Åbo Akademi University

Fredrik Le Bell

The Time Series Convergence of Dispersion in Financial Analysts' Forecasts

Fredrik Le Bell (born 1980)

- M.Sc. (Econ. & Bus. Adm.), 2008, Åbo Akademi University
- B.Sc. , 2008, Åbo Akademi University (Natural Sciences)
- B.Sc. (Econ. & Bus. Adm.), 2007, Åbo Akademi University

Åbo Akademi University Press
Tavastgatan 13, FI-20500 Åbo, Finland
Tel. +358 (0)2 215 3478
E-mail: forlaget@abo.fi

Sales and distribution:
Åbo Akademi University Library
Domkyrkogatan 2–4, FI-20500 Åbo, Finland
Tel. +358 (0)2 -215 4190
E-mail: publikationer@abo.fi

THE TIME SERIES CONVERGENCE OF DISPERSION IN FINANCIAL ANALYSTS' FORECASTS

The Time Series Convergence of Dispersion in Financial Analysts' Forecasts

Fredrik Le Bell

Åbo Akademis förlag | Åbo Akademi University Press
Åbo, Finland, 2014

CIP Cataloguing in Publication

Le Bell, Fredrik.
The time series convergence of
dispersion in financial analysts'
forecasts / Fredrik Le Bell. - Åbo :
Åbo Akademi University Press, 2014.
Diss.: Åbo Akademi University.
ISBN 978-951-765-762-4

ISBN 978-951-765-762-4
ISBN 978-951-765-763-1 (digital)
Painosalama Oy
Åbo 2014

Abstract

The standard Bayesian learning model under asymmetric information due to Barry and Jennings (1992) and Barron, Kim, Lim and Stevens (1998), shows how dispersion in the forecasts made by financial analysts emerges as a function of the uncertainties that the analysts are facing. A key implication arising from these models is that dispersion in forecasts can only increase due to increased information asymmetry.

This study expands the models of Barry and Jennings (1992) and Barron et al. (1998) by explicitly considering the role that releases of common information, interpreted as annual earnings releases, have on belief convergence. The study shows that learning from common information in a fixed learning regime, causes rapid convergence of subjective beliefs. The convergence of beliefs on common information in turn dictates the magnitude of maximum dispersion that asymmetric information can cause, resulting in a monotonically decreasing maximum amount of dispersion.

Empirical estimations using data on all US listed companies between 1995-2010 confirm that observed levels of forecast dispersion exceed theoretically implied maximums when taking into account the amount of commonly observed information that has become available through earnings announcements. Exceedance of theoretically implied maximums for forecast dispersion is prominent when a company experiences negative earnings.

The joint evidence of the study suggest that levels of dispersion in analysts' forecasts are too high to find theoretical support. Consequently asymmetric information alone cannot yield observed levels of forecast dispersion.

The results have intuitive appeal, since dispersion increasing from asymmetric information as in Barry and Jennings (1992) and Barron et al. (1998), implies increased certainty. The evidence in this study instead suggests that dispersion resulting from asymmetric information alone is not possible, and thus opens up possibilities for interpreting increased dispersion as increased uncertainty. The study finally discusses potential pathways for such interpretations, involving agents restarting their learning procedures, or agents acting as if the conditioning distribution is non-fixed.

Svensk sammanfattning

Spridning i prognoser gjorda av finansanalytiker anses vanligen kunna förklaras med en Bayesiansk modell, utvecklad av Barry och Jennings (1992) samt Barron, Kim, Lim och Stevens (1998). Modellen visar hur spridningen i prognoser uppstår som ett resultat av den osäkerhet som analytikerna står inför. Den huvudsakliga förutsättningen för att spridning mellan analytikers prognoser skall kunna öka, är att informationsasymmetri ökar.

Denna avhandling utvidgar modellerna av Barry och Jennings (1992) samt Barron et al. (1998) genom att explicit beakta hur publikt tillgänglig information, information som tolkas uppkomma i form av företags resultatrapporter, måste påverka konvergensen i analytikernas övertygelser (beliefs). Avhandlingen visar att inlärandet som sker p.g.a. publik information i en fixerad Bayesiansk miljö leder till en snabb konvergens i subjektiva övertygelser. Konvergensen i övertygelser som sker p.g.a. publik information visar sig i sin tur diktera den maximala mängd spridning som den asymmetriska informationen kan orsaka, och detta leder till att den maximala spridningsmängden måste sjunka monotont över tiden.

De empiriska estimationerna i avhandlingen visar att den mängd spridning i prognoser som observeras i data överstiger den teoretiska maximimängden för spridning i prognoser, då man beaktar den mängd publik information som blivit tillgänglig genom företagens resultatrapporter. Det visar sig även att överträdelser i förhållande till de teoretiska maximinivåerna för spridning i prognoser förekommer speciellt i situationer där ett företag rapporterar förluster. Data som används är Amerikanska listade bolag under tidsperioden 1995 -2010.

Den sammanlagda bevismängden i avhandlingen tyder på att de empiriskt observerade nivåerna för spridning i analytikerprognoser är för höga för att kunna finna stöd i de modeller som utvecklats för att förklara dem. Därmed kan inte den asymmetriska informationen som vanligtvis används som en förklaring för en ökning i spridningen av prognoserna, ensam åstadkomma de nivåer i prognosspridning som de facto observeras i data.

Resultaten är intuitivt tilltalande då en ökning i prognosspridning till följd av asymmetrisk information, så som i Barry och Jennings (1992) samt Barron et al. (1998), samtidigt även innefattar att analytiker blir säkrare i sina övertygelser då spridningen i prognoser ökar. Resultaten ur denna avhandling hävdar istället att prognosspridningen inte enbart kan bero på asymmetrisk information, och detta öppnar samtidigt möjligheten till att i själva verket tolka en ökning i prognosspridning som ökad osäkerhet. Slutligen diskuterar avhandlingen möjliga banor för tolkningar av ökad prognosspridning som involverar ökad osäkerhet. Detta kan ske genom att analytiker tvingas starta om sin inlärningsprocess eller genom att analytikerna använder icke-fixerade sannolikhetsfördelningar i sin informationsmängd.

Dedicated to the memory of my father

Acknowledgments and Preface

I want to express my deepest gratitude to my supervisor, Professor Lars Hassel, for not only believing in my work, but also for providing the financial support that made it all possible. Even as I changed course in the choice of topic, and decided to attend the PhD program at Aalto University, you always saw to it that it was possible. I also want to thank you for giving me the time and freedom to pursue my dissertation topic to a depth and extent such that I can feel proud of my work, and at the same time for knowing exactly when to push me to wrap it up. Thank you.

I also really want to thank my second supervisor, Dr Rickard Olsson (Umeå School of Business and Economics). Through our working together, I learned valuable lessons about, among other things, being meticulous and really understanding what is going on in empirical data. Most importantly however, you really helped me believe in my research, and for this I am thankful. I highly value the time you took to help me out when we e.g. used to work late nights at the office in Stockholm.

I also want to thank my two pre-examiners, Professor Jörgen Hellström (Umeå School of Business and Economics) and Professor Mika Vaihekoski (University of Turku, Turku School of Economics) for time and effort put in to reading the manuscript and for providing valuable comments and suggestions for improvement. A special thank you goes to Professor Jörgen Hellström for reading and commenting on an earlier (and messier) draft of my dissertation, comments that (I hope) improved the dissertation but also gave me more confidence in the work I had done.

I am forever grateful to Dr. Mikko Leppämäki at Aalto University who accepted me into the graduate program GSF (Graduate School of Finance, Aalto University). Without it, the work in this thesis would have been utterly impossible. Naturally, this is very much due to the all the great people who taught the coursework at the GSF at Aalto University, and to whom I owe a great intellectual debt. These are, in order of teachings:

Associate Professor Patrik Sandås (McIntire School of Commerce, University of Virginia), Assistant Professor Peter Nyberg (Aalto University), Professor Markku Lanne (University of Helsinki), Professor Klaus Munck (Copenhagen Business School), Dr. Mikko Leppämäki (Aalto University), Professor Martin Ruckes (Karlsruhe Institute of Technology), Professor Renee Adams (UNSW Australia Business School), Professor Luigi Guiso (Einaudi Institute for Economics and Finance), and Professor Harrison Hong (Princeton University).

The PhD program, GSF, was probably the most intellectually difficult, but at the same time the most intellectually stimulating and rewarding process I have ever gone through. I don't think anything in my future will provide the same level of intellectual growth that this period of my life ended up providing. The knowledge that I gained not only made the work in this thesis possible, but in some ways even ended up transforming the way I view life in general.

I also want to thank Andrew Conlin and Jyri Kinnunen, my fellow PhD students who I met and worked through the GSF program with. Getting to know you guys made the whole process a memorable part of my life.

I want to thank all my colleagues at the Åbo Akademi University School of Business and Economics where the work on the dissertation has been (mostly) carried out, for providing a very nice atmosphere to work in. I also really want to thank my colleagues in the Accounting department specifically, for picking up the teaching load and allowing me to focus on my research when needed.

I want to add a specific thank you to my friend M.Sc. (Tech) Thomas Björklöf for the great help you provided in the beginning of my graduate school studies. You helped me out immensely with getting my algebra skills back on track, as well as getting me started with programming.

I am greatful for financial support from the MISTRA foundation, as well as the following foundations: Dagmar och Ferdinand Jacobssons Fond, Stiftelsen för Åbo Akademis Forskningsinstitut, and Klingendahls Fond.

Lastly, an enormous thank you to my mother and sister, as well as my friends, not only for your support, but for understanding and putting up with me prioritizing work first more often than not.

Finally, I also want to take the opportunity to say a few words about the process that eventually has lead up to this doctoral dissertation.

Working on a quantitative topic with a theoretical emphasis is in retrospect rewarding, but at times it can be anything but easy, and it certainly is not as linear as it might look given the end result.

At times, the frustration of months of work leading apparently nowhere can really test one's character. However the knowledge that others have been there before, and successfully overcome it, can provide hope and inspiration. This is eloquently pictured[1] in the following quotes by Emanuel Derman in his description of his PhD process in his book "My Life as a Quant":

"The rest of the time I came into my office each day and tried to work on various extensions of one of my models. [...] As time went by, I hit an impasse in the problem I was working on, and there were no colleagues with similar enough interests off whom I could bounce ideas. For months, every new attempt I made to circumvent the sticking point failed; I needed hope most of all. If early on a given day, I had a sudden idea for a strategy that might solve my problem, I would quit work immediately and go home. In this way I postponed disappointment for one more day. I preferred to do nothing for a little while, and then go to bed savoring the little dollop of optimism that my new strategy would work the following day. Usually, my joy lasted only for a few hours; each new method failed pretty quickly."

While the above can give a sense of what the more difficult times during the process can be like, thankfully the following quote from Emanuel Derman's book "My Life as a Quant", applied much more often:

"I still remember the excitement of working late into the night [...] I recall best the unspoiled joy of spontaneously waking early, tired but driven, and then rushing off to work because I wanted to go to work and couldn't sleep any more. I was excited to see what came next."

I can personally still vividly remember that exact Saturday when I finally came up with the solution on how to solve the main problem that my dissertation is concerned with. After that very day, I feel that the dissertation almost wrote itself.

While the work of my PhD studies culminates in the approximately 150 pages of this doctoral dissertation, I would like to shed some light onto the amount of, in some sense invisible, work that can go into the exploration of a quantitative topic.

[1] I thank Dr. Peter Nyberg for originally bringing me on to these quotes.

Some indication can be found in the amount of algebra/calculations that this work has involved. When I include the work done in graduate school to the derivations of the results in this dissertation, the number of pages of hand written calculations/algebra I have in my office easily exceeds 1000 pages (I stopped counting after I reached 1000 pages). There are also thousands after thousands of lines of code written for various programs, endless hours of simulations that have been carried out, and finally a research diary used for keeping track of results and ideas that has ballooned to close to five hundred pages at the time of writing.

This makes the dissertation itself, the proverbial tip of the iceberg.

Åbo, 4.11.2014

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Introduction

Professional financial analysts, who forecast firm earnings and issue buy and sell recommendations, tend not to agree in their forecasts about the future. The disagreement in their views about the future is readily observable in that the forecasts they generate exhibit dispersion. This disagreement is illustrated in Figure 12[1]. To a layman disagreement about the future hardly comes as a surprise - since forecasting the future is fraught with uncertainty, it is not difficult to imagine forecasters having divergent views on the future. In rationally anchored theoretical models for how financial analysts formulate their forecasts such as Barry and Jennings (1992) and Barron et al. (1998), the link between disagreement and uncertainty is more complex.

The models of Barry and Jennings (1992) and Barron et al. (1998) provide the standard setting for how analysts are thought to formulate their forecasts. Here, in the context of the model, an agent is faced with two sources of uncertain information; common and privately observed. Common, or public information is information available to all agents, whereas private information is individual specific. All information is normally distributed. The model assumes Bayesian learning, and implies that an agent has a prior (uncertain) belief on the true value of the parameter he/she is trying to learn. A noteworthy general feature of the Bayesian learning model is that beliefs are updated over realizations. Upon receiving new information, the agent uses the new information in conjunction with his/her prior belief, to arrive at

[1]See page 80.

an updated, sharper belief on the parameter he/she is estimating. This is called the posterior belief. When agents have different information sets, due to differential or asymmetric information, their (posterior) beliefs differ. When the assumed setting of the model is that of financial analysts, disagreement manifests itself as dispersion in forecasts.

In the models of Barry and Jennings (1992) and Barron et al. (1998), serving as the standard theoretical framework for analyzing forecasts, dispersion does not directly measure the underlying uncertainty or risk. There will, for example, exist no disagreement no matter how high the uncertainty is, if all agents share the same information[2]. Consequently, in order for disagreement to exist within the context of the model, agents need to be endowed with privately observed information and dispersion in forecasts can only *increase* if there is an increase in private information (information asymmetry).

Private information is the only factor that can yield increases in forecast dispersion in the models of Barry and Jennings (1992) and Barron et al. (1998) and empirical studies that observe increases in dispersion typically conjecture that private information is the cause for increased dispersion. Examples where inferences are drawn based on a rationale similar to the above, where dispersion in forecasts is assumed to result from information asymmetry, include e.g. Lang and Lundholm (1996), Adut, Sen and Sinha (2008), Ali, Liu, Xu and Yao (2009), Barron, Stanford and Yu (2009). Careful theoretical analysis of the Barry and Jennings (1992) and Barron et al. (1998) models however reveals that the aforementioned interpretation, where dispersion increases due to increased information asymmetry, nests a somewhat counterintuitive consequence. When agents gain access to more privately observed information, resulting in an increase in dispersion, the agents themselves become more *certain*. In a Gaussian Bayesian learning model under asymmetric information, *increases* in disagreement and increases in certainty are inherently two sides of the same coin. The realism of this effect, where increases in dispersion imply more certainty, an inherent feature of the Bayesian model[3], is usually not specifically addressed.

Theoretically, dispersion can only increase if agents are endowed with more private information and as a result, intuitively linking increases in dispersion to increases in uncertainty is contradicted by theory. Further, the Bayesian learning present in the models of Barry and Jennings (1992) and Barron et al. (1998), not only limits

[2]This is of course in no means limited to the models of Barry and Jennings (1992) and Barron et al. (1998), rather this is a standard result whenever agents are rational and information is common.

[3]This is a feature of the Gaussian Bayesian learning model under asymmetric information.

increases in dispersion to increased private information, but in a strict sense agents can never become uncertain in their subjective beliefs. Increased certainty on new information holds for all information, common and private and is a general feature of Gaussian Bayesian learning. Yet, it seems that in reality, people can become more uncertain in their beliefs.

In the context of financial analysts forecasting firm earnings, it appears that in some states of the world, analysts do seem to believe that it is possible for the future of a company to become more difficult to forecast, and thus more uncertain.

The suspicion that dispersion is linked to views of increased uncertainty can be exemplified anecdotally by analysts' written reports[4] for companies with poor[5] performance, where analysts often refer to the futures of these companies with expressions along the lines of "decreased visibility", "weak outlook" or "with visibility relatively low". These verbal descriptions, at least heuristically, seem to describe states where the future has, simply put, become more uncertain.

The above, citing analysts descriptions on the future of weakly performing companies, thus provides a hint at the possibility that analysts *in reality* can become subjectively more uncertain about the variable they are forecasting. This however is not possible under the theoretical models of Barry and Jennings (1992) and Barron et al. (1998). In cases where these low visibility environments are associated with increases in dispersion, the interpretation resting on the standard theory would be that these increases in dispersion would have resulted from analysts having acquired more private information. Again however, increased private information implies increased *certainty*, not uncertainty.

It is this somewhat contradictory issue, where theory suggests increases in dispersion imply more certainty but intuition and anecdotal evidence suggest dispersion implies increased uncertainty, that this study aims at providing further evidence on.

1.2 Background

The literature using data on financial analysts is voluminous, especially in the context of capital markets research in Accounting. This is exemplified in the preface to the "I/B/E/S Research Bibliography" by Brown (2000), where he notes that the

[4]Content analysis of selected, undisclosed analysts' earnings reports. Analysts typically provide these extensive written analyses along with their estimates and recommendations.

[5]This refers to companies either in loss territory or with a an outlook that is deteriorating, where losses are subsequently observed ex post.

I/B/E/S Research Bibliography gathers over 500 studies on analysts expectations that make use of I/B/E/S data[6]. Studies surveying the literature on the use of analyst data in various settings can be found in e.g. Givoly and Lakonishok (1984), Brown (1993), Ramnath, Rock and Shane (2008).

Why are analysts forecasts important and widely used in research? Prices, including those of stocks, theoretically equal discounted future payoffs (Cochrane, 2005). The view that it is only the future prospects, not the past, that is important in the purchase decision of a stock, is already present in work such as Graham and Dodd's Security Analysis from 1934. Anticipation, or expectations of the future are thus key in the price or value of an asset. These fundamental views are also mirrored in Cragg and Malkiel (1968): "The price of a share (of common stock) is − or should be − determined primarily by investors' current expectations about the future value of variables that measure the relevant aspects of a corporation's performance and profitability, particularly the anticipated growth rate of earnings per share." In the introduction to the I/B/E/S Research Bibliography" (2000), Lawrence Brown, also mentions the views of both Graham and Dodd (1934) and Cragg and Malkiel (1968). The first words of the introduction to the I/B/E/S Research Bibliography" (2000) however embodies the fundamental importance of expectations in one short sentence: "Expectations drive share price".

It is thus expectations of the future that drives stock prices. Since expectations are generally unobserved, the rationale for using financial analysts as objects of study, is that data on financial analysts' forecasts represent a unique proxy for the otherwise unobservable investor expectations (e.g. Givoly and Lakonishok (1984), Brown (1993), Bradshaw, Drake, Myers and Myers (2012) Keane and Runkle (1998)).

While analysts' forecasts are important on a general level, as they serve as proxies for unobserved investor expectations, this study specifically focuses on the *variability* between the forecasts of analysts. Generally, measures capturing the spread or variability of a variable are tied to notions of risk or uncertainty. Consequently forecast dispersion, measuring the spread of reported forecasts by analysts, at first glance corresponds well with an idea of uncertainty or risk. Thus, through studying the dispersion in forecasts, one hopes to gain an understanding of the uncertainties or perceived risks that analysts are facing. In the models for how analysts form their expectations such as Barry and Jennings (1992) and Barron et al. (1998), it turns out that dispersion in forecasts is not necessarily a direct function of the uncertainty that analysts are confronted with.

[6] I/B/E/S is short for the Institutional Brokers Estimate System, gathering data on analysts' earnings estimates since 1971.

There is also ample evidence that forecast dispersion is related to various empirical phenomena. For example, dispersion in forecasts has been shown to be related to the accuracy of the forecasts that analysts produce (Lang and Lundholm, 1996). Forecast dispersion has also been shown to be related to both higher credit spreads on corporate bonds (Guntay and Hackbarth, 2010), as well as worse credit ratings (Avramov, Chordia, Jostova and Philipov, 2009). Furthermore, there is growing evidence that dispersion in forecasts has implications for asset pricing. This follows the Diether, Malloy, and Scherbina (2002) result that disagreement between financial analysts leads to lower cross-sectional stock returns. If one views disagreement as a proxy for risk, this result is somewhat puzzling from a classical risk return standpoint[7]. Follow up studies by Park (2005) and Yu (2011) establishes the negative relation between forecast dispersion and return on a portfolio level. Recent (unpublished) work by Ali et al. (2009) provide evidence that the negative return relation is driven by negative earnings surprises, whereas (unpublished) work by Xu and Zhao (2010) find evidence that the dispersion negative return relation is driven by idiosyncratic volatility.

The study of analysts' forecasts is a well established field. This is exemplified by the sheer volume of studies of analysts' forecasts, seen in both the I/B/E/S Research Bibliography (2000) and the survey papers by Givoly and Lakonishok (1984), Brown (1993), Ramnath, Rock and Shane (2008) covering research in the field. Expectations are integral for the pricing of financial assets, hence studying analysts' forecasts (proxies for expectations) can provide answers on the link between information, expectations and valuation. Dispersion in forecasts is superficially linked to uncertainty or risk, and dispersion in forecasts has been shown to be empirically linked to various phenomena, seen in the short review in the preceding paragraph. Yet, there exists something of a disconnect between the predictions from the main theoretical models that model the emergence of forecast dispersion, Barry and Jennings (1992) and Barron et al. (1998), where dispersion implies certainty, and interpretations of dispersion as risk or uncertainty.

1.3 Formulation of the problem

Forecasts exhibit dispersion. This can be seen in Figure 12[8]. In theoretical models for how financial analysts formulate their forecasts such as Barry and Jennings

[7] Diether et al. (2002) therefore reach the conclusion that dispersion cannot be used as a proxy for risk. Instead, the authors infer that the result is driven by a Miller (1977) type of effect.

[8] See page 80.

(1992) and Barron et al. (1998), dispersion in forecasts can only increase when analysts gain access to more private information. Indeed, Barron et al. (2009) revisit this idea in a recent study, conjecturing, in an earnings announcement setting, that "[E]arnings announcements increase information asymmetry, on average, because analysts and investors develop new private information in conjunction with the announcement". This notion of increased production or acquisition of private information being the driver behind increases in dispersion, will henceforth be labelled the *private information acquisition hypothesis*.

The anecdotal evidence from analysts' written reports finds analysts citing the outlook for poorly[9] performing companies in particular, as having low or weak visibility. The fact that analysts find the futures of some companies, especially those performing poorly, as having low or weak visibility, points toward analysts seeing increased uncertainty about the future of the companies. It also turns out that companies reporting losses have associated levels of dispersion that are on average four times higher than companies reporting positive earnings[10]. If the private information acquisition hypothesis is valid, and since dispersion is associated with losses, the implication is that agents do not become more uncertain around negative earnings, rather they develop more private information and become more certain about the future of the firm. Such a hypothesis would, however, contradict both a heuristic idea of negative earnings indicating more uncertainty and a notion of low or weak visibility.

The problem is thus that the theory that describes analyst behavior, such as Barry and Jennings (1992) and Barron et al. (1998), states that dispersion in forecasts can only increase if information asymmetry increases, simultaneously implying a reduction in uncertainty. At the same time, dispersion is high in environments characterized by negative earnings, where analysts describe the futures of poorly performing companies as uncertain. This represents an obvious contradiction, since analysts say they become more uncertain, but theory suggests they become *less* uncertain. This contradiction between theory and evidence is exacerbated by the fact that empirical work, especially in Accounting Research, uses the prediction of increased information asymmetry as an explanation for observed dispersion.

On a more specific theoretical level, it seems that an important factor affecting models such as Barry and Jennings (1992) and Barron et al. (1998), has been largely left unconsidered. This factor consists of how information must affect learning over

[9]A poorly performing company is loosely defined at this point as one that will report negative earnings ex post.

[10]See e.g. Table 3, Panel (A) on page 87.

time. An explicit application of the Barry and Jennings (1992) and Barron et al. (1998) style models to a time series setting has, to the best knowledge of the author, not been previously addressed in research. The models of Barry and Jennings (1992) and Barron et al. (1998) are very general and could in some sense be thought to already encompass a time dimension, in terms of how agents update over a sample as in DeGroot (1970). The models of Barry and Jennings (1992) and Barron et al. (1998) however contain no clear mentioning that they should be interpreted as such. Neither is there any indication of how and at what rate, information should become available to agents.

This study identifies a theoretical gap, in that the models for how analysts formulate their forecasts, such as Barry and Jennings (1992) and Barron et al. (1998), do not explicitly take into account the effect of convergence in beliefs that occurs sequentially over annual earnings announcements (common information). The study consequently develops a representation of the asymmetric information Gaussian Bayesian learning model of Barry and Jennings (1992) and Barron et al. (1998), a model that specifically accounts for the release of common information over time. The model yields restrictions on the magnitude of forecast dispersion that can be used to test the underlying theory. The compatibility of observed levels of dispersion with theoretically implied levels has implications for whether increases in private information can be thought to unambiguously explain increases in forecast dispersion.

Testing for whether forecast dispersion conforms to the theoretical bounds in essence equates to testing the underlying model itself. Due to the general inherent difficulty of identifying what information agents are facing, or what information is contained in their information sets, such direct tests in these settings are not common. There thus exists an empirical gap in testing if the observed level of dispersion are compatible with theory. Testing the compatibility of observed levels of dispersion with those of theory is possible through the novel idea in this study, where levels of dispersion can be compared to the amount of common information that agents have updated over.

While some of the tests performed in this study seem superficially related to those of other authors (e.g. Adut et al., 2008), the tests are very different in spirit and in the set of questions they set to answer. E.g. Adut et al. (2008) use the underlying theory of dispersion increasing in information asymmetry, in their case the Barron et al. (1998) version, in *explaining* empirically observed phenomena, i.e. forecast dispersion. The stance of this study on the other hand is to test the performance of the Bayesian learning model, even though the variables used in the tests are similar to e.g. those used by Adut et al. (2008). Ideologically (in terms of testing aspects of underlying theory), the tests in this study are much more similar in spirit to e.g. Keane and Runkle (1990), and Keane and Runkle (1998), who generally assess rationality of forecasts and the unbiasedness of predictions. This study however not only differs from e.g. Keane and Runkle (1990), and Keane and Runkle (1998) in terms of methodology, but takes rationality as given, and tests whether observed levels of forecast *dispersion* can be generated by the asymmetric information Gaussian Bayesian learning model.

Thus, the ultimate empirical question in the study is to test the performance of the underlying rational learning model in order to provide answers on the theoretical question of whether (increases in) forecast dispersion must imply increases in information asymmetry and individual certainty.

1.4 Aim and Methods

Increases in the dispersion of analysts' forecasts are typically attributed to increases in information asymmetry. In a Bayesian learning model under asymmetric information, such as that of Barry and Jennings (1992) and Barron et al. (1998), increases in information asymmetry occur due to more privately observed information which leads to individual agents becoming more certain in their beliefs. Consequently, upon observing increases in dispersion, the implicit (standard) conclusion is that agents on an individual level become more certain in their beliefs. This study aims at analyzing in detail if levels of dispersion can be explained by asymmetric information as the theory of Barry and Jennings (1992) and Barron et al. (1998) predicts.

The purpose of this study is twofold, and consists of an interlinked theoretical and empirical part.

The theoretical purpose of the study is to perform an extension of the model of Barry and Jennings (1992) and Barron et al. (1998) for how financial analysts are thought to formulate their forecasts. The contribution lies in explicitly taking into account the amount of common information that financial analysts become endowed with through earnings announcements. The idea is to keep track of the amount of learning (convergence of individual beliefs) that occurs from common information and subsequently check what magnitudes of disagreement due to private information the model can sustain.

The empirical purpose of the study is to test the developed model extension that takes into account the amount of common information that financial analysts must become endowed with through earnings announcements. Specifically, the study aims at empirically assessing whether derived maximum bounds for dispersion implied by the model extension are violated in the data. If the magnitude of disagreement breaches theoretically predicted values, it rejects the learning model and implies that financial analysts do not necessarily always have to become more certain in their beliefs.

To address the issue on whether private information can yield sufficient magnitudes of disagreement, the study starts by developing a model[11], that retains all the elements of Barry and Jennings (1992) and Barron et al. (1998) and shows how uncertainties and forecast dispersion evolve over time as analysts must gain access to increasing amounts of common information through annual earnings announcements. Since private information is unobservable, cumulative common information constitutes a benchmark for the size of disagreement. Here, the study is guided by the intuition of Brown (1993), who advocates that " Bayesian theory suggests that earnings announcements generally increase the precision of individual's estimates of *future* earnings", emphasis added.

While the model aims at preserving the elements of Barry and Jennings (1992) and Barron et al. (1998), the nature of the public information is such that it is disseminated slowly over time[12]. This has implications for the evolution of forecast dispersion over time, and a secondary objective of this study is to illustrate this evolution in detail.

More importantly, the model, or representation, yields restrictions for how large forecast dispersion can become, at maximum, for each period. Through the restrictions for maximum dispersion, it is possible to evaluate whether the private information acquisition hypothesis holds. This has direct implications for whether increases in dispersion can be interpreted as increases in certainty, as is the case under the private information acquisition hypothesis, or whether the model is too restrictive and interpretations involving increased uncertainty are more likely candidates.

A detailed analysis of forecast dispersion in a setting where common information is released over time through earnings announcements, along with the implied maximum bounds for forecast dispersion, define the theoretical contributions of the study.

[11]it is more fitting to use the term representation or application, since the intention is not to state the model in *mathematical* terms is a new model. The innovation lies in in the way the model is interpreted to apply in a setting of a sequence of earnings announcements.

[12]This can be seen as a piece by piece construction of the *sample* used in reaching the posterior in e.g. DeGroot (1970)

In order to evaluate the performance of the model, and consequently that of the private information acquisition assumption, empirical tests are carried out for the derived maximum bounds for dispersion. The evaluation of the maximum bound rests crucially on identification of the uncertainty in the public information. Considerable attention is given to this matter such that confidence in identification is reached.

Estimations are carried out on a sample consisting of all US exchange listed companies from 1995-2010 that have data available in Worldscope and I/B/E/S. The contradiction between heuristically perceived increases in uncertainty and increases in dispersion is epitomized in negative earnings environments. The empirical assessment of the maximum bounds therefore additionally employs negative earnings as causes for potential exceedance of the maximum bounds for dispersion.

The empirical estimations themselves consist of standard pooled OLS regressions, but the novelty lies in a direct mapping of theoretical concepts into observable quantities in the form of an uncertainty ratio. The derived maximum for dispersion is a function of the exogenous uncertainty agents are assumed to face and the number of periods. Dividing through by exogenous uncertainty yields the uncertainty ratio, forecast dispersion in relation to exogenous uncertainty, which is a simple function of the number of periods. The uncertainty ratio, as measured in the data and used as a dependent variable in the empirical estimations, has a theoretically dictated exact maximum threshold value. By controlling for the (minimum) amount of periods that can be observed in the data, the theoretically implied maximum value for the uncertainty ratio is trivially computed. By performing post estimation Wald tests for implied threshold values, the significance of the potential exceedance of the theoretical bounds can be assessed.

The empirical aspiration of the study is thus to test whether observed levels of forecast dispersion are compatible with the maximum bounds that take into account the amount of information that agents have gained access to through annual earnings announcements. The conformity of dispersion in forecasts with the maximum bounds, provides answers on the applicability of the Bayesian learning model under asymmetric information when common information dissemination is accounted for. Failure of the maximum bounds to hold implies that the learning model nesting increased subjective certainty does not hold, thus giving space to explanations involving actual increased subjective uncertainty.

Finally, the study aims at providing a supplementary mechanism that is able to explain larger increases in forecast dispersion, a mechanism nesting a notion of increased uncertainty without sacrificing rationality.

