and official estimates would give forecasters 24 GDP observations (12 prediction markets data and 12 *statistical agencies data*) instead of 12. Consequently, forecasters would be able to use and exploit a monthly flow of GDP data, instead of relying only on quarterly data subjected to at least three annual revisions.

Chapter 3: Empirical applications

In the last section of the previous chapter, simple methods on how to incorporate prediction market information in macroeconomic forecasting model were introduced. This chapter will use the available data of prediction markets to test them. The structure of the chapter is the following. First, it will be analysed the available dataset, which is the one used by Wolfers (2004). Second, the ideas proposed in the last chapter will be analytically formalized. In this way, it will be possible to carry out proper tests to judge their value. Finally, the results will be discussed.

3.1 Data

Before going through the dataset used for the analysis performed in this chapter, it is worth to spend some words on economic prediction markets data. To begin with, contrary to other economic variables that are freely and easily available, economic prediction markets data are not so common. Only for the period that goes from October 2002 to July 2005, which corresponds to the first three years of the Macroeconomic derivatives market run by Goldman Sachs, available data exists. So the biggest issue concerning these data is the absence of a large database that can be used to perform robust statistical analysis about these markets. Because the sample size is not so wide, attention must be the guide to follow in the analysis of the results. Mistaking *noise* with *signal* or causation with correlation, are the most likely errors that could be made.

The database used in this study is exactly the same of the one of Wolfers (2006). It consists of 153 observations of four different macroeconomic derivatives exchanged on prediction markets in the period between October 2002 and July 2005. The variables are the following: non-farm payrolls[69], initial unemployment claims, retail sales and business confidence index. Each variable will be briefly discussed in the following lines.

Non-farm payrolls statistic reports the number of paid workers in the US. On the first Friday of each month, it is published by the Bureau of Labour Statistics. Since it reports how many jobs have been created in a given month, forecasters and economic agents closely look at this indicator to have an idea of the wealth of the economy.

Another statistic that can be used to assess the situation of the US job market is the initial unemployment claims. These are reported on a weekly base by the U.S. department of Labour and represent the number of documents filled by individuals requiring unemployment benefits. While

[69] To be fair, the data in the database is the change in the non-farm payrolls.

the first two statistics deal with unemployment, the other two can be used to portray the state of the industrial production in the US.
The first indicator is retail sales. It is compiled by the Census Bureau and Department of commerce and is published on monthly basis. It reports the number of in-store sales taking place in the month before the release. Since retail sales are an important fraction of consumption – and, in turn, of the GDP -, they give an idea of the financial wealth of families. Increasing retail sales might indicate an increase in family income or a higher propensity to consumption.
Finally, the last indicator is the ISM purchasing manager index (PMI). As the other statistics, it is released each month. Broadly speaking, it reflects the overall sentiment of the manufacturing sector. High number – bigger than 50 - suggests more confident companies, while low number reflects more pessimistic or conservative companies. Sentiment influences companies' future plans or actions. The more optimistic companies are, the more they will invest and hire new workers.
Although the total number of observations is not wide, the database is quite deep. For each auction, there are numerous variables that can be scrutinized. However, only the data used in the analysis will be introduced. Along with the economic indicator actual value, prediction market expectation calculated from the probability distribution function and consensus forecast are reported. Moreover, standard deviation for both prediction market and consensus predictions is computed. These four variables constitute the bulk of the database I have used to perform my studies.

3.2 Methodology

3.2.1 Comparing prediction markets and survey forecasts

At the end of the last chapter, two interesting points of prediction markets have been evidenced: higher prediction accuracy relatively to surveys and the use of probability distribution function as a measure of economic uncertainty. Here, these two elements will be analytically formalized and, then, tested.
Prediction markets have been successful in providing predictions about future events. As the time goes by, their forecasts tend to be more accurate than pools or surveys. Efficient aggregation of information and economic incentives can, in part, explain this edge in accuracy. Indeed, prediction markets force their participants to put money where their mouths are. On the other hand, surveys or pools do not face the same constraint. Their predictions can be thought as snapshots of the current economic situation. As it has been reported in previous chapters, private and public forecasters might not always value prediction accuracy over other goals. Starting from these findings, I have

decided to answer the following question: Which model is able to outperform, in terms of accuracy, other models? To rephrase the question: Which data is more informative?
Three simple models have been set up: a first order autoregressive model and two linear multivariate regression models. These two differ one from the other in terms of explanatory variables. Although both models employ as input lagged observation of the variable to be forecasted, the first one uses as the other predictor the consensus prediction made one week before the release, while the second one utilizes the one-day-ahead prediction market forecast. Analytically speaking the three models are the following:

- **Naïve Model**: $y_t = \phi y_{t-1} + \epsilon_t$
- **Consensus Model**: $y_t = \phi y_{t-1} + \beta x_t + \epsilon_t$, where x_t is the consensus forecast.
- **Prediction Market Model**: $y_t = \phi y_{t-1} + \beta z_t + \epsilon_t$, where z_t is the prediction market forecast.

The approach that I have used slightly differs from the one followed by Wolfers and Gurkaynak (2006) to compare the relative prediction accuracy between prediction markets and surveys. In their paper, the authors firstly regressed separately the actual realization on prediction market and survey forecasts. Then they regress again the released value on the two variables simultaneously. I have tried to expand this analysis a little bit further.
First, I have introduced a lagged variable of the economic release to assess if the same results found by the authors were confirmed.
Second, instead of using raw data, I have differenced all the variables in order to deal with stationary variables. Although this approach can reduce the level of information contained in the variables, the same regression analysis exposed earlier has been carried out.
Third, it has been studied the performance in terms of accuracy of the 1-step-ahead predictions of the previously exposed models and of a new model[70].

$$y_t = \alpha + \beta x_{t-1} + e_t$$

y_t represents the change in the release macroeconomic variable between t and $t-1$, while x_{t-1} is equal to difference in the prediction market forecasts between $t-1$ and $t-2$. As explained earlier in the text, I have chosen to focus only on the non-farm payrolls data because they were the prediction market having the highest participation and so the most liquid.

---

[70] Namely, $y_t = \alpha + \beta_1 IProduction_{t-1} + \beta_2 BConfidence_{t-1} + \varepsilon_t$

In order to have a better picture of the accuracy of each of these models, in and out of sample analyses have been carried out. As for the in-sample-analysis, all the models have been estimated and the predictions have been compared. However, in sample analysis tends to not be the best way to pin down the best model for what concerns the accuracy[71]. Because of this, it has been divided the sample in two sub-samples. The first one has been used to estimate the models and the second one has been used to carry out a pseudo-out of sample analysis. For each new observation after the sample period, a forecast has been made. Then, the forecast error has been stored and the model, with a new observation, has been estimated again. This procedure has been followed until the last observation available.

3.2.2 Forecasting unemployment with prediction market data

The unemployment rate is one of the most important economic variables. Although it is considered lagging indicator, policy-makers economists and analysts closely look at his dynamics. These give information about the state of the job market as well as the economic wealth of countries. Forecasters have been trying to make predictions about the future level of unemployment rate mainly using two different techniques (Barnichon and Reqarda, 2012). The first method consists in using time series data of the unemployment to extract relevant past information and then use them to forecast future value. The other method uses the relationship between GDP and unemployment growth, namely the Okun's law, to make predictions about the future unemployment rate. However, owing to the application of new technology and the explosion of Big Data, slightly different methods have been advanced to forecast the unemployment rate. One of the most intriguing is the one proposed by Varian and Choi (2012). The authors have built a model in order to exploit daily Google Trend Data[72] and combine them with past unemployment observation to make forecasts. Obviously, it is not the goal of this section to review all the immense literature about unemployment forecasting. However, at least to my limited knowledge, no models have been used prediction markets data to forecast the unemployment rate yet. Because of this, I have decided to set up a simple model to study whether prediction market forecasts could enhance the prediction accuracy of the model.

---

[71] Over-fitting and the tendency to mistake noise for signal are two of the issues that can boost in sample accuracy, but at the same time having poor out of sample accuracy.

[72] That is, research queries such as "*Jobs, Initial Claims, welfare*" that people are supposed to look for on Google after being laid off.

As explanatory variables I have used the one-month lag unemployment rate and the non-farm payrolls forecast of the prediction market.

$$u_t = \alpha + \varphi u_{t-1} + \beta x_t + e_t$$

Non-farm payrolls have been used for several reasons. First, it is the unique available prediction market time series with the longest span of time. Second, it is theoretically tied to the unemployment rate. It is legitimate to think that, *ceteris paribus*, an increase in the non-farm payrolls should lead to a decrease in the unemployment rate. This reduces, at least to some extent, the risk to mistake correlation with causation in the analysis of the relationship of two variables movements. Of course to be sure that the relationship is economically meaningful, more observations are needed. But unfortunately, this is something we are short of. In addition, non-farm payrolls are one of most important economic indicator. Since it tends to move closely with the job market[73] and business conditions, economists look at it to gauge the prevailing economic situation.
Finally, non-farm payrolls seem to be the most liquid variable of all the set of auctions (Beber and Brandt, 2008). This should ensure that information contained in the auction equilibrium prices reflect the most diverse economic agent beliefs.
The other economic derivative variable that could have been used in this analysis was the initial unemployment claims. However, it has not been possible to exploit the pieces of information embedded in this variable. The number of observation was not enough to guarantee statistically significant results. It could have been as well interesting study the combined effect of initial claims and non-farm payrolls predictions on the unemployment rate.

3.2.3 Prediction Markets and Macroeconomic Uncertainty

The relationship between prediction markets data and macroeconomic uncertainty has been object of several studies. Wolfers and Gurkanyak (2006) is one of the first papers dealing with this issue. The authors regressed the standard deviation of a given economic derivative auction on the value of VIX three months, two months and one day before the release day. Performing this regression, the authors have tried to study if previous levels of implied volatility could explain the uncertainty about the macroeconomic release. By noticing how the value of the VIX one-day prior the auction was not statistically correlated with the standard deviation of the auction, the authors have reached

[73] Data about the number of jobs created comes from different industries and companies.

the conclusion that the auction's uncertainty is more tied to specific release event than to the underlying economic uncertainty.

Beber and Brandt (2008) have addressed the same argument with a different approach. They have investigated the relationship between the ex-ante uncertainty computed from the prediction markets and the ex-post uncertainty from the implied volatility of the financial markets[74]. In order to compute an ex-ante uncertainty measure, the authors have used the price of the vanilla options derived from the prediction market instead of the binary options price. Several findings have been identified. First, high levels of ex-ante uncertainty are followed by greater reductions in the ex-post implied volatility[75]. The authors have found that standard deviation of the non-farm payrolls can explain roughly the 40% of the change in the volatility of Treasury bond futures during the announcement days. Differences in volatility during the non-farm payrolls release days, it is a common pattern that has been found also in other markets. For instance, federal funds futures are as twice as volatile on the Friday in which non-farm payrolls are released than other Fridays (Gadanecz et. al (2007)).

Second, bonds have evidenced a tendency to be more sensitive to prediction market news than stocks. This difference, according to the authors, can be explained by the presence of cyclical and anti-cyclical stocks responding in different ways to macroeconomic releases.

Finally, the trading volume in the bond options market tends to decrease before the announcement day and then, once the macroeconomic figure is released, it rallies. This implies that when the uncertainty about the macroeconomic release is large, economic agents wait the announcement day to carry out financial strategies. Huang (2015) has confirmed these results and has proved how bond prices tend to jump on the same day of the macroeconomic release. Non-farm payrolls have been proved to be the most influential economic derivative in explaining the market uncertainty.

Starting from these findings I have tried to study a little bit deeper the relationship between macroeconomic uncertainty and the information provided by prediction markets. I have set up a simple linear regression model, in which the difference between the opening price of the VIX on the announcement day and the closing price on the day before the announcement, has been regressed on the *forecast surprise*, namely:

---

[74] Concerning the bond market, implied volatility measures have been computed from the 30-year treasury bond futures, 10-year treasury note futures, 5-year Treasury note futures and Eurodollar futures. For the stock market, the implied volatility has been derived from the options on the S&P500 index with a 30 day maturity.

[75] Same results, but with survey forecasts, have been found by Kim and Vernecchia (1991), Ederington and Lee (1996) and Nosfinger and Prucyk (2003).

$$\Delta VIX_t = \alpha + \beta\xi_t + \epsilon_t$$

The approach I have followed slightly differs from the methods I have reported earlier. Instead of using the standard deviation of the auction as a measure of the auction uncertainty, I have employed the *forecast surprise*. Namely:

$$\xi_t = \frac{(x_t - \hat{x}_{t|t-1})}{\bar{\sigma}} \tag{11}$$

The rationale behind this choice is the following. If all the economic agents taking part to the prediction market efficiently use all the information available and aim at being as accurate as possible[76], they should not be able to forecast only the uncertain part of that economic variable. So the forecast surprise should reflect, at least to some extent, the underlying uncertainty of the macroeconomic release.

The larger is the surprise, the bigger is supposed to be the economic uncertainty. The next question is: Does this measure help to explain uncertainty in the financial markets? How economic agents react to this surprise? In order to answer to these questions, an index reflecting the economic sentiment must be identified. In line with the previous literature, I have used the VIX. Defined by several economists as the *fear index*, the VIX index is an investor's measure of the expected S&P 500 volatility over the next 30 days[77]. The VIX index is computed from put and call options of S&P 500 stocks. All other things equal, an increase in the market volatility is associated with an increase in the option prices because the probability of the option to be in the money is higher. This helps to understand the relationship between economic uncertainty and VIX. The more agitated is the market, the higher is the VIX index because the higher is the option prices.

Instead of past observations of the VIX, I have used the change in the VIX closing price occurred between the opening price on the day of the announcement and the closing price on the day before the announcement. This has helped to circumscribe[78] a span of time in which it is resonate to look for the impact of the *forecast surprise* effect.

---

[76] Assume the inverse in prediction markets is contradictory to say at least. Why participants that bet real money should be interested in not being accurate and winning?

[77] If the VIX has a value of 20, investors believe that, with a 68% probability, the S&P 500 index will move upward or downward in a range of 20% over the next year.

[78] The prediction market auction takes place two hours before the opening of the other financial markets, so I have thought that the surprise effect is more likely to be factored in the opening price than in the closing price on the announcement day.

3.3 Results

3.3.1 Prediction Markets and Consensus Forecast Results

Tables 1.1 to 4.3 display the results of the regression analysis of all the prediction market economic derivatives. Although simple in his structure, the prediction market model tends to outperform the other models both in terms of R^2 and root mean square error (RMSE) for all the variables. As for the fitness of the model, using prediction markets data instead of surveys increases always the R^2. As one might expect, in all the economic derivatives the naïve model is the one with the lowest degree of fitness. Although the initial claims derivate contract is the category with the highest number of observations, it displays the lowest R^2 in all the dataset. Conversely, retail sales prediction market model is able to explain the highest level of variance.

As for the RMSE that is used as a measure of forecast accuracy, the overall picture described earlier does not change. By comparing the different values of the RMSE, it can be noticed how the naïve model is the one with the lowest performance in terms of accuracy. More interesting is the comparison between prediction market and surveys. Although the differences between the two models for all the variables are not huge, the ratio between the RMSE of prediction market forecast and surveys forecast is smaller than one for all the economic derivatives. Retail sales is the category in which the edge in accuracy is the most evident with an RMSE ratio of roughly 0.92. However, it must be noticed how this is the prediction market for which the number observation is the lowest.

Moving from the RMSE to the estimated coefficients, prediction market parameters are significative at 1% level throughout the analysis. Survey market parameters are statistically significative for all the regression but the initial claims. Prediction market parameters tend to be lower than survey predictions for all the economic derivatives.

Tables 5.1 to 5.3 display the results of the regression analysis carried out with differenced (*stationary)* non-farm payrolls data. Introducing stationary data into the models brings about slightly different results. The naïve model goes from being the one with the lowest level of prediction accuracy to the most accurate. Prediction Market model still tends to be more precise than the consensus model.

Tables 6.1 to 6.4 report the regression analyses of the 1-step-ahead forecasting models. The naïve model displays the highest level of accuracy while prediction market and consensus models are outperformed. The *general model*, namely the one with industrial production and business confidence as regressors, is the one with the lowest level of accuracy.

3.3.2 Forecasting unemployment with prediction market data

Regression analysis has been conducted to study the effect of prediction market data on unemployment forecasting. Tables 7.1 to 7.4 summarize the main results of this study. The first one reports regression data of the prediction market model, the second shows data of the consensus model, the third one displays the results of the naïve model and the last one the combined effect of prediction market, consensus, unemployment lagged observation and non-farm payroll release data. As it can be seen, prediction market model outperforms in terms of accuracy the consensus model. The RMSE ratio between prediction market model and consensus model is 0.93, while the one between prediction market model and the naïve is 0.95.

Moreover, from table 7.4, it results that prediction market forecasts about non-farm payrolls (*damean)* are negatively and statistically correlated with the unemployment rate. Coherent with the economic theory, this suggests that an increase in the number of non-farm payrolls is supposed to be followed by a decrease in the unemployment. It is interesting to notice how the actual release of the non-farm payrolls (*dnfp)* does not a have statistical effect on the unemployment rate. This could reflect a higher degree of information embedded in the prediction market data. This argument will be analysed deeply in the discussion section.