1.5 Theoretical Landscape

The standard (rational) theoretical approach to model how financial analysts formulate their earnings forecasts, typically utilizes a Bayesian learning approach, where an agent learns about a parameter through sampling from a normal distribution. This approach originates in DeGroot (1970). See Eq. (1) on page 15. In the context of financial analysts' forecasts, an updating equation such as Eq. (1) can be found in e.g. Barry and Jennings (1992), Barron et al. (1998), and Ottaviani and Sorensen (2010). When the signal agents are receiving is understood to be private, in the sense that other agents do not observe its realization, forecasts differ. Barry and Jennings (1992) are the first to derive explicit formulas for the dispersion in forecasts in setting equivalent to the above, whereas Barron et al. (1998) augment the analysis by adding the role of the forecast error. While this study does not focus on the role of the forecast error, the study of Barron et al. (1998) is added as a second study that will be referred to as "benchmark" studies regarding forecast dispersion.

Naturally there exists other ways of approaching forecast dispersion or disagreement from a theoretical point of view. The study proceeds by reviewing some of these approaches below. The model of Abarbanell, Lanen and Verrecchia (1995) is a rationally based equilibrium model, ultimately aimed at analyzing volume and price behavior around earnings announcements. While the model contains forecast dispersion as an effect, the model is distinct since forecasts are drawn exogenously, a fact emphasized in Barron et al. (1998). Also, Abarbanell et al. (1995) are explicit about the fact that the model specifically contains two periods.

An ideologically contrasting starting approach is taken in the models of Harris and Raviv (1993) and Kandel and Pearson (1995), analyzing volume and trade behavior resulting from differences in beliefs/opinions. While the models are not focused on explaining forecast dispersion, they are related in the sense that they incorporate a theoretical mechanism that results in differing beliefs. This mechanism is however different from the one employed in this study, since the distinguishing assumption in Harris and Raviv (1993) and Kandel and Pearson (1995) is that agents are allowed to interpret public information differently.

A similar mechanism for belief diversity is employed in Morse, Stephan and Stice (1991). Morse et al. (1991) is an early study close to the empirical setting of this study in that Morse et al. (1991) find that dispersion in forecasts tends to increase following earnings announcements[13]. Theoretically, the proposed model by Morse et al. (1991) is however different from the rational Bayesian setting of this study where common information leads to belief convergence. Agents in the Morse et al. (1991) model are allowed to interpret a single common realization differently and this makes the model of Morse et al. (1991) to be ideologically related to the approach of Harris and Raviv (1993) and Kandel and Pearson (1995).

Furthermore, studies such Diether et al. (2002) and Yu (2011), invoke the rationale of the model of Miller (1977) in explaining empirical results regarding forecast dispersion. Miller (1977) presents the argument that as long as the entire supply of a security can be absorbed by a minority of potential purchasers (and pessimists face short sale constraints), the market price will be above the mean evaluation of the potential investors, and thus represents the views of the most optimistic investors[14]. However, since the Miller (1977) model conditions on the existence of heterogeneity, and lacks a clear quantitative belief diversity mechanism that depends on information, it is also distinct from the rational, asymmetric information approach invoked in this study. It is also somewhat unclear how disagreement among *analysts* would increase from them being constrained from absorbing the supply of securities.

Finally, Johnson (2004) develops a model for the relationship between forecast dispersion and expected returns in a continuous time setting. The model rests on the idea that fundamentals are unobservable, and information about the fundamentals adds a separate layer of (idiosyncratic) uncertainty. This, in combination with a contingent claims analysis of capital structure, leads to decreasing expected returns for higher levels of idiosyncratic (parameter) risk. One important differentiating factor can be found in Johnson (2004): "Second, I do not explicitly model information production/acquisition or capital structure choice, which are both clearly endogenous aspects of the problem". The former is exactly the purpose of the modeling here.

It is worth noting that in the "benchmark" models, and consequently in the representation developed in this study, it is assumed that forecasters honestly report their expectations in the forecasts they publish. Indeed, in using financial analysts forecasts as proxies for expectations, this assumption would be implicit, unless the

[13]The evidence from Morse et al. (1991) is subsequently questioned in Brown and Han (1992), who argue that dispersion only increases following earnings announcements for the largest decile of earnings surprises.

[14]A further implication of the Miller (1977) model is thus that an increase in the divergence of opinion will increase price, and thus as is standard, and noted by Yu (2011), lead to a subsequent lower return.

converse is pointed out. An example can be given in the model of Ottaviani and Sorensen (2010), where forecasters act strategically.

A final differentiating factor, that unfortunately makes comparing approaches somewhat difficult, stems from distributional assumptions regarding uncertainty. Many models, representing both the rational and the behavioral approach, use a (at least partial) binary setup. In a binary setup, even where rationality is maintained, new information can actually reduce the confidence in an estimate, whereas the (rational) fully Gaussian case never allows for this - new information must, by construction increase the precision on the subjective confidence in the estimate (Chamley, 2004).

Different approaches to model disagreement exist, as the short theoretical review demonstrates. Since the purpose in this study is to explore issues relating to the private information acquisition hypothesis and the embedded predictions regarding subjective uncertainty, the model in this study is developed to closely resemble those of Barry and Jennings (1992) and Barron et al. (1998). The underlying model is thus a rational Bayesian model, subject to asymmetric information where all information is distributed normally and the informative distribution remains fixed throughout what is interpreted as the time dimension. A discussion on the applicability of alternative explanations driving forecast dispersion is left for the concluding section when the performance of the model has been investigated empirically.

1.6 Outline of Study

Chapter 2 commences by introducing the standard theory of Bayesian belief updating. The analysis is then augmented with asymmetric information, after which results pertaining to the variance of beliefs (forecast dispersion) are presented. The study then proceeds by showing how the standard Bayesian belief updating model under asymmetric information can be applied in an earnings announcement setting that spans multiple periods. The problem is modelled in a specific setup where both public and private signals are drawn each period. Putting structure on private information in addition to that of public information aids in analyzing the behavior of forecast dispersion, which is carried out in Chapter 3, but for the most important results, private information is treated as a free parameter.

Chapter 3 aims at illustrating the dynamic behavior that forecast dispersion can take. This is best illustrated by anchoring the behavior of forecast dispersion to the specific setup where both public and private signals are drawn each period as a benchmark. Additional changes to private information endowments are then carried out. By calibrating the model to reasonable values that can be observed from empirical data, the chapter illustrates more intuitively the maximum bounds for forecast dispersion, as well as under what conditions dispersion in forecasts can be thought to increase.

Chapter 4 is aimed at empirical assessment of the multiperiod representation developed in Chapter 2. More specifically, the maximum bounds for dispersion, derived in Chapter 2, are tested in the data. Estimations are carried out on a sample consisting of all US exchange listed companies from 1995-2010, where companies are required to have data available in Worldscope and I/B/E/S. Considerable effort is exerted towards providing proper identification of the uncertainty that agents in the model are assumed to condition on. Chapter 4 also considers and tests and alternative hypothesis for the observed levels of forecast dispersion.

Chapter 5 summarizes both the theoretical and empirical findings. Chapter 5 also discusses what implications the results have on agents' subjective uncertainties. In addition, the study assess how the empirical results are related to both other theoretical approaches, as well as that of the alternative hypothesis of Chapter 4. Chapter 5 also discusses how the evidence of the study can be positioned in the context of the asset pricing evidence in the literature that is related to the dispersion in forecasts. Finally, the chapter concludes and makes explicit the contributions that this study makes.

Chapter 2

Theory

2.1 Bayesian Belief updating

This section starts by introducing the standard "closed"[1] *general* Bayesian framework for updating beliefs, to illustrate in its simplest form how agents filter information through a belief updating equation to arrive at an updated expectation and how subjective uncertainty is affected. All the analysis in this study will be set in a Gaussian framework. Since agents are assumed to honestly report their expectations, expectations of the agents in the model are analogous with forecasts made by analysts.

The general form of how a Bayesian agent processes and updates his beliefs of an unknown scalar parameter θ over new information is as follows. Initially, the agent starts out with a prior belief. The prior belief about θ is normal, and summarized by $\theta \sim N(\mu_0, \tau_0^2)$. Next, the agent receives a signal (observation/realization) from a distribution informative on θ, parameterized as $y \sim N(\theta, \sigma^2)$. This is analogous to stating that $y = \theta + \eta$, where $\eta \sim N(0, \sigma^2)$. Note that the agent is assumed to know σ^2. The agent now updates his/her belief in accordance with Baye's rule as:

$$\mu = E[\theta|y] = \frac{\frac{1}{\tau_0^2}\mu_0 + \frac{1}{\sigma^2}y}{\frac{1}{\tau_0^2} + \frac{1}{\sigma^2}}. \tag{1}$$

This is the general form of the posterior mean, resulting from updating a prior (normally distributed) belief with a realization from a normal distribution centered on

[1] Standard models of Bayesian belief updating will be referred to as "closed" or "fixed". With the risk of not giving a fully rigorous definition, what is implied by closed or fixed is the standard assumption that the parameter to be estimated is unknown, but can be represented by a probability distribution. This distribution is assumed to remain "fixed" over time. This is not the same thing as the parameter, θ, being fixed in the Frequentist sense.

the true mean[2]. It can be noted that the posterior is actually a posterior *distribution* of θ, where the posterior distribution in this case is normal, DeGroot (1970). The posterior variance, following DeGroot (1970), is given by:

$$\tau_1^2 = \left(\frac{1}{\tau_0^2} + \frac{1}{\sigma^2}\right)^{-1} \tag{2}$$

Thus, the agent's full posterior distribution of θ, conditional on y, is:

$$\theta \sim N\left(\frac{\frac{1}{\tau_0^2}\mu_0 + \frac{1}{\sigma^2}y}{\frac{1}{\tau_0^2} + \frac{1}{\sigma^2}}, \left(\frac{1}{\tau_0^2} + \frac{1}{\sigma^2}\right)^{-1}\right) \tag{3}$$

This can of course be summarized by: $\theta \sim N(\mu, \tau_1^2)$.

Generally, the decision rule to forecast the mean can be shown to minimize squared error loss (See e.g. Ottaviani and Sorensen, 2010; DeGroot, 1970). The decision rule to forecast the mean is implicit and standard in e.g. the models of Barry and Jennings (1992) and Barron et al. (1998). The Bayesian forecasting procedure, or decision rule, invoked here implicitly uses squared error loss as a loss function and results in the minimization of posterior risk. As a consequence, it is therefore implicitly assumed that utility functions are quadratic (Berger, 1985; Särkkä, 2013).

Since some authors referred to in this study (e.g. Prendergast and Stole, 1996) prefer a variance representation instead of a precision representation, and uncertainties have a natural analog in variances, it is perhaps useful to note that an equivalent representation of equation (3) in terms of variances, is given by:

$$\theta \sim N\left(\frac{\tau_0^2 y + \sigma^2 \mu_0}{\tau_0^2 + \sigma^2}, \frac{\tau_0^2 \sigma^2}{\tau_0^2 + \sigma^2}\right)$$

The uncertainty about θ, after receiving a signal (and conditioning on the prior), decreases mechanically with the amount $1/\sigma^2$, DeGroot (1970). This property of the Bayesian updating regime says that the *uncertainty* about the parameter that each agent faces, can never increase[3]. Another way of stating this is that the variance of the posterior is always smaller than the variance of the prior, that is $\tau_0^2\sigma^2/(\tau_0^2 + \sigma^2) < \tau_0^2$, as made explicit by Prendergast and Stole (1996). As the

[2]The study does not venture further into *why* the belief updating equation takes this specific form. For a thorough analysis of how this (standard) result arises through the use of the likelihood function, see eg. DeGroot (1970), Berger (1985) or Gelman, Carlin, Stern and Rubin (2004).

[3]This is the case if agents are pure Bayesians, they know exactly the distributional properties of their signals and the true mean of the process exists. These assumptions are relaxed later when explaining the empirical results.

study moves on to discuss disagreement, it is important to note that at this point the analysis specifically concerns *individual* beliefs.

A useful feature concerning the signals is their linear additivity. If an agent receives an additional signal from the same distribution as his/her prior, in the context of equation (2), his/her updated uncertainty about θ would be $\tau_1^2 = (1/\tau_0^2 + 1/\tau_0^2)^{-1} = (2/\tau_0^2)^{-1} = \tau_0^2/2$. Receiving additional signals of the same magnitude, results in an evolution of the subjective belief as depicted in Figure 1 on page 44. This also makes clear that multiple signals can be summarized by just one signal. This is a useful property in dealing with private signals in the model development later: Analysts could also be thought to receive any number of private signals during e.g. any interim period between annual earnings announcements, but the total information content of these signals can be represented by just one signal.

2.2 The standard model

This section presents the standard model of belief updating with common and private information that leads to differing beliefs between agents. The results are very general and do not take a specific stance on the exact evolution of information over time. The development closely follows that of Barry and Jennings (1992) and Barron et al. (1998).

There are N agents in the economy (indexed by i:j,k,l,...,N) who forecast an unknown earnings variable θ. Prior beliefs about θ are summarized by $\theta \sim N(\mu_0, \tau_0^2)$. As the model development proceeds with using precision[4] (inverse of variance) instead of variance, it can be noted that $\tau_0^2 = 1/h \Leftrightarrow h = 1/\tau_0^2$. Initially, all available information is common and thus assumed to be contained within the prior. The average, or consensus, belief is $E[\theta] = \mu_0 = u_{tj} = u_{tk}$, that is, everyone makes the same forecast from the commonly available information.

Apart from observing commonly available information, agents also have access to private information (costlessly). Private information is introduced in the standard fashion[5]; that is, as a signal, z_i, informative on θ. In particular, $z_i \sim N(\theta, \nu_i^2)$. This

[4]The choice of variance or precision is obviously one of preference, however the precision formulation has a somewhat more direct intuitive interpretation; an observation is directly weighted by its corresponding signal precision, see e.g. Chamley (2004). Throughout this study, both precisions and variances will be used.

[5]Here, the signal shows a dependence on i, that is the signal variance is allowed to be specific to each agent. However, in the later analysis of the dispersion of forecasts, a simplifying assumption is made where all agents have access to the same distribution. The consequences of this is discussed

is of course equivalent with stating, in the more commonly seen form: $z_i = \theta + \varepsilon_i$, where $\varepsilon_i \sim N(0, \nu_i^2)$. Working with precisions instead of variances as before: $\nu_i^2 = 1/s_i \Leftrightarrow s_i = 1/\nu_i^2$.

After observing their signals, agents update their beliefs in accordance with Bayes' rule, and agent i's belief is given analogously to equation (1), as:

$$u_i = E[\theta|z_i] = \frac{\frac{1}{\tau_0^2}\mu_0 + \frac{1}{\nu_i^2}z_i}{\frac{1}{\tau_0^2} + \frac{1}{\nu_i^2}} \tag{4}$$

The posterior variance of agent i is again analogous to equation (2) and is given by:

$$\tau_{1i}^2 = \left(\frac{1}{\tau_0^2} + \frac{1}{\nu_i^2}\right)^{-1} \tag{5}$$

Thus, agent i's posterior distribution of θ, conditional on z_i, is:

$$\theta \sim N\left(\frac{\frac{1}{\tau_0^2}\mu_0 + \frac{1}{\nu_i^2}z_i}{\frac{1}{\tau_0^2} + \frac{1}{\nu_i^2}}, \left(\frac{1}{\tau_0^2} + \frac{1}{\nu_i^2}\right)^{-1}\right) \tag{6}$$

Since agents receive different signal realizations, their individual posterior means (or forecasts) will differ, that is: $u_{t_2j} \neq u_{t_2k}$. This is the mechanism that leads to the divergence of beliefs. However, as most of the results will be derived in a setting where all private information is of equal precision[6], that is $\nu_i^2 = \nu^2 \forall i$, it is useful to note that the existence of differential beliefs does not depend on differential precision of private information.

2.3 Properties of agents' beliefs

The consensus, or mean belief $\bar{u}$, is simply defined, as in Barron et al. (1998), by the sample average of individual beliefs u_i:

$$\bar{u} = \frac{1}{N}\sum_{i=1}^{N} u_i$$

later. (The signal *realizations* will naturally differ in all setups in order to induce information asymmetry).

[6]Barron et al. (1998) derive many of their predictions on forecast dispersion under this assumption. Furthermore, in their footnote 8, they motivate this by stating that "most empirical test for differences in the magnitude of forecast error across analysts have failed to find significant differences ".

The extent to which these beliefs vary, is naturally measured by the sample variance of the beliefs around their mean, as in Barron et al. (1998):

$$d = \frac{1}{N-1}\sum_{i=1}^{N}(u_i - \bar{u})^2 \tag{7}$$

This represents the diversity of beliefs.

Note that equation (7) is the sample variance. In order to make predictions about what the expected diversity of beliefs will be, it is natural to define a diversity measure, V, to be the expected value of the sample variance[7]. Thus:

$$var(u_i) = V = E[d] = E\left[\frac{1}{N-1}\sum_{i=1}^{N}(u_i - \bar{u})^2\right] \tag{8}$$

Defining diversity in terms of sample variance has an advantage in cases where one wishes to analyze the effect of correlated draws from the signal distribution.

2.4 Properties of belief diversity

Assuming all agents face the same uncertainty, that is $\nu_i^2 = \nu^2 \forall i$, diversity of beliefs can be shown to take on the following specific form,

Proposition 1:

$$V = \frac{\nu^2\tau_0^4}{(\nu^2+\tau_0^2)^2} \tag{9}$$

Proposition 1 is the specific form of the expected diversity of beliefs, or dispersion in forecasts, after agents have observed their signals, expressed in terms of τ_0^2 and ν^2, the prior and signal variance, respectively. This result is derived by both Barron et al. (1998), and Barry and Jennings (1992). This corresponds to equation (19) on page 427 in Barron et al. (1998), and equation (5) on page 172 in Barry and Jennings (1992). See Appendix B for a lengthy proof of Proposition 1.

[7]See Barron et al. (1998) for a discussion of this and note that Barry and Jennings (1992) also define diversity this way. As a side note, it is perhaps worth pointing out that V as defined here and in Barry and Jennings (1992) corresponds to the definition of "D" in Barron et al. (1998). V, as defined in Barron et al. (1998) in turn corresponds to the posterior variance, given here in e.g. Eq. (2) and Eq. (5).

Appendix B derives the expected variance of beliefs (u_i's) using the precision representation given in equation (4). Thus, the final result,

$$E[var(u_i)] = \frac{s}{(h+s)^2}, \tag{10}$$

is exactly as equation (19) on page 427 in Barron et al. (1998). Since the above representation uses precisions instead of variances, rewrite in terms of variances, where $h = 1/\tau_0^2$ and $s = 1/\nu^2$, yielding:

$$E[var(u_i)] = \frac{\frac{1}{\nu^2}}{(\frac{1}{\tau_0^2} + \frac{1}{\nu^2})^2}$$

Since τ_0^2 and ν^2 are positive, the above expression, after some algebra, reduces to equation (9).

In order to see the equivalence between Proposition 1, and equation (5) on page 172 in Barry and Jennings (1992), set $\tau_0^2 = \sigma^2/n_0$, and $\nu^2 = \sigma^2/n$. Then, simply plug into equation (9):

$$V = \frac{(\frac{\sigma^2}{n})(\frac{\sigma^2}{n_0})^2}{[(\frac{\sigma^2}{n}) + (\frac{\sigma^2}{n_0})]^2},$$

which after some algebraic manipulations, yields:

$$V = \frac{n\sigma^2}{(n_0+n)^2} = \frac{n}{n}\frac{n\sigma^2}{(n_0+n)^2} = \frac{\sigma^2}{n}\left[\frac{n}{(n_0+n)}\right]^2,$$

where the last result, $V = (\sigma^2/n)[n/(n_0+n]^2$, equals equation (5) on page 172 in Barry and Jennings (1992), QED.

2.5 The role of the prior mean

Proposition 1 shows that belief diversity (and consequently forecast dispersion, in the cases where forecasts are assumed to be honestly reported expectations/beliefs by agents who act in a non-strategic manner), is completely determined by the two sources of uncertainty facing the agents; common uncertainty (prior variance), τ_0^2, and private uncertainty (signal variance), ν^2. A consequence of Proposition 1, that becomes clear immediately, is that the diversity of beliefs does not depend on the prior mean.

This is interesting since a subset of the empirical tests analyzes the association between forecast dispersion and negative realizations of earnings, the latter which is information that reasonably constitutes commonly available information. The above makes clear that there exists no *direct* mechanism that links the two. This can also be illustrated by a situation where agents first observe a draw from common information (earnings) and a private signal, which is shown in the example of Appendix A.

The example in Appendix A provides the intuition for the fact that common information, i.e. earnings, cannot directly affect the diversity of beliefs. This can also be seen in a more thorough context in Appendix B, in the midst of the derivation of Proposition 1. Appendix A also demonstrates some of the modeling elements that are used in the following sections, in the modeling of the behavior of belief diversity applied to earnings releases in a time series setting.

To re-iterate, the diversity of beliefs depends on the perceived uncertainty in the information that agents have (will have). This is characterized by the variance. The particular feature of the Bayesian setting however, is that an agent, after having constructed a prior, updates over a (one at a time in a sense) *realization* that arrives from the informative distribution, the variance of which is known to the agent. The "value" of the realization however, does *not* affect how his/her updated perceived variance of the estimated parameter evolves. The dispersion in forecasts is again affected through the agent's (agents') updated belief certainty, which in turn is a direct input into the formula for the dispersion in beliefs in the aggregate.

2.6 New dynamics

The analysis up until this point has been purposely vague on the exact nature of information dissemination and uncertainty dynamics over time. The result of Barry and Jennings (1992) (and Barron et al., 1998) in Proposition 1 describes how diversity of beliefs depends on the two sources of uncertainty facing the agents, namely common and private information. The representation of Barry and Jennings (1992), follows in their words, "the suggestion by Raiffa and Schlaifer (1961) that a natural conjugate prior or posterior distribution has an interpretation in terms of "equivalent sample information." ". Since private information in Barry and Jennings (1992) is specified by[8] $x_i \sim N(\tilde{\mu}, \sigma^2/n)$, taking the derivative of V with respect to n, (and increasing n) has the interpretation of increasing the quality of private

[8]In their notation. There is actually a i subscript on n, implying that the correct form would have n_i instead of n, but since the analysis of the derivatives assumes that all agents have access to the same private information, the subscript is left out at this point already.

information.

$$\partial V/\partial n = \sigma^2(n_0 - n)/(n_0 + n)^3$$

Since this is positive for $n < n_0$, it is the case that diversity can increase if more private information is added (Common, or public information in Barry and Jennings (1992) is given by $x_0 \sim N(\tilde{\mu}, \sigma^2/n_0)$). It is this feature that is behind the "private information acquisition hypothesis". However, taking the derivative of V wrt n_0 is strictly negative for all values/combinations of common and private information.

$$\partial V/\partial n_0 = -2n\sigma^2/(n_0 + n)^3$$

This again has the interpretation of increasing common information (from the same distribution), by adding observations, through increasing n_0.

It thus seems that indeed, diversity can increase if agents gain access to more private information [9]. Simultaneously however, there exists an effect of diversity decreasing monotonically with common information. Since over time, a minimum requirement should be that agents receive more and more common (public) information, it appears to be too early to state whether increasing private information can yield increasing diversity in *sufficient* amounts, without simultaneously considering the restriction of continuously increasing common information. Thus, in order to fully specify the evolution of forecast dispersion over time, there is a need for a model that takes into account the speed of convergence on common information over time. Whether private information can give rise to required (observed) levels of forecast dispersion, can only be fully assessed after controlling for the rate of convergence on public information.

This is the very issue that this study sets out to explore next, that is, to develop a model for the dynamics of belief diversity over time, given reasonable constraints on common information dissemination.

2.7 Multiperiod convergence

The intuition is to model the information dissemination process in a way that matches how information about company earnings is released to the market. For simplicity, it will be will assumed that this happens only annually, that is through

[9] A requirement for this is also that the starting private information is vague.

companies' annual earnings releases (naturally, this could be extended to quarterly information as well). At the same time however, this also seems to be a minimum requirement for realism - since new information de facto is released through earnings announcements, it would be unrealistic in this setting to start by endowing agents with commonly available information in the priors, but then proceeding by only adding private information (For an example of such a model, in a completely different setting, see e.g. Prendergast and Stole, 1996). In periods smaller than one year, this could be the case, if agents start off with all historical information, including the latest earnings, and then receive only private signals throughout the interim period (this can of course be represented by simply one signal). Once the next year's earnings announcements are made available however, this information must be thought to have at least a common component, and should be included in the model as an addition of common information that agents take into account in formulating their forecasts.

There exist further reasons, apart from the above, that seem to suggest that the correct way to analyze the learning behavior and the resulting (potential) disagreement, is through a model that encompasses multiple periods. To exemplify, Brown (1993) notes that " Bayesian theory suggests that earnings announcements generally increase the precision of individual's estimates of *future* earnings" (emphasis added). It is difficult to see how this intuition would apply to the learning process, unless it is implied that learning is connected to over time[10].

The idea of the model is then that at each period, agents get endowed with both common and private information. In the (arbitrary) starting point, there exists common information represented by the prior, and some private information. There is no need to add a separate earnings announcement that constitutes a draw from common information in the beginning, since as illustrated in the example in Appendix A, the prior can be assumed to already contain this information. Then, for each year (period), both common information and private information will be added, which agents will subsequently update over. Common information, y, will be assumed to be drawn from a distribution, the moments of which (specifically the second moment) are given by those of the historical record[11]. In the subsequent analysis, this will be treated (by the economist) as a fixed and observable parameter.

[10]This could of course be done easily by simply stacking a sequence of disconnected learning periods after one another. The problem however then becomes that learning only occurs within a period, whereas agents would forget the information just learned when starting the following period.

[11]Obviously, the model in no theoretical way depends on this assumption. It is mentioned here only in the sense of "setting the scene", that is giving the intuition for how the model maps to reality.

Private information, z, will be the variable that will remain as a free parameter[12], where changes in the parameter will be used to study the sensitivity of the model in later stages.

2.8 A model for agents learning over time

The study now proceeds by modelling how agents, after having performed a first update, described below, construct subsequent updates over new common information that arrives at each period, in addition to private information. In this setup, agents will receive both public (common) and private signals at each period. In subsequent sections the study refers to this particular setup as simply "the model". The need for specifying the arrival of one public signal for each period is important for keeping track of the convergence that occurs over multiple periods. In the specific model below however, a private signal is additionally added at each step. While this is not perhaps necessary, it adds intuition and realism to analyzing the behavior of forecast dispersion in the following chapter.

The idea is now that at the outset, agents start with differing priors since each individual's prior is the result of the history of common information and his/her personal (history of) private signal(s). In other words, the prior that each agent uses at each period, is his/her own individual (previous) forecast. Thus, initially, agents are assumed to have performed an updating procedure such as in section 2.2, and each agent now has an individual forecast of θ, u_{1i}. The study follows the previous notation to the extent that $h = 1/\tau^2$ refers to the precision on the prior, and $s = 1/\nu^2$ refers to the precision on the privately observed signal. The study also introduces $r = 1/\sigma^2$, which refers to the precision of the common information, y, and is a a constant throughout time.

It will now be assumed that for the first belief update that agents have performed, yielding u_{1i}, agents view firm earnings as coming from a distribution $y \sim N(\theta, \sigma^2)$. Before performing the first update, agents then have access to the full history of a firm's earnings, and consequently use the variance in these earnings to assess σ^2 (naturally, they can use the historical mean of the earnings series as an initial estimate of u_{0i}, which is common information and thus equal to $u_0 = \mu_0$). The estimate of σ^2 $(1/r)$ will thus be used as the variance of the prior in the first update[13].

[12]Throughout, the study uses the term parameter both in a general sense, as in this case, as well as in the statistical sense. The definition is not made explicit at each point, rather the context is allowed to define the meaning.

[13]Note that although this perhaps might seem like a strong assumption at first glance, this is not

The implication is thus that in constructing their *first* forecast (u_{1i}), agents use $r = 1/\sigma^2$ as their prior $h = 1/\tau^2$.

Recall that the first update, in terms of the previously used time indices, was given by $u_{1i} = (h_{(0)}u_0 + sz)/(h_{(0)} + s) = (h_{(0)}\mu_0 + sz)/(h_{(0)} + s)$. In light of the above, this is equivalent with $u_{1i} = (r_{(0)}u_0 + sz)/(r_{(0)} + s) = (r_{(0)}\mu_0 + sz)/(r_{(0)} + s)$. The fact that for the first update agents use r as their prior h will be made use of later.

The analysis will now proceed by using n for time indexing and n can thus be thought of as a time period, where for example $n = 1$ implies that the updating from 0 to 1 has already occurred.

Notice that there is a slight abuse of notation that comes with the change, with the advent of using n to denote the period when updating has already occurred, instead of using the strict t indices to denote time, as in the previous "one period" models. Previously, $1/\tau_0^2$ (or h_0), was used to describe the information that existed prior to the first updating over new information at $t = 0$. Time post updating was, in reference to the above, referred to as $t = 1$. Thus, the information, in terms of time indexing, going in to the forecast, was kept separate from the time indexing of the forecast after updating had occurred. Now however, indexing time by n, notationally bundles up all the information under one index (*except* for the prior mean, u_0), such that it is implied that information going in to the forecast, actually predates the information that is the result of the forecast. So what would have been formulated as $u_{1i} = (h_{(0)}u_0 + sz)/(h_{(0)} + s)$ under indexing by t, now becomes $u_{1i} = (h_{(1)}u_0 + sz)/(h_{(1)} + s)$ under the change to n-indexing. Actually, in the cases where also the private signals require indexing, under n-indexing, it is the case that $u_{1i} = (h_{(1)}u_0 + s_{(1)}z_{(1)})/(h_{(1)} + s_{(1)})$[14]. Furthermore, for the first belief update, the fact that agents are assumed to use the distributional characteristics of (observed) earnings, y, as their initial prior, implies in terms of the switch to n-indexing that: $u_{1i} = (r_{(1)}u_0 + s_{(1)}z_{(1)})/(r_{(1)} + s_{(1)})$. The reason for performing this change in indexation, is that it results in a gain in terms of keeping track of the number of periods that have been used for updating, which will provide to be useful in some of the later analysis. Simultaneously, this comes with a slight cost in terms

the case - the assumption is mainly being done to keep the notation clear. To see why this does not affect the results, note that as long as the common information that is observed by the agents in the form of earnings, y, remains (somewhat) fixed, and empirically matches the observed history of earnings, the effect of an (initial) prior of much higher variance will quickly be overshadowed by the learning from the new, much "sharper" information that is the earnings series. On the other hand, if the initial prior would be *sharper* than the commonly observed earnings, y, beliefs, and subsequently dispersion in forecasts converge at a much faster rate than assumed here. This is explored in the empirical section.

[14]The idea being that the signal, in this case the private signal z, "belongs" to first (nth) updating event.

of reduced clarity of what information is actually used in the updating. For the subsequent analysis however, r_1 will be labeled simply r, since the true variance of earnings, y, are assumed to be a constant over time at this point (But the omitted subindex on r would be 1, as in the above example).

Having clarified the notation, the forecasting behavior for the second period, $n = 2$ can now be analyzed. The main feature is that now, all agents will, for each and all subsequent periods, receive private information[15] from z and common information from y.

The agent will now make the following forecast, which, after having been carried out, results in what is labeled as the $n = 2$ forecast:[16]

$$u_{2i} = E[\theta|y,z] = \frac{\frac{1}{\tau^2_{(2)}}u_{(1)i} + \frac{1}{\nu^2_{(2)}}z_{(2)i} + \frac{1}{\sigma^2}y}{\frac{1}{\tau^2_{(2)}} + \frac{1}{\nu^2_{(2)}} + \frac{1}{\sigma^2}} = \frac{h_{(2)}u_{(1)i} + s_{(2)}z_{(2)i} + ry}{h_{(2)} + s_{(2)} + r}. \quad (11)$$

While Eq. (11) analyzes the period 2 update specifically, the general form of all subsequent updates for $n \geq 2$ is given by:

$$u_{ni} = \frac{h_{(n)}u_{(n-1)i} + s_{(n)}z_{(n)i} + ry}{h_{(n)} + s_{(n)} + r}.$$

Using the forecast(s) above, Appendix C derives the following proposition, which is an expression for the variance of forecasts, resulting from agents updating over (any) normally distributed, unequal priors.