3.3.3 Prediction Markets and Macroeconomic Uncertainty

Table 7 reports the results of the regression analysis between the *surprise* in prediction market and the difference between the value of the VIX before and after the release of the macroeconomic data. A quick glance at the table reveals how the forecast surprise is negatively correlated with the change in the closing price of the VIX index between the two days before and after the announcement date. Although the R^2 is 0.15, the estimated parameter of the *surprise* variable is statistically significant. This suggests that the wider is the distance between the expected value and the release the bigger is the decrease or increase in the financial market volatility.

3.4 Discussion of results

3.4.1 Prediction Markets and Consensus Forecast

Regression analysis has evidenced mixed results. Adding a lagged observation of the economic variable object of the prediction market has confirmed all the previous findings of Wolfers and Gurkanyak (2006). As expressed earlier in the text, both the closeness to the release date and the efficient aggregation of information can account for this advantage. Although this study confirms that prediction markets are able to provide forecasts qualitatively better than surveys, it must be noticed how the difference in accuracy between survey and prediction market model is not so conspicuous.

Let's now move to the interpretation of the regression results obtained with differenced data. As explained earlier, this analysis has been performed only with non-farm payrolls data. All the variables have been differenced in order to deal with stationary variables. Although the prediction market model still outperforms the consensus model, results have shown how the naïve model is the one with the higher level of accuracy. One possible explanation for this finding could lie in the variables used to perform the analysis. Since prediction market forecasts focus on predicting the *next future*, by taking the difference between the previous and actual prediction the overall level of information contained in the predictions decreases[79]. The same pattern has been also evidenced by consensus data. It seems that by rendering the variable stationary hinders the prediction power of both the prediction markets and consensus forecasts.

For what concerns the 1-step-ahead forecasts, results have shown that the naïve model tends to better than other models in terms of accuracy both in the in-sample and out-sample analysis. This, to a certain extent, can be explained by the fact that both consensus and prediction market forecasts have lost their prediction power once they have been rendered stationary. However, it must be noticed how Consensus forecasts data show the tendency to provide better 1-step-ahead predictions with respect to prediction market data.

I am aware that this study is far from being complete. However, the structure of prediction markets and the auction mechanism have prohibited carrying out several other investigations. First, it could have been interesting to test the accuracy of the two models at an equal distance from the release day. In this way, the potential bias arising from the distance between the two forecasts would have been solved. As a consequence, the results of this analysis could have been more robust. Second,

[79] This is also might be suggested by the fact that both prediction and consensus data turns from being statistically significant to insignificant when they are differenced.

had run one or two auctions, at least, one or two weeks from the release day, it could have been possible to study the behaviour of prediction market forecasts over time. For instance, it could have been investigated how the prediction accuracy evolves over time. Finally, it could have been possible to study how prediction markets incorporate new information into prices and compare them with other economic markets.

Although the results are statistically significant, attention is needed in deriving strong conclusions from these analyses. First, the number of observations available for each economic derivative should not be overlooked. All the variables in the dataset, except initial unemployment claims, have an amount of observation lower than 40. Obviously, a larger dataset could have helped to obtain stronger results both in the economic and statistical sense. Second, auction liquidity is another element to keep in mind in the analysis of the data. The more lively and liquid is a market, the stronger are the conclusions that can be extrapolated from the moments of the variable we are interested in. According to Beber and Brandt (2006), the most liquid economic derivative throughout the period taken in the analysis is non-farm payrolls[80]. Because of this, caution is needed in the examination of the other economic derivatives regression results.

3.4.2 Forecasting Unemployment with Prediction Market Data

Prediction markets have been proved successful in providing valuable information for forecasting the release of the economic data underlying the auction. However, it could be interesting to study if prediction markets forecasts can be used to forecast other macroeconomic variables.

To my limited knowledge of the relevant economic literature, this is the first time that economic derivative contracts have been used to forecast other macroeconomic variables. Because of this, it is difficult to compare the results obtained here with other precedent studies.

As it has been discussed above, the regression analysis seems to suggest a relationship between the prediction market forecast for non-farm payrolls and the unemployment rate. Here, the same precautions evidenced earlier about the interpretation of the data are needed. The low number of observations calls for great attention in suggesting any form of causality between these two variables. It must be also noticed that the regression results could be in some way biased by the span of time involuntarily taken into consideration. Indeed, a simple look at the data reveals how over the period of analysis the unemployment rate was relatively stable[81]. It would have been

[80] The authors have reached this conclusion by having conversations with the management of Goldman Sachs.

[81] The standard deviation was close 0,3%.

interesting to test this relationship in a more unstable and agitated economic context, but due to the closure of the Economic Derivative market, this has not been possible.
However, some intriguing aspects arise in this analysis. First, prediction market forecasts for non-farm payrolls give information to the dynamics of the unemployment that the actual release about non-farm payrolls do not seem to give. This is evidenced by the fact that prediction market parameter is statistically significant and negatively correlated with the unemployment, while the actual release is not. Second, even adding the survey forecast to the regression does not affect the estimation of the prediction market parameter.
One possible explanation could be the following. Since prediction market forecasts come out from the aggregation of different economic agents' beliefs, they can convey information about the underlying sentiment of the economy. Prediction market forecasts cannot be statistically accurate, namely forecast errors averaging around the zero, but they can give information about the direction of the economy. This, in turn, can help to explain the dynamics of the economic activity and of the unemployment.
Once again, it is important to gauge these results with the highest level of precaution. More data and research are fundamental to deepen the comprehension of this phenomenon.

3.4.3 Prediction Markets and Macroeconomic Uncertainty

The regression analysis carried out to study the relationship between prediction markets data and uncertainty in the financial market differs from the previous works under two aspects. First, instead of the standard deviation of the auction as a measure of uncertainty, the forecast *surprise effect* has been used. Second, as a dependent variable has been employed the change of the VIX index between the opening price on the day of the announcement and the closing price the day before the announcement[82]. The rationale behind this approach is the following.

The larger is the gap between the prediction market forecast and the release figure, the more uncertain[83] is supposed the economic environment surrounding the auction. It seems legitimate to think that, even if forecast errors will always occur because of the complexity of the economic system, *surprise effects* will be more limited during periods of economic stability[84]. Because of this, the forecast surprise can be thought as a loose proxy for the macroeconomic uncertainty.

How does this variable affect financial market volatility? Does the volatility blow up or plummet? The answer to these questions depends on the sign of the *surprise effect*. We can have two separate cases. First, the prediction market forecast underestimates the release value. The released figure is better than what market participants thought. This yields a positive *surprise effect*. Because the estimated coefficient of the regression is negative, the final impact on the change of the VIX index is negative as well. This means that the financial market volatility tends to decrease in response to positive macroeconomic *surprise effects*. If the VIX index is thought as a measure of the fear underlying financial markets, a positive *surprise effect* could reassure investors during economic turmoils as well as increase their confidence during expansionary economic periods.

Second, the prediction market forecast overestimates the actual release. The published figure is worst than what prediction market participants thought. In this case, the forecast error provokes a negative *surprise effect*. Since the estimated parameter is negative, the compound effect on the VIX change is positive. The index value increases. The overall economic context is perceived a little bit more agitated by financial investors. As a consequence, there is an increase in the expected volatility of the financial markets over the next 30 days.

---

[82] Since the prediction market forecasts are available one hour before the opening of the VIX negotiations, it is possible to suppose that the forecast surprise is factored in the opening VIX price.

[83] Here, uncertainty does not have a negative connotation. Roughly speaking, uncertainty is tied to the process of discovering the result of an unknown event. However, the result can be either positive or negative. Uncertain environment does not imply negative outcomes. For instance, a student can be uncertain about the outcome of an exam. The fact of being uncertain does not mean that the final grade will be negative.

[84] Compare survey of professional forecasters errors during the 1968-1995 period and 1996-2006.

The results obtained in this analysis are in line with previous findings in the literature. Although all the past works have used consensus forecast to compute the surprise effect, non-farm payrolls prediction market forecasts show the same tendency to negatively affect the implied volatility (See Nofsinger and Prucyk (2003)).

Although the model and technique used in the analysis is far from being perfect, this study has analysed under a different lens the relationship between macroeconomic uncertainty and financial volatility. By using the *surprise effect* instead of the standard deviation of the prediction market auction as a predictor, it has been possible to study separately the effect of positive and negative *surprises* on the financial market volatility.

As it has been explained in the previous regression analyses, it is important to gauge this result with caution. The same precautions suggested earlier hold here. More data and accurate research are needed to state strong conclusions.

Besides the possible biases deriving from the data, there is room to improve the analysis. The assumption of symmetry in response to different *surprise effects* stemming from the linear relationship analysis could not be appropriate in this setup. Indeed, it is difficult to believe that economic agents will react in the same way to a positive or negative surprise[85]. Moreover, the economic context surrounding the surprise is important as well. For instance, in a depressed economy, a released figure that is higher than the prediction market forecast could spur more the economic sentiment of the financial investors leading to a more consistent reduction of the VIX index.

[85] Regarding this point, Nofsinger and Prucyk (2003) have shown how financial markets react more actively following negative surprise, while positive surprise effects have milder effect on the market volatility.

Conclusion

The relationship between prediction markets data and macroeconomic forecasting has been the object of this text. Starting from the idea that forecast errors have been frequent since the use of prediction models, prediction markets have been investigated to analyse if the information they convey can be used in forecasting models. Three different but simple analyses have been carried out.

First, as it has been done in Wolfers and Gurkanyak (2006), prediction markets and surveys forecasts have been compared to study which of the two provides better predictions. Results were mixed. On one hand, adding a lagged observation of the economic variable object of the prediction market has confirmed all the previous findings of Wolfers and Gurkanyak (2006). On the other hand, both the analysis with the stationary variables and the one-step-ahead forecasts have evidenced that the naïve model tends to be more accurate than the other two models. This, to a certain extent, can be explained by the fact that both consensus and prediction market forecasts have lost their prediction power once they have been rendered stationary.

Second, it has been investigated if prediction market information can be used to forecast other macroeconomic variables. Although it would have been interesting extending this analysis to several macroeconomic indicators, the limited available dataset has hindered this study. Because of this, it has been possible only to investigate the relationship between non-farm payrolls data derived from prediction markets and the unemployment rate. Regression results have shown that simple forecasting models employing prediction market forecasts as predictors tend to slightly outperform other models in terms of accuracy. Moving from the *forecasting analysis* to the *economic analysis,* regression results have evidenced another interesting feature. Non-farm payroll prediction market data has been the only parameter to be statistically significant and negatively correlated with the unemployment rate. This might suggest that prediction markets convey information with better qualitative contents relatively to other variables.

Finally, it has been studied how macroeconomic surprises affect the volatility of financial markets. The *surprise effect* has been computed by simply subtracting the prediction market forecast from the actual release of the macroeconomic variable. Since prediction market participants have an economic incentive in forecasting the actual release value, the *surprise effect*, namely the forecast error of the prediction market adjusted for the standard deviation of the previous forecast errors, has been used as a measure of economic uncertainty. The VIX index used here as a financial volatility proxy has been then regressed on the *surprise effect*. Coherently with the previous literature, the latter has shown to negatively affect the former. Negative *surprise effects* tend to bring about positive changes in the VIX index while positive *surprise effects* cause negative changes in the VIX. This implies that financial markets tend to be more volatile in the occurrence of negative *surprises* and to be less volatile in response to positive *surprise*. As it has been reported in the discussion section, it is important to gauge these results with extreme caution. A small number of observations, liquidity issue for some prediction markets and the methods used to carry out these studies could potentially bias some results. More research is needed to achieve stronger results in this field.

The reopening of the economic derivative market and the negotiation of new contracts could spark the interest in these markets both of economic agents and researchers. For instance, the creation of a continuous auction market, similar to the one used to predict political outcomes, about the quarter GDP growth would enhance the use of prediction markets. By having a continuous forecast market in which investors bet on different outcomes, several analyses could be made. It would be possible to study what are the effects of economic news on the forecasts provided by the prediction market participants.

That is, how participants' beliefs about future economic outcome vary as new and unexpected information becomes available. In addition, it would be possible to study how the forecast accuracy of the prediction market evolves over time and compare it with the consensus forecast. These are just a few analyses that could be implemented if this kind of market was available. In conclusion, liquid, dynamic and lively prediction markets could provide pieces of information that forecasters could use in their models.

Bibliography

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Vega, C. (2002). *Micro effects of macro announcements: Real-time price discovery in foreign exchange* (No. w8959). National bureau of economic research.

Arrow, K. J., Forsythe, R., Gorham, M., Hahn, R., Hanson, R., Ledyard, J. O., ... & Neumann, G. R. (2008). The promise of prediction markets. *SCIENCE-NEW YORK THEN WASHINGTON-*, *320*(5878), 877.

Aruoba, S. B. (2008). Data revisions are not well behaved. *Journal of money, credit and banking*, *40*(2‐3), 319-340.

Batchelor, R. (2001). How useful are the forecasts of intergovernmental agencies? The IMF and OECD versus the consensus. *applied economics*, *33*(2), 225-235.

Batchelor, R., & Dua, P. (1992). Conservatism and consensus‐seeking among economic forecasters. *Journal of Forecasting*, *11*(2), 169-181.

Beber, A., & Brandt, M. W. (2009). Resolving Macroeconomic Uncertainty in Stock and Bond Markets*. *Review of Finance*, *13*(1), 1-45.

Berg, J., Nelson, F., & Rietz, T. (2003). Accuracy and forecast standard error of prediction markets. *Tippie College of Business Administration, University of Iowa*.

Black, F. (1986). Noise. *The journal of finance*, *41*(3), 529-543.

Bloom, N. (2009). The impact of uncertainty shocks. *econometrica*, *77*(3), 623-685.

Bowles, C., Friz, R., Genre, V., Kenny, G., Meyler, A., & Rautanen, T. (2007). The ECB survey of professional forecasters (SPF)-A review after eight years' experience. *ECB occasional paper*, (59).

Bowles, C., Friz, R., Genre, V., Kenny, G., Meyler, A., & Rautanen, T. (2011). An evaluation of the growth and unemployment forecasts in the ECB survey of professional forecasters. *OECD Journal: Journal of Business Cycle Measurement and Analysis*, *2010*(2), 1-28.

Broadbent B., (2013). *Forecast Errors*. Transcript of the speech given at The Mile End Group of Queen Mary, University of London.

Clements, M., & Hendry, D. (1998). *Forecasting economic time series*. Cambridge University Press.

Clements, M. P., & Hendry, D. F. (Eds.). (2008). *A companion to economic forecasting*. John Wiley & Sons.

Croushore, D. (2006). Forecasting with real-time macroeconomic data. *Handbook of economic forecasting*, *1*, 961-982.

Croushore, D., & Stark, T. (2003). A real-time data set for macroeconomists: Does the data vintage matter?. *Review of Economics and Statistics*, *85*(3), 605-617.

D'Agostino, A., Gambetti, L., & Giannone, D. (2013). Macroeconomic forecasting and structural change. *Journal of Applied Econometrics*, *28*(1), 82-101.

Ederington, L. H., & Lee, J. H. (1996). The creation and resolution of market uncertainty: the impact of information releases on implied volatility. *Journal of Financial and Quantitative Analysis*, *31*(04), 513-539.

Fildes, R., & Stekler, H. (2002). The state of macroeconomic forecasting. *Journal of macroeconomics*, *24*(4), 435-468.

Gadanecz, B., Moessner, R., & Upper, C. (2007). Economic derivatives. *BIS Quarterly Review, March*.

Gurkaynak, R., & Wolfers, J. (2006). *Macroeconomic derivatives: An initial analysis of market-based macro forecasts, uncertainty, and risk* (No. w11929). National Bureau of Economic Research.

Hamilton, J. D. (1994). *Time series analysis*, Princeton: Princeton university press.

Hendry, D. F. (1997). THE ECONOMETRICS OF MACROECONOMIC FORECASTING*. *The Economic Journal*, *107*(444), 1330-1357.

Hendry, D. F., & Clements, M. P. (2003). Economic forecasting: Some lessons from recent research. *Economic Modelling*, *20*(2), 301-329.

Huang, X. (2015). Macroeconomic news announcements, systemic risk, financial market volatility and jumps. *Finance and Economics Discussion Series*, *97*.

Kim, O., & Verrecchia, R. E. (1991). Market reaction to anticipated announcements. *Journal of Financial Economics*, *30*(2), 273-309.

Knotek II, E. S. (2007). How useful is Okun's law?. *Economic Review-Federal Reserve Bank of Kansas City*, *92*(4), 73.

Lamont, O. A. (2002). Macroeconomic forecasts and microeconomic forecasters. *Journal of economic behavior & organization*, *48*(3), 265-280.

Laster, D., Bennett, P. B., & Geoum, I. S. (1997). Rational bias in macroeconomic forecasts. *FRB of New York Staff Report*, (21).

Leitch, G., & Tanner, J. E. (1991). Economic forecast evaluation: profits versus the conventional error measures. *The American Economic Review*, 580-590.

Loungani, P. (2001). How accurate are private sector forecasts? Cross-country evidence from consensus forecasts of output growth. *International journal of forecasting*, *17*(3), 419-432.