Proposition 2:

$$var(u_n i) = \frac{s_n^2[var(z_{ni})] + h_n^2[var(u_{n-1i})]}{(h_n + s_n + r)^2} \quad (12)$$

Proposition 2 shows how the variance of forecasts depends on the (variance of the) latest signal and the variance of the priors. Note that while the specific characteristics of information dissemination are important for the dynamics of earnings studied here, one could easily change the model interpretation by removing the latest realization of common information, r, since this only affects the denominator (and no other parameters depend on it), yielding: $var(u_n i) = (s_n^2[var(z_{ni})] + h_n^2[var(u_{n-1i})])/((h_n +$

[15]time indices on the private signals, z, are tracked, since the private information acquisition hypothesis will lead to an analysis of changes in the private signals.

[16]$n = 2$ since two updating events have been performed (tracked by the model).

$s_n)^2$). By further assuming a constant prior, that is $var(u_{n-1i}) = 0$, equation (12) reduces to:

Corollary 1:

$$var(u_n i) = \frac{s_n^2[var(z_{ni})]}{(h_n + s_n)^2} \tag{13}$$

Which is exactly equal to Proposition 1, and thus shows that Proposition 2 is a different specification of the general result of Barry and Jennings (1992).

Proposition 2 however, will transpire to be a useful specification in analyzing situations where changes to the private signal (distributions) are assumed to occur at certain points in time. Proposition 2 also serves a starting point for developing the time series behavior of the proposed model, introduced next.

The introduction of the common signal y at each period, in addition to the private signals, has some interesting consequences for the rate of convergence of forecast dispersion over time. Using equation (11) as the starting point, Appendix D derives an expression for diversity as a function of time, and is given in the following proposition:

Proposition 3:

Given the uncertainty dynamics above, it can be shown that belief diversity, as a function of time, n, takes the following form[17]:

$$var(u_n i) = \frac{ns_{(1)}}{(nr + ns_{(1)})^2} \tag{14}$$

The strength of Proposition 3 lies in the fact that diversity of beliefs, at any point in the future, can be expressed in terms of "starting" variances, that is the initial precisions (or uncertainties) of common and private information. In comparison to Barry and Jennings (1992), who do show *generally* that dispersion can increase for increasing private information precisions, Proposition 3 *simultaneously* considers the evolution/convergence of common uncertainty over time, and allows for analyzing whether increases in private information can actually yield observable dynamics

[17]Being fully rigorous, the model was defined for $n \geq 2$. However, setting $n = 1$ is equal to using the "one-period" belief updating model, where the resulting variance of forecasts equals setting $n = 1$ in Proposition 3. Thus, Proposition 3 is also defined for $n = 1$. Also, Appendix D derives the result using the assumption that signals are constant over time (s). Here however the time-subscript is added to indicate that this is the signal that agents used in the formation of their first update, but nevertheless remains a constant going forward (n) periods.

(increases) in belief diversity, *given* the rapid convergence on common information over time. Note however that the dynamic is still exactly as in Barry and Jennings (1992), and can thus be found by specifying Proposition 1 correctly. The result in Proposition 3 however, explicitly shows how uncertainty evolves over time in the model of this study and has an advantage in forcing one to consider the continuous convergence on common information being added.

2.9 Properties of belief diversity under convergence on common information

For the subsequent analysis, the result in Proposition 3 will be labeled V', that is:

$$V' = var(u_n i) = \frac{ns_1}{(nr + ns_1)^2},$$

as it is constrained the same uncertainty dynamics of V, derived in Barry and Jennings (1992), but here makes the constraints put forth by the model of the convergence on common information over time, explicit. The partial derivatives of V' are:

$$\frac{\partial V'}{\partial s_1} = \frac{r - s_1}{n(r + s_1)^3}, \qquad \frac{\partial V'}{\partial r} = \frac{-2s_1}{n(r + s_1)^3}, \qquad \frac{\partial V'}{\partial n} = \frac{-s_1}{n^2(r + s_1)^2}$$

The leftmost and middle derivatives share the exact dynamic as the corresponding derivatives in Barry and Jennings (1992)[18]. The rightmost derivative of V' with respect to time (n), which is new in comparison with Barry and Jennings (1992), is overall decreasing in n but since the derivatives are partials, keeping the other variables constant, one cannot make statements about any form of dynamic evolution of private signal variance over time, as this case refers to s_1 being a constant only. As for the first (leftmost) derivative, it shows that dispersion is initially increasing in s_1 while $r > s_1$, but this is *within one time period only.* Thus, the derivatives are not sufficient for studying the full scale dynamic evolution of belief diversity, if one wishes to study whether increases in private information can yield increases in belief diversity while allowing for the convergence of beliefs on new common information.

In order to achieve this, that is study whether increases in private information can yield increases in belief diversity while allowing for the convergence of beliefs on

[18]Notice that n in the above derivatives is *not* the same as n in Barry and Jennings (1992)

new common information, one can use a combination of Proposition 2 and Proposition 3: specifically, for any sequence of periods where private information, s, is kept constant, diversity evolves according to Proposition 3, whereas if the private information is allowed to change at some period n^*, diversity is analyzed with Proposition 2, using the result from Proposition 3 at $n^* - 1$ as the prior, $var(u_{n-1i})$ in Proposition 2. This is shown in the following:

Consider some time period n^*. At n^*, agents are endowed with a more informative signal z', $(s' > s)$ than they have updated over up until n^*. Thus, at n^*, diversity is given by Proposition 2:

$$var(u_{n^*i}) = \frac{s'^2_{n^*}[var(z'_{n^*i})] + h^2_{n^*}[var(u_{n^*-1i})]}{(h_{n^*} + s'_{n^*} + r)^2},$$

and[19] $var(u_{n^*-1i}) = [(n^* - 1)s_1]/([(n^* - 1)r + (n^* - 1)s_1]^2)$. As Appendix D makes clear, any h_n can be represented by the square root of the denominator of the variance of the forecasts from the previous period. Here, $h_{n^*} = [(n^* - 1)r + (n^* - 1)s_1]$. Furthermore noting that $s'^2_{n^*}[var(z'_{n^*i})] = s'_{n^*}$, and inserting into Proposition 2 gives:

$$var(u_{n^*i}) = \frac{s'_{n^*} + [(n^* - 1)r + (n^* - 1)s_1]^2 \frac{[(n^* - 1)s_1]}{[(n^* - 1)r + (n^* - 1)s_1]^2}}{([(n^* - 1)r + (n^* - 1)s_1] + s'_{n^*} + r)^2}$$

$$var(u_{n^*i}) = \frac{s'_{n^*} + [(n^* - 1)s_1]}{[(n^*)r + (n^* - 1)s_1 + s'_{n^*}]^2}$$

This shows the variance of forecasts at n^*, given the new, signal s'_{n^*}, expressed in terms of the beliefs that have evolved up until $n^* - 1$, from the initial values at $n = 1$, s_1 and r. Thus. if the interest lies in seeing whether s'_{n^*} can increase dispersion at n^*, a value for s'_{n^*} must be found, such that the following condition is satisfied:

$$\frac{s'_{n^*} + [(n^* - 1)s_1]}{[(n^*)r + (n^* - 1)s_1 + s'_{n^*}]^2} > \frac{[(n^* - 1)s_1]}{([(n^* - 1)r + (n^* - 1)s_1]^2)} \tag{15}$$

Unfortunately, finding a "nice" analytical expression, on the sought after solution space, for any n^*, treating all parameters as free, is not possible. Nevertheless, the condition in equation (15), provides the intuition for checking whether belief

[19]The variance comes from Proposition 3, since $n^* - 1$ periods have passed before the current period n^*. The fact that $var(u_{n^*-1i})$ contains and i index is due to the fact that it is a variance of individual forecasts, not that the variance is agent specific.

diversity increases from one period to the next, if private information is allowed to change within the context of the model.

While Eq. (15) shows the general condition for increasing dispersion from any n^*-1 to n^* given a private signal s'_{n^*}, solving still involves tracking the full evolution of signals from 1 to n^*-1. To exemplify, if $n^*=2$, it corresponds to a situation where the prior (from $n^*=1$) sums up everything that has happened in the past[20]. Here, since the analysis does not concern maximums specifically, the choice of using n^* in place of n is arbitrary. In terms of equation (15):

$$\frac{s'_2+s_1}{[2r+s_1+s'_2]^2} > \frac{s_1}{(r+s_1)^2} \tag{16}$$

Alternatively, if n^* is general, Eq. (15) is:

$$\frac{s'_{n^*}+s_{n^*-1}}{[h_{n^*}+r+s'_{n^*}]^2} > \frac{s_{n^*-1}}{(h_{n^*-1}+s_{n^*-1})^2}$$

$$\frac{s'_{n^*}+s_{n^*-1}}{[h_{n^*-1}+s_{n^*-1}+r+s'_{n^*}]^2} > \frac{s_{n^*-1}}{(h_{n^*-1}+s_{n^*-1})^2} \tag{17}$$

It becomes evident that the representation (Eq. [17]), in itself, is not sufficient in describing increases in forecast dispersion from any n^*-1 to n^*, since in order to fully specify h_{n^*-1}, one needs to know exactly how h_{n^*-1} has evolved from r. This in turn depends on the length of the period, n^*. Thus the evolution of h_{n^*-1} and s_{n^*-1} from $n=1$ to $n=n^*-1$ needs to be fully specified, as is done in the derivation leading up to[21] Eq. (15). Thus the difference between Eq. (15) and Eq. (17) is that Eq. (15) specifies the evolution from starting values, while the notation in Eq. (17) implies general forms for the priors. The problem that becomes apparent is that regardless of formulations, the full evolution of uncertainties must always be tracked.

Simultaneously, it is possible to see from Eq. (16) and Eq. (17) how Proposition 2 and 3 are equivalent of the Barry and Jennings (1992) model, given the exact evolution of signals over time. One can set $s'_2=s_1$ in Eq. (16), to end up with $2s_1/(2r+2s_1)$. Substituting forward yields Proposition 3.

Notice however that it is always possible to find a signal s'_{n^*max}, such that belief diversity at n^* is maximized, using the fact that the derivative of V is always equal

[20]By definition of course the prior always sums up all past information, but what is meant here is that going back only 1 period leads back directly to starting values or uncertainties, in this case r and s_1.

[21]Where the assumption is a static evolution from $n=1$ to $n=n^*-1$ from starting values.

to zero for $s = h$. This is the $\partial V/\partial n$ in Barry and Jennings (1992), in *their* notation, and $\partial V'/\partial s$ in the notation of this study, respectively.

One could thus try to set the signal equal to the prior, implying $s'_{n^*max} = h_{n^*}$, where $h_{n^*} = [(n^*-1)r + (n^*-1)s_1]$, implying that $s'_{n^*max} = [(n^*-1)r + (n^*-1)s_1]$. However, the maximum in $\partial V/\partial n$ in Barry and Jennings (1992) actually applies to *total information*, so the maximum will be found at the point where all common information is of equal precision to private information. Thus the signal s'_{n^*max} that maximizes dispersion at n^*, is given by: $s'_{n^*max} = [(n^*)r - (n^*-1)s_1]$. To see this, simply gather all private information on the left-hand side and set it equal to common information: $s'_{n^*max} + (n^*-1)s_1 = (n^*)r$, which is total s equal to total h. Using this implies that for each n, the maximum variance that can be achieved given the variance for the previous period, is:

$$var_{max}(u_{n^*i}) = \frac{[(n^*)r - (n^*-1)s_1] + [(n^*-1)s_1]}{[(n^*)r + (n^*-1)s_1 + [(n^*)r - (n^*-1)s_1]]^2}$$

$$var_{max}(u_{n^*i}) = \frac{(n^*)r}{[(n^*)r + (n^*)r]^2} = \frac{(n^*)r}{[2(n^*)r]^2} = \frac{1}{4(n^*)r}$$

The derivations above become somewhat convoluted by the fact that the evolution of private information in the model follows a particular predefined pattern, and is disseminated at each period. If one completely relaxes the evolution of the private signals over time, and simply assumes one cumulative private signal[22], one can simply set the precision of this cumulative private signal, say z'' (precision s''), equal to the common information in the prior[23] *at* n, nr ($s'' = nr$). This yields simply:

$$var_{max}(u_{ni}) = V_{max} = \frac{nr}{[nr + nr]^2} = \frac{nr}{[2nr]^2} = \frac{1}{4nr} \tag{18}$$

Equation (18) is an expression for the maximum variance of forecasts as a function of the amount of common information that has become available from $\sigma = 1/r$, over n periods. The most important part of Eq. (18) is the subtle but significant difference compared to a single period setup, where one would have h instead of nr in the denominator. In other words, nr tracks the amount of convergence that must have occurred from agents receiving one realization from a distribution with variance $\sigma = 1/r$, per period n.

[22] This can be done due to the linearity of precisions.

[23] In exact mathematical terms this should read "in the information set at n, implying that n signals have been received an updated over from r.

2.10 Analytical properties of the model

This section, along with subsections, delves deeper into some analytical properties of the model. These results are used to some extent in the illustrations in Chapter 3. However, for the reader more interested in the intuition of the model, it is perhaps preferable to skip to the concluding section (2.12 on page 37) of the chapter.

The study now proceeds by analyzing in detail some of the dynamics of the model. This analysis lays the mathematical foundation for some of the results in Chapter 3. Using the precursor to the general result in Eq. (18) (and Eq. [15]) , the (maximum) amount by which the variance can increase from (any) $n^* - 1$ to n^* is:

$$\frac{1}{4(n^*)r} / \frac{[(n^* - 1)s_1]}{([(n^* - 1)r + (n^* - 1)s_1]^2)}$$

Here the numerator (left) is the maximum amount of dispersion at n^* (given r), and the denominator (right) is current dispersion at n^*. Cleaning up yields the following proposition:

Proposition 4:

The maximum change in dispersion, labelled potential, $\Delta_{n^*,max}$, going from *any* $n^* - 1$ to n^*, where $n^* \geq 2$, can be found by using the starting precisions, r and s_1, given by the following equation:

$$\Delta_{n^*,max} = \frac{(n^* - 1)(r + s_1)^2}{4(n^*)rs_1} \tag{19}$$

Eq. (19) is the factor or multiplier by which dispersion can change at maximum from (any) $n^* - 1$ to n^*, given initial starting values for s_1 and r. Thus, if dispersion at $n^* - 1$ can increase going to n^*, Eq. (19) must be > 1. Since all variables in Eq. (19) are positive, Eq. (19) can never become negative.

The first implications of Eq. (19) can be seen by differentiating. The analysis of the derivatives (w.r.t. r and s_1) is akin to analyzing the starting conditions for the model.

The derivative of the factor (Eq. [19]) w.r.t. the private signal is given by:

$$\frac{\partial \Delta_{n^*,max}}{\partial s_1} = -\frac{(n^* - 1)(r - s_1)(r + s_1)}{4(n^*)rs_1^2}$$

The derivative is initially negative, for values of $s_1 < r$, implying that increasing s_1 yields smaller increases in potential, up until the extremum at $r = s_1$, where the derivative becomes zero. Beyond this point, The derivative becomes positive as $s_1 > r$, implying that increasing s_1 will yield larger increases in potential (going forward).

The derivative of Eq. (19) w.r.t r is given by:

$$\frac{\partial \Delta_{n^*,max}}{\partial r} = \frac{(n^* - 1)(r - s_1)(r + s_1)}{4(n^*)r^2 s_1}$$

This shares the intuition of the derivative w.r.t. s_1, where the extremum is attained for $r = s_1$, and is shown here for completeness. The derivative is initially negative for values of $r < s_1$, implying that increasing r yields smaller increases in potential, up until the extremum at $r = s_1$. Beyond this point, the derivative becomes positive as $r > s_1$, implying that increasing r will yield larger increases in potential. The fact that the potential grows as r grows in precision beyond that of s_1 is a somewhat unrealistic effect that will be labelled reverse potential, and is discussed further in the following sections.

The model, and consequently Eq. (19), only refer to an evolution of the system over time in exact accordance with Proposition 3 specified by the starting conditions. Manipulating s_1 and r above refers to an analysis of different starting conditions for the model, and is *not* the same thing as changing signals over time.

The most important part of the analysis, in terms of providing intuition for the dynamics over time, is concerned with what happens to the increases in potential regarding changes in n^* :

$$\frac{\partial \Delta_{n^*,max}}{\partial n^*} = \frac{(r + s_1)^2}{4(n^{*2})r s_1}$$

The derivative of Eq. (19) is always positive, implying that for every n^* going forward, the potential increases, regardless of r and s_1. If for example starting from a value of $r > s_1$, then the larger n^* becomes, the larger the potential or the multiplier on the dispersion at $n^* - 1$ becomes. Recall that this still implies r and s_1 are fixed in the beginning.

Even though the derivative w.r.t. n^* remains positive for all values of n^*, the condition for increasing *dispersion* from $n^* - 1$ to n^*, requires that Eq. (19) > 1. Setting $r = s_1$ implies that $\partial \Delta_{n^*,max}/\partial n^*$ reduces to $1/(n^{*2})$. The derivative is still positive however, implying that the potential is growing, even though it was stated that dispersion was already at its maximum. Any $s_{n^{*\prime}} > s_{n^*}$ (implying $n^{*\prime} > n^*$) will thus technically have a bigger potential in comparison with any previous signal.

The increases are relative to the sizes of previous increases and this is overshadowed by the rate of decline in dispersion, seen in setting $r = s_1$ in Eq. (19), yielding: $(n^* - 1)/(n^*)$, which is always < 1, showing that dispersion is decreasing for all $n > n^*$. However, for $r \neq s_1$, one can always find a sufficiently large n^* such that Eq. (19) > 1.

2.11 Higher precision private signals

This subsection analyzes how the arrival of more precise private signals affects increases in dispersion and particularly looks at potential. This cannot be analyzed directly with the derivatives above, since the evolution of signals is fixed from the beginning (the evolution of signals is fixed from the beginning in Proposition 3 alone, implying that agents will keep getting signals of size r and s_1). Overall it would be of interest to find the exact conditions for when Eq. (19) > 1. Analogously to Eq. (15) however, there exists no "nice" general solution for this inequality[24].

The analysis now proceeds by considering changes in private signals. This analysis unfortunately becomes somewhat unintuitive analytically and consequently the mechanism of only a few examples are considered.

If the signal s'_{n^*max} arrives at any n^*, dispersion will jump to its maximum possible value at that point, the size of the jump (in terms of a multiplier on previous dispersion) from the previous period being described by Eq. (19), while the size of the (maximum) dispersion is described by Eq. (18) . This signal, s'_{n^*max}, disrupts the steady evolution of the system described by Proposition 3, and the signal s'_{n^*max}, by definition maximizes dispersion, and now places the system at what can be viewed as a "new starting point".

After the arrival of the signal s'_{n^*max}, the system has used up all its potential in increasing dispersion for any periods going forward, and the future evolution of the system is now analogous[25] to $r = s_1$, already analyzed previously. Thus, from this point onward, after the arrival of a signal s'_{n^*max}, the system is already on a trajectory of maximum dispersion for all periods going forward, and no new signal

[24]In order to be clear, there does exists an analytical solution, however it is not particularly informative/intuitive and is thus excluded.

[25]The evolution is similar in that future private signal are equal to common signals. In order to model the evolution of the system analytically, the previous learning must be controlled for.

can increase this (since by assumption more common information enters each period going forward). As shown already in the last paragraph of the previous subsection, this trajectory (changes in dispersion), is in itself declining (Setting $r = s_1$, Eq. (19) simply reduces to: $(n^* - 1)/(n^*)$ which is always < 1). Staying on the maximum trajectory assumes that private signals after the arrival of s'_{n^*max} , are equal to common signals, $s_{n>n^*} = r$.

The "Calibration" section will illustrate this effect (see e.g. Figure 7 on page 51), namely after the system has reached a stage of (compounded) common signals = (compounded) private signals, the maximum dispersion going forward will be for private signals being equal to common signals, that arrive from the (fixed) distribution with variance σ^2. The study now demonstrates analytically how the model can be solved forward and what is implied by the "new starting point".

Previous to the arrival of the signal s'_{n^*max}, the system evolves according to Proposition 3. If signals remain stable after n^*, the system can again be solved forward using Proposition 3. The interest thus lies in finding dispersion and potential for some future point in time m, that occurs after n^*, so that $m > n^*$. Thus, the sought after idea is Proposition 3 for $m > n^*$. Up to the point n^*, there have been n^* common signals of size r, and $(n^* - 1)$ signals of magnitude s_1 and the signal s'_{n^*max}. The learning over this information must be accounted for in the prior in a formulation for the system for t $m > n^*$. The equation below, where $s_{n>n*}$ indicates private signals after n^*, uses the fact that precisions add linearly, and the variance of beliefs can be represented as:

$$\frac{(n^* - 1)s_1 + s'_{n^*max} + (m - n^*)s_{n>n*}}{[n^*r + (m - n^*)r + (n^* - 1)s_1 + s'_{n^*max} + (m - n^*)s_{n>n*}]^2}$$

The corresponding potential at m (using the precursor to Eq. [19]) is given by:

$$\frac{1}{(4[n^*r + (m - n^*)r])} \Big/ \frac{(n^* - 1)s_1 + s'_{n^*max} + [(m - 1) - n^*]s_{n>n*}}{[n^*r + [(m - 1) - n^*]r + (n^* - 1)s_1 + s'_{n^*max} + [(m - 1) - n^*]s_{n>n*}]^2}$$

First, it can be noted that by definition $s'_{n^*max} = [(n^*)r - (n^* - 1)s_1]$. Secondly, the maximum trajectory for dispersion required that private signals after the arrival of s'_{n^*max} are equal to common signals, $s_{n>n^*} = r$. Inserting, the expression for the variance reduces to $[n^*r + (m - n^*)r]/[(n^*r + (m - n^*)r) + (n^*r + (m - n^*)r)]^2 = (mr)/(2mr)^2 = 1/4mr$, implying that dispersion is at its maximum for any m. For this trajectory, potential reduces to $(1/4mr)/([(m-1)r]/[2(m-1)r]^2) = (m-1)/m$, which is always less than 1 and shows that dispersion cannot increase for any m going forward.

If instead the private signals at the "new" starting point continue to be of the same magnitude as the signal s'_{n^*max}, implying that $s_{n>n^*} = s'_{n^*max}$, the private signals will start adding information faster than common information. The equation for the variance of beliefs however can now only be reduced to:

$$\frac{(n^*-1)s_1 + s'_{n^*max} + (m-n^*)s'_{n^*max}}{[n^*r + (m-n^*)r + (n^*-1)s_1 + s'_{n^*max} + (m-n^*)s'_{n^*max}]^2}$$

$$\frac{(n^*-1)s_1 + ((m+1)-n^*)s'_{n^*max}}{[n^*r + (m-n^*)r + (n^*-1)s_1 + ((m+1)-n^*)s'_{n^*max}]^2}$$

Inserting the definition of the maximizing signal $s'_{n^*max} = [(n^*)r-(n^*-1)s_1]$, yields:

$$\frac{(n^*-1)s_1 + ((m+1)-n^*)[(n^*)r-(n^*-1)s_1]}{[n^*r + (m-n^*)r + (n^*-1)s_1 + ((m+1)-n^*)[(n^*)r-(n^*-1)s_1]]^2},$$

the dynamics of which become difficult to analyze analytically. Similarly, the formula for potential (which is excluded) becomes equally difficult to interpret, as the above expression is one part of that result. Figure 8 on page 52 in Chapter 3 however provides a more intuitive illustration of this situation. What happens is that private signals start adding information faster than common signals after the maximum at n^* has been reached. This will add "reverse potential" to the model, in the sense that now, dispersion over time is not on track for maximal dispersion, and as n grows beyond n^*, at some point the difference grows large enough that dispersion in the future can increase to its maximum level by suddenly becoming *less* informative. This type of behavior where the private signal precision jumps back and forth will generally be ruled out. Also, in this case where after s'_{n^*max} signals continues to be of size s'_{n^*max}, dispersion actually leaves its maximum trajectory such that any "reverse potential" will at maximum take dispersion back up to the maximal trajectory of $r = s$ after s'_{n^*max}.

The formula for the multiplier on previous dispersion to reach its maximum value, potential in Eq. (19), and the formulas for dispersion, become somewhat cumbersome to use analytically once changes in private signals are introduced. In fact, the sizes of *increases* in dispersion from one period to the next become unbounded in the limit (Look at Eq. (19) when s_1 approaches zero [from the positive side]). Analytical results become even less transparent in the situations depicted in Figures 10 and 11 in Chapter 3, where the private signal grows at each period. As such, analytical solutions to these cases are omitted.

While analytical solutions increase in complexity due to the need for tracking exact signal evolutions, the formulas are easy to apply numerically. Dispersion can always be solved recursively from initial values using Proposition 2, or alternatively through accumulating the total cumulative private signal sequence $(s_1 + s_2 + ... + s_m)$ from initial values. Also, it is important to note that the maximum absolute *size* of dispersion is always given by Eq. (18) , i.e.: $var_{max}(u_{n^*i}) = 1/4nr$, regardless of assumptions on $r - s_1$. This implies that the maximum dispersion is a constant at each n, given r. Consequently, the multiplier on previous dispersion can always be found through solving for the implied maximum value for dispersion, and the previous value for dispersion.

The intuition of the multiplier and the dynamics for dispersion become clearer in the illustrations in the following chapter.

2.12 Discussion of theoretical results

The emphasis of this study is to rigorously consider the private information acquisition hypothesis. The private information acquisition hypothesis rests theoretically on the fact that disagreement is increasing private information in a Bayesian learning model where learning occurs from normally distributed information sources, common and private. As has been shown (Appendix B), disagreement results from the cross sectional variance of the posterior expectations of agents, where the variance of the posteriors turns out to depend only on the variances (precisions) of the signal distributions.

The rational Bayesian learning model under asymmetric information where all information is normally distributed is very general, and the resulting disagreement from such a setting can always be traced back to a representation such as that of Barry and Jennings (1992). In fact, even the information dynamic in the very different study by Prendergast and Stole (1996) can be shown to conform to the restrictions of the representation of Barry and Jennings (1992).

The point being made here is that *mathematically* the model here is the same model as Barry and Jennings (1992), but in terms of *application* there is a difference. The key distinction lies in that common information is specifically constrained, forcing information to enter agents' information set at a specific rate. While this can be seen as simply a different interpretation or application of the Barry and Jennings (1992) result, Barry and Jennings (1992) do not state that their model should be interpreted over a time dimension. The model here thus arrives at the ultimate posterior expectation (from cumulative information), through a piece by piece construction of the *sample* used in reaching the posterior expectation. DeGroot (1970) indeed derives the posterior expectation over a *sample* of information, whereas Barry and Jennings (1992) similarly derive results for dispersion using a sample of information. The insight of this study is to explicitly trace the *construction* of that sample as it evolves over time, through controlling for the minimum amount of common information that becomes available from the common distribution. While the results of Barry and Jennings (1992) are general, the stance of this study is that to apply the model to the real world, a specific interpretation of the model is needed. The study argues that to achieve realism in application of the model, the setting is that of annual earnings releases that endow agents with more information over time.

The assumption made in this study is that common signals arrive from a fixed distribution (with σ^2) that represents earnings, earnings which occur annually (n). Additionally, agents have access to private information. The most important idea is now that by assuming that private information is already maximized, the model only depends on two parameters. One can now differentiate the (general) maximum of Eq. (18) wrt n.

$$\frac{\partial V_{max}}{\partial n} = -\frac{1}{4n^2 r} \tag{20}$$

Importantly, this differentiation does not lose generality, since in the above r is fixed, as r is supposed to be fixed by assumption. Furthermore, the above allows for agents having access to any (full) private information. The above thus shows that as more common information arrives (through increasing n) from the distribution with precision r, the maximum dispersion becomes smaller for each period, even if agents have access to private information such that disagreement is maximized.

In summary, the theoretical results help in keeping track of the cumulative common signal that agents have received over n periods. The variance in forecasts is a function of this amount of information.

The main take-aways from the model are the following:

1) For agents receiving signal realizations from a common information distribution that remains fixed over time, there exists a private signal, equal to the *cumulative* common signal, that maximizes the dispersion in beliefs (forecasts) at each time n. This maximum dispersion is monotonically decreasing in n, implying that the maximum dispersion in forecasts always becomes smaller as n grows, i.e. time moves on.

2) Once maximum dispersion has been reached, a *more informative private signal* that arrives can only act as to *reduce* the dispersion in forecasts. Consequently, the private information acquisition hypothesis, and the resulting increase in dispersion, applies only to a situation preceding the situation in 1), that is a situation where private information, cumulatively, is less informative than common information.

A corollary to the above is that once private information has become more informative than common information, dispersion could actually be increased if agents "lose" some of their private information. Since this is relation to the total amount of available common information, one possibility for this is of course the arrival of more common information. By choosing *not* to receive private signals, dispersion could increase in this scenario. This was labelled the "reverse potential" result and was ruled out here on the grounds of realism. Nevertheless, in this setting it represents an effect not covered in previous literature.

In summation, as long as the distribution from which agents receive more information (realizations) over time remains fixed, the maximum amount of dispersion must become smaller for each period. Furthermore, increases in dispersion can only be a transitory effect. Once maximum dispersion for a specific point in time has been reached, more information in the next period results in the maximum amount of dispersion in that period being smaller than the maximum amount of dispersion in the previous period.

Chapter 3

Calibration

The aim of this chapter is to illustrate how the dispersion in forecasts in the model of Chapter 2 evolves over time when agents receive more and more common information in the form of earnings realizations. The convergence on common information constitutes the benchmark evolution of beliefs and dispersion, whereas different endowments of private information are considered in order to illustrate how they affect dispersion. To achieve a realistic sense of the bounds implied by the model, the model is calibrated to values that can be observed in empirical data.

It is worth emphasizing that the calibrations to data values are performed only in order to illustrate some of the dynamic behavior that can occur within the model. The idea here is to pick values that at least to some extent are representative of what occurs in the data. Whereas the model in reality applies to learning about the earnings of individual firms (distributions) this section performs calibrations on an aggregate level. As such, some statements in this section apply only to this section specifically and not to assumptions made in the empirical or theoretical section. A full description of the data is found in the empirical analysis in Chapter 4 (section 4.4 on page 68 and section 4.8 on page 77).

To begin the calibrations, start with Eq. (14), reproduced below. Even though private information remains private, and thus unobservable, private information is the only variable that cannot be observed.

$$V' = var(u_n i) = \frac{ns_1}{(nr + ns_1)^2}$$

In the equation above, there are 4 variables. By first considering an (arbitrary) first period[1], $n = 1$, the amount of variables reduces to 3, out of which 2 can be observed

[1] This is essentially the u_{1i} leading up to Eq. (11).

or proxied. V', or V, is the variance of the forecasts[2] ($\sqrt{V}$ being the standard deviation) that can be observed and r, the precision of common information, can be extracted from the historical variance[3] of earnings, since $1/r = \sigma^2$. The empirical proxy for forecast dispersion ($\sqrt{V}$) used in this section is the standard deviation of forecasts for the full year ending, measured June 30th divided by the corresponding consensus (mean) forecast (for each company year), arriving at a coefficient of variation measure. This variable is Winsorized at the 1st and 99th percentile, since close to zero mean values will artificially inflate the coefficient of variation metric. This is also pointed out by e.g. Minton, Schrand and Walther (2002). To find V, the measured dispersion metric[4] can simply be squared.

Under the theoretical model of the previous chapter, the Bayesian agent constructs his/her forecast through observing common and private information. Expectations are updated over realizations from a distribution of both assumed private information and common information, where the latter is firm level earnings. Theoretically the variance of the common information distribution σ^2 is assumed to be fixed. The implication of a fixed σ^2 is that with more realizations from the distribution with variance σ^2, an agent's subjective uncertainty decreases, even though the signal continues to be of the same magnitude. Simultaneously, a fixed (and observable) σ^2 implies that it is possible to measure.