Mankiw, N. G., & Shapiro, M. D. (1986). *News or noise? An analysis of GNP revisions* (No. w1939). National Bureau of Economic Research.

Mankiw, N. G., Runkle, D. E., & Shapiro, M. D. (1984). Are preliminary announcements of the money stock rational forecasts?. *Journal of Monetary Economics*, *14*(1), 15-27.

Manski, C. F. (2006). Interpreting the predictions of prediction markets. *economics letters*, *91*(3), 425-429.

McNees, S. K. (1992). How large are economic forecast errors. *New England Economic Review*, *1*, 25-33.

Nofsinger, J. R., & Prucyk, B. (2003). Option volume and volatility response to scheduled economic news releases. *Journal of Futures Markets*, *23*(4), 315-345.

Montgomery, A. L., Zarnowitz, V., Tsay, R. S., & Tiao, G. C. (1998). Forecasting the US unemployment rate. *Journal of the American Statistical Association*, *93*(442), 478-493.

Mork, K. A. (1987). Ain't behavin': Forecast errors and measurement errors in early GNP estimates. *Journal of Business & Economic Statistics*, *5*(2), 165-175.

O'Leary, D. E. (2011). Prediction markets as a forecasting tool. *Advances in business and management forecasting*, *8*, 169-184.

Öller, L. E., & Barot, B. (1999). *Comparing the accuracy of European GDP forecasts* (No. 64). National Institute of Economic Research.

Öller, L. E., & Teterukovsky, A. (2007). Quantifying the quality of macroeconomic variables. *International journal of Forecasting*, *23*(2), 205-217.

Ormerod, P., & Mounfield, C. (2000). Random matrix theory and the failure of macro-economic forecasts. *Physica A: Statistical Mechanics and its Applications*, *280*(3), 497-504.

Ottaviani, M., & Sørensen, P. N. (2007, May). Aggregation of information and beliefs in prediction markets. In *London Conference on Information and Prediction Markets*.

Plott, C. R., Wit, J., & Yang, W. C. (2003). Parimutuel betting markets as information aggregation devices: experimental results. *Economic Theory*, *22*(2), 311-351.

Reifschneider, D., & Tulip, P. (2007). Gaging the Uncertainty of the Economic Outlook from Historical Forecasting Errors.

Schuh, S. (2001). An evaluation of recent macroeconomic forecast errors. *New England Economic Review*, 35-36.

Servan‐Schreiber, E., Wolfers, J., Pennock, D. M., & Galebach, B. (2004). Prediction markets: Does money matter?. *Electronic Markets*, *14*(3), 243-251.

Silver, N. (2012). *The signal and the noise: Why so many predictions fail-but some don't*. Penguin.

Snowberg, E., Wolfers, J., & Zitzewitz, E. (2012). *Prediction markets for economic forecasting* (No. w18222). National Bureau of Economic Research.

Snowberg, E., Wolfers, J., & Zitzewitz, E. (2005). 18 Information (in) efficiency in prediction markets. *Information efficiency in financial and betting markets*, 366.

Snowberg, E., & Wolfers, J. (2010). *Explaining the favorite-longshot bias: Is it risk-love or misperceptions?* (No. w15923). National Bureau of Economic Research.

Stark, T. (2010). Realistic evaluation of real-time forecasts in the Survey of Professional Forecasters. *Federal Reserve Bank of Philadelphia Research Rap, Special Report*, *1*.

Stark, T., & Croushore, D. (2002). Forecasting with a real-time data set for macroeconomists. *Journal of Macroeconomics*, *24*(4), 507-531.

Stekler, H. O. (2007). The future of macroeconomic forecasting: Understanding the forecasting process. *International Journal of Forecasting*, *23*(2), 237-248.

Stock, J. H., & Watson, M. W. (1989). New indexes of coincident and leading economic indicators. In *NBER Macroeconomics Annual 1989, Volume 4* (pp. 351-409). MIT press.

Stock, J. H., & Watson, M. W. (1993). A procedure for predicting recessions with leading indicators: econometric issues and recent experience. In *Business cycles, indicators and forecasting* (pp. 95-156). University of Chicago Press.

Swann, G. P. (2006). *Putting econometrics in its place: a new direction in applied economics*. Edward Elgar Publishing.

Teschner, F., Stathel, S., & Weinhardt, C. (2011, January). A prediction market for macro-economic variables. In *System Sciences (HICSS), 2011 44th Hawaii International Conference on* (pp. 1-9). IEEE.

Tetlock, P. (2004). How efficient are information markets? Evidence from an online exchange. *Social Science Research Network*.

Zarnowitz, V. (1992). Consensus and uncertainty in economic prediction. In *Business Cycles: Theory, History, Indicators, and Forecasting* (pp. 492-518). University of Chicago Press.

Zarnowitz, V., & Braun, P. (1993). Twenty-two years of the NBER-ASA quarterly economic outlook surveys: Aspects and comparisons of forecasting performance. In *Business cycles, indicators and forecasting* (pp. 11-94). University of Chicago Press.

Wallis, K. F. (1989). Macroeconomic forecasting: a survey. *The Economic Journal*, 28-61.

Wolfers, J., & Zitzewitz, E. (2006). *Five open questions about prediction markets* (No. w12060). National Bureau of Economic Research.

Wolfers, J., & Zitzewitz, E. (2004). *Prediction markets* (No. w10504). National Bureau of Economic Research.

Wolfers, J., & Zitzewitz, E. (2006). *Interpreting prediction market prices as probabilities* (No. w12200). National Bureau of Economic Research.

Wolfers, J., & Zitzewitz, E. (2006). *Prediction markets in theory and practice* (No. w12083). national bureau of economic research.

Tables

```
Linear regression model:
    nfp_t ~ 1 + nfp_t_1 + amean

Estimated Coefficients:
                   Estimate      SE         tStat        pValue
    (Intercept)      -39658      28005     -1.4161       0.16739
    nfp_t_1        -0.24793    0.21663     -1.1445       0.26177
    amean            1.2254    0.29028      4.2214    0.00021864


Number of observations: 32, Error degrees of freedom: 29
Root Mean Squared Error: 1.04e+05
R-squared: 0.509,  Adjusted R-Squared 0.475
F-statistic vs. constant model: 15, p-value = 3.29e-05
```

(**Table 1.1**: Prediction Market Model - Non-farm payroll)

```
Linear regression model:
    nfp_t ~ 1 + nfp_t_1

Estimated Coefficients:
                   Estimate      SE         tStat      pValue
    (Intercept)     42212       25238      1.6725     0.10482
    nfp_t_1         0.47009     0.16759     2.805     0.008745


Number of observations: 32, Error degrees of freedom: 30
Root Mean Squared Error: 1.29e+05
R-squared: 0.208,  Adjusted R-Squared 0.181
F-statistic vs. constant model: 7.87, p-value = 0.00874
```

(**Table 1.2**: Naïve Model - Non-farm payroll)

```
Linear regression model:
    nfp_t ~ 1 + nfp_t_1 + mmean

Estimated Coefficients:
                   Estimate      SE         tStat        pValue
    (Intercept)      -53317      31718      -1.681       0.10351
    nfp_t_1        -0.23598    0.22464     -1.0505       0.30217
    mmean            1.3505    0.34019      3.9697    0.00043438


Number of observations: 32, Error degrees of freedom: 29
Root Mean Squared Error: 1.06e+05
R-squared: 0.487,  Adjusted R-Squared 0.451
F-statistic vs. constant model: 13.7, p-value = 6.32e-05
```

(**Table 1.3**: Consensus Forecast Model - Non-farm payroll)

```
Linear regression model:
    napm_t ~ 1 + napm_t_1 + amean

Estimated Coefficients:
                   Estimate      SE         tStat       pValue
    (Intercept)     -6486.8     3317.2     -1.9555      0.061347
    napm_t_1       -0.12811    0.14739    -0.86921       0.39269
    amean             1.242    0.17172      7.2325    1.1106e-07


Number of observations: 29, Error degrees of freedom: 26
Root Mean Squared Error: 1.4e+03
R-squared: 0.935,  Adjusted R-Squared 0.93
F-statistic vs. constant model: 187, p-value = 3.81e-16
```

(**Table 2.1:** Prediction Market Model - ISM Index)

```
Linear regression model:
    napm_t ~ 1 + napm_t_1

Estimated Coefficients:
                   Estimate      SE          tStat      pValue
    (Intercept)     6953.8       4679.5      1.486        0.14886
    napm_t_1       0.87728     0.083416     10.517     4.7638e-11


Number of observations: 29, Error degrees of freedom: 27
Root Mean Squared Error: 2.38e+03
R-squared: 0.804,  Adjusted R-Squared 0.797
F-statistic vs. constant model: 111, p-value = 4.76e-11
```