In the empirical estimations in Chapter 4, information can be controlled for on a firm by firm level and the empirical estimations therefore endow agents with information available up to the time of forecasting. In contrast, in order to achieve the calibrations in this chapter, there is a need to represent earnings volatility by a single number that aggregates across all firms in the sample. The estimate of earnings volatility is thus measured on a coefficient of variation basis. To arrive at a meaningful estimate, the earnings volatility measure aggregates both cross sectionally and over time, until the end of the sample period. This causes the earnings volatility measure that proxies for σ^2 to include forward looking information, but as the point of this section is to calibrate the model to reasonable values that represent the data and give an illustration of the behavior of the model, this is deemed not to

[2]While in Chapter 2 a distinction between V and V' was made such that the latter, V', referred to the variance of beliefs in the context of the model representation developed in this study, generally both V and V' both refer to the variance of beliefs. Since the models are mathematically equivalent under similar assumptions, the (empirical) variance of beliefs will denoted simply V, when a distinction is not needed.

[3]Naturally, the *volatility* is σ, i.e. the standard deviation.

[4]Here an assumption that the standard deviation can be proxied by the coefficient of variation is made. While the empirical section also introduces more refined measures, the aggregate analysis performed here requires the use of the coefficient of variation, since aggregation is performed across all firms.

constitute a problem here.

On a general level it is difficult to make statements on when the model "begins", that is given σ^2, how many updates has the Bayesian agent carried out. The approach taken in this study, discussed at length in section 4.6, is that the number of observable realizations from common information i.e. annual earnings announcements, defines the minimum amount of learning that has occurred as in the model. The issue is somewhat more prominent in this section compared to Chapter 4, since by using the aggregate measure here it is a) not possible to track the exact number of earnings realizations per firm and b) σ^2 is measured from ending data where possible. The latter involves the problem whether *estimates* of σ^2 are different if measured e.g. at the beginning of the sample period[5] in 1995 or at the end in 2010. As mentioned earlier however, the aim of this section is to illustrate the behavior of the model for reasonable parameter values, and as such σ^2 in this chapter is measured through equal weighted aggregation of the full sample period coefficient of variation of earnings for each firm. In terms of exact identification in the empirical estimations, the important part is that tracking the amount of realizations that have occurred from the commonly observed informative distribution, dictates the *minimum* amount of convergence of beliefs that must occur.

3.1 Static calibration

The first important property of the model, individual subjective belief convergence, can now be illustrated. Starting by assuming that there exists no private information, the convergence of individual (and the common) belief could be illustrated as in Figure 1 (σ^2 normalized to 1). Notice that this is not something that could be observed, since there would exist no dispersion. Nevertheless, this illustrates the point on how (individual) beliefs converge on information. If one would ask for reported confidence intervals for the strength of beliefs, this is how theory suggests they evolve. What is important is that in Figure 1, at each point in time, the agent gets a new realization from the distribution with variance σ^2. As the analysis moves to introduce private information, the magnitude of σ^2 is always assumed to be known. In the above, the assumption is that at 0, agents combined their prior, consisting of σ^2, with a realization from σ^2 at 1, ending up with the forecast at 1.

The above demonstrates the important property of belief convergence on new infor-

[5]Naturally, without using additional data it is not possible to estimate σ^2 from 1995 data. Robustness tests at the end of Chapter 4 however deal with the issue as it relates to the empirical estimations, by making use of a longer period of data.

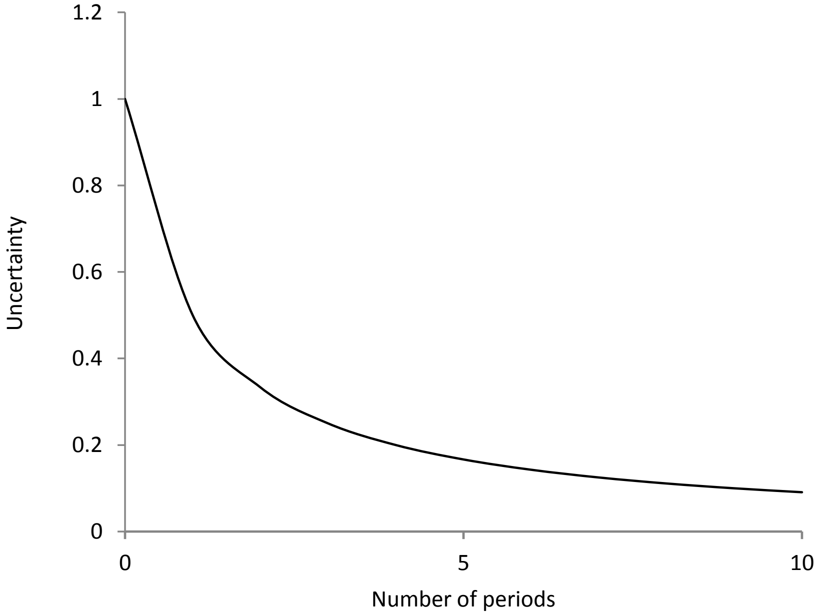

Figure 1: Evolution of uncertainty about θ, where τ_0^2 is normalized to 1. The convergence pattern is the result of signals being of strength $1/\tau_0^2$. The rate of convergence of the belief in the figure is on a variance basis.

mation. Next, the model is calibrated to values that can be observed in the data. The idea is to obtain an overall estimate of σ^2 and assuming that this is the uncertainty of the commonly observed distribution that agents start with. Going forward agents will continue receiving realizations from this commonly observed distribution at each period.

As described previously, σ^2 is found by measuring for each company the average standard deviation of earnings for all available years, scaled by the corresponding (absolute value) of the mean, and subsequently averaging this coefficient of variation measure over all companies. To convert the (average) coefficient of variation of earnings in to the precision metric, r, the coefficient of variation is squared and inverted, the assumption being that the coefficient of variation proxies for σ. The average[6] coefficient of variation of all company earnings in the sample is 1.76, while the average coefficient of variation for dispersion is 0.1434.

Since V and r are now known, it is possible to solve for the precision (variance) of the private information, by simply finding the roots to V, that is:

[6]This number puts no restrictions on the data, such as enforcing December fiscal year ends. See Section 4.7. The exact number used in the calculations below is 1.7625.

$$s_1 = \frac{-2rV \pm \sqrt{1-4rV}+1}{2V}$$

, where in this case the *lower* precision number is of interest (the higher precision number is the reverse potential result).

Plugging into the preceding quadratic equation (implicitly assuming $n = 1$), it is the case that the private information, required to produce the observed dispersion, given that σ^2 is 1.76^2 (≈ 3.1), has to have a variance of 463 ,approximately 149 times higher than that of common information ($0.0206 = s/(s+0.3219)^2$). In other words, the private signal is 149 times less precise than common information (on a variance basis). Notice that the private and common information distributions can be directly compared, since by assumption they are centered on the same value, θ. Naturally, there always exists another solution since the equation is quadratic. However, as noted previously, this root is generally ruled out since it assumes that private information is much more informative than common information. The choice of ruling out these solutions are done only on grounds of realism - these "reverse potential" results would affect the analysis in a symmetric way, and will still yield the same result.

The first calibrations are static in the sense that they are only aimed at finding appropriate starting values for the uncertainty components in order to produce approximate levels for observed forecast dispersion. Finding these "starting values" assumes that only one updating event has occurred (to arrive at the initial values), using a prior of σ^2, estimated at 1.76, and a private signal (or aggregation of), 149 times less precise (in variance terms). In standard deviation terms this implies a factor of 12.21. Assuming now for simplicity that these would have been the starting values, yielding an initial dispersion of the observed 0.1434, it is possible to see what happens if the model is "simulated" forward using Eq. (14) (Proposition 3) and the values obtained for the parameters. Adding signals of 1.76^2 and 463 for each period, in accordance with Eq. (14), yields Figure 2, which tracks how dispersion would behave over time.

If the initial value for dispersion, from which the private signal variance is backed out, instead would have been the average value in 1995, the starting value for dispersion equals 0.1206. Plugging in to the quadratic formula implies that the variance of the private signal would have to be ($0.0145 = s/(s+0.3219)^2$), or approximately 659, or 212 times the variance of common information. In terms of standard deviations, this implies that the standard deviation of the private information distribution should be approximately 14.57 times higher than the standard deviation of common

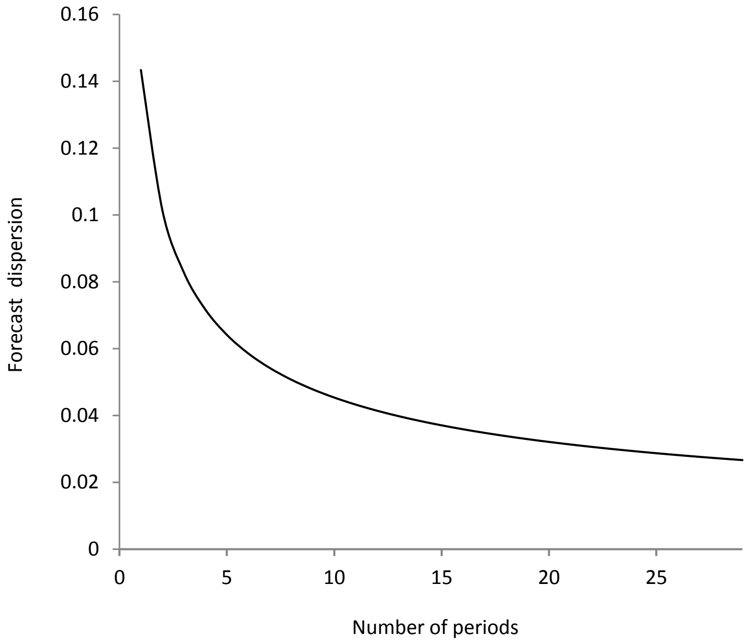

Figure 2: The line in Figure 2 represents the evolution of forecast dispersion from starting values of σ^2, (estimated at 1.76^2), and ν^2 estimated at 463, where signals of respective magnitude are observed each period. Forecast dispersion is portrayed on a standard deviation basis.

information. The evolution of forecast dispersion from the above starting values is depicted in Figure 3, along with the evolution from the starting values of Figure 2, included for comparison. As Figure 3 shows the evolution of forecast dispersion is similar to that of using the total sample average for forecast dispersion (0.1434), versus using the 1995 average for forecast dispersion (0.1206), the difference being that forecast dispersion for the latter is on a slightly lower trajectory.

Naturally, it is quite simplifying to assume that this procedure would locate the correct starting point for learning. Nevertheless, this shows what convergence would look like if the average measures for the data would be used as starting values for the learning process, and agents would continue receiving signals of the same size for each period.

Keeping the assumption of the existence of a true, time invariant σ^2, but instead assuming that learning has already occurred in the past before the sample period, the situation can be depicted as in Figure 4.

The choice of a previous learning period, in this case 5, is arbitrary, but illustrates

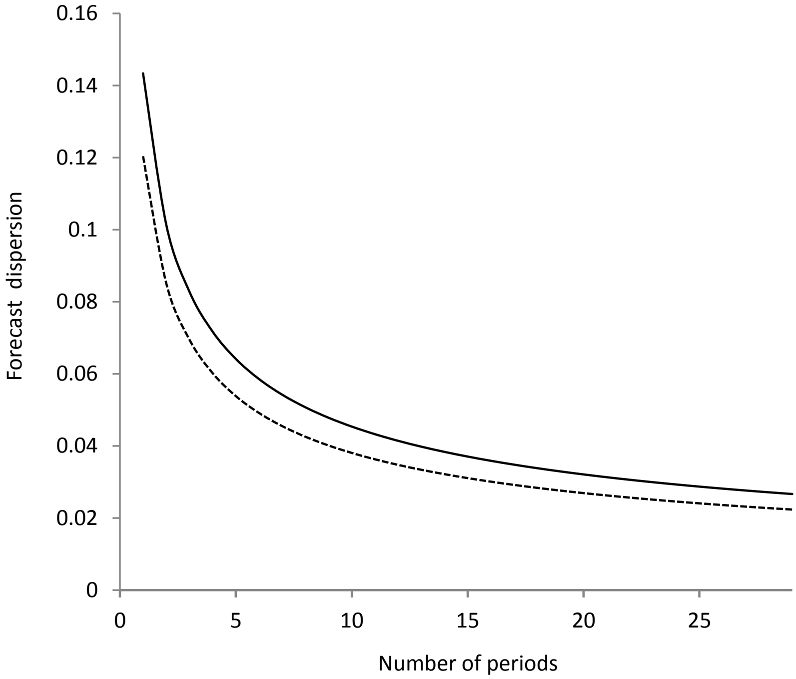

Figure 3: The solid line in Figure 3 represents the evolution of forecast dispersion from starting values of σ^2, (estimated at 1.76^2), and ν^2 estimated at 463, where signals of respective magnitude are observed each period. The dotted line in turn depicts the evolution of forecast dispersion where ν^2 is estimated at 659. Forecast dispersion is portrayed on a standard deviation basis.

the main point. If some learning has occurred in the past, the evolution of dispersion flattens out (all the while keeping in mind that at each period agents get new signal realizations from σ^2 and ν^2, here illustrated for the same values as previously, that is 1.76^2 and 463, respectively). Obviously, the average dispersion now does not match well with the observed levels of dispersion previously estimated, that is 0.1434. In fact, Figure 12 on page 80, which tracks the average dispersion (measured on a coefficient of variation basis as here), shows that dispersion does not contain neither a decreasing trend, nor values such as those in Figure 4. The flattened part of the curve in Figure 4 is approximately 2 to 3 times too low in comparison with actual observed values.

One solution to this would be to go back to the pre-analysis learning period and stipulate that the private information distribution actually was of higher precision than the suggested (variance) of 463. This would be the case in Figure 5.

In Figure 5, dispersion better matches the observed values (which are 0.1434 on

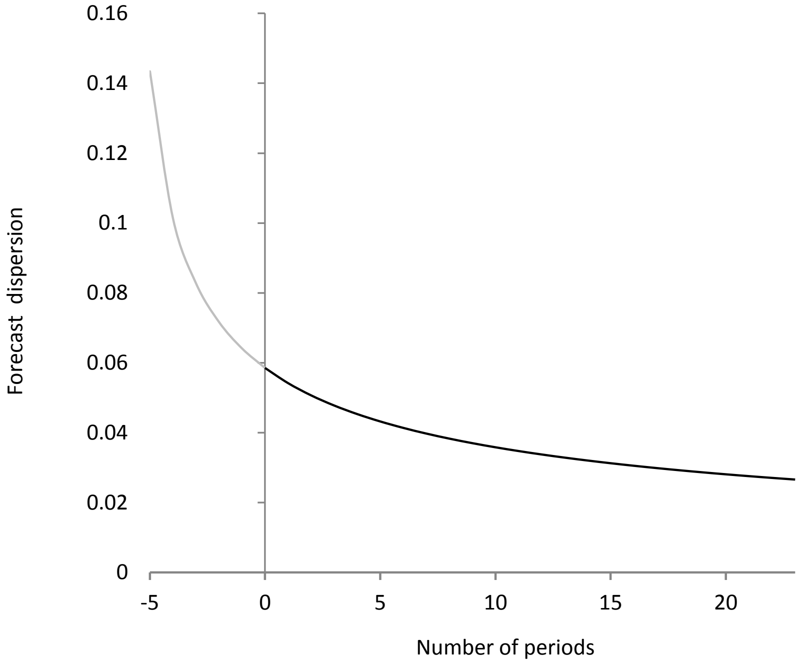

Figure 4: The line in Figure 4 represents the evolution of forecast dispersion from starting values of σ^2, (estimated at 1.76^2), and ν^2 estimated at 463, where signals of respective magnitude are observed each period. The (arbitrary) first 5 periods, positioned at the negative side of the x-axis, are illustrated as a grayed out part of the line. Forecast dispersion is portrayed on a standard deviation basis.

average), even though the decline is somewhat steeper. In order to produce this scenario however, the starting variance for private information is sizably lower, in this case 65.9. This on the other hand would act to reduce the potential in the system, since $r - s_1$ is now smaller. Even this is not enough to match the observed trend for dispersion observed in Figure 12 on page 80, since in the learning with fixed private signals in Figure 5, dispersion continues to show a decreasing trend. To be able to produce a flat to increasing trend, private signals must be of increasing precision. This is introduced in the following section, after the effect of large private signals that maximize dispersion directly have been considered.

3.2 Dynamic calibration

The previous calibrations were concerned with finding appropriate starting values for the uncertainty constituents, and studying the evolution of the system in accordance with Proposition 3. Now instead changes to the variance of the private signal

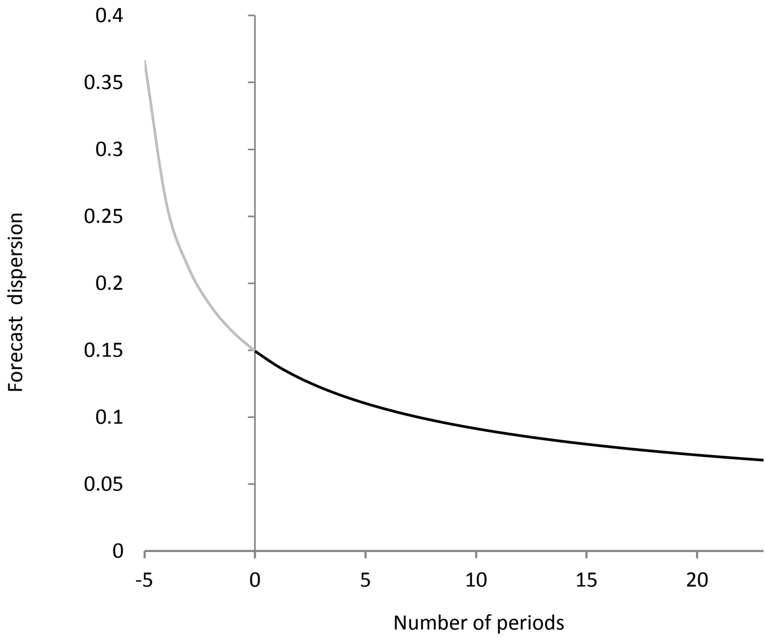

Figure 5: The line in Figure 5 represents the evolution of forecast dispersion from starting values of σ^2, (estimated at 1.76^2), and ν^2 estimated at 65.9, where signals of respective magnitude are observed each period. The (arbitrary) first 5 periods, positioned at the negative side of the x-axis, are illustrated as a grayed out part of the line. Forecast dispersion is portrayed on a standard deviation basis.

distribution are considered. These are interpreted as the arrival of more informative private signals. The subsequent figures illustrating the behavior of dispersion[7], are augmented with lines representing maximum dispersion (Eq.[18]) and the potential, or multiplier from previous to maximum (current) dispersion (Eq.[19]).

Since potential measures the multiplier on dispersion at $n-1$ that yields maximum dispersion at n, analyzing potential *going forward* implies looking at potential for $n+1$. In the illustrations, if analyzing the situation at some point e.g. $n=7$, potential for increasing dispersion (next period) is potential at $n=8$.

Figure 6 tracks the learning process with values of σ^2 estimated at 1.76^2, and ν^2 estimated at 463, the same values as those in Figure 2. Signals of corresponding magnitude are drawn each period. This corresponds to a case where the difference between r and s_1 is somewhat large. The somewhat large difference between the

[7]Values for dispersion at each point can most easily be found by using Proposition 2 at each point. Similarly, the potential or multiplier to maximum dispersion is easily found using the precursor to Eq. (19).

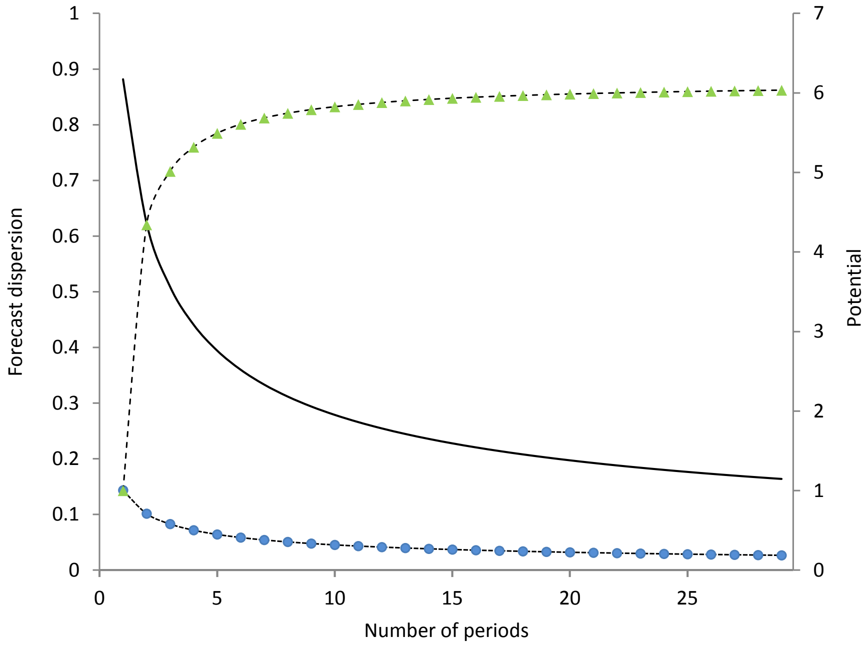

Figure 6: The dotted line (blue circles) represents the evolution of forecast dispersion from starting values of σ^2, (estimated at 1.76^2), and ν^2 estimated at 463. The solid line represents the maximum amount of dispersion that can be achieved at any time period. Simultaneously, this corresponds to the case where $\nu^2 = \sigma^2$. The dashed line (green triangles), describes the maximum potential for increasing dispersion from its previous value at $n-1$. Here, if a signal, s'_{n^*max}, arrives that raises dispersion to its maximum value, potential for raising dispersion is used up.

uncertainty of the common and private signals implies that there exists potential to increase dispersion substantially. The line in Figure 6 representing potential visually illustrates the size by which current dispersion (at any point) can be increased by a more informative private signal.

To illustrate how the potential is used up with a signal of strength s'_{n^*max} (a private signal that maximizes dispersion by definition), Figure 7 shows the arrival of a signal, s'_{n^*max} at $n = 7$ and dispersion jumps to its maximum value at $n = 7$. Since dispersion is at its maximum for the amount of information at $n = 7$ (i.e. conditioning on σ^2), and the *maximum* dispersion is always decreasing in n, the potential drops to below one. Here, for $n > 7$ the private signal is equal to r. A slightly technical implication that follows is that the private signals for $n > 7$ are essentially of slightly lower precision[8] than s'_{n^*max} (at n^*). This corresponds to the

[8]Why is it the case that private signals for $n > 7$ are of lower precision than s'_{n^*max} (at n^*)? For $n > 7$, private signals will equal r. The statement above thus implies that r is of lower precision (less) than s'_{n^*max}. Since s'_{n^*max} is essentially $n * r - (n-1) * s_1$, the implied inequality is $n * r - (n-1) * s_1 > r$. Solving yields $r > s_1$, which is the case since starting private precisions

first scenario in section 2.11 in Chapter 2.

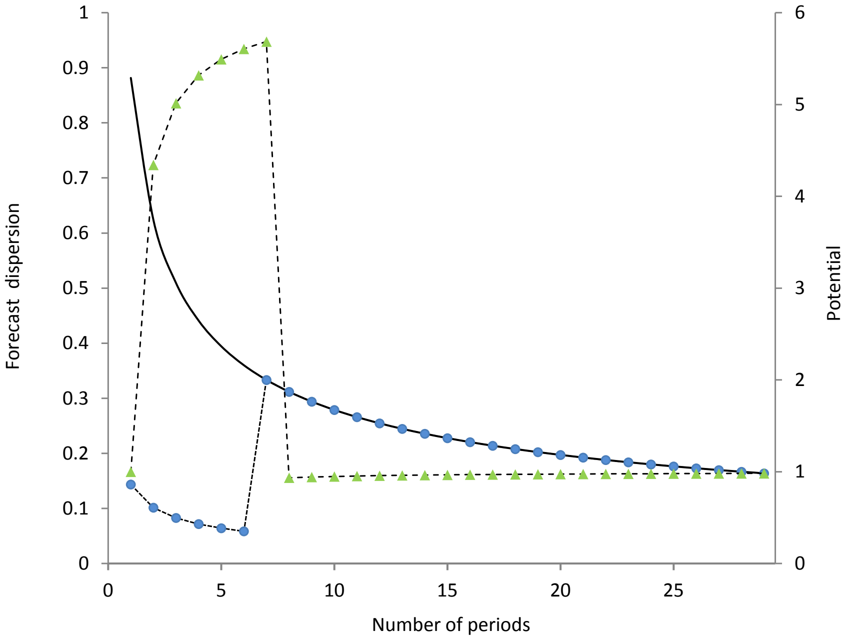

Figure 7: The dotted line (blue circles) represents the evolution of forecast dispersion from starting values of σ^2, (estimated at 1.76^2), and ν^2 estimated at 463. The solid line represents the maximum amount of dispersion that can be achieved at any time period. Simultaneously, this corresponds to the case where $\nu^2 = \sigma^2$. The dashed line (green triangles), describes the maximum potential for increasing dispersion from its previous value at $n-1$. The private signal, s'_{n^*max}, arriving at $n = 7$, raises dispersion to its maximum value and consequently uses up all potential. Since here, for $n > 7$ private signals $s = r$, the system is on a trajectory for maximum dispersion. This is illustrated by the fact that after $n = 7$, potential never exceeds 1 .

Figure 8 illustrates a similar case, but where signals following $n = 7$ continue to be of strength s'_{n^*max}, describing the analysis in the second scenario in section 2.11 in Chapter 2. This illustrates two additional aspects. First, it demonstrates the fact that as private signals continue to be more precise than r for $n > 7$, the system is not on track for maximal dispersion, and this builds up potential in the system going forward (seen in the fact that the dotted line representing potential (triangles) eventually rises above 1 in Figure 8). The second, and essential point, is that if the system has reached the maximum dispersion at $n = 7$, after the arrival of s'_{n^*max}, then the system has used up its potential for increasing dispersion as a result of increased private information. It can now be readily seen that after this point the proposed dynamics under the private information acquisition hypothesis fail to

are by assumption less precise than common information.

hold. That is, after s'_{n^*max} at $n = 7$, increasing private information (the standard explanation for forecast dispersion), now only accelerates convergence and is not capable of producing higher levels of dispersion. Now the system deviates from the trajectory of maximum dispersion for $n > 7$, and potential starts adding up going forward. This however is "reverse potential" in the sense that in order to utilize this potential, and cause subsequent increases in dispersion, private signals would have to become less informative.

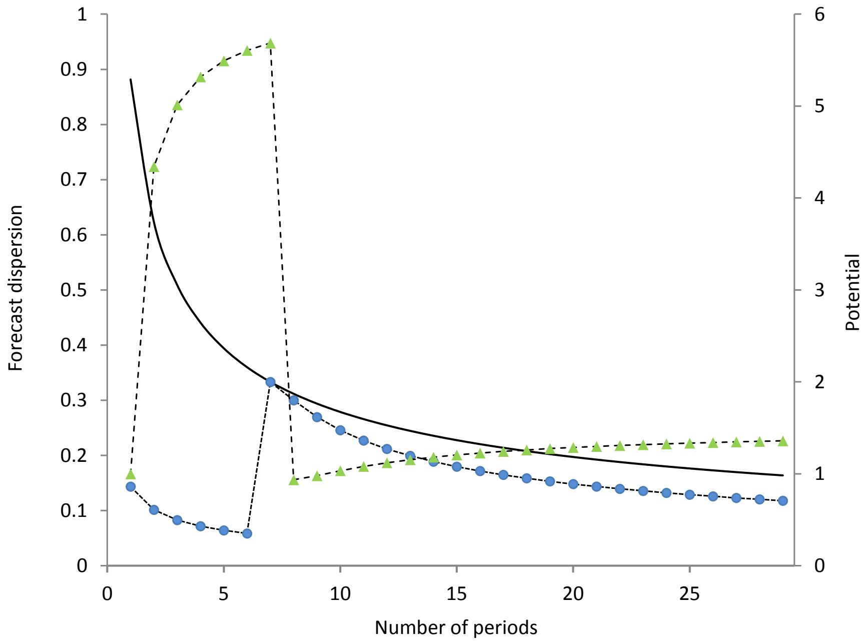

Figure 8: The dotted line (blue circles) represents the evolution of forecast dispersion from starting values of σ^2, (estimated at 1.76^2), and ν^2 estimated at 463. The solid line represents the maximum amount of dispersion that can be achieved at any time period. Simultaneously, this corresponds to the case where $\nu^2 = \sigma^2$. The dashed line (green triangles), describes the maximum potential for increasing dispersion from its previous value at $n - 1$. The private signal, s'_{n^*max}, arriving at $n = 7$, raises dispersion to its maximum value and consequently uses up all potential. Since here, for $n > 7$ private signals continue to be s'_{n^*max}, the system is not on a trajectory for maximum dispersion. This is illustrated by the fact that after $n = 7$, the potential exceeds 1 and that dispersion (dotted line, blue circles) starts deviating from the track of maximum dispersion (solid line).

The notion of used up potential is interesting, since it shows that even if some instances of jumps in dispersion could be explained by increases in private signals, at some point the same increases in private signals start having the opposite effect, that is *decreasing* dispersion.

The illustrations up until this point show how the model can be calibrated to values observed from empirical data and solved forward. The role of informative private signals that increase dispersion (to its maximum) in the context of the model have also been considered. Excluding the private signals that temporarily increase (maximize) dispersion, the trend of dispersion in the model is decreasing over time. The empirical trend for dispersion in forecasts, seen in Figure 12 on page 80, shows a different tendency, where the overall trend is flat to increasing, with the additional large increases occurring in the years 2008-2009. To match the empirical flat average trend in dispersion, small amounts of increases in private information must be added to the model each period. The implication of adding private information at each period however is that potential is reduced each period. Thus, in a scenario that is capable of yielding a flat to increasing trend in dispersion, potential for (large) increases in dispersion disappear fast. In Figure 9 constant (fraction) increases in private information, from the starting values $\sigma^2 = 1.76^2$ and $\nu^2 = 463$, are depicted for 10, 20, and 30 percent annually (on a variance basis).

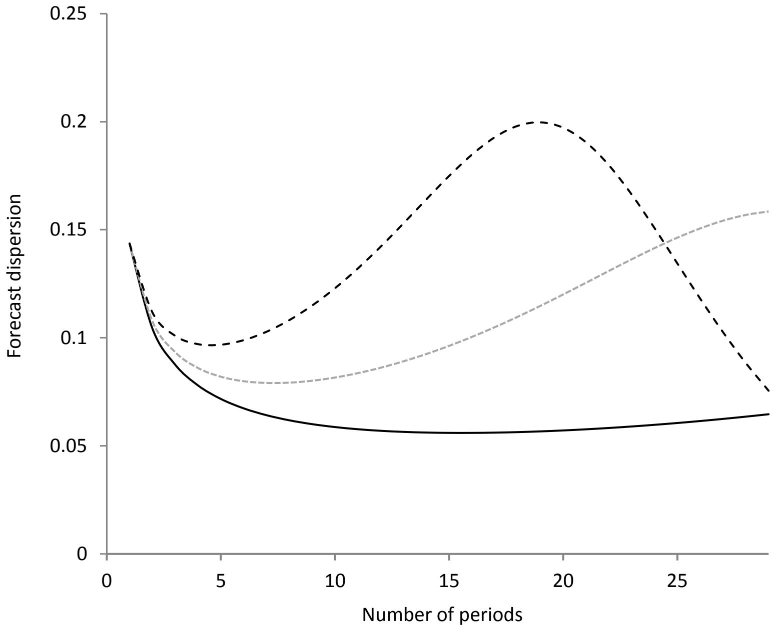

Figure 9: The solid line depicts the evolution of dispersion where private signals become more informative (reduction in variance) by 10 percent each year. The gray dotted line depicts the analogous situation but with private signals becoming more informative by 20 percent each year. Finally, the dashed line shows the evolution of dispersion where private signals become more informative by 30 percent each year. The starting values are: $\sigma^2 = 1.76^2$ and $\nu^2 = 463$.