(**Table 2.2:** Naive Model - ISM Index)

```
Linear regression model:
    napm_t ~ 1 + napm_t_1 + mmean

Estimated Coefficients:
                   Estimate      SE        tStat       pValue
    (Intercept)     -8397.5        3698    -2.2708      0.031675
    napm_t_1       -0.37282     0.19431    -1.9186      0.066074
    mmean            1.5194      0.2277     6.6731    4.4392e-07


Number of observations: 29, Error degrees of freedom: 26
Root Mean Squared Error: 1.47e+03
R-squared: 0.928,  Adjusted R-Squared 0.922
F-statistic vs. constant model: 167, p-value = 1.48e-15
```

(**Table 2.1:** Consensus Forecast Model - ISM Index)

```
    iclm_t ~ 1 + iclm_t_1 + amean

Estimated Coefficients:
                   Estimate        SE         tStat       pValue
    (Intercept)    1.2092e+05      58834      2.0553      0.044211
    iclm_t_1         0.018229     0.1687     0.10805       0.91431
    amean             0.61841    0.24992      2.4744       0.01619


Number of observations: 63, Error degrees of freedom: 60
Root Mean Squared Error: 1.49e+04
R-squared: 0.18,  Adjusted R-Squared 0.153
F-statistic vs. constant model: 6.59, p-value = 0.00259
```

(**Table 3.1:** Prediction Market Model - Initial Claims)

```
Linear regression model:
    iclm_t ~ 1 + iclm_t_1

Estimated Coefficients:
                   Estimate        SE         tStat       pValue
    (Intercept)    2.2838e+05      41325      5.5266      7.1785e-07
    iclm_t_1          0.31503    0.12351      2.5506         0.01328


Number of observations: 63, Error degrees of freedom: 61
Root Mean Squared Error: 1.55e+04
R-squared: 0.0964,  Adjusted R-Squared 0.0816
F-statistic vs. constant model: 6.51, p-value = 0.0133
```

(**Table 3.2:** Naive Model - Initial Claims)

```
    iclm_t ~ 1 + iclm_t_1 + mmean

Estimated Coefficients:
                   Estimate        SE         tStat       pValue
    (Intercept)    1.4128e+05      71698      1.9705      0.053396
    iclm_t_1         0.090461    0.19493     0.46406       0.64429
    mmean             0.48743    0.32944      1.4796       0.14422


Number of observations: 63, Error degrees of freedom: 60
Root Mean Squared Error: 1.54e+04
R-squared: 0.128,  Adjusted R-Squared 0.0991
F-statistic vs. constant model: 4.41, p-value = 0.0163
```

(**Table 3.3:** Consensus Forecast Model - Initial Claims)

```
Linear regression model:
    rsx_t ~ 1 + rsx_t_1 + amean

Estimated Coefficients:
                   Estimate     SE         tStat       pValue
    (Intercept)     0.10329    0.14375    0.71854       0.47966
    rsx_t_1        -0.41741    0.13574    -3.0751      0.0053535
    amean            1.2086    0.27312     4.4249     0.00019521


Number of observations: 26, Error degrees of freedom: 23
Root Mean Squared Error: 0.356
R-squared: 0.597,  Adjusted R-Squared 0.562
F-statistic vs. constant model: 17.1, p-value = 2.86e-05
```

(**Table 4.1:** Prediction Market Model – Retail Sales)

```
Linear regression model:
    rsx_t ~ 1 + rsx_t_1

Estimated Coefficients:
                   Estimate     SE         tStat       pValue
    (Intercept)     0.60478    0.11778     5.1349     2.9546e-05
    rsx_t_1        -0.51132    0.17857    -2.8634      0.0085651


Number of observations: 26, Error degrees of freedom: 24
Root Mean Squared Error: 0.474
R-squared: 0.255,  Adjusted R-Squared 0.224
F-statistic vs. constant model: 8.2, p-value = 0.00857
```

(**Table 4.2:** Naive Model – Retail Sales)

```
Linear regression model:
    rsx_t ~ 1 + rsx_t_1 + mmean

Estimated Coefficients:
                   Estimate      SE         tStat       pValue
    (Intercept)    0.036088     0.1858     0.19423        0.8477
    rsx_t_1        -0.48239    0.14639     -3.2952     0.0031669
    mmean            1.4395    0.40205      3.5804     0.0015842


Number of observations: 26, Error degrees of freedom: 23
Root Mean Squared Error: 0.388
R-squared: 0.521,  Adjusted R-Squared 0.48
F-statistic vs. constant model: 12.5, p-value = 0.000209
```

(**Table 4.3:** Consensus Forecast Model – Retail Sales)

```
mdl_naive_diff =


Linear regression model:
    dnfp_t ~ 1 + dnfp_t_1

Estimated Coefficients:
                   Estimate      SE         tStat       pValue
    (Intercept)       12240      22620      0.5411       0.59257
    dnfp_t_1       -0.56755    0.15672     -3.6214     0.0011063


Number of observations: 31, Error degrees of freedom: 29
Root Mean Squared Error: 1.26e+05
R-squared: 0.311,  Adjusted R-Squared 0.288
F-statistic vs. constant model: 13.1, p-value = 0.00111
```

(**Table 5.1:** Non-farm-payrolls differenced naïve model)

```
Linear regression model:
    dnfp_t ~ 1 + damean

Estimated Coefficients:
                   Estimate      SE         tStat       pValue
    (Intercept)     7381.3       27239     0.27099     0.78832
    damean         0.42805     0.56168     0.76209     0.45216


Number of observations: 31, Error degrees of freedom: 29
Root Mean Squared Error: 1.5e+05
R-squared: 0.0196,  Adjusted R-Squared -0.0142
F-statistic vs. constant model: 0.581, p-value = 0.452
```

(**Table 5.2:** Non-farm-payrolls differenced prediction market model)

```
Linear regression model:
    dnfp_t ~ 1 + dmmean

Estimated Coefficients:
                   Estimate      SE         tStat       pValue
    (Intercept)     9462.1       27539     0.34359     0.73363
    dmmean         0.14322      0.7662     0.18693     0.85302


Number of observations: 31, Error degrees of freedom: 29
Root Mean Squared Error: 1.52e+05
R-squared: 0.0012,  Adjusted R-Squared -0.0332
F-statistic vs. constant model: 0.0349, p-value = 0.853
```

(**Table 5.3:** Non-farm-payrolls differenced consensus model)

	Prediction Market Model	Naive Model	Consensus Model
Prediction Market Model	1	0.8005	0.977
Naive Model	1.249	1	1.221
Consensus Model	1.022	0.8187	1

(**Table 5.4** : RMSE Comparison for Non-farm payrolls forecasts)

	Prediction Market Model	Naive Model	Consensus Model
Prediction Market Model	1	1.19	0.9907
Naive Model	0.8381	1	0.8303
Consensus Model	1.0094	1.204	1

(**Table 5.5** : RMSE Comparison for Non-farm payrolls forecasts with differenced variables)

	Prediction Market Model	Naive Model	Consensus Model
Prediction Market Model	1	0.61	0.943

(**Table 5.6**: Out of sample RMSE Comparison for Prediction Market Model (Non-farm payrolls raw data))

	Prediction Market Model	Naive Model	Consensus Model
Naïve Model	0.734	1	0.758

(**Table 5.6**: Out of sample RMSE Comparison for Naive Market Model (non-farm payrolls differenced data))

```
mdl_naive =


Linear regression model:
    y ~ 1 + x1

Estimated Coefficients:
                   Estimate     SE         tStat      pValue
    (Intercept)      12194       23408     0.52093     0.60651
    x1             -0.56744    0.15977    -3.5516    0.0013782


Number of observations: 30, Error degrees of freedom: 28
Root Mean Squared Error: 1.28e+05
R-squared: 0.311,  Adjusted R-Squared 0.286
F-statistic vs. constant model: 12.6, p-value = 0.00138
```

(Table 6.1: 1-step-ahead-forecast- naïve model (Non-farm payrolls))

```
mdl_mkt =


Linear regression model:
    y ~ 1 + x1

Estimated Coefficients:
                   Estimate     SE         tStat      pValue
    (Intercept)      25018       24331     1.0282      0.31263
    x1             -2.2639     0.68186    -3.3201    0.0025079


Number of observations: 30, Error degrees of freedom: 28
Root Mean Squared Error: 1.31e+05
R-squared: 0.282,  Adjusted R-Squared 0.257
F-statistic vs. constant model: 11, p-value = 0.00251
```

(Table 6.2: 1-step-ahead-forecast- consensus model (Nfp))

```
mdl_pm =


Linear regression model:
    y ~ 1 + x1

Estimated Coefficients:
                   Estimate     SE         tStat      pValue
    (Intercept)      20419       25019     0.81614     0.42131
    x1             -1.4724     0.50899    -2.8927    0.0073126


Number of observations: 30, Error degrees of freedom: 28
Root Mean Squared Error: 1.35e+05
R-squared: 0.23,  Adjusted R-Squared 0.203
F-statistic vs. constant model: 8.37, p-value = 0.00731
```

(Table 6.3: 1-step-ahead-forecast-prediction model (Nfp))

```
mdl_gen = 


Linear regression model:
    y ~ 1 + x1 + x2

Estimated Coefficients:
                   Estimate      SE             tStat       pValue 
    (Intercept)    3297.9             31342     0.10522     0.91698
    x1              26182             59362     0.44105     0.66269
    x2              17857        1.1376e+05     0.15698     0.87643


Number of observations: 30, Error degrees of freedom: 27
Root Mean Squared Error: 1.57e+05
R-squared: 0.00817,  Adjusted R-Squared -0.0653
F-statistic vs. constant model: 0.111, p-value = 0.895
```

(**Table 6.4:** 1-step-ahead- Industrial Production and Business Sentiment model (Nfp))

```
Linear regression model:
    du_t ~ 1 + damean

Estimated Coefficients:
                   Estimate         SE            tStat       pValue
    (Intercept)    -0.00023238      0.00019326    -1.2024     0.23894
    damean           -8.72e-09      3.9852e-09    -2.1881     0.03687


Number of observations: 31, Error degrees of freedom: 29
Root Mean Squared Error: 0.00107
R-squared: 0.142,  Adjusted R-Squared 0.112
F-statistic vs. constant model: 4.79, p-value = 0.0369
```

(**Table 7.1:** Unemployment Prediction Market Forecast Model)

```
Linear regression model:
    du_t ~ 1 + dmmean

Estimated Coefficients:
                   Estimate         SE            tStat       pValue
    (Intercept)    -0.00026595      0.00020671     -1.2866    0.20842
    dmmean         -4.5709e-09      5.7512e-09    -0.79477    0.43321


Number of observations: 31, Error degrees of freedom: 29
Root Mean Squared Error: 0.00114
R-squared: 0.0213,  Adjusted R-Squared -0.0124
F-statistic vs. constant model: 0.632, p-value = 0.433
```

(**Table 7.2**: Unemployment Consensus Forecast Model)

```
md_naive =


Linear regression model:
    y ~ 1 + x1

Estimated Coefficients:
                   Estimate         SE            tStat       pValue
    (Intercept)    -0.00039437      0.00020884    -1.8883     0.069383
    x1                -0.22887         0.18086    -1.2654      0.21614


Number of observations: 30, Error degrees of freedom: 28
Root Mean Squared Error: 0.00111
R-squared: 0.0541,  Adjusted R-Squared 0.0203
F-statistic vs. constant model: 1.6, p-value = 0.216
```

(**Table 7.3**: Unemployment Naive Forecast Model)

```
Linear regression model:
    du_t ~ 1 + damean + dmmean + dnfp

Estimated Coefficients:
                   Estimate       SE            tStat      pValue
    (Intercept)    -0.00027056    0.00018202    -1.4865     0.14873
    damean         -2.1475e-08    6.9067e-09    -3.1093    0.004388
    dmmean          1.9338e-08    9.2477e-09     2.0911    0.046062
    dnfp              1.94e-09    1.2504e-09     1.5515     0.13241


Number of observations: 31, Error degrees of freedom: 27
Root Mean Squared Error: 0.001
R-squared: 0.296,  Adjusted R-Squared 0.218
F-statistic vs. constant model: 3.79, p-value = 0.0217
```

(**Table 7.4:** Forecasting Unemployment with prediction market and survey non-farm payroll forecast, unemployment lagged observation and non-farm payroll released value.)

```
Linear regression model:
    delta_vix ~ 1 + surp

Estimated Coefficients:
                   Estimate     SE          tStat       pValue
    (Intercept)    -0.075004    0.090546    -0.82835    0.41402
    surp            -0.20154    0.087703      -2.298    0.02871


Number of observations: 32, Error degrees of freedom: 30
Root Mean Squared Error: 0.488
R-squared: 0.15,  Adjusted R-Squared 0.121
F-statistic vs. constant model: 5.28, p-value = 0.0287
```

(**Table 8:** Non-farm payrolls surprise and Financial Volatility)

Summary of the thesis

Thesis object and Context

The broad object of the following thesis is to study the prediction power of information embedded in economic prediction markets and to test if this information can be used in forecasting models. The specific objects of this thesis are: (i) to pin down the features that can make prediction markets interesting to economic forecasters; (ii) to compare prediction market forecasts and professional forecasts to assess which one is most accurate; (iii) to apply prediction market data to simple forecasting models to test whether an increase in the accuracy arises; (iv) to use prediction market data to study if they can, to a certain extent, reflects the underlying sentiment in the economy.

I do believe that any type of analysis would result flawed if no attention is paid to the context that surrounds the object of the study. This is the rationale behind my willingness to begin the thesis with an overview of what are the issues affecting the forecasting science.

The economy is unarguably one of the most complex social systems in the world. His intrinsic and inner complexity stems, to a certain extent, from the subjects that inhabit it, namely human beings. Driven by different beliefs and goals, they interact between them giving rise to economic transactions. The results of these interactions are economic variables. In addition, human beings create social and political institutions in order to regulate themselves and the system they dwell in. Because of the myriads of interactions occurring between these agents, the system itself is quite *unstable.* By simply reading

the economic news that every day is in the newspapers or on TV, one could be tempted to say that stability is not the normality in the economic system. On top of that, the economic system is also full of uncertainties deriving from this instability. If everything were stable and steady, there would not be the need for predictions. Unfortunately, this is not the case. Economic forecasts are pivotal in every economic field. Countries rely on economic predictions to assess the impact of their policies. Financial institutions use forecasts in order to extrapolate relevant and valuable information from the markets and companies to identify the most remunerative investments. Needless to say, being able to provide forecasts as accurate as possible would be invaluable. However, due to the complexity of the underlying economic system, it is inevitable to incur in some form of error. These errors might assume several forms.
First, they could be the result of model miss-specification. Forecasters can mistake the form of relationship between variables, exclude relevant observations as well as include irrelevant variables.
Second, judgemental biases can account for forecast errors. Literature has shown that professional might value other goals more important than accuracy depending on the field in which they operate.
Finally, errors can be also explained by the involuntary use of non-informative (or *noisy*) economic data. If the ratio between the signal and the noise contained in the data is low, forecasters face several issues in extracting valuable information from the variables they deal with. The higher is the level of noise included in the data, the more likely forecasters might mistake correlation for causation. As a

consequence, forecasting accuracy is negatively influenced. Since economic forecasting is not entirely a *hard science,* it is not always possible to disentangle one source of error from the other. In order to solve these issues and enhance the accuracy performances of their models, forecasters constantly develop new techniques and employ different data **(chapter 1).**

Preliminary considerations on prediction markets

Prediction markets are venues in which participants can bet on future events. In order to do that, they interact with other participants by exchanging contracts tied to the event they are betting on. The most common contract negotiated on prediction markets is the *winner-takes all.* All the participants that bet on the outcome that turns out to be the actual realization are remunerated while the others lose the amount of money they had betted. For instance, let's assume that just two future states exist, *x* and *y.* Participants taking place to the prediction market, do not know which state will occur. However, they have some beliefs about it (i.e., past experience, information) and the trade accordingly to them. Once the market is open, participants start to exchange contracts between them. Suppose that, at a given point in time, *t*, before the closure of the market, the price of the contract tied to the event *x* is equal to 0.60$.

If the state turns out to be exactly *x*, an investor that buys this contract will be remunerated with an amount of money equal to 1-0.6$, for each contract he/she stipulates. Obviously, all the participants with contracts tied to the *y* state will lose the money they had betted. Since contracts exchanged on prediction markets are similar to binary options, it is

possible to interpret the contract price as a *market probability* that the underlying event will occur. Going back to the example previously presented if the price of the contract tied to the state *x* is 0.60$, market participants believe that the event *x* will take place with a probability of 60%.
Prediction markets display some interesting features that can be used in forecasting models. First, conversely to professional forecasters or polls, prediction market participants are forced to *"put their money where their mouths are".* By being remunerate if they are correct, participants tend to provide forecasts as accurate as possible. As it has been evidenced earlier, this is not always the case for professional forecasters. They can adjust their predictions to the *consensus* or value other goals more important than accuracy (i.e., publicity). Second, since individuals have economic incentives in finding new information and trade on it, prediction markets are able to efficiently aggregate different beliefs. Rapid incorporation of new information and resistance to manipulations are two consequences stemming from the efficient aggregation of information that, in turn, can increase the prediction accuracy. Finally, because participants are allowed to bet on different outcomes, a probability distribution function of the entire event can be derived. As one might expect, prediction markets are not invulnerable to drawbacks. To begin with, the correct balance between interest and contractibility of the argument is one of the main issue affecting prediction markets. Indeed, the ultimate goal of a prediction market is to attract the most disperse, in terms of beliefs, public. If prediction markets were able to allure only participants with the same beliefs, no transactions would take

place. This is why, the presence of *noise traders* is fundamental for having prediction markets that are liquid and, hence, efficient.