While a slow increase in private information precision (reduction in variance) over time can yield a constant rate of dispersion over time (for a finite period), Figure 9 shows that this is sensitive to the exact rate at which this happens (this also depends on $r - s_1$, not shown here, i.e. different values yield different dynamics). Importantly, the evolution of dispersion for the 20 and 30 percent rates hints at the effect at play; the increases in dispersion are the result of the system eating up potential over time, thus leaving less and less potential left over for increases in dispersion following the arrival of higher precision signals.

Figure 10 illustrates what happens to potential in the case of 30 percent (variance) annual increases in private information precision each period.

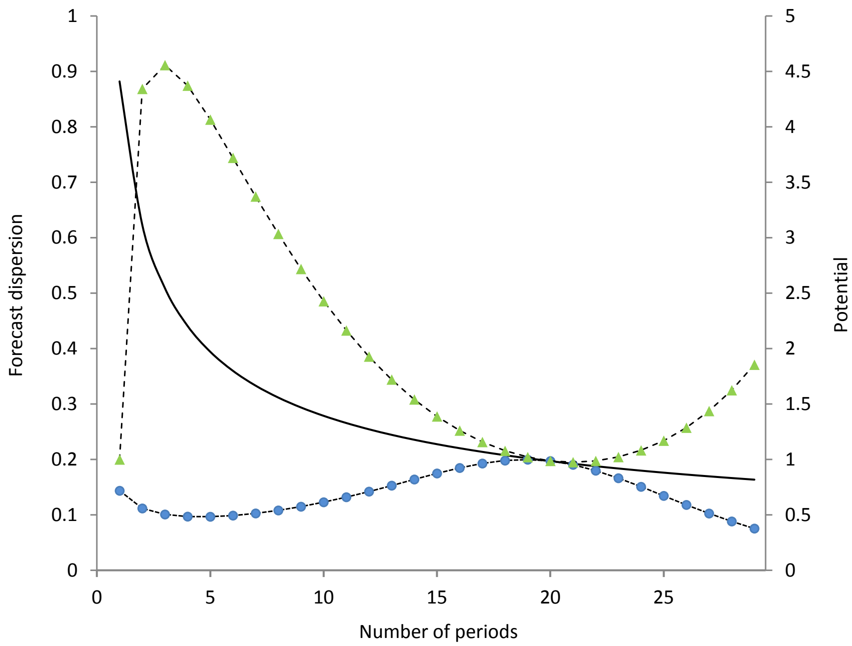

Figure 10: The starting values are again: $\sigma^2 = 1.76^2$ and $\nu^2 = 463$. The dotted line (blue circles) is the same as the dashed line in Figure 9, and shows the evolution of dispersion where private signals become more informative by 30 percent each year (in variance terms). The solid line represents the maximum amount of dispersion that can be achieved at any time period, while the dotted line (green triangles), describes the maximum potential for increasing dispersion from its previous value at $n - 1$.

In Figure 10, it is apparent that the private signals that increase over time, erode the potential from the system up until the point at approximately 20. From here on, the increasing private signals act to reduce dispersion. The fact that the potential starts growing results from the fact that the system is now acquiring reverse potential such that in principle dispersion could be increased by reversing the evolution of the private signals such that they suddenly turn to become less informative.

Now, withholding the private information acquisition hypothesis, given the assumption of private information increases of 30 percent annually, any (more informative) signal arriving after the system has used up its potential, will now lead to a *drop* in dispersion, as depicted in Figure 11.

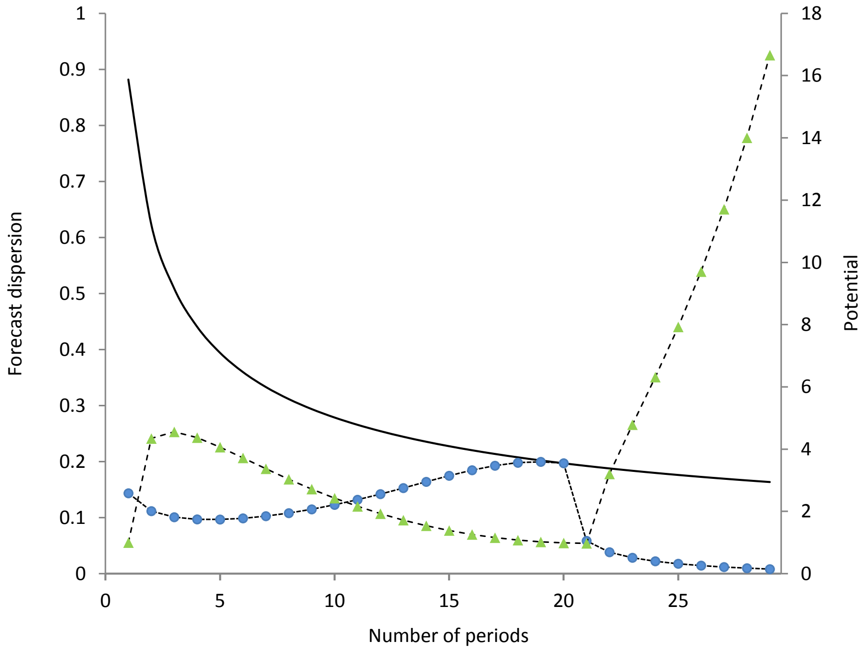

Figure 11: The starting values are again: $\sigma^2 = 1.76^2$ and $\nu^2 = 463$. The dashed line (blue circles) is the same as in Figure 10 (and the dashed line in Figure 9), and shows the evolution of dispersion where private signals become more informative by 30 percent each year. The solid line represents the maximum amount of dispersion that can be achieved at any time period, while the dotted line (green triangles), describes the maximum potential for increasing dispersion from its previous value at $n-1$. At $n = 21$, a more informative signal (more private information) arrives and now leads to a substantial *drop* in dispersion.

The more informative signal arriving at $n = 21$, in this case a signal that is 10 times[9] more informative than the previous signal on a standard deviation basis, now causes a reduction in dispersion. This is in contrast to the situation portrayed in Figures 7 and 8, where, as the standard theory predicts, the increased private signal yielded an increase in dispersion. (Note that again, the reason for the large increase in

[9]The value of the chosen private signal is arbitrary. It is chosen such that the effect can be clearly illustrated in Figure 11; a private signal of *any* size will reduce dispersion beyond this point.

potential (not dispersion) that occurs after the signal at n=21 consists of "reverse potential").

This exemplifies how the private information acquisition hypothesis is heavily dependent on the amount of potential the system has. Also, it shows that this potential is finite.

3.3 Concluding remarks

This section aimed at illustrating some of the main dynamics of the model of chapter two, where adding information over time affects convergence properties.

The key aspects are naturally the same[10] as those from Chapter 2, only here the illustrations will hopefully provide for better intuition. Simultaneously however, it is important to keep in mind that each illustration only captures a limited view of the model, where the exact dynamics are valid only for the chosen values. These values (starting values) however are chosen to represent the data, whereas manipulation of the values are chosen as to give as much intuition for the possible dynamic behavior that the model allows for.

As noted earlier, the calibrations are carried out for illustrative purposes, and the calibrations to the data are made on an *aggregate* level. There are however numerous ways that e.g. an average pattern of a flat to increasing dispersion year by year, as in Figure 9, could be produced - this could for example be the result of high starting variances for private information for all companies, combined with subsequent small increases in private information precision for each period. Alternatively, the same average aggregate pattern could be produced if a fraction of all companies experience large increases in dispersion each year (following large increases in private information precision, given a low starting precision), whereas the remaining fraction of companies could evolve in accordance with fixed common and private information signals, as in e.g. Figure 2, showing a steady decline in dispersion over time, the aggregate effect being an observed average amount of dispersion which remains fairly constant year over year.

It is worth emphasizing that all theoretical predictions of the model are on a firm level, where the model of learning is for Bayesian agents who receive common and private signals from the (true) distribution of company earnings, that allows the agent to learn the true mean of the process. As the model does not make predictions

[10]See section 2.12.

on aggregate effects for all companies, the empirical tests in the following chapter will construct test that are fine-grained enough to allow for isolating the required firm-level dynamics.

The empirical section will therefore be concerned with various tests of the model (Proposition 3), analyzing specifically whether forecast dispersion conforms to the (maximum) theoretical bounds proposed by the model.

Chapter 4

Empirical Analysis

This chapter is aimed at empirically testing the implications that come out of the model developed in Chapter 2. The rationale for expanding the Barry and Jennings (1992) and Barron et al. (1998) model is that the predictions from the (their) model(s) are widely used in the literature, especially in the form of the private information acquisition hypothesis, where increases in forecast dispersion are associated with agents acquiring more private information. This theoretical prediction is usually taken more or less at face value.

Even though the Barry and Jennings (1992) and Barron et al. (1998) model can be viewed as nesting a multiperiod setting, proper operationalization of the model to fit empirical data spanning multiple periods becomes difficult. This is due to the fact that *prima facie*, the model only allows for one aggregated input for common information. It is with respect to this aggregate common information that analyses in the literature are being made. Specifically, the private information acquisition hypothesis comes out of the first derivative of the variance of the forecasts (forecast dispersion) w.r.t. private information, implying that common information is kept fixed. Common information here however is total common information, and cannot be compared across time periods without simultaneous consideration of information release/dissemination dynamics. It is exactly this comparison that the model representation of this study achieves, simultaneously allowing for testing whether observed levels of forecast dispersion seen in the data are compatible with the learning model, given that one can assume that the earnings series reasonably represents common information.

There are two major predictions that directly come out of the model extension developed in Chapter 2. These are:

1) For *reasonable* values of existing private information, there exists a maximum amount (factor) by which dispersion can increase. This is Eq. (19).

2) *Regardless* of the amount of private information, there exists for each time period a maximum amount of dispersion in forecasts that is allowed by the model, given that one can reasonably measure common information. This is Eq. (18) .

Unfortunately, as private information is not measurable *by definition*, mathematically, as private information goes to zero, the potential increases in the form of a factor become infinite (in the limit). While this could be considered as perhaps not being reasonable, or at least realistic, using prediction 1) would still involve the use of additional assumptions on private information, assumptions that would not be very transparent.

Thus instead, the main tests in this chapter will initially use prediction 2), which puts no limits on private information. Here the strong assumption instead involves common information, but this is advantageous since assumptions regarding common information remain transparent.

4.1 Hypothesis formation

While the first tests will be concerned with "setting the scene", in the sense of looking at dispersion in forecasts in general, the main interest lies in evaluating the performance of the model. The evaluation of the performance of the model simultaneously tests the embedded private information acquisition hypothesis. Assessment of the model is to be carried out through evaluating the boundaries in Eq. [18] .

Eq. [18] states how large the dispersion in forecasts can be at a maximum at each point in time, and is given by $var_{max}(u_{ni}) = 1/4nr$. To be able to test the performance of the model, the theoretical constructs in Eq. [18] have to be mapped to empirically observed counterparts. The LHS in the above is of course simply the variance of forecasts, and thus the emphasis lies on finding a proxy for r in the denominator of the RHS, the distribution from which agents are thought to receive earnings realizations. While this is in no means neither an easy, nor completely unambiguous task, the subsequent sections (4.7, 4.17, 4.18) provide an in depth discussion on how this is achieved. Once a reasonable proxy for r is constructed however, the maximum magnitude of dispersion going forward is known.

A combination of the intuition from Chapter 1, relating dispersion to uncertainty, along with the attributes of the model from Chapter 2, particularly the rapid convergence of learning, leads to the following hypotheses:

H1: The magnitude of forecast dispersion is too large to be supported by agents' acquisition of private information.

H2: The magnitude of forecast dispersion is too large to be supported by agents' acquisition of private information, and this is prominent in "uncertain environments" categorized by negative earnings realizations.

H1 can be seen as more of a general test of the model, and subsequently on the performance of the private information acquisition hypothesis, whereas H2 relates more directly to the potential link between dispersion and uncertainty. Under the null of both hypotheses, where the fixed learning model is withheld, the implication of observing increases in dispersion is *always* that agents have become better informed. When dispersion increases, there is more information in the economy and more information will always lead to a sharper belief on the level of the individual analyst (given that the learning regime remains fixed). Consequently, when dispersion increases, agents become *more* certain. This is the opposite of the intuition of dispersion being related to increased uncertainty, since increased dispersion under the model implies increased *certainty*.

4.1.1 Background for H2

The starting *intuitive* hypothesis is that negative earnings possibly proxy for (idiosyncratic) environments where forecasting is difficult or, in other words, uncertainty is high. In contrast, theoretically disagreement (V or V') is completely characterized by uncertainty components only, that is common uncertainty (prior variance) and private uncertainty (signal variance). The example in Appendix A makes explicit that there exists no direct (theoretical) mechanism through which any realizations could affect disagreement, other than through the mechanical effect of information addition that causes disagreement and uncertainty to decrease in Bayesian learning environments (each common draw reduces uncertainty). This uncertainty decrease, or convergence, occurs, and is fixed/given, regardless of observed common values. The location of the signal realization does not affect subjective uncertainty, the only factor that affects the subjective uncertainty is the variance of the distribution from which the signal is drawn. Thus, negative earnings (or any realizations) themselves do not provide a direct mechanism for affecting disagreement.

If the null hypothesis under H2 holds, that is that agents are learning from realizations from a time-invariant process, then the mechanism that could cause the hypothesized negative earnings to induce increases in disagreement is an increase in private information (private information acquisition) around losses. In other words, around losses, agents acquire more private information[1]. The link between uncertainty (disagreement) and losses is thus indirect; private information merely acts as a catalyst.

As long as one assumes that learning is rational and indeed occurs over time, it is not theoretically correct to assume that subjective uncertainty could have increased. To validate an exploration in to explanations involving increased uncertainty for increased dispersion, one must first rule out that predictions from the standard model are possible.

4.2 Related empirical research

The empirical research covered in this section can be seen to be most closely related to H2, namely that dispersion is hypothesized to be linked to heuristically viewed uncertain environments,[2] especially when the indicator for such uncertain environments is negative earnings.

While the recent interest toward disagreement in Finance is in one form or other closely tied to the Diether et al. (2002) "puzzle"[3] of forecast dispersion and lower returns, empirical work in Accounting Research has been more geared towards studying dispersion as it relates to information in terms of notions of uncertainty or information revelation, typically around earnings announcements. While some empirical work in Accounting Research is similar to tests performed in this study, the *interpretation* of the results are different, since inference is typically drawn solely based on the idea of dispersion increasing as a result of agents having access to more private information.

Morse et al. (1991) is an early study on the convergence or divergence of forecasts around earnings announcements, where the authors show that forecast dispersion, following large earnings surprises, tends to increase. The stance of Morse et al. (1991) is however different from this study in some key aspects, in that they are analyzing earnings surprises, and that they measure dispersion *following* earnings surprises. Brown and Han (1992) also question the empirical results in Morse et al. (1991), suggesting that the results in Morse et al. (1991) are driven by the use of I/B/E/S summary data and that the effect found in Morse et al. (1991) can

[1]This requires the assumption that private information is initially vague, in order for potential to exist. Additionally, the predictions under the null rest on the somewhat restrictive assumption of the existence and availability of private information.

[2]This is covered in greater detail in the following section "uncertainty and disagreement, revisited"

[3]This is a puzzle in the standard risk-return world.

only be identified for the largest decile of earnings surprises. Brown and Han (1992) further conjecture that the combined evidence of (1) Bayesian belief revision theory and (2) the time series properties of earnings, suggests that earnings announcements should increase the convergence of analysts' beliefs. Because the theoretical position in Morse et al. (1991) is very different from a standard rational Bayesian one, the similarity of this study to Morse et al. (1991) is mostly limited to the empirical approach of analyzing forecast dispersion around earnings announcements.

This study is not the first to propose a link between disagreement and lower future earnings, and some authors have, for example, used this notion as a potential resolution for the dispersion-lower returns puzzle documented by Diether et al. (2002). Indeed, Ali et al. (2009), and to a lesser extent Xu and Zhao (2010), are two examples of such studies.

Ali et al. (2009) hypothesize that the forecast dispersion anomaly is driven by: "firms' tendenc[ies] to disclose good news about future earnings on a timely basis and to delay the disclosure of bad news". Ali et al. (2009) thus base their starting point in the selective disclosure hypothesis in eg. Verrecchia (1983). Ali et al. (2009) further refer to Lang and Lundholm (1996) in that "firms that provide less public disclosure about future earnings exhibit greater forecast dispersion, presumably because financial analysts reliance on their private source of information is greater for these firms". Yu (2011) confirms that firms that provide earnings guidance have lower dispersion, but Yu (2011) subsequently excludes firms that provide guidance, focusing on only on firms that do not provide earnings guidance. Nevertheless Ali et al. (2009) show that "After controlling for the relation between forecast dispersion and future earnings, we find that forecast dispersion is no longer negatively related to future stock returns".

While drawing a connection between dispersion and negative earnings, Ali et al. (2009) use this empirical finding to show that it is related to the Diether et al. (2002) pricing puzzle of forecast dispersion leading to lower future returns. As such, Ali et al. (2009) do not test what drives dispersion, rather their results show that controlling for earnings surprise, the forecast dispersion - negative return anomaly disappears. Importantly, the theoretical stance of Ali et al. (2009) is that of private information acquisition driving dispersion, the very issue that this study aims to provide evidence on.

Xu and Zhao (2010) is another study aiming at explaining the dispersion-lower returns puzzle, in their case through analyst coverage and idiosyncratic volatility. The idiosyncratic volatility "puzzle" is a relatively new negative return anomaly, due to Ang, Hodrick, Xing, and Zhang (2006), who show that stocks with high idiosyncratic volatility earn low average returns. The robustness of this evidence has however more recently been questioned by Bali and Cakici (2008).

In assessing the role of idiosyncratic volatility on the dispersion - negative return relation, Xu and Zhao (2010) find that the negative return effect is absent for stocks with high idiosyncratic risk (and are followed by more analysts), implying that idiosyncratic risk is the culprit behind the negative relation between dispersion and returns. Conversely Yu (2011) points out that the disagreement effect on (market) returns is orthogonal to the Ang, Hodrick, Xing, and Zhang (2006) idiosyncratic volatility effect, and that the disagreement - negative returns relation is *not* driven by idiosyncratic volatility. Xu and Zhao (2010) find that the dispersion negative return relation is still present for firms with low idiosyncratic risk (and low analyst coverage), and argue that the Ali et al. (2009) result where lower future earnings explain the dispersion return relation is driven by these firms.

As this study is not aimed at providing direct explanations for the dispersion - lower returns puzzle, the extent to which negative earnings relate to the dispersion - negative return anomaly, and interacts with the other predictors, such as those in Yu (2011) or Xu and Zhao (2010), is not addressed in this study.

Adut, Sen and Sinha (2008) is a study that is empirically related to the tests performed in this study, as the authors hypothesize and subsequently test the idea that the variance of analysts' forecasts is larger in "bad-news" environments. Adut et al. (2008) also specifically show that forecast dispersion is related to ex post negative earnings. The greatest similarity between the tests in this study and those of Adut et al. (2008) is their use of the same notion of bad-news, that is negative earnings and they also study this ex post. While Adut et al. (2008) show that that this hypothesis indeed holds empirically, they base their theoretical stance, as does Ali et al. (2009), in the selective disclosure hypothesis (Verrechia (1983)). Furthermore, Adut et al. (2008) use the measures in Barron et al. (1998) to arrive at the conclusion that the amount of private information of analysts are always higher in a bad-news environment.

In summary, the previous evidence on dispersion and future earnings closest to this study are Ali et al. (2009) and Adut et al. (2008). The important distinction between this study and those of Ali et al. (2009) and Adut et al. (2008) are that while studying the same empirical relation between dispersion and future earnings, Ali et al. (2009) and Adut et al. (2008) base their explanations on the private information acquisition hypothesis, the very hypothesis this study aims to test.

4.3 Uncertainty and disagreement, revisited

This section aims to illustrate how this study relates to the predictions of Barron et al. (2009), as well as clarifying the meaning of uncertainty in context of the model of this study in comparison to both Barron et al. (1998) and Barron et al. (2009).

First, as long as the learning model in the Bayesian setting remains fixed, more information leads to sharper subjective belief on θ. A sharper belief means less variance or uncertainty (in the normal sense of the word). This uncertainty decrease occurs regardless of information source, that is private or common information. Also, since increases in private information is the required condition for *increasing* dispersion, the implication here is that when dispersion increases, uncertainty *decreases*, on the level of the individual agent. As long as the model holds, subjective uncertainty, the uncertainty that the agent uses to assess θ can never *increase*.

The "external" common uncertainty that agents in the model of this study are facing is represented by the variance, σ^2, of the distribution representing common information. The uncertainty from the common distribution is filtered through the learning equation, and agents' beliefs reflect and are a function of the external uncertainty σ^2, and the number of periods.

Since Eq. (18) denotes the maximum for dispersion, regardless of private information,[4] this study looks at the uncertainty that agents are facing in terms of common uncertainty, σ. Consequently, this is the relevant definition of (external) uncertainty that is analyzed, since private information is always assumed to already be maximized. Thus, when this study refers to the uncertainty that agents are facing, for the above reason, this usually only denotes common uncertainty. By separately keeping track of common uncertainty at all times, the link between external uncertainty and the uncertainty in the information sets remains straightforward. This is in contrast to the uncertainty measure used in Barron et al. (1998) and Barron et al. (2009). Below an attempt is made to explain the difference.

In their attempt at separating the drivers behind dispersion in forecasts, Barron et al. (2009) conjecture that *"changes in dispersion primarily reflect changes in information asymmetry whereas levels of dispersion primarily reflect levels of uncertainty"*.

The compatibility of the first part of the above statement ("changes in dispersion

[4] The term regardless is slightly ambiguous since the maximum is attained for a specific value of private information, namely when cumulative private information is equal to cumulative common information.

primarily reflect changes in information asymmetry") with the expanded model developed here, or the Barron et al. (1998) model, is easily demonstrated. This can be seen by taking the partial derivative of the variance of forecasts with respect to private information, $\partial V'/\partial s_1$ (or $\partial V/\partial s_1$). Keeping common information fixed, implying the analysis is within one period, increases in information asymmetry indeed yield increases in dispersion (given that private information is initially vague or non-existent). Thus, a necessary condition for increasing dispersion is an increase in private information, and can be seen to describe what Barron et al . (2010) are referring to in the first part of the above statement. However, it is important to remember that while increases in private information are necessary for dispersion to increase (within a fixed learning regime), it is *not* necessarily the other way around; increases in private information do *not* unconditionally imply that dispersion will increase. This is very much a point made in this study and can for example be seen in Figure 11; as private information increases in precision beyond that of common information, private information acquisition or production instead starts decreasing dispersion. Furthermore, in a setting as that of this study, the convergence over multiple common signals will also yield decreases in dispersion if private information is kept fixed. Over time changes (decreases) in dispersion are theoretically expected that occur due to common information.

The second part of the statement in Barron et al. (2009), "... levels of dispersion primarily reflect levels of uncertainty", is somewhat more complex to analyze. First, the term uncertainty that Barron et al. (2009) (and Barron et al., 1998) are using, is somewhat distinct from the measures used here. This study at all times distinguishes between the two sources of uncertainty, namely private and common uncertainty, and consequently keep track of their evolution separately. In Barron et al. (2009) (and Barron et al., 1998) on the other hand, the term uncertainty actually denotes the precision (or variance) of the posterior distribution of the parameter being estimated[5]. As such, it is a mixture of *both* sources of uncertainty. The perspective of this study is that analyzing uncertainty as a mixed measure of two sources of uncertainty complicates the analysis and makes assumptions/implications less transparent. Hence this study does not use a mixed measure of uncertainty.

Returning to the assessment of the second part of the statement, start by looking at how the *common* part of uncertainty affects dispersion. This can be seen by taking

[5]The posterior variance in e.g. "The standard model" given in Eq. 6) is given by $\left(\frac{1}{\tau_0^2} + \frac{1}{\nu^2}\right)^{-1}$. Since $1/\tau_0 = h$ and $1/\nu^2 = s$, the posterior variance in terms of precisions can be written as $\frac{1}{h+s}$, (Here $\nu_i^2 = \nu^2 \forall i$). This is exactly the definition of V in Barron et al. (1998), found in their Eq. (19).

the partial derivative[6] of V' with respect to r, where $\partial V'/\partial r$ is overall decreasing in r. Also, since $r = 1/\sigma^2$, the implication is that that V' is increasing in σ^2, and thus the higher the common uncertainty, the higher the disagreement, keeping private information fixed. (This can also be seen by taking the derivative of V or V' in the variance representation w.r.t. τ^2 or σ^2, respectively). Stating that the higher the common uncertainty, the higher the disagreement, or to some extent saying that levels of disagreement reflect common uncertainty, are however only useful statements in comparing different starting scenarios or companies. Because a static analysis does not realistically describe the problem at hand, analyzing increases in uncertainty in the information sets does not happen *within* a learning regime for a given company.

The connected learning regime dictates that the uncertainty in the information set always decreases. As σ^2 characterizes the objective distribution that remains fixed, a discussion on increases in r or σ^2 alone is not correct considering the representation in V' (unless the consideration concerns starting values). Instead, increases in certainty in the V' representation should be thought about as occurring through the combined value of σ^2, or r multiplied by n. In the context of the representation in V, if h is thought of as cumulative measure consisting of both n and r, then it is correct to state that h, cumulative common certainty, can never decrease.

Of course, as the term uncertainty used in Barron et al. (2009) (and Barron et al., 1998), the precision (or variance) of the posterior distribution, depends on common information, it also follows that uncertainty, defined as in Barron et al. (2009) (and Barron et al., 1998), is reflected in the levels of dispersion. However, without further specification of common information and especially how much convergence has occurred, simply stating that levels of dispersion reflect common uncertainty seems somewhat incomplete.

The main take away is that the Barron et al. (2009) notion that the levels of disagreement reflect the levels of uncertainty is indeed correct, but that the uncertainty that is reflected is the mixture of common and private uncertainty. Furthermore, regardless of whether "uncertainty" refers to the mixed measure as in the previous statement, or the common source of uncertainty, the uncertainty reflected in the levels of dispersion is specific to a chosen point in time, and cannot remain constant if earnings are released to the market.

[6] see section 2.9.

4.4 Sample description

The initial sample consists of all US exchange listed companies from 1995-2010, dead and alive, whose main exchange is listed as US. This does not include American depositary receipts or pink sheets. A company is determined to be a US company if Datastream lists the exchange on which the company is listed as "United States" and currency as "United States Dollar" and that the company's ISO code is US. A further screen requires that the I/B/E/S country code is also US. (This removes 55 companies). Companies are furthermore required to have data available in Worldscope and I/B/E/S.

The above requirements yield an initial sample of 8633 companies. As a result, the total starting sample consists of 138128 individual firm-year observations.

All company data is retrieved from Worldscope (through Datastream) and all estimate data is obtained from I/B/E/S. However, some additional data that is directly related to analysts, is also taken from I/B/E/S. The most important example is earnings per share, which is taken from I/B/E/S, and constitute the so called "street" earnings, which do not match the GAAP earnings reported in e.g. Worldscope. The reason for using "street" earnings instead of GAAP earnings is that this puts the estimates and the realized values on the same page. (Consequently, the standard deviation of realized earnings is also based on the "street" measure). The number of analysts is also taken from I/B/E/S.

Since dead equities are included, the final estimation sample is reduced from the starting sample. In addition, data is missing for many firm-year observations, and as some constructed measures require many observations per firm, which further reduces the working sample size.

The initial sample (panel) is unbalanced. Since the main empirical variable of interest is forecast dispersion, the availability of data on forecast dispersion dictates the maximum number of observations that are available. The number of defined firm-year observations for forecast dispersion is 46125.

4.5 Construction of variables

The study is concerned with analyzing increases in disagreement between analysts. Disagreement comes out directly from theory, in the form of V or V', the variance of agents' expectations (forecasts).

The basis for the calculation of disagreement between analysts is the standard deviation of (annual) EPS estimates, obtained from I/B/E/S. This is per firm year.

The standard deviation of EPS estimates needs to be scaled properly, in order to make comparisons across firms possible. Different authors use different deflators for the standard deviation of forecasts. E.g. Duru and Reeb (2002), in their study of international diversification and forecasting accuracy, and Guntay and Hackbarth (2010), in their study of forecast dispersion and corporate bond spreads, both use price as a deflator. Here however, the choice of deflator needs to be specified beyond simply making forecast dispersion comparable across firms. Since the interest lies in the size of perceived uncertainty by analysts in relation to commonly observed uncertainty (earnings volatility), both variables need to be specified in terms of the coefficient of variation. This makes it possible to compare the spread, or dispersion of two distributions directly, since the coefficient of variation is a normalized measure. Consequently, deflating the standard deviation of EPS estimates with the mean of the estimates (for the same firm year), yields the desired coefficient of variation metric.

Diether et al. (2002), in their widely cited paper on the relation between forecast dispersion and (negative) stock returns, use the exact same measure for forecast dispersion as the one used in this study, that is the standard deviation of forecasts is scaled by the absolute value of the mean forecast. The distinction is that they use I/B/E/S detail data whereas this study uses I/B/E/S summary data.

The main concern with using the I/B/E/S summary data is perhaps that there are potential stale forecasts in the data set. This is also noted by Brown (1993). If there are stale forecasts and EPS undergoes a large change, a stale forecast can artificially inflate the standard deviation of forecasts. This would be especially problematic if this is due to a systematic effect, for example, it might be possible that analysts stop following firms that report losses. Then, loss firms could see a biased standard deviation to the upside. There are strong reasons to doubt that this is the case however. First, in the setup in this study, dispersion is measured *prior* to observing losses, thus making it unlikely that analysts systematically drop out before losses occur. Secondly, and more importantly, the final estimations of this study are performed on a sub-sample that consists of firms that have complete data for all years for the variables of interest, causing the firms to be on average larger and having more analysts estimates.

Even though there are some concerns and (potential) inconsistencies with using the

I/B/E/S summary data, these concerns are alleviated in that e.g. Diether et al. (2002) document that: "[T]he mean and standard deviation values calculated from the Detail History file data closely track the values in the Summary History file.". This view is affirmed by Yu (2011) who subsequently also uses I/B/E/S summary data.

Using the absolute value of the mean estimate as a deflator however, is potentially problematic, as this can lead to extreme observations when the variable (absolute value of the mean) is close to zero, as mentioned by Minton et al. (2002). To deal with these potential problems, the variables are Winsorized at the 1st and 99th percentile values, as in Minton et al. (2002).

The forecast dispersion metric is thus the standard deviation of forecasts around the mean forecast, deflated by the absolute value of mean forecast. This will be used as a one key dependent variable in the empirical estimations.

The standard deviation of EPS estimates and the mean EPS estimate used in the forecast dispersion metric, are measured on the 30th of June each year. Robustness tests are also performed by measuring the aforementioned on the 30th of April each year, but unless otherwise noted, forecast dispersion refers to the measure constructed from data on the 30th of June. The robustness tests using 30th of April data are unreported and available upon request.

Finally, in terms of the timing of events, the time periods in both the model and the data are on an annual level. To allow for information being incorporated into analysts' information sets already before (annual) earnings are announced, dispersion is measured *before* earnings occur. This allows for the possibility of analysts being aware of uncertainties relating to a firm that earnings data does not necessarily capture, uncertainties that subsequently manifest themselves in terms of realized negative earnings.

4.6 Uncertainty ratio

Instead of merely looking at sizes of the increases in dispersion (measuring potential under prediction 1), more can be said about the magnitude of dispersion in forecasts if the magnitude of dispersion can be conditioned on the uncertainty, or variance of the assumed distribution of common information.

The predictions for forecast dispersion arise endogenously from the model and are company specific. For each firm, as long as a proxy for the uncertainty in common information (the inverse of r in Eq. [18]) can be found, it is possible to put maximum bounds on the size of disagreement. Under the assumption of fixed signals, the maximum magnitude of disagreement depends on the variance of the commonly observed signal distribution and the number of updating events that have occurred.

The uncertainty of the assumed conditioning distribution (in variance terms σ^2 $(1/r)$) will be proxied by the standard deviation of historical earnings. In particular, the standard deviation of earnings will be measured recursively, for each firm, such that at each point maximum information is used to estimate σ, without using forward looking information. Firm specific uncertainty estimated for the maximum amount of available information will then constitute an estimate of σ in such a way that it is assumed that σ has been the conditioning distribution for agents throughout the measured historical period. While another possibility would be to estimate signal standard deviations as separate estimates at each period, such an estimation procedure would be based on too few data points and would bias results[7]. See section 4.14.1 and 4.17 for a thorough discussion on the implications of such an assumption.

In the simple setup of the model, each (Bayesian) agent constructs a forecast of earnings. The forecast is constructed from (all) the information that the agent has. Mathematically, the agent does this by observing (earnings) realizations from the distribution(s), the mean of which he/she is trying to estimate. All relevant information is thus summarized by the variance of the distribution in his information set, and the (sequence of) earnings realizations. Since the agent forecasts *earnings*, it is only natural to assume that distribution he/she uses is that of the earnings as well.[8] In keeping the empirical mapping as close to the theoretical quantities as possible, the mapping assumes that the agent only uses the variance of earnings in his assessment. This is again *not* an assumption in terms of theory. It is for this reason that the empirical mapping uses the simplifying assumption that the earnings variance, or uncertainty, is sufficient to describe the agents' information. (since this is what the theory says he/she does).

Naturally, there may exist other information that agents use in their information sets, apart from the history of earnings. However, to the extent that this additional

[7] While the main empirical estimations estimate standard deviations based on a comparatively small number of observations, Section 4.16 confirms that the standard deviations used in the empirical estimations lie close to values estimated for a much longer sample using data going back to 1978. Also, it is worth noting that the cross-sectional standard deviations used as the basis for the calculation of dispersion in forecasts in many studies uses much fewer cross-sectional observations than time series observations for earnings used here.

[8] Note that this is not an assumption in terms of theory, it becomes an assumption only in the parameterization of the model. The simple Bayesian model is a single parameter model.

unobserved information is common, and as long as the estimate of the uncertainty of the earnings distribution is correctly identified, additional common information will only act as to *reduce* the subjective variance of the parameter being estimated. Additional unobserved common information, to the extent that is informative on earnings beyond that of observed earnings information, will make the actual variance of common information that agents condition on, *smaller* than the observed variance of earnings that are used as a proxy for common uncertainty in the empirical tests. Accordingly, if agents have access to any additional common information, the strength of their beliefs (the subjective "confidence interval" for the parameter which constitutes the strength of the belief) will actually be narrower than the observed variance of earnings used as a proxy here[9]. Thus, using only observable common information as measured by the historical volatility in company earnings, which has the potential of underestimating the amount of information that agents actually have[10], gives an advantage to the null hypothesis which maintains the rational Bayesian learning model sustained in the literature.

Having identified $\sigma^2(1/r)$, proxied by historical firm level earnings variance, all ingredients for analyzing Eq. (18) are in place. Empirical uncertainty ratio measures can now be constructed. The uncertainty ratios scale measures of forecast dispersion with earnings volatility (that proxies for theoretical common uncertainty) and the uncertainty ratios directly proxy the theoretical construct $(\sqrt{V}/\sigma)$ that can be extracted from Eq. (18). The rationale for using the ratio $(\sqrt{V}/\sigma)$ is the following. Eq. (18) gives the maximum magnitude of dispersion in forecasts, as a function of the variance of the common conditioning distribution and the number of signal realizations. Using Eq. (18) , where V_{max} is shorthand for $var_{max}(u_{ni})$, $V_{max} = 1/4nr$ $\Leftrightarrow \sqrt{V}_{max} = \sqrt{1/4nr} \Leftrightarrow \sqrt{V}_{max}/\sqrt{1/r} = \sqrt{1/4n}$. Since $(1/r) = \sigma^2$ it follows that $\sqrt{V}_{max}/\sigma = \sqrt{1/4n}$, or in the notation that will be used $(\sqrt{V}/\sigma)_{max} = \sqrt{1/4n}$. The left-hand side of the final identity is the maximum magnitude of dispersion in forecasts scaled by earnings volatility, and turns out to be a simple number that is a function of n, the number of earnings realizations. Given the number of earnings realizations, the number of observed annual earnings announcements in the data, one can easily find the maximum value for $(\sqrt{V}/\sigma)$. Through constructing an empirical measure which directly maps to $(\sqrt{V}/\sigma)$, the performance of the model can be judged by analyzing whether $(\sqrt{V}/\sigma)$ in the data conforms to the value implied

[9] Another way to see this is to think of other unobservable common information as additional draws, alternatively viewing the series as being "longer" than what can be observed, i.e. the measured n understates true n.

[10] As per the previous footnote, if agents have received more common draws than what can be observed, their subjective uncertainty would be smaller (alternatively again one can view n as being larger) than what could be estimated from the commonly observed information together with observed n.

by n. The empirical estimations will assess the performance of the model by testing whether the uncertainty ratio $(\sqrt{V}/\sigma)$ stays in line with the *maximum* magnitude of the ratio, given by n.

To provide some intuition for the sizes of maximum disagreement implied by the model, start by considering $n = 1$. The case where $n = 1$ gives the maximum fraction of disagreement in relation to the variance of a commonly observed informative distribution, when the common distribution is thought to be the Bayesian prior and updating is performed over one private signal realization for each agent. This case gives the absolute maximum fraction of disagreement that can occur, and is *not equal* to the variance of common information, rather it is 1/4 th of the variance of the common uncertainty $(1/[4 \times 1])$, or 1/2 of the standard deviation $(\sqrt{1/[4 \times 1]})$ of the common uncertainty. As n increases, that is as more (common) information is incorporated into the agents' information sets, the maximum amount of dispersion declines, resulting in a decline in the maximum value for $(\sqrt{V}/\sigma)$. For example, for a period corresponding to $n = 10$, where it is implied that common information has arrived from the same (fixed) distribution each period, the maximum for the uncertainty ratio $(\sqrt{V}/\sigma)$ is 1/40 in variance terms or $\sqrt{1/40}$ in standard deviation terms. In the latter case, using standard deviations, the maximum value for $(\sqrt{V}/\sigma)$ where $n = 10$ is thus approximately equal to 0.158. This means that a for a company with $n = 10$, the maximum that the ratio $(\sqrt{V}/\sigma)$ can attain is 0.158, regardless of values for private information. Thus, even if the starting point for the company at hand would have been a very low amount (low precision, high variance) of private information, implying that there was sizeable potential to raise dispersion from previous levels, the maximum amount dispersion in relation to common uncertainty could rise to, is 0.158

The study considers two different parameterizations of the theoretical relation $(\sqrt{V}/\sigma)$.

The first parameterization of dispersion in forecasts to common uncertainty is performed on a coefficient of variation basis. This variable is labelled Disptosigma, referring to "dispersion to sigma", where sigma, σ is the standard deviation of the common information distribution. This variable is constructed by dividing forecast dispersion (coefficient of variation) with the corresponding coefficient of variation of earnings for each firm.

The coefficient of variation for earnings is calculated by first calculating the standard deviation of EPS, constructed on a recursive (growing) basis for each firm. Note that this is the "street" EPS measure, obtained directly from I/B/E/S. If a firm has a minimum of two (consecutive) year EPS observations, the first recursive standard deviation estimate is constructed, and is used as an estimate for the volatility of earnings, corresponding to the latter year. If there are n observations, the recursive estimate is based on 1...n (trailing observations). The recursive standard deviation thus uses the full available sample information, up to the current point. Missing observations are skipped and not interpolated. A similar recursive estimation is performed for the mean of the earnings series, for each firm (and year).

The coefficient of variation of earnings is calculated at each step, using the total sample information up to the current point, by dividing the standard deviation estimate with the absolute value of the mean. Having constructed the coefficient of variation of EPS for each firm year, Disptosigma is constructed by dividing each firm year observation of forecast dispersion with the coefficient of variation of EPS (lagged). Naturally, this estimate loses one (year) observation for each firm, reducing the working sample.

Disptosigma is constructed by dividing each firm year observation of forecast dispersion with the lagged value of coefficient of variation of EPS. Since forecasts are measured on the 30th of June each year, e.g. 1997, and the realized EPS for the fiscal year 1997 is not known until early 1998 (all realized EPS values in the data file are already lagged such that realized values for 1997, which are reported in 1998, are shifted to occur at the time identifier 1997), the full year results for 1997 are obviously not available to analysts in June of 1997. Thus, the information sets of analysts in 1997 (recall that annual observations are used), only contain the full year realized results of 1996 and earlier. Thus, the dispersion of forecasts measured in June of 1997, is compared to the information that analysts have available at the time, namely the full year results of 1996. Therefore the variable Disptosigma uses current dispersion scaled by lagged earnings volatility (σ).

The theoretical restrictions map one to one with standard deviations (variances to be completely rigorous). Proxying standard deviation with the coefficient of variation has an advantage in terms of comparability, and is the only way to construct direct tests for measuring direct responses of dispersion, but it is not necessarily a direct proxy for theoretical values. When using measures based on the coefficient of variation, it is assumed that theoretical restrictions on standard deviations carry over to coefficient of variation measures. This is theoretically well motivated since in the Bayesian learning environment, both forecasts and realizations center on the true mean of the process, θ. On empirical grounds however, this is slightly more complicated. Since the coefficient of variation of earnings is based on the recursive standard deviation scaled by the recursive mean, the deflator for the coefficient of variation for historical earnings is the (total) time series mean of the earnings series.

The dispersion measure on the other hand, is based on the cross-section of earnings forecasts at one point in time, constructed by deflating the standard deviation of forecasts with the mean forecast. To the extent that analysts in reality are not pure Bayesian agents, constructing unconditional mean forecasts, these two different deflators (mean estimates) may not coincide.

The study therefore constructs a more robust measure using standard deviations directly, thereby ensuring that identification is correct. This robust measure, DTSdirect ($\sqrt{V}/\sigma$), is analogous to Disptosigma, except that it uses standard deviations directly. The numerator is the raw standard deviation of forecasts for a firm for a particular year, deflated by the (lagged) recursive standard deviation for past earnings. The recursive standard deviation for past earnings is calculated exactly as above in the construction of the variable Disptosigma.

Both variables, Disptosigma and DTSdirect ($\sqrt{V}/\sigma$), are subject to the same restrictions. The direct theoretical restrictions $(\sqrt{V}/\sigma)_{max} = \sqrt{1/4n}$ however have a one to one correspondence to the variable DTSdirect ($\sqrt{V}/\sigma$). In the variable DTSdirect ($\sqrt{V}/\sigma$), the numerator $\sqrt{V}$ exactly equals the standard deviation of forecasts, and the denominator σ is the standard deviation of earnings. Thus, to the extent that σ corresponds to the estimated volatility of past earnings, the uncertainty ratio ($\sqrt{V}/\sigma$) is exactly the variable DTSdirect ($\sqrt{V}/\sigma$).

4.7 Control variables

The general tests of the model (H1) are performed through evaluating the maximum bound $(\sqrt{V}/\sigma)_{max}$. In terms of H1, if the maximum theoretical bounds for forecast dispersion are violated, it does not in effect matter which variable (or an aggregate, cumulative effect of all variables) is the culprit. Any variable (or combination of) that would cause an exceedance of the theoretical bounds, would itself be a novel result. This is because the main theoretical prediction that is being tested, is the maximum allowed dispersion allowed by the model - the theory is silent on what could cause such effects. (Strictly speaking the theory does not allow for such effects, the only allowable cause is an effect that increases private information).

The empirical tests of H2 evaluate the performance of the bound $(\sqrt{V}/\sigma)_{max}$ specifically around negative earnings. Negative earnings are the key hypothesized trigger for increases in dispersion and consequently the effect must be separated from other effects.

Lang and Lundholm[11] (1996) and Duru and Reeb (2002) find that forecast dispersion is negatively related to analysts' forecast accuracy. The hypothesis (H2) is that negative earnings are related to uncertainty and dispersion. Negative earnings have been found to carry with them an unexpected component, and Hwang, Jan and Basu (1996) (in Duru and Reeb, 2002), find that analysts' forecasts of losses are, on average, less accurate than forecasts of profits. Consequently, accuracy is controlled for in an attempt to disentangle the potentially different effects on dispersion that accuracy and negative earnings have, especially as the two are related to each other. In summation, the study controls for accuracy in the empirical estimations, due to the fact that at least heuristically forecast dispersion, accuracy and losses, all tend to be related to uncertainty about future profits. While the term used is accuracy, it can equally well be interpreted as either earnings surprise or forecast error.

The accuracy variable is constructed in the following way. Using the year 1997 as an example, forecast accuracy (or surprise) is measured as the (absolute value of the) difference between realized earnings in 1997 (that become known in early 1998), minus the consensus estimate in 1997 (June), scaled by stock price in 1997 (June) multiplied by 100. The timing of measures, as well as scaling by price is analogous to the construction of the accuracy variable in eg. Duru and Reeb (2002). In contrast, Duru and Reeb (2002) and Lang and Lundholm (1996) both multiply the accuracy measure[12] with (-1). In this study, the accuracy measure is *not* multiplied by (-1) and thus the accuracy measure in this study has a direct interpretation in terms of surprise or forecast error, the higher the value on the accuracy measure, the greater is the surprise in earnings as compared to consensus[13]. Earnings are earnings per share.

Since this study is interested in the size of the regression coefficients, specifically the size of the effect on the relation of dispersion to common variance, the accuracy variable is additionally scaled by 100 so that sizes are easier to compare. The raw accuracy variable (without scaling by 100), has a mean of 0.023 and a median of 0.0066, with a 95 percentile value of 0.0997 and a maximum value of 0.351. Interpreting a regression coefficient on the raw accuracy variable would entail the use of additional consideration, since accuracy could in the maximum case, only move from 0 to 0.351. So a coefficient on accuracy of e.g. 0.2 would imply that a 1 "unit" move in the independent variable (in this case accuracy), would correspond to a 0.2 move in the dependent variable. However, such a move is not possible since

[11] Based on significant correlation.

[12] Note that in the Lang and Lundholm (1996) study, all variables in the accuracy metric are indexed by (t)

[13] Some caution might thus be warranted in comparing signs of the accuracy relation to studies by other authors.

no such values exist, so in other words the absolute maximum move in the dependent variable in this example would be 0.0702, where this would be achieved only between the two extreme end values of the accuracy distribution. Consequently, to facilitate easier comparisons in the size of coefficients, the accuracy metric is multiplied by 100.

Lang and Lundholm (1996) find that larger changes in earnings are associated with less accurate forecasts. While this study is not directly focused on the forecast error (accuracy) of analysts' forecasts, changes in earnings are controlled for due to their general association with the forecasting environment and the accuracy of forecasts, the latter which in turn is associated with negative earnings environments (Hwang et al.,1996, in Duru and Reeb, 2002). Change in earnings is measured simply as realized earnings for the current year minus realized earnings for the previous year, scaled by realized earnings for the previous year, multiplied by 100. Earnings are earnings per share.

Other controls used are the natural logarithm of market capitalization to control for size, and Following, controlling the number of estimates or analysts following the firm. Industry effects are controlled for by industry dummies, based on ICB industry classification codes[14].

Finally, it is important to keep the length of the forecasting window fixed. Some authors such as Duru and Reeb (2002) approach the issue by controlling for forecast horizon by using a measure of forecast months. In the context of this study, this approach is deemed insufficient due to a suspicion that there are possible mismatches between realized earnings and forecasts for those earnings. Furthermore, as more data becomes available throughout a year, e.g. 3/4 of annual earnings are known in Q4, there is a concern that this affects the estimates on dispersion. To keep the forecast horizon fixed, to ensure forecasts match earnings, and to make sure that it can be reasonably deduced what (common) information analysts have available to them, this study instead opts to use only firms whose fiscal year end is December. Consequently, the length of the estimate window is fixed to six months.

4.8 Descriptive statistics

Table 1 reports descriptive statistics for the main variables used. Panel A reports the "raw" variables used in the construction[15] of the main dependent uncertainty

[14] The results in the tables use the term "sector" for the industry dummies.

[15] The variables used depend on specification. See the section "Construction of variables".

variables: the standard deviation of the EPS forecasts, the mean (consensus) forecast, and the rolling (recursive) standard deviation of realized EPS. While Panel A in Table 1 reports descriptive statistics for earnings per share, the measure used in the uncertainty measure, in this case Disptosigma, is the absolute value of EPS (not shown). The variables *as reported* in Panel A are not Winsorized. This is due to the fact that Winsorizations are performed only *after* the uncertainty variables are constructed.

Table 1: Descriptive statistics. Observations reported in Table 1 have to have corresponding observations on earnings per share (EPS), as well as December fiscal year ends. **Panel A**: Sd forecasts is the raw standard deviation of EPS forecasts, on a firm year basis. Mean forecast is the average of forecasts of analysts for a given firm year. All forecasts are measured June 30th each year. Recursive sd EPS is the standard deviation of realized EPS firm by firm, estimated recursively (growing estimation window). EPS is the earnings per share for each available firm year. **Panel B**: All variables in Panel B are winsorized at the 1st and 99th percentile, except for Loss and Following. Forecast dispersion is the firm year standard deviation of forecasts, scaled by the (absolute value of the) corresponding mean forecast. Disptosigma is forecast dispersion scaled by the recursively estimated coefficient of variation of firm historical EPS. The recursively estimated coefficent of variation of firm historical EPS is estimated from (full year) information available up to the current point, effectively implying a half year lag for the denominator. DTSdirect ($\sqrt{V}/\sigma$) is the firm year standard deviation of forecasts, scaled by the recursively estimated standard deviation of earnings. The recursively estimated standard deviation of earnings (EPS) is estimated from (full year) information available up to the current point, effectively implying a half year lag for the denominator. Accuracy is (the absolute value of) the difference between the realized earnings for a given firm year, and the consensus estimated earnings June 30th of the same year, scaled by year end price. Change in earnings is defined as the simple net return of firm EPS, multiplied by 100. Size is the natural logarithm of market capitalization. Following is the number of analysts covering a firm at the time of forecasting. Loss is a binary indicator variable taking the value 1 if EPS is less than zero.

Variable	**mean**	**sd**	**min**	**max**	**skew.**	**kurt.**	**N**
Panel (A)							
Sd forecasts	0.26	7.69	0.00	1003.80	99.19	11268.02	31166
Mean forecast	1.95	90.06	-8128.00	5825.16	9.41	4227.87	35835
Recursive sd EPS	2.54	76.63	0.00	6169.76	64.45	4530.64	34756
EPS	1.18	107.15	-12736.00	6774.00	-25.97	7037.32	40282
Panel (B)							
Forecast dispersion	0.15	0.35	0.00	2.50	4.93	30.03	31118
Disptosigma	0.29	0.56	0.00	3.67	4.07	21.51	23686
DTSdirect ($\sqrt{V}/\sigma$)	0.25	0.38	0.00	2.54	3.85	20.62	25488
Accuracy	0.02	0.05	0.00	0.35	4.30	23.83	35348
Change in earnings	-1.34	171.34	-860.71	812.50	-0.31	15.95	33570
Size	13.33	1.75	8.41	17.69	0.24	2.78	39171
Following	7.18	6.39	1.00	47.00	1.61	5.86	35835
Loss	0.20	0.40	0.00	1.00	1.47	3.16	40282

The measures in Panel A however do not necessarily provide much intuition in themselves. E.g. comparing the standard deviation of forecasts across firms is not particularly meaningful since firm-specific EPS varies heavily, and furthermore the unwinsorized measures are heavily affected by the extreme EPS measures that are represented by just a handful of firms.

Panel B in Table 1 reports descriptive statistics for both the three uncertainty variables, forecast dispersion, Disptosigma, and DTSdirect ($\sqrt{V}/\sigma$), as well as for the control variables, accuracy[16], change in earnings[17], size, and following.

All variables[18] in Panel B are Winsorized at the 1st and 99th percentile. Notice that while the variables in Panel B are still not necessarily behaving "nicely" in terms of normality, the most important measure used in this study, DTSdirect ($\sqrt{V}/\sigma$), behaves better than forecast dispersion. Recall that forecast dispersion was defined as the standard deviation of EPS forecasts scaled by the consensus forecast, and is the exact measure used in e.g. Diether et al. (2002). While other authors, e.g. Duru and Reeb (2002) use price as a deflator for the EPS forecasts, constructing a dispersion measure analogously[19] using the data in this study still yields a skewness of 4.75 and a kurtosis of 28.32. Thus, the accepted approach in the literature in terms of distributional characteristics, gives no reason to doubt inferences drawn from the DTSdirect ($\sqrt{V}/\sigma$) measure.

The non-normality of the variables is of course driven by the fact that that the variables involving standard deviations (as well as accuracy) are truncated at zero.

The data in Table 1 are subject to a few conditions that has shrunken the data from its initial size. Apart from the requirement of companies having December fiscal year ends, the data in Table 1 are matched based on the availability of earnings data. Note however that the empirical estimations also require forecast data to be available and thus the descriptive statistics may not be fully representative of samples used in estimations. (Also the subsequent estimations will add additional requirements that will further shrink the data set).

[16] The accuracy measure here refers to the "raw" measure in the sense that it is before being multiplied by 100.

[17] It is slightly interesting that the average change in earnings (earnings growth) has a mean that is negative. Forcing the companies to have earnings observations for all years throughout the sample however, as is the case with main estimations performed in the following sections, yields a mean for the variable equal to 4.34.

[18] Except Loss and Following.

[19] That is deflating the standard deviation of forecasts by price.

4.9 Empirical estimations

Figure 12 illustrates how forecast dispersion behaves in the data on aggregate. Due to the lack of scaling, and the fact that data is aggregated across all firms in the sample, the dispersion measure is based on the coefficient of variation.

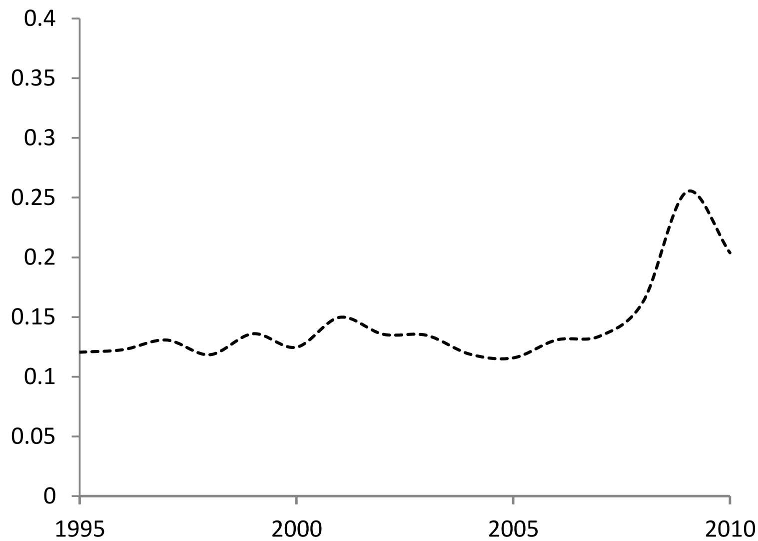

Figure 12: Bottom up aggregated forecast dispersion for US exchange listed companies for the years 1995-2010. Aggregation is performed through equal weighting. Forecast dispersion is the firm year standard deviation of forecasts, scaled by the (absolute value of the) corresponding mean forecast. See Table 1 for a further description.

The magnitude of the dispersion in forecasts, as well as the overall behavior, is similar[20] to that of Figure 1 in Park (2005), which is a predecessor to Yu (2011), in showing that disagreement leads to lower portfolio returns. On the other hand, Figure 1 in Yu (2011), which also aims at illustrating the dynamics of forecast dispersion, looks markedly different from Figure 12 and the corresponding figure in Park (2005). The reasons for this can be found in that firstly, Yu (2011) uses the dispersion in long term growth forecasts, *not* EPS estimates. Secondly, Yu (2011) aggregates dispersion through *value weighting* instead of equal weighting as in this study and in Park (2005).

[20]Park (2005) however scales the standard deviation of forecasts with actual earnings. More importantly, earnings used in Park (2005) are top down aggregate forecasts for the S&P 500. Also, Park (2005) utilizes a weighting scheme for the standard deviation in forecasts in an attempt to fix the forecasting horizon while taking advantage of monthly data.

Next, Figure 13 illustrates how dispersion in the aggregate behaves in relation to the maximum bounds derived from theory. This provides an initial visual examination of H1. In Figure 13, the grey line is $(\sqrt{V}/\sigma)_{max}$, that is the maximum magnitude forecast dispersion can have in relation to common uncertainty, and here constitutes the maximum bound. The dotted line is $(\sqrt{V}/\sigma)$, that is the variable DTSdirect, and illustrates what is implied by the data.

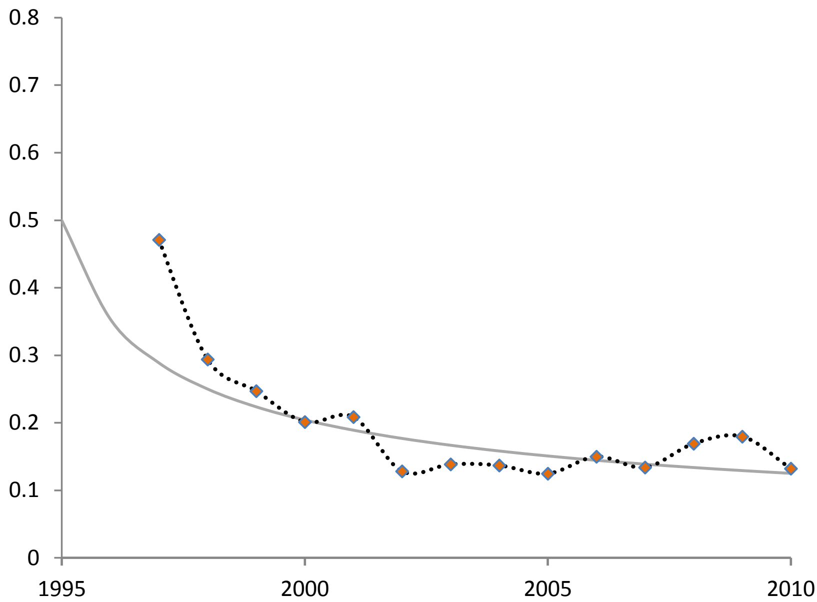

Figure 13: Comparison of the maximum bound for dispersion, $(\sqrt{V}/\sigma)_{max}$, and the actual value for DTSdirect $(\sqrt{V}/\sigma)$. The solid grey line is the maximum bound $(\sqrt{V}/\sigma)_{max}$, which constitutes the maximum value for DTSdirect $(\sqrt{V}/\sigma)$ for each period. The dotted line with red squares shows the actual value for DTSdirect $(\sqrt{V}/\sigma)$. See Table 1 for a further description of the variable DTSdirect $(\sqrt{V}/\sigma)$. Firms must have full earnings histories to be included in the data that DTSdirect $(\sqrt{V}/\sigma)$ is based on in the figure.

While the data for the initial years (up to 2000) are difficult to draw conclusions from[21], something interesting occurs in 2008 and 2009; forecast dispersion in the data, on aggregate, exceeds its maximum theoretical value. To examine whether the exeedance of the implied maximum values illustrated in Figure 13 is significant, empirical estimations of H1 are carried out.

The equation below formally tests H1, where DTSdirect $(\sqrt{V}/\sigma)$ is regressed on a constant. Results are reported in Table 2.

[21]For the initial years there is not enough data to get meaningful estimates for the realized standard deviation of earnings.

$$DTSdirect = \alpha + (\beta_1 \cdot Controls) + \epsilon \tag{21}$$

Specification (1) in Table 2 is for the full sample 1995-2010 and shows that the average value for DTSdirect is 0.256. As can be deduced from Figure 13 however, this is somewhat problematic since in the beginning of the sample period DTSdirect is poorly estimated (the pre 2000 values for DTSdirect are in all likelihood driven by biased standard deviation estimates for realized earnings). Secondly, recalling that the threshold level for DTSdirect, $(\sqrt{V}/\sigma)_{max}$, is guided by n, it is difficult to compare the value for DTSdirect in specification (1) against the maximum threshold level, since the sample spans a large period[22].

Table 2: Table 2 displays the results from pooled OLS regressions of the variable DTSdirect ($\sqrt{V}/\sigma$) on an intercept and controls (specification 4 only), as described below. Specification/column (1) is a univariate regression of the variable DTSdirect ($\sqrt{V}/\sigma$) on an intercept for the years 1995-2010. Specification (2) repeats the estimation in (1), but on a sample that enforces the availability of a full earnings (1995-2010) history for a firm to be included in the sample. Specification (3) repeats (2), but is estimated on a subsample form 2006-2010, enforcing the full availability of earnings observations as above. Specification (4) again repeats (3), with the addition of controlling for sector effects based on ICB industry classification codes. The full availability of earnings observations are enforced as above. Reported standard errors are White (1980) standard errors, robust to heteroskedasticity, and are found in parentheses. All variable definitions are described in Table 1.

	Dependent Variable: DTSdirect			
	(1)	(2)	(3)	(4)
Constant	0.256***	0.214***	0.170***	0.227***
	(0.00231)	(0.00304)	(0.00309)	(0.0160)
Observations	27,603	9,277	3,338	3,338
R-squared	0.000	0.000	0.000	0.053

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Consequently, in order to keep better track of n, and to ensure that sufficient converge has occurred, that is enough updates have been carried out, specification (3)[23] is estimated for a sub-sample from 2006-2010. The requirement for inclusion in the sub-sample, is that firms have complete earnings and estimate histories throughout

[22]The only threshold that could be used with certainty is 0.5.

[23]specification (2) is simply as specification (1), that is on the full sample but with the added requirement of full sample earnings histories for each individual firm.

the whole sample[24] (1995-2010). While this induces a large-cap bias, it ensures that results are not driven by small firms and guarantees that the minimum amount of updates that must be assumed to have occurred are known with certainty. Thus, a firm present in the sub-sample, in e.g. 2006, will have had earnings from 1995-2005 (Estimates are made on June 30 th of 2006. Consequently, analysts' information sets contain full year earnings up to, and including, 2005).

Agents should then, according to theory, have had *at least* [25] 10 annual observations of commonly observable information, causing them to have updated over common information 10 times. If agents are forecasting earnings as coming from a fixed distribution, the maximum fraction of forecast dispersion in relation to common "uncertainty", ($\sqrt{V}/\sigma$), (for any values on private information) that the learning model allows for, is 0.158. Thus, in 2006 (and beyond), the maximum theoretical value that ($\sqrt{V}/\sigma$) can take is 0.158, given that the underlying learning model is true.

Naturally, as with all threshold values used in this study, they are based on *observable* n, which is the number of periods from the beginning of the sample up to the point of measurement. This is the reason for enforcing the full availability of firms' earnings histories described above. Conversely, the observed values of the variable DTSdirect ($\sqrt{V}/\sigma$) are however not directly dependent on n. Thus, if one would use a longer history of earnings, the value of the variable DTSdirect ($\sqrt{V}/\sigma$) does not change[26], while the threshold value $(\sqrt{V}/\sigma)_{max}$ changes. Given that the value for DTSdirect ($\sqrt{V}/\sigma$) remains constant (subject to the caveat in the previous footnote), one would effectively use a lower threshold value for judging exceedance of the model bounds in the regression results. In terms of visualization, this would amount to the curve for the threshold value, $(\sqrt{V}/\sigma)_{max}$, in Figure 13 being adjusted downward. See also the final paragraphs of section 4.16 for a discussion of the probable magnitude of true n.