Second, participants with private or inside information can hinder the functioning of prediction markets. If other participants believe that with respect to the argument of the prediction market someone could have information that is not publicly available, they will unlikely enter in the market. Finally, also the absence of information regarding a specific argument might obstacle prediction markets. If there is no information that can be used by participants to make forecasts, obviously none will be able to make predictions (**chapter 2**).

Macroeconomic prediction markets

Although not as common as political prediction markets, two macroeconomic prediction markets operated in the last years. The first one denominated *"Economic Derivative Market"* was run between October 2002 and September 2005 by Goldman Sachs and Deutsche Bank operating as counterparts. Four economic variables were the object of separate auctions: Non-farm payroll, initial unemployment claims, retail sales and ISM index. The second one, called *"Economic Indicator Exchange"* (*"EIX"*), was run online with play-money during 2009. It allowed participants to forecast the released value of 5 variables regarding the German economy: GDP, inflation, investments, export and unemployment. The EIX design differed from the *Economic Derivative Market*, which shared the same features evidenced previously for prediction markets. To begin with, instead of holding auctions the day before the release day, the EIX worked as a continuous double auction. In this way, participants could constantly interact one with the other. This has helped to spur the liquidity of the market that is one of the main issues of prediction markets. Secondly, instead of several binary contracts, a single stock with a linear pay-off was exchanged. Finally, to maintain the participants active and motivated in the market, monthly and yearly prizes were hand out to the participants that provided the best forecasts. Literature has shown that for both the macroeconomic prediction markets, forecasts made by the participants were more accurate than the ones of professional forecasters. However, it must be noticed that the difference between the two is not statistically significant. Moreover, for what concerns the *economic derivative market*, professional forecasters and prediction market participants did not share the same information set. This is an element that, at least to a certain extent, could account for the different accuracy performances of the two forecasts **(chapter 2).**

Empirical approach and methodology

In order to investigate the thesis objects exposed at the beginning of this summary, three different but simple studies have been carried out. First, prediction markets data has been used to study if they provide accurate forecasts about the economic variable object of the prediction market. A relative comparison, in terms of prediction accuracy, has been performed. Three models have been set up. The first one is an autoregressive process of order one (*naïve model)* that has been used as a benchmark. The last observation of the economic release has been used to forecast the next one. The second one (*consensus model*) employs as predictors the consensus forecast and the one-period lagged observation of the economic release. The last one (*prediction market model*) differs from the *consensus model* in that prediction market forecasts have been used instead of consensus forecasts.

Second, prediction markets data have been used to study if they can convey useful information about the unfolding of other macroeconomic variables. Although it would have been interesting extending this analysis to several macroeconomic indicators, the limited available dataset, in terms of variables and number of observations, has hindered this study. Because of this, it has been possible only to investigate the relationship between non-farm payrolls data derived from prediction markets and the unemployment rate. Another variable present in the dataset that could have been used was initial unemployment claims. However, due to the limited number of monthly observations, it has not been possible to employ it as a predictor. This analysis has been set up to answer to two separate questions:(i) Does the prediction market model display the better accuracy performance? (ii) What does happen to the prediction market forecast when the actual release of the data and the survey prediction are used? Do they lose all the predictive power?

Third, it has been investigated if prediction market forecasts could, at least, to some extent, reflect or explain the underlying sentiment in the economy. Macroeconomic uncertainty measures are hard to come by. Although not completely satisfying, literature has evidenced that the time series of the standard deviation extrapolated from prediction market is a better uncertainty measure than disagreement between professional forecasters. However, I have focused on the *macroeconomic surprise or news* [1] instead of the standard deviation as a measure of uncertainty. The rationale behind this choice is the following. If all the economic agents taking part to the prediction market efficiently use all the information available and aim at being as accurate as possible, they should not be able to forecast only the uncertain part of that economic variable. So the forecast surprise should reflect, at least to some extent, the underlying uncertainty of the macroeconomic release. The larger is the surprise, the bigger is supposed to be the economic uncertainty. The VIX index, namely a measure of the S&P500 expected volatility, has been used to study the effect of the *macroeconomic surprise* on the implied volatility of financial markets. Instead of past observations of the VIX, I have used the change in the VIX closing price occurred between the opening price on the day of the announcement and the closing price on the day before the announcement. This has helped to circumscribe a span of time in which one might look for the impact of the *forecast surprise* effect **(chapter 3).**

Results

Before going through the results of the analyses and their discussions, it is necessary to spend few lines on the attention needed in deriving conclusions from these data.

First, the number of observations available for each economic derivative should not be overlooked. All the variables in the dataset, except initial unemployment claims, have a number of observations between 30 and 50. Obviously, a larger dataset

[1] Defined as the difference between the released value and the forecast.

could have helped to obtain stronger results both in an economic and statistical sense. Second, auction liquidity is another element to keep in mind in the analysis of the data. The more lively and liquid is a market, the stronger are the conclusions that can be extrapolated from the moments of the variable we are interested in. Finally, the regression results could be in some way biased by the span of time involuntarily taken into consideration.

As for the comparison in terms of accuracy between prediction market, consensus and naïve models, the results are mixed. Prediction market model, according to the past literature, outperforms (both in the in and out sample analysis) the others when raw variables are used. However, when stationary, hence differenced, data has used the results are different. Although the prediction model still offers a slightly better accuracy with respect to the consensus model, the naïve model largely outperforms the two aforementioned models. One possible explanation for this finding could lie in the variables used to perform the analysis. Since prediction market forecasts focus on predicting the *next future*, by taking the difference between the previous and actual prediction the overall level of information contained in the predictions decreases. Same results have been found for the 1-step ahead forecasts. Even in this context, naïve model outperforms the rest of the models.

The unemployment analysis presents interesting results. The prediction market model slightly outperforms in terms of accuracy both the naïve and the consensus model. The RMSE ratio between prediction market model and consensus model is 0.93, while the one between prediction market model and the naïve is 0.95. Moreover, it is interesting to notice that non-farm payroll forecasts have been the only parameter to be statistically significant and negatively correlated with the unemployment rate. This might suggest that prediction markets convey information with better qualitative contents relatively to other variables.

Finally, the relationship between macroeconomic uncertainty and financial volatility has been studied. Results display that the *macroeconomic surprise* is negatively and statistically correlated with the VIX price variation. Coherently with the literature, negative *surprise effects* tend to bring about positive changes in the VIX index, while positive *surprise effects* cause negative changes in the VIX. This implies that financial markets tend to be more volatile in the occurrence of negative *surprises* and to be less volatile in response to positive *surprise* **(chapter 3).**

Conclusion

Prediction markets have been proven to efficiently aggregate information and provide forecasts that tend to outperform other methods (such as poll or surveys). This is especially true for prediction markets about political events. However, for what concerns economic prediction markets, attention should not only be put on the accuracy of the forecasts made by the participants. Indeed, both the literature and the analyses carried out in this thesis have evidenced that the advantage in terms of prediction precision is limited with respect to consensus forecasts. Moreover, due to the small available dataset, it is difficult to prove that the difference between the two forecasts is statistically significant and relevant.

Because of this, more attention should be reserved for the pieces of information that can be extrapolated from the economic prediction markets and how to use them to enhance the comprehension of the economic system. For instance, as it has been done in this context, information contained in prediction markets can be used to mimic, at least to some extent, the underlying sentiment of the economy.

Unfortunately, further research about economic prediction markets has been obstructed by the closure of the *Macroeconomic Derivative Market* due to the lack of liquidity of the market. Perhaps, the creation of a large and lively new economic prediction market about the Euro area or the United States could spark the interest of both researchers and investors in this market. For instance, by having a continuous forecast market, it would be possible to study what are the effects of economic news on the forecasts provided by the prediction market participants. In this way, it would be possible to analyse the impact of not only economic but also politics news on the beliefs of economic agents. Then, it would be possible to compare the reaction of *common* financial markets and prediction markets to study if there are differences. Moreover, professional forecasters could get monthly flows of information about the GDP growth that could be exploited through now-casting models.

In summary, economists should not look at economic prediction markets only in terms of prediction accuracy but they should consider them as small-scale experimental economy where they can study how the economic agents beliefs evolve and unfold over time.