In specification 3, the intercept is equal to 0.170 and highly significant. The significance level of the test however is for the constant being different from *zero*. In order to be able to test the restrictions of the theoretical model, the test must be

[24]Notice that Figures 13 and 16 are also subject to this constraint.

[25]Since the earnings for the year 1995 are included, in terms of the model one could use a threshold value of $n = 11$. However, using what is effectively $n - 1$ in reality allows for the prior in the model to be set very freely. In, fact the only restriction on the prior is that it is normally distributed. While this is not explicitly made use of in the model (in order to keep the formulas more compact), this essentially just relaxes the assumption of setting $h = r$ in the first period. Footnote 11 in section 2.8 also discusses the issue.

[26]The numerator remains the same but since the denominator consists of the estimated standard deviation, a longer sample could affect the value.

for the constant being larger than the threshold value $(\sqrt{V}/\sigma)_{max}$, which for $n = 10$ is equal to 0.158.

A post estimation Wald test on the restriction that the intercept is equal to 0.158 confirms that this is indeed the case. The F-statistic equals 14.06 (1, 3337), and the corresponding Prob $>$ F $=$ 0.0002. This validates H1 and confirms that the dispersion in forecasts is too large for the fixed Bayesian learning model to hold.

Specification (4) repeats the above estimation but includes sector dummies based on ICB industry classification codes. The findings here corroborate those of specification (3) in that H1 is accepted. Repeating the Wald test after running specification (4) confirms the result. The F-statistic equals 18.89 (1, 3328) and the corresponding Prob $>$ F $=$ 0.0000.

These findings are very interesting. The implication is that the dispersion in forecasts cannot simply be explained by private information, since the threshold value $(\sqrt{V}/\sigma)_{max}$ is the maximum amount of dispersion in forecasts (in relation to earnings variance) that the fixed Bayesian learning model allows for. Recall that the maximum was derived for a signal, s'_{n^*max}, defined in such a way that it is the amount of private information that by definition maximizes dispersion. Consequently, there can exist no private signal that can cause an exeedance of this bound. Yet the bound is exceeded in the data, as the results in Table 2 confirm. What this implies is that there must be other explanations that are driving the results.

4.10 Illustration of the negative earnings - dispersion relation.

Before proceeding with testing H2, the following example company provides some intuition on the dynamics of uncertainty around pending losses, and to what extent theoretical maximum bounds are violated in the data. This is depicted in Figure 14, which also shows how negative earnings seem to be a large contributor to increases in dispersion. Note that this is only an example which is to serve as an illustration.

In Figure 14 it is clear that DTSdirect $(\sqrt{V}/\sigma)$, readily exceeds its maximum value[27], which is approximately 0.158, since DTSdirect takes the value 0.87 in 2008. The large rise in the standard deviation of forecasts, the driver behind the rise in DTSdi-

[27]notice that DTSdirect (and Disptosigma) measures exceedance w.r.t. to the latest estimate of the (recursive) standard deviation. See section 4.17 for a thorough analysis of this.

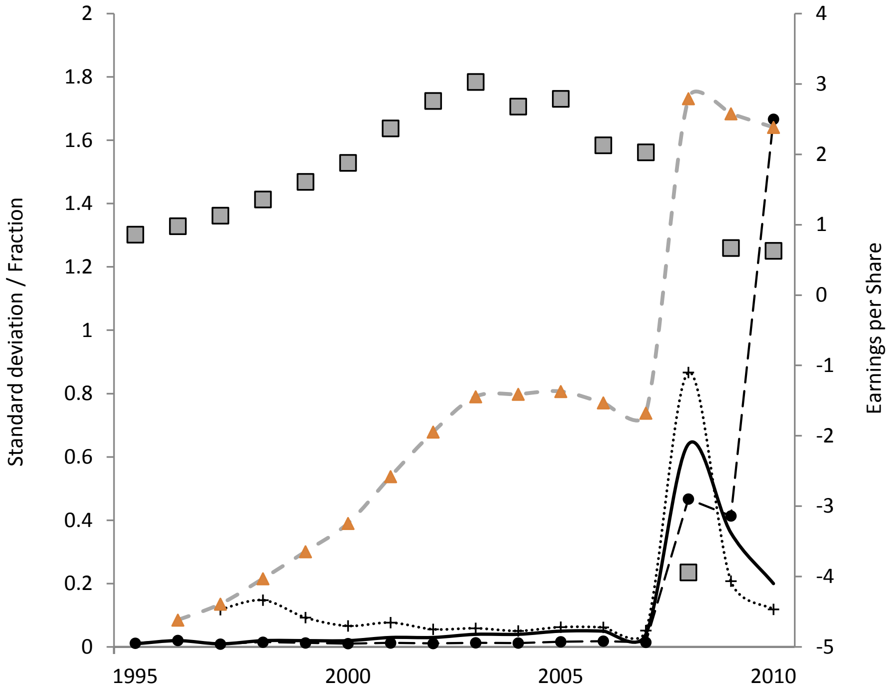

Figure 14: Fifth Third Bancorp. The grey boxes are earnings per share (right axis). The black dashed line with circles shows forecast dispersion, defined on a coefficient of variation basis. The grey dashed line (triangles) shows the estimated volatility (standard deviation) of earnings (recursive estimate). The solid line is the standard deviation of forecasts. The dashed line (plusses) shows the variable DTSdirect, which is the ratio of forecast standard deviation to earnings volatility, the maximum theoretical value of which is less than 0.158 (for the year 2007 and onwards). Here, the value of DTSdirect is 0.87, implying that DTSdirect exceeds its theoretical bound by a factor of over 5X.

rect, can also be readily observed. Interestingly, forecast dispersion, as it is usually measured, makes its largest increase in 2010, even though the standard deviation of forecasts has already decreased significantly at this point. The reason for this is that here forecast dispersion is defined on a coefficient of variation basis, and as the mean forecast for 2010 is 0.12, this measure becomes artificially inflated due to the low absolute value of the mean forecast (the standard deviation of forecasts in 2010 is 0.2, implying that the forecast dispersion measure based on the coefficient of variation becomes 1.67). This shows the potential caveats of measuring forecast dispersion as a coefficient of variation. Note that this caveat applies well beyond this study. Simultaneously however, this shows the strength of the robust DTSdirect measure.

Figure 15 provides some additional intuition for the magnitude of the standard

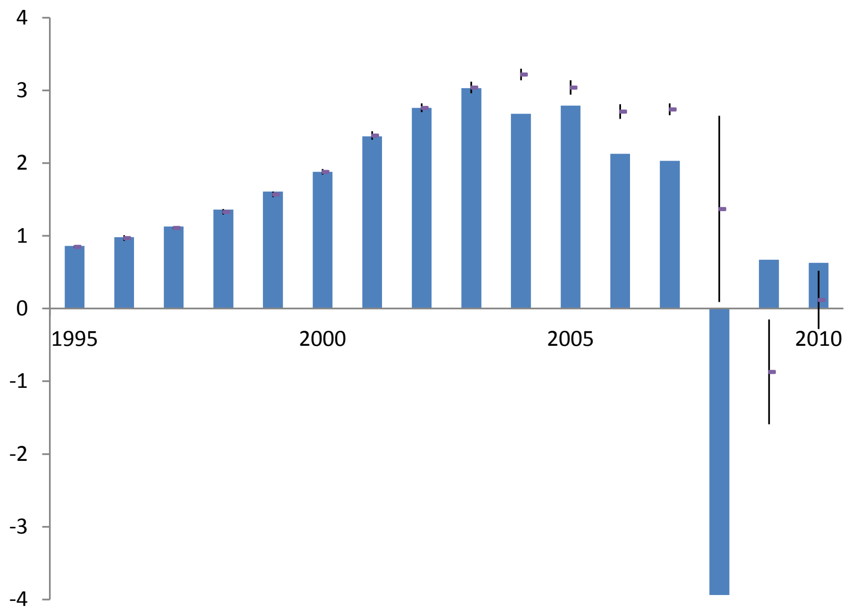

Figure 15: Fifth Third Bancorp. In the figure, the thick (blue) candles represent realized earnings. The vertical bars in turn represent the 2 standard deviation interval around the mean forecast, which in turn is represented by the dot on the right-hand side of each bar. The 2 standard deviation interval is based on the (one) standard deviation of earnings forecasts.

deviation of forecasts in the example of Fifth Third Bancorp. In the figure, the thick (blue) candles represent realized earnings. The vertical bars in turn represent the 2 standard deviation interval around the mean forecast, which in turn is represented by the dot on the right-hand side of each bar. The 2 standard deviation interval is based on the standard deviation of earnings forecasts.

In this illustration it is easier to directly compare the magnitude of the disagreement in forecasts, since it on the same level as the realized earnings. In the years up to 2008, the spread in forecasts is very small, whereas in 2008 the spread widens considerably. In fact, the standard deviation in forecasts increases $16X$ from the previous year.

Viewed from another angle: in order for the standard deviation in forecasts in 2008 to be compatible with the null of the Bayesian learning model, the standard deviation of the distribution over which agents updated in 1995, would have to have been over 4 (as the maximum standard deviation in 2008, in this case 0.64, can be at most 0.158 of the standard deviation of the common information 10 periods earlier, implying 0.64/0.158, which is > 4). In the figure, this would imply a *one* standard deviation interval in 1995 that encompasses virtually the whole vertical axis, and the corresponding two standard deviation interval, as in the figure, would extend far beyond the axes of the figure.

4.11 Empirical estimations of H2

The previous example demonstrated how large increases in dispersion tend to specifically occur around *negative earnings*, apart from the general evidence provided in the results in Table 2, that the magnitude of dispersion on aggregate exceeds its theoretically implied bounds. The combined evidence so far leads up to analyzing H2, namely that the driver behind large increases in dispersion are negative earnings.

For initial illustration, Table 3 shows the results of a mean-comparison test (with unequal variances), comparing the average values for the three different uncertainty measures, forecast dispersion, Disptosigma and DTSdirect, for years of negative earnings vs. positive earnings.

Table 3: Table 3 reports results of mean-comparison tests using the full sample of data, 1995-2010. The t-tests assess whether the uncertainty measures have the same mean for when a firm reports negative earnings vs. when a firm reports positive earnings. The results are for the variables forecast dispersion, Panel (A), Disptosigma, Panel (B), and DTSdirect ($\sqrt{V}/\sigma$) , Panel (C). The degree of freedom correction resulting from unequal variances is due to Satterthwaite (1946). All variable definitions are described in Table 1.

	Panel(A)	Panel(B)	Panel(C)
	Dispersion	Disptosigma	DTSdirect
Mean (no loss)	0.100	0.228	0.232
Diff. Mean (no loss) - Mean (loss)	-0.317***	-0.404***	-0.145***
t-stat	(-38.95)	(-27.46)	(-18.30)
N	33506	25607	27603

t statistics in parentheses
* p<0.05, ** p<0.01, *** p<0.001

Panel A in Table 3 reports the results for forecast dispersion, where on average the dispersion in forecasts for firm years associated with losses is 4 times[28] higher than the dispersion in forecasts for firms reporting positive earnings.

[28] Since the difference between Mean (no loss) and Mean (loss) is -0317, and Mean (loss) is equal to 0.100, it follows that Mean (loss) must equal 0.417.

Although the effect is sizeable, this alone is however not sufficient to make statements on the performance of H2, since there is always the possibility that dispersion increased from a large acquisition of private information, from previous levels of vague private information. This is exactly the difficulty with using prediction 1) from Eq. (19); without being able to control for previous levels of information, it is not possible to make statements about the magnitude of the increases in dispersion. The preliminary evidence on the relation between forecast dispersion and negative earnings however confirm the results in Adut et al. (2008).

In order to facilitate an anchoring of the increases in forecast dispersion associated with losses to their theoretical bounds given by the uncertainty of the commonly available historical information, Panel B in Table 3 reports the difference in means test on the disagreement (forecast dispersion) to common variance ratio, Disptosigma, between years with negative and positive earnings.

The outcome is large and highly significant and shows that while on average for non-loss firm years, the dispersion is 22.8 percent of the size of the cumulative past historical earnings variation, for firm years associated with losses, this fraction rises to 63.1 percent. To the extent that the measures (forecast dispersion and the coefficient of variation for past earnings) can be interpreted on the same basis as the restriction on standard deviations that come directly out of theory, the exceedance of the level of 50 percent implies a violation of the maximum bounds for disagreement, even for one single updating event[29]. However, some caution is needed in the interpretation of results using the Disptosigma measure, since it is based on the coefficient of variation.

Panel C in Table 3 reports the same test for the more robust variable DTSdirect, which behaves better. Here the difference between firm years with positive and negative earnings is smaller, and also the value for DTSdirect under losses is a more reasonable 0.377.

Notice however that using measures spanning the whole sample is subject to the same caveat as the results in Table 2; it is impossible to be sure how many updating events agents have performed. Also, there exists a possibility that losses are heavily concentrated in the early part of the sample. As such the only threshold one can be sure that applies is the one for one updating event, namely 0.5. While the Disptosigma measure in Panel (B) of Table 3 does exceed this threshold, the robust measure DTSdirect does not, and thus more refined tests are needed.

[29]The restriction comes out of Eq. (18) , where, if common uncertainty, without loss of generality, is normalized to 1, the maximum disagreement in variance terms is 1/4 th, yielding 1/2 in standard deviation terms.

Table 4 repeats the intuition of Table 3, through estimating the following regression specification(s):

$$ForecastDispersion = \alpha + \beta_1 \cdot loss + \beta_2 \cdot Controls + \epsilon \tag{22}$$

$$Disptosigma = \alpha + \beta_1 \cdot loss + \beta_2 \cdot Controls + \epsilon \tag{23}$$

$$DTSdirect = \alpha + \beta_1 \cdot loss + \beta_2 \cdot Controls + \epsilon \tag{24}$$

Table 4: Table 4 displays the results of pooled OLS regressions of dispersion/uncertainty variables on negative earnings and firm-level control variables 1995-2010. All estimations enforce the availability of full earnings (1995-2010) histories for firms to be included in the sample. In Panel (A) the dependent variable is (forecast) Dispersion, in Panel (B) Disptosigma, and in Panel (C), DTSdirect ($\sqrt{V}/\sigma$), respectively. All regressions include industry controls (ICB industry classification) and year controls. Reported standard errors are White (1980) standard errors, robust to heteroskedasticity, and are found in parentheses. All variable definitions are described in Table 1.

Independent Variable	Dependent Variable		
	Panel (A)	Panel (B)	Panel (C)
	Dispersion	Disptosigma	DTSdirect
Loss	**0.383*** **	**0.461*** **	**0.139*** **
	(0.0280)	**(0.0413)**	**(0.0187)**
Accuracy	0.00410**	0.00140	0.00458***
	(0.00164)	(0.00229)	(0.00117)
Change in earnings	-0.000179***	-0.000278***	5.48e-05**
	(3.44e-05)	(4.84e-05)	(2.44e-05)
Size	-0.0129***	-0.0145***	-0.00637**
	(0.00313)	(0.00506)	(0.00314)
Following	0.000833	0.000644	0.000198
	(0.000651)	(0.00105)	(0.000691)
Constant	0.369***	0.769***	0.308***
	(0.0514)	(0.0772)	(0.0534)
Year fixed effects	Yes	Yes	Yes
Sector fixed effects	Yes	Yes	Yes
Observations	9,742	8,932	9,121
R-squared	0.228	0.168	0.168

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

In Table 4, Equation (22) refers to Panel (A), Equation (23) refers to Panel (B), and Equation (24) refers to Panel (C). The estimations are carried out for the full sample (1995-2010). The reason for the seemingly low number of observations (compared to the initial sample) derives from enforcing the restriction that firms must have full earnings histories to be included in the sample.

The results in Table 4 are qualitatively similar to those of Table 3, except now various other variables related to forecast dispersion are controlled for. While the results are indeed stronger for all three measures, one can still not make definitive inference on the performance of the theoretical bound $(\sqrt{V}/\sigma)_{max}$ (Panel (A) and Panel (B)), since the estimation caveat for the standard deviation of earnings still applies.

4.12 Subsample estimations of H2

While the evidence in Table 3 and Table 4, combined with Figure 13, not only shows indications of forecast dispersion increasing with losses, but that the association potentially points at breaches of the theoretically allowed maximum bounds (H2), it is however not fine grained enough to resolve the issue. Additional partitioning of the data however, manages to resolve the remaining ambiguities. This is shown in Figure 16.

Figure 16 splits the data for DTSdirect measure used in Figure 13, between negative and positive earnings. The thick dotted line with green triangles represents DTSdirect $(\sqrt{V}/\sigma)$ for negative earnings, while the light dotted line with blue squares represents DTSdirect $(\sqrt{V}/\sigma)$ for positive earnings. The grey line is again the maximum boundary $(\sqrt{V}/\sigma)_{max}$. In Figure 16, it is evident that not only are negative earnings associated with much higher values for DTSdirect $(\sqrt{V}/\sigma)$, but these values for negative earnings heavily breach $(\sqrt{V}/\sigma)_{max}$. The corresponding figure for Disptosigma is similar, but more exaggerated, but the Disptosigma measure is left out due to the problems with the coefficient of variation measures. While the visual evidence is compelling, the significance of the effect seen in Figure 16, which simultaneously coincides with H2, needs to be assessed empirically.

As in the tests for H1, definitive answers can be provided by tests on a sub-sample for the later years in the sample. Again, the requirement for inclusion is that firms have full earnings histories within the whole period of 1995-2010. Testing on the sub-sample achieves two things: First, it allows for having more meaningful estimates of the standard deviation of earnings, and secondly allows for a smaller window under which to assess the threshold $(\sqrt{V}/\sigma)_{max}$, as it depends on n. In some ways, it would perhaps be more fitting to simply view the period 2006-2010 as the main sample on which estimations are performed, whereas 1995-2005 would constitute a pre-estimation sample.

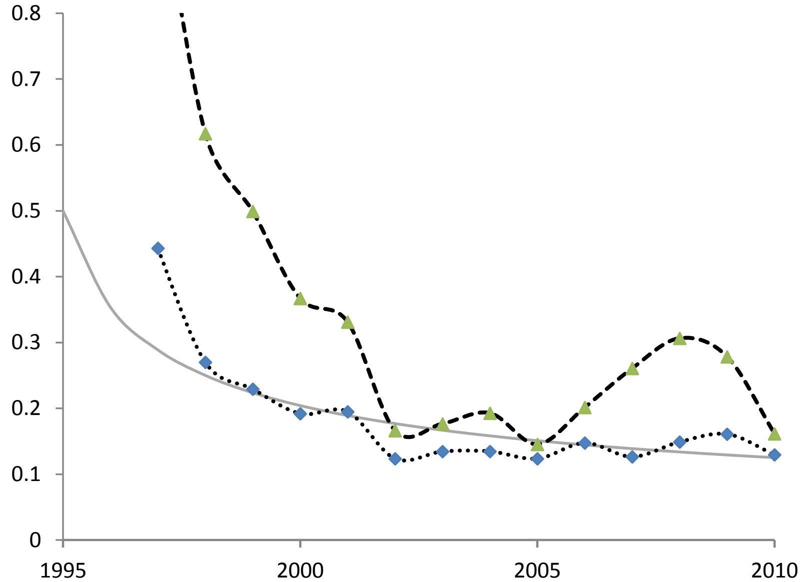

Figure 16: Comparison of the maximum bound for dispersion, $(\sqrt{V}/\sigma)_{max}$, and the actual values for DTSdirect $(\sqrt{V}/\sigma)$ for positive and negative earnings. The solid grey line is the maximum bound $(\sqrt{V}/\sigma)_{max}$, which constitutes the maximum value for DTSdirect $(\sqrt{V}/\sigma)$ for each period. The dotted line with the blue squares shows the actual value for DTSdirect $(\sqrt{V}/\sigma)$ for years with positive earnings. The thick dotted line with green triangles shows the actual value for DTSdirect $(\sqrt{V}/\sigma)$ for years with negative earnings. See Table 1 for a further description of the variable DTSdirect $(\sqrt{V}/\sigma)$. Firms must have full earnings histories to be included in the data that DTSdirect $(\sqrt{V}/\sigma)$ is based on in the figure.

Table 5 reports the results of a re-estimation of equations (22-24), this time for a sub-sample period of 2006-2010. Panels (A) and (B) are included for completeness, but as the primary goal it to test H2 (and H1), that is the theoretical bounds of the model, the analysis is contained to Panel (C).

Agents should again, according to theory, have had *at least* 10 annual observations of commonly observable information, causing them to have updated over common information 10 times. If agents are forecasting earnings as coming from a fixed distribution, the maximum fraction of forecast dispersion in relation to common uncertainty, $(\sqrt{V}/\sigma)$, (for any values on private information) that the learning model allows for, is again 0.158. Thus, in 2006 (and beyond), the maximum theoretical value that $(\sqrt{V}/\sigma)$ can take is 0.158, given that the underlying learning model is true.

In Panel (C) in Table 5, the coefficient on the negative earnings (Loss) is highly significant and large in magnitude (both in absolute terms and compared to the

Table 5: Table 5 displays the results of pooled OLS regressions of dispersion/uncertainty variables on negative earnings and firm-level control variables for a sub sample 2006-2010. All estimations enforce the availability of full earnings (1995-2010) histories for firms to be included in the sample. In Panel (A) the dependent variable is (forecast) Dispersion, in Panel (B) Disptosigma, and in Panel (C), DTSdirect ($\sqrt{V}/\sigma$), respectively. All regressions include industry controls (ICB industry classification) and year controls. Reported standard errors are White (1980) standard errors, robust to heteroskedasticity, and are found in parentheses. All variable definitions are described in Table 1.

Independent Variable	Dependent Variable		
	Panel (A)	Panel (B)	Panel (C)
	Dispersion	Disptosigma	DTSdirect
Loss	**0.303*** **	**0.257*** **	**0.0494*** **
	(0.0488)	**(0.0592)**	**(0.0145)**
Accuracy	0.00769***	0.00634**	0.00618***
	(0.00243)	(0.00308)	(0.00130)
Change in earnings	-0.000193***	-0.000373***	9.59e-06
	(5.73e-05)	(8.49e-05)	(3.23e-05)
Size	-0.0119**	-0.0159**	-0.00548*
	(0.00577)	(0.00749)	(0.00289)
Following	-0.000297	0.000320	-0.000179
	(0.00119)	(0.00160)	(0.000635)
Constant	0.296***	0.543***	0.323***
	(0.101)	(0.175)	(0.0512)
Year fixed effects	Yes	Yes	Yes
Sector fixed effects	Yes	Yes	Yes
Observations	3,315	3,255	3,318
R-squared	0.208	0.155	0.139

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

control variables). Performing a post estimation Wald test on the restriction that the intercept + the effect on negative earnings is equal to 0.158, yields a rejection (F [1, 3299] = 17.14, Prob > F = 0.0000), implying that the overall effect around negative earnings on DTSdirect, exceeds the maximum bound $(\sqrt{V}/\sigma)_{max}$, which implies accepting H2 as true[30]. Thus, in 2006 (and beyond), the value of $(\sqrt{V}/\sigma)$, exceeds its theoretical maximum value of 0.158, and thereby rejects the underlying model, implying in turn that it is not possible for an acquisition of private information to yield the observed increases in forecast dispersion.

Simultaneously, the results in Table 5 can be seen as an additional test of H1, both

[30]The magnitude on the coefficient of Loss alone in Panel (C) of Table 5 does not exceed the threshold value, but the correct judgment is of the overall effect, including the intercept. If one only looks at the effect of the coefficient on Loss, it excludes the "starting" level for dispersion.

in terms of the above, but also since performing a post estimation Wald test on the restriction that the intercept alone is equal to 0.158, also leads to a rejection, F(1, 3299) = 10.38, Prob > F = 0.0013.

Notice that while some empirical results here, taken at face value, are similar to those of the previous studies of Ali et al. (2009) and Adut et al. (2008), the fact that the private information acquisition hypothesis is rejected, questions the interpretations these authors make. Ali et al. (2009) and Adut et al. (2008) both conclude that the negative earnings environment that is associated with increased dispersion, is a result of selective disclosure in the sense that firms withhold bad news and therefore analysts rely more on private information, resulting in higher levels of dispersion in forecasts. However, regardless of how much private information analysts have access to, the resulting levels of forecast dispersion are too large for this interpretation to hold, since the model itself is rejected.

These findings are interesting. The implication is that the dispersion in forecasts cannot simply be explained by private information, since the threshold value $(\sqrt{V}/\sigma)_{max}$ is the maximum amount of dispersion in forecasts (in relation to earnings variance) that the fixed Bayesian learning model allows for. Recall that the maximum was derived for a signal, s'_{n^*max}, defined in such a way that it is the amount of private information that by definition maximizes dispersion. Consequently, there can exist no private signal that can cause an exeedance of this bound. Yet the bound is exceeded in the data, as the results in Table 5 confirm. What this implies is that there must be other explanations that are driving the results. These explanations, that have been briefly touched upon previously, can now be formalized.

4.13 Possible explanations

The combined evidence thus far rejects the model nesting the private information acquisition hypothesis. Since the tests of H2 confirm that bound exceedance is especially prominent around "uncertain" environments categorized by negative earnings, an explanation close at hand rests on the intuitive idea of negative earnings actually increasing the uncertainty of a firm's future. Anecdotal evidence for this could be inferred from the analysts' reports in the introduction, where around negative earnings, the outlook for the future is typically described in terms such as decreased or low visibility. As the fixed Bayesian model(s) do not allow for such effects, further speculation along the lines above at an earlier stage would have been somewhat premature. Now, as the results show that the model in the fixed form does not hold, it is motivated to suggest an alternative explanation that would be consistent with the observed increases in dispersion, which simultaneously has a more direct interpretation in terms of actual increased uncertainty.

This explanation entails that losses, as the intuitive hypothesis puts forth, actually do increase the uncertainty about the firm, that the agents face. As has been discussed previously, the standard Bayesian learning categorized as a time-invariant fixed process, as in the model of Chapter 2, does not permit increases in uncertainty (about the parameter, on the level of the individual), and therefore the only explanation for disagreement is based on asymmetric information.

Instead, the study proposes that agents are *not* always treating firm earnings as coming from a time-invariant distribution; for example when losses occur, in a very real way, agents become more uncertain about the distribution from which the earnings arrive, and this disrupts the learning process. This is consistent with agents viewing the process that they are estimating as (potentially) containing structural breaks, where a new regime requires learning to start over. Under such a scenario, agents, when faced with losses, become uncertain about the earnings process of the firm and "take a step back" in their estimation of the firm earnings process, effectively forcing the agents to restart their learning process from the beginning.

Theoretically, a mechanism that would allow for such a dynamic, which implies actual increased uncertainty on an individual level, is one where agents simply must formulate new priors. In essence, agents, after becoming faced with a firm entering loss territory, now view the firm as a "new" firm, or a new process, whose earnings process must be estimated from the beginning. This allows for both much higher levels of allowable dispersion, as well as the appealing implication that individual uncertainty actually increases.

4.14 Additional tests/hypothesis

Before analyzing the matter presented above further, it is worth returning to an interesting feature that comes out the expanded learning model in Chapter 2, that is following "large enough" increases in dispersion, subsequent private information acquisition should lead to *decreases* in dispersion. Since increases in dispersion in a fixed regime are the result of private information acquisition, one should reasonably expect that environments where the private information acquisition notion is invoked, should always be related to private information acquisition. So for example, if one can show that increases in dispersion are related to some identifiable effect or setting (e.g. losses), and one uses private information as an explanation for the increase in dispersion in this setting, then one should continue to expect this setting (losses) to be associated with private information acquisition, unless one has a believable story on why this would not be the case. This becomes interesting after dispersion reaches its "saturation point" since beyond this point, private information will speed up convergence and lead to dispersion declining at fast rate.

Having identified negative earnings as a driver for not only increases in dispersion, but more importantly that dispersion in forecasts exceeds its theoretically implied bounds, now allows for indirectly assessing the theoretical prediction that private information under some circumstances should result in *decreased* dispersion. While it was deemed difficult to judge what factor dispersion could increase by, since this depends on the previous cumulative amount of (unobservable) private information, it is now possible to replace the use of the cumulative past private signal with the empirical results of Table 5. To see why this is so, notice that at the point where dispersion reaches the $(\sqrt{V}/\sigma)_{max}$ bound, private information has reached its maximum capacity for increasing dispersion, implying that further private information acquisition can now only lead to decreases in dispersion.

By disregarding for the moment that the Eq. (18) bound is actually exceeded[31] in the results of Table 5, if one entertains the idea that dispersion only barely reaches the bound, theory predicts that beyond this point, dispersion will only decrease. In order to gain more insight into the findings, negative earnings are partitioned into those that occur in a firms history for the first time (in the sample) vs negative earnings that occur subsequently. If it is the case that that already first time occurring negative earnings (b)reach the bound, then it is evident that subsequently occurring negative earnings must lead to *decreases* in dispersion if the private information acquisition hypothesis holds.

Returning now to the idea of increased uncertainty, if agents instead view the futures of firms transitioning into negative earnings territory as truly having become more uncertain, dispersion in forecasts would be allowed to increase for subsequent earnings. The reason for this is that now the initial losses that breach the maximum bound for dispersion, force the agents to restart their learning procedures, i.e. the

[31] Note that this is only for the sake of argument. Since dispersion exceeds the bound it is the case that had it been private information that drove the increase in dispersion, dispersion could at most have reached the bound. If the private signal was more informative, dispersion would have instead "bounced back". So in cases where one assumes large acquisitions of private information, one would actually expect to see smaller increases in dispersion, implying that dispersion is already on its way down, or even decreases in dispersion.

parameter being estimated is now one from a new distribution. Furthermore, agents subjective certainty would no longer be constrained by the number of realizations received from a distribution with variance σ^2, and thus the threshold for maximum dispersion would be different both due to the higher "new σ", as well as the fact that the threshold to be compared against would now be higher, since n is now the distance between the first and the "other" loss.

This yields the additional hypothesis:

H2.1: Conditional on initial negative earnings reaching the Eq. (18) bound $(\sqrt{V}/\sigma)_{max}$, if agents truly become more uncertain, subsequent negative earnings will yield increases in forecast dispersion.

In order to test H2.1, the following regression specifications are estimated:

$$ForecastDispersion = \alpha + \beta_1 \cdot firstlossdum + \beta_2 \cdot otherlossdum + \beta_3 \cdot Controls + \epsilon \quad (25)$$

$$Disptosigma = \alpha + \beta_1 \cdot firstlossdum + \beta_2 \cdot otherlossdum + \beta_3 \cdot Controls + \epsilon \quad (26)$$

$$DTSdirect = \alpha + \beta_1 \cdot firstlossdum + \beta_2 \cdot otherlossdum + \beta_3 \cdot Controls + \epsilon \quad (27)$$

Table 6 reports the results of running equations (25-27), and are reported in Panels A to C, respectively. As is the case with the previous setups, the results in Panel (C) (and Panel [B]) suffer a bias resulting from the fact that the standard deviation of earnings does not (necessarily) have enough observations for convergence to correct values. Nevertheless, the results in Panel (A) indicate that the magnitude of the response in forecast dispersion is similar, whether the loss occurs for the first time or not.

Table 7 reports the results of running equations (25-27) for the sub-sample period 2005-2010, and are reported in Panels A to C, respectively.

In panel C in Table 7, the coefficient on first time negative earnings (First Loss) is again highly significant and large in magnitude (both in absolute terms and compared to the control variables). A post estimation Wald test on the restriction that the intercept + the effect on first time negative earnings is equal to the threshold level for $n = 10$ of 0.158, yields a rejection ($F(1, 3298) = 27.57$, Prob $> F = 0.0000$). This implies that now the overall effect on DTSdirect around negative earnings that

Table 6: Table 6 displays the results of pooled OLS regressions of dispersion/uncertainty variables on first time occurring negative earnings and subsequently occurring negative earnings 1995-2010. First time occurring negative earnings (First Loss) is defined as a negative EPS observation that occurs for the first time in a firm's earnings history in 1995-2010. Other Loss is any other negative EPS observation that occurs in a firm's earnings history in 1995-2010, excluding the negative EPS observation accounted for in First Loss. Estimations include firm-level control variables. All estimations enforce the availability of full earnings (1995-2010) histories for firms to be included in the sample. In Panel (A) the dependent variable is (forecast) Dispersion, in Panel (B) Disptosigma, and in Panel (C), DTSdirect ($\sqrt{V}/\sigma$), respectively. All regressions include industry controls (ICB industry classification) and year controls. Reported standard errors are White (1980) standard errors, robust to heteroskedasticity, and are found in parentheses. All variable definitions not found above are described in Table 1.

Independent Variable	Dependent Variable		
	Panel (A) Dispersion	Panel (B) Disptosigma	Panel (C) DTSdirect
First Loss	**0.387*** **	**0.963*** **	**0.200*** **
	(0.0508)	**(0.0926)**	**(0.0380)**
Other Loss	0.382***	0.307***	0.121***
	(0.0305)	(0.0375)	(0.0196)
Accuracy	0.00409**	-0.000861	0.00432***
	(0.00166)	(0.00221)	(0.00115)
Change in earnings	-0.000178***	-0.000155***	6.94e-05***
	(3.39e-05)	(4.34e-05)	(2.35e-05)
Size	-0.0129***	-0.0157***	-0.00652**
	(0.00314)	(0.00497)	(0.00315)
Following	0.000830	0.000281	0.000151
	(0.000648)	(0.00103)	(0.000690)
Constant	0.370***	0.788***	0.311***
	(0.0516)	(0.0762)	(0.0534)
Year fixed effects	Yes	Yes	Yes
Sector fixed effects	Yes	Yes	Yes
Observations	9,742	8,932	9,121
R-squared	0.228	0.202	0.169

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

occur for the first time, exceeds the maximum bound $(\sqrt{V}/\sigma)_{max}$.

While the result in Panel (C) of Table 7 that other losses seem to indicate a muted response on DTSdirect, first turn to look at the corresponding result in Panel (A). Notice that the effect of other losses on forecast dispersion is even larger than that of first time occurring losses. This implies that even as the first time occurring losses,

Table 7: Table 7 displays the results of pooled OLS regressions of dispersion/uncertainty variables on first time occurring negative earnings and subsequently occurring negative earnings in a subsample 2006-2010. First time occurring negative earnings (First Loss) is defined as a negative EPS observation that occurs for the first time in a firm's earnings history in 1995-2010. Other Loss is any other negative EPS observation that occurs in a firm's earnings history in 1995-2010, excluding the negative EPS observation accounted for in First Loss. Estimations include firm-level control variables. All estimations enforce the availability of full earnings (1995-2010) histories for firms to be included in the sample. In Panel (A) the dependent variable is (forecast) Dispersion, in Panel (B) Disptosigma, and in Panel (C), DTSdirect ($\sqrt{V}/\sigma$), respectively. All regressions include industry controls (ICB industry classification) and year controls. Reported standard errors are White (1980) standard errors, robust to heteroskedasticity, and are found in parentheses. All variable definitions not found above are described in Table 1.

Independent Variable	Dependent Variable		
	Panel (A)	Panel (B)	Panel (C)
	Dispersion	Disptosigma	DTSdirect
First Loss	**0.210****	**0.745*****	**0.172*****
	(0.0823)	**(0.137)**	**(0.0448)**
Other Loss	0.324***	0.142***	0.0211
	(0.0516)	(0.0528)	(0.0150)
Accuracy	0.00814***	0.00385	0.00560***
	(0.00248)	(0.00305)	(0.00126)
Change in earnings	-0.000221***	-0.000225***	4.63e-05
	(5.89e-05)	(7.70e-05)	(2.88e-05)
Size	-0.0113*	-0.0191***	-0.00626**
	(0.00581)	(0.00720)	(0.00284)
Following	-0.000288	0.000298	-0.000192
	(0.00119)	(0.00155)	(0.000628)
Constant	0.288***	0.576***	0.331***
	(0.102)	(0.171)	(0.0504)
Year fixed effects	Yes	Yes	Yes
Sector fixed effects	Yes	Yes	Yes
Observations	3,315	3,255	3,318
R-squared	0.210	0.190	0.151

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

as evidenced[32] in Panel (C), breach the maximum bound $(\sqrt{V}/\sigma)_{max}$, other losses continue to yield increases in dispersion (Panel A), even if the model clearly predicts that once the bound is reached, a further acquisition of private information should only act as to reduce dispersion. Thus, even when dispersion reaches its maximum level, subsequent losses continue to give rise to increases in dispersion, even if this is not possible under the model. This now confirms H2.1.

[32]This refers to the Wald test of the effect of the intercept + the coefficient on Firstlossdum.

Now turn to the result in Panel (C), where other losses seem to have negligible effect on DTSdirect. In order to appreciate potential reasons for this, a further analysis of what is happening in the data and the variables must be carried out. There are two different effects at work. One is mechanical, and has to do with the conditioning in the variable DTSdirect, and the other is how the conditioning is related to theory.

Mechanically, losses can be outliers in firms' earnings series (as in the somewhat extreme example in Figures 14 and 15), and can therefore in some instances increase the variance of the firms' earnings series[33] (going forward). For ease of understanding, this section will refer to the variance of the earnings series as σ' under an assumption that a (first) loss induces a higher mechanical variance for the earnings series going forward. The variable DTSdirect is constructed to only take into account all information available up to the current point in time (not forward looking), as agents are not assumed to know about potential future changes in firms' earnings distributions before they actually occur; agents are assumed to use the data available to them at the current point in time. Therefore, the variable DTSdirect does not account for potential increased variance of the earnings series that occurs after a (first) loss occurs.

If the historical variance σ is somewhat stable, an increase in dispersion around a (first) loss, *relative* to the agents' current information sets/beliefs about the earnings series (in essence $(\sqrt{V}/\sigma)$), can be substantial. Simultaneously, if the variance of the earnings series has become larger following a first loss that occurred previously, then a similar increase in dispersion on absolute terms as the previous (first) loss, now on a relative basis (to the new σ') can be smaller in magnitude. If such a mechanical effect is present, where a first time occurring loss yields a larger σ' going forward, this effect can be a contributing factor to a small response to other losses in Panel (C) in Table 7 even if the response in dispersion to both the first and other losses is of similar magnitude in Panel (A) in Table 7.

It is however not sufficient to study the mechanical effect of potential increases in the variance of a firm's earnings series going forward in the variable DTSdirect in isolation. Instead, one must consider the effect in conjunction with predictions from theory.

If the assumption about a break in learning holds, then following a break, agents restart the estimation process using the new series. Thus in relation to this new series, one that has potentially become more volatile, a loss that follows, which

[33]The study does not perform specific tests for whether there are actual breaks in the earnings series variance, rather the point here is to discuss what effect on DTSdirect such breaks would have.

would yield a similar increase in dispersion on absolute terms as the previous loss, now on a relative basis (to the new series) can be smaller in magnitude. Thus, following an initial loss, dispersion would again be theoretically compatible with the new earnings series, and this effect can be a contributing factor to the small response to other losses in Panel (C) in Table 7. Withholding the assumption of a break in the series, it is not clear exactly how or why agents know a break has occurred, since there is no direct theoretical mechanism that allows for this[34]. It is thus possible that the variance (σ') in the new conditioning set is even larger, making the effect of other losses on DTSdirect in Panel (C) in Table 7 biased to the upside.

If instead the private information acquisition hypothesis held, regardless of the conditioning set[35] for other losses, the effect should be negative, which would show up as a negative response on other losses in *both* Panel (A) and Panel (C) in Table 7, since (b)reaching the $(\sqrt{V}/\sigma)_{max}$ bound means dispersion on an *absolute* level can only decrease going forward. The analysis however is complicated by the fact that the variable DTSdirect, following an initial loss that increases earnings variance in combination with the assumption of the null, uses the wrong conditioning. Why is this the case? First, the construction of the DTSdirect variable assumes that σ is constant. While the *estimated values* for σ are not constant in the variable DTSdirect, there is a substantial margin of error in favor of the null, since this estimation acts as if true σ is equal to the latest estimate of σ, even though estimated σ tends to grow[36].

Potential breaks in the earnings series induced by losses however, are not compatible with the rationale for the construction of DTSdirect. Since agents do not know the losses before hand, a breach of the bound that uses the conditioning set (a function of σ) up until the loss, implies the learning model is compromised. The only way to assume that the new increased earnings variance (σ') would be compatible with the learning up to the point of the breach at the loss, is if the increased variance going forward actually consists of noisier signals - the following subsection analyses this scenario in greater detail. If this is the case however, the correct way to assess the maximum size of the dispersion continues to be σ from before the loss. This is the case because if one entertains the learning process as continuing, new noisier signals do not affect the amount of learning that has occurred in the past. New noisier signals will at most affect individual uncertainty, and consequently the maximum bound, in such a way that it is equal to the period before[37].

[34]The alternative explanation provided is not analytically formalized.

[35]The completely rigorous term here would be *subsequent* conditioning set. These conclusions implicitly condition on the preceding information set being correctly identified.

[36]Both the section 4.17 as well as Appendix E deal with this issue further.

[37]Since the Bayesian agent never loses information. Consequently, whatever uncertain signals

The *variable* DTSdirect however, compares the dispersion to the latest estimate of earnings variance. In the case of a large previous loss, it is possible that the estimated variance (σ') has increased. This new variance (σ') is not the "σ" in the information set anymore, in the sense assumed in the construction of the variable DTSdirect. The σ, or variance in the information set[38] *cannot* increase if learning remains fixed. However if the learning is breached, then DTSdirect is again correct[39], since the new σ' that DTSdirect uses, is the correct σ if learning has restarted. The implication here is that the first loss that already exceeded the maximum, cannot be followed by private information acquisition or any other effects that could increase dispersion. Instead, withholding the private information acquisition assumption, the only effect one could expect under the model is for dispersion to start decreasing. Yet, the evidence in Panel A in Table 7 suggests that the contrary is true.

The conclusion is therefore that while it might seem on first glance from Panel (C) of Table 7 that other losses are within the limits of the model, this can only be the case if the model has undergone a break.

4.14.1 The implication of a noisier signals interpretation

Before drawing definite conclusions from the results of Table 7, a somewhat more detailed discussion on the role of viewing losses as noisier signals, briefly touched upon in the previous section, is warranted.

Nothing prohibits one from viewing earnings realizations that follow increases in earnings variance (induced by losses) as signals, centered on the same expected value as before, that now simply are signals with more noise (higher variance). This essentially equates to treating σ as being non fixed. While section 4.17 deals with this issue in a more general manner, the following section delves deeper into the issue as it relates here.

What happens is that if agents would treat the new signal as a noisy signal, but from the same underlying process, the common part of agents subjective beliefs do not change by much. If the signal would be assumed to be unchanged, then learning would proceed as in the model. However, no matter how uncertain the new signal is, the agents' subjective uncertainty can at most remain at (close to) the same level as previously. What this in turn implies is that the common part

the agent receives, his/her past signals/information are never weakened.

[38]Naturally, the (common part of) uncertainty in agents' information sets is *a function of* σ.

[39]Perhaps a better term is compatible. As discussed on the previous page, the "new" earnings variance might be higher than what is implied in the estimate in the variable DTSdirect.

of information that is included in the agents' information sets (cumulative r or σ), which ultimately defines the threshold for dispersion, can at maximum remain at (close to) the previous level. So if the threshold under a fixed model for $n = 10$, is equal to 0.158, then for a signal that is completely uninformative (high variance), the threshold for maximum dispersion at $n = 11$ can never exceed 0.158 (in relation to the previous σ that DTSdirect conditions on[40]).

The reason for constructing the DTSdirect measure in such a way that it acts as if the latest estimate for the standard deviation of earnings is correct, is that it is directly compatible with the theory, which assumes that agents are getting signals from $y \sim N(\theta, \sigma^2)$.

While theoretically and agent can receive signals of varying precisions that are centered on the same mean, it is not possible to estimate this from data, since this would entail using each estimate of the standard deviation for each firm. One would have to assume that at each point, the estimate of the signal variance is exactly equal to the uncertainty of the signal, and one would consequently have to base the standard deviation on very few observations. Furthermore, one could not use the threshold levels for DTSdirect that are a function of n, since this usage rests on the signals being of equal size at each period. One would instead have to make individual threshold calculations for *each* firm.

As an example, in the case of the company in the Figures 14 and 15, the standard deviation of forecasts for the corresponding (first) loss year is 0.64. The variable DTSdirect compares this against the standard deviation of previous earnings, which since the beginning of the sample period equates to 0.74. The assumption in using the threshold levels for DTSdirect is that the latest standard deviation estimate for earnings, 0.74, constitutes σ. Thus, in assuming that agents have updated over a σ of 0.74, one would use the number of updates that agents should have received, in evaluating the bound. Since the empirical estimations do not keep track of the exact number of updates, true n in this case is actually 12 instead of 10 that is used as the threshold in the empirical estimations. Thus the true rejection bound is actually *lower* than 0.158.

If one instead suggests that agents are not using a time invariant σ, or that it cannot be estimated correctly, and one instead argues that agents are receiving signals of varying precision at each period, one would in the previous case accumulate signals estimated at each step[41]. Because the estimates for the standard deviation at the

[40]Since learning must be assumed to have occurred w.r.t. a distribution of lower variance in the past, conditioning on a new higher variance signal would be incorrect.

[41]See section 4.17 for specifics of the signal accumulation procedure.

first periods are low, accumulation of the assumed signals of differing precisions as observed at each step would imply that the maximum allowed dispersion at the time of the negative earnings is 0.032. As noted, there now exists no general threshold level that can be used. This is in contrast to the previous calculation where σ is assumed to remain fixed, where the threshold levels could be used, where the 0.158 fraction of 0.74 implies that the maximum allowed dispersion that the estimations assume is 0.12. Thus, in this case for example, the threshold level that DTSdirect employs, is approximately four times higher than what it would be under the above rationale.

In conclusion, the results of Table 7 can be summarized as follows: The first time negative earnings occur, dispersion increases to a level that is too high to be supported by the underlying theoretical model. Subsequent losses for a firm tend to increase dispersion to the same extent or more, even if theoretically the effect should turn negative. Furthermore, while the effect on DTSdirect for other losses seems to point toward a nonexistent outcome, this is a mechanical effect due to the interplay of both the construction of the DTSdirect variable, as well as that of the break in the standard deviation of the earning series that negative earnings tend to produce. Since the first time negative earnings breach the theoretical bounds implied by the model, private information acquisition alone cannot be the driver be the increases in dispersion. Consequently, another explanation is required. The one hypothesized is one where analysts truly become more uncertain about the firm earnings, and have to restart learning. If this is correct then the observed effect of other losses on DTSdirect is consistent with an explanation where the levels of dispersion lie well within the theoretical bounds implied by the uncertainty in a new learning regime.

Finally, the results in panel C of Table 7 can also be seen as an additional test of H1, both in terms of the above, but also since performing a post estimation Wald test on the restriction that the intercept alone is equal to 0.158, also leads to a rejection, $F(1, 3298) = 11.78$, $Prob > F = 0.0006$.

4.15 Discussion of results/Summary

Overall the empirical results show that dispersion in forecasts breaches the theoretical bounds implied by $(\sqrt{V}/\sigma)_{max}$. Apart from also occurring in general, exceedance is triggered especially around negative earnings. This implies that the increases in dispersion cannot solely be driven by an acquisition of private information. Instead, the evidence corresponds well with an idea of increased uncertainty. As the theoretical model does not allow for such effects however, and since it is indeed losses that trigger exceedance, the evidence supports the hypothesis that agents, when faced with losses can actually become more uncertain. By viewing losses as indicators of regime shifts in the firms' earning series, agents, through restarting their learning process, actually become more uncertain, on an individual level, about firm earnings.

While these results are not compatible with increases in private information being the only driver for dispersion, they do not eliminate the need for private information. Furthermore the results do not compromise Bayesian learning itself, they merely questions agents acting as if there exists one time invariant distribution over which they update.

Since dispersion in forecasts is too large to solely be supported by acquisition of private information, the conclusion is that it must be the case that especially when negative earnings occur, agents, though remaining fully rational, actually become more uncertain about the future of the firm.

This is easy to accept on heuristic grounds, supported by anecdotal evidence from analysts reports where the future of firms facing negative earnings is typically denoted by expressions such as decreased visibility. The proposed theoretical mechanism for this is one that maintains rationality, but instead suggests that analysts are not treating the firm earnings as coming from a stable, fixed distribution. When a firm suddenly becomes unprofitable - it is due to a inherent change of the firm (or its earnings process) itself. The rational Bayesian agent who becomes aware of this, now must restart his/her learning. Instead of continuing in using his/her past information - under which he/she could eventually learn the new parameter, it is more efficient for him/her to restart his/her learning by constructing a new prior. This causes a temporary surge in uncertainty, as the agent in a very real way becomes more uncertain, which also leads to a much larger allowed level for forecast dispersion. However, the temporary uncertainty increase, leads to the agent being more efficient in estimating the new mean for the series, than what would have been the case if he/she had continued updating/conditioning on the wrong/old variable.

4.16 Robustness

One potential concern is if the sample period standard deviation does not correctly identify the true standard deviation of the earnings series. This could happen for general reasons (the sample period is less volatile) or specific reasons, e.g. that the pre-sample period earnings histories contain losses that deviate from the rest of the (sample) earnings series[42]. Most importantly however, there is the question to what extent the estimated standard deviation has enough observations for convergence.

In all three instances there is a concern that the standard deviation in the sample period is downward biased from the true standard deviation. This would inflate the DTSdirect ($\sqrt{V}/\sigma$) measure, since the real value for σ is higher and therefore true DTSdirect is lower. This would in turn be problematic in terms of inference, since if the uncertainty (standard deviation) in the agents' information sets is higher (than estimated), consistency with the standard model allows for higher levels of dispersion.

In order to assess this potential problem, robustness tests are carried out where all companies that have complete earnings histories in the sample (DTSdirect observations) are sampled and where their complete earnings histories are compared to their earnings histories used in the sample 1995-2010. The maximum earnings length histories used in the robustness tests are available from 1978, and consequently full sample histories vary from 32 years (starting in 1978) to 17 years (starting in 1995, the main sample) (Due to the fact that full earnings histories are enforced in the sample, the minimum sample spans 17 years). The study constructs measures for standard deviation and coefficient of variation, and compares the measures for the sample period (1995-2010) to the corresponding full histories.

The results of the robustness tests indicate that on average, the full sample standard deviation is only 2.56 percent higher in the full earnings history of the companies, than in the sample used. One can thus be fairly safe in assuming that the estimated in-sample common uncertainty, as measured by the variability in historical earnings, is a robust proxy for the true standard deviation of earnings that the theory posits that agents are using, and that the uncertainty ratios are well identified. Viewed differently, this supports one of the main assumptions that here, on average, earnings series' have well defined, time invariant second moments. This is important since the learning under the model assumes that agents condition their expectations on a fixed, true variance for the DGP from which they observe the series of realizations.

[42] In this case earnings volatility would be *under* estimated in sample.

(The analogous results for the coefficient of variation shows that on average the full histories have a coefficient of variation that is on average 10 percent higher, yielding a potential upward bias on coefficient of variation comparisons of roughly equal magnitude).

There is still a concern that actual convergence of the estimated standard deviation values has taken place only in 2010. Furthermore, another potential concern is that even if on average, the estimated in-sample standard deviation is only off by 2-3% from true, full sample values, outliers might be driving the result or that there is a conditional effect at play: companies with (first time occurring) negative earnings might be the ones with the most biased standard deviation estimates. Since the largest estimated standard deviation difference (between the estimated and the true value) is approximately 100%, this could potentially bias the dispersion/sigma measure upward by an equal amount. To ensure this is not the case, all observations where the exeedance is over 50% are deleted. Here the midpoint of the sub-sample estimation, namely 2008, is used. Performing the adjustments confirm that the results are not driven by wrongly estimated standard deviations: regression results in all specifications remain unaffected by the deletions.

Matters become slightly more complicated when the earnings series undergo potential larger breaks, especially as whether they have occurred or not are important for the evaluation of the performance of the theoretical model. Consequently it is not enough to identify an unconditional forward looking standard deviation for the earnings series in the case that the series is made up of different regimes.

Due to the fact that the DTSdirect measure is ill-equipped in dealing with these breaks in the earnings series of companies following initial losses, additional robustness tests are performed. Recall that under a break in the earnings series following an initial loss, regardless of whether one relaxes the assumption of a fixed σ, or that of a fixed learning regime, DTSdirect uses the wrong conditioning. One thus wants to make sure that the estimates for the standard deviations are correct up until the first loss where DTSdirect still conditions correctly. In order to do this, the study instead samples all companies that have complete earnings histories in the sample (DTSdirect observations) and measures full standard deviations up to 2006, and compares these full earnings history standard deviations to standard deviations estimated in the sample between 1995-2006. Due to the shorter estimation window for the standard deviation, results are off to a slightly higher extent - the true standard deviation is on average 8 percent higher from the estimates carried out for the years 1995-2006.

In order to ensure that this does not affect results, an analogous removal of company id's as above is performed. The criterion for deletion is that the in-sample standard deviation between 1995-2006 compared to full sample standard deviation (-2006) is over 50 percent. This effectively removes the corresponding risks that the previous deletion procedure deals with, i.e. that outliers might be driving the result or that companies with (subsequent) first time occurring negative earnings are the ones with the most biased standard deviation estimates. Additionally, this ensures that standard deviation estimates for the earnings series are robust going in to the sub-sample estimations. Performing the adjustments again confirm that the results are not driven by wrongly estimated standard deviations as regression results in all specifications remain unaffected by the deletions. Furthermore, after having carried out the deletions, the full standard deviation drops to being on average only 5 percent higher from the estimates carried out for the years 1995-2006.

Additional note: One of the companies dropped in the above procedure is Activision. In the sample period (1995-2010), earnings per share are between -0.05 and 0.69. In the years preceding the sample period, 1994 and earlier, there are large gaps in the data, with earnings available only for 1989-1990, 1986-1987 and 1984. The 1984 earnings per share are reported at 16.17, which obviously heavily affects the full sample standard deviation in comparison to the sample period standard deviation. The gaps in the early data, combined with inconsistently large earnings per share observations, warrants a suspicion that there are potential adjustment errors in the (early) data. This shows that the sample period standard deviation estimate is not necessarily biased in itself, in terms of underestimating the true value, rather the difference between the estimated and true standard deviation values, are potentially driven by data errors.

Thus the combined evidence that the sample standard deviation estimates that are potentially biased on the downside (which potentially overstate the results) a) do not affect the results, and b) that the standard deviation estimates that seem too low in comparison to full sample values, are biased because of data errors, point to the fact that the estimates for standard deviation used in the study are correct.

Note finally that there is also an opposite effect at play here. As discussed earlier, it is somewhat naive to assume that firm earnings histories that start in 1995 represent the beginning of all firm histories. As the earnings histories date back to a maximum of 32 years here, the convergence that has to occur on common information implies that the maximum fraction of dispersion to σ, is lower than the threshold used. In fact, the average earnings history for firms in the robustness sample is 25.44 years. Since the regressions (using the year 2006 and onwards sample) used n=10 to arrive at the maximum fraction of dispersion that the model allowed, 0.158 (from Eq. [18]), the correct n on average would actually be approximately 20. This in turn implies that on average, the maximum fraction of dispersion to common uncertainty, and subsequently the threshold to be used in judging exceedance of theoretical bounds, is in reality as low as 0.0913.

4.17 Robustness 2: A note on growing standard deviation over time

If the standard deviation grows over time, the predictions from the model, which assumes a constant distribution for common information, have to be interpreted slightly differently.

This however is not very problematic, as has been noted previously, if the relaxation of a fixed distribution only refers to the variance. In this case, where it is implied that the learning regime continues to remain fixed, differing variances across time can be interpreted simply as signals of different precisions where some signals are noisier than others. This section investigates how DTSdirect is affected under such an assumption.

Note that the empirical measure of common uncertainty uses the latest estimate of the standard deviation for the earnings series. If the (true) standard deviation is growing, the estimate of the standard deviation of earnings is inflated (Since this estimate is used in the variable DTSdirect as a fixed quantity over time). As the uncertainty ratio, DTSdirect ($\sqrt{V}/\sigma$), compares current dispersion in forecasts to this measure, the recursive standard deviation of earnings, DTSdirect acts as if the learning would have occurred w.r.t. to this latest estimate of common variability throughout history. The fact that in reality, learning has occurred w.r.t. lower variance "distributions" (i.e. signals) in the past, there has been more convergence than what the current estimate of the standard deviation suggests, and this implies that the maximum bound for dispersion at the current point in time (if using the bound from the non-growing case) is actually overstated. This is illustrated in the following, somewhat lengthy example.

It is true that in absolute dispersion terms, the maximum amount of dispersion in, say, period 10 is higher if the (true) standard deviation of earnings has grown over time (noisier signals), vs. a constant case, *given* that the starting point was the same. The reasons for this should be clear at this point: noisier signals equal less learning, and as the maximum dispersion is a function of the amount of convergence, slower convergence supports higher dispersion.

If this starting point is, without loss of generality, normalized to 1, then first in the constant case, after 10 learning periods, the maximum dispersion is equal to 0.158 in magnitude. This however is not the correct case to compare to. The correct comparison is for a value that is equal to what σ has grown to at the point of measurement. Thus, if the standard deviation has grown annually, on average (in variance terms 1.14), then at point 10, the signal is 3.25 in variance terms (1.8 in standard deviation terms). Thus, the correct case of comparison, is one where signals of $\sigma = 1.8$ have been received for 10 periods.

Using the above where the variance has grown by an approximate factor of 1.14 (this is the corresponding value to the average annualized growth of the average standard deviation of earnings in the sample), the variance starting at 1, reaches 3.25 at n=10. The cumulative sequence of precisions is thus $1+1/1.14+1/14^2+...+1/1.14^9$, equaling 5.946 at n=10. Inserting the cumulative precision into Eq. (9), yields 0.042 (0.205 in standard deviations terms), which is the maximum absolute dispersion at this point (where it is implied that the private signal is of equal strength).

In order to see how one can accumulate the signals in the above, conditional on knowing the exact evolution of signals over time, notice that it is in this case possible to arrive at Eq. (18) directly, using equation 14 (V'). Set $r_1 = s_1$ in order to arrive at $nr_1/(2nr_1)^2$, which yields $1/4nr_1$. This refers to the model presented in the study, where the common information distribution remains fixed. One can also in this case arrive at Eq. (18) by using the Barry and Jennings (1992) intuition, instead using the Barron et al. (1998) representation given in Equation (9). Setting $h = s$ yields $1/4h$. Now using an equivalent sample size argument similar to the one referenced in Barry and Jennings (1992), h, can now be seen to represent the cumulative sequence of (common) signals that have occurred up to the current point in time.(If the sequence of signals has been constructed from a fixed distribution at each point in time, then h corresponds to $n \times r_1$.)

The earlier paragraph compares two cases where σ is known in both cases. Empirically however, DTSdirect cannot distinguish between the two. Instead, DTSdirect acts as if in this case there have been 10 realizations of a signal with variance 3.25 (1.8 in standard deviation terms). Consequently, DTSdirect, the way it is measured, suggests that the maximum amount of allowed dispersion is $\sqrt{V}_{max} = \sqrt{1/[4 \times 10 \times (1/3.25)]} = 0.285$, whereas the true maximum dispersion, knowing the true evolution of signals over time is equal to $\sqrt{V}_{max} = \sqrt{1/[4 \times 5.946]} = 0.205$.

To state the above differently, what is observed in the data is that $\sigma = 1.8$. DTSdirect thus assumes that σ has been equal to 1.8 for $n = 10$, which would imply a threshold for dispersion equal to $1.8 \times 0.158 = 0.285$. Thus, if a level of dispersion of 0.285 for the company in question is observed, then DTSdirect compares this to $\sigma = 1.8$, which results in dispersion being exactly at threshold level of 0.158. However, if σ has de facto grown, the correct rejection level would be for a dispersion in the data of 0.205 (In the above). In this case the *correct* comparison would be for $(\sqrt{V}/\sigma) = 0.205/1.80$, yielding a rejection threshold of 0.11. However, DTSdirect does not take into account potential growth of σ, and thus DTSdirect gives a large advantage in favor of the null.

The above correction applies if the real variance has grown, i.e. in a sense agents receive increasingly noisier signals. What the calculations show is that in this case the DTSdirect measure generally has the potential to overstate estimated σ, which in turn leads to using a threshold value that is compared against that is too high. So when e.g. a threshold value of 0.158 is used, it is likely that the true threshold that should be used is lower, (e.g. 0.111 in the numerical example), since the DTSdirect *understates* the amount of convergence that has actually occurred.

On the other hand, if there is an estimation error regarding the standard deviation of earnings, such that only in the later part of the sample will the estimated standard deviation have converged to its true value, then DTSdirect is unbiased. The robustness tests in the previous section deal with this specific issue, and the results indicate that on average σ is well measured in sample, even before the beginning of the later part of the sample that constitutes the space for the sub sample estimations.

4.18 Robustness 3: Estimating the standard deviation for growing EPS

There is still one more issue that warrants discussion. Careful consideration of the estimation procedure for the (recursive) estimate of the standard deviation, shows that the standard deviation is the spread around the time series mean[43]. When the variable (EPS) is growing, the time series mean is generally lower than the current level of earnings, around which the standard deviation of forecasts is centered (the forecasts are of course centered on the mean *estimate* of earnings but this lies, as a

[43]This is speaking loosely. The standard deviation of course is in units of EPS, but centered on time series average.

general rule, close to the level of current earnings).

This suggests that perhaps a better way to analyze earnings would be in the form of a return series, where the mean is better behaved. The theoretical implication here is of course that agents are estimating the growth rate of earnings, which is what they do, at least implicitly. Observations on the standard deviation in forecasts is however not available on this basis[44], and in order to carry out an analysis on this level, the data has to be transformed. This is carried out below.

In the case of the example company Fifth Third Bancorp, earnings pre 2008 have grown on average 8.4 percent annually, with a standard deviation of 14.7 percentage points annually (implying that a one standard deviation interval for the growth in earnings is 8.4 $\pm$ 14.7). If this would be considered to be the true (common signal) distribution, then in accordance with the rate of convergence, after 10 periods the maximum amount of dispersion would be given by 2.32 percentage points (0.158 $\times$ 14.7) around the average return of 8.4 percent growth. Since the level of earnings in 2007 is 2.03, the average growth that is to be expected from the data,[45] equals 2.201, obtained simply from 2.03 $\times$ 1.084 (the level of earnings multiplied by the average growth rate).

Since learning has occurred for 10 (in reality of course >10) periods, the aforementioned standard deviation of 2.32 percentage points around this value, applies. The lower one standard deviation bound[46] is given by 2.03 $\times$ (1.084 - 0.0232) = 2.1534. Analogously for the upper one standard deviation bound 2.03 $\times$ (1.084 + 0.0232) = 2.2476. Consequently, in units of the *current earnings*, the implied standard deviation, corresponding to $(\sqrt{V}/\sigma)_{max}$, is 0.047 (2.201 $\pm$ 0.047). This is contrast to the acceptance level that DTSdirect uses, which is equal to 0.115. (0.158 $\times$ 0.73, which is the threshold value for $n = 10$, although the true threshold equals 13, and 0.73, which is the standard deviation estimate for 2007).

What this shows is that the structuring of the variable DTSdirect gives a large advantage in favor of the null.

[44]one could argue that LTG (long term growth) estimates are exactly this measure, but there are reasons to suspect that apart from LTG having much fewer observations, analysts are nowhere else putting as much effort as in getting the front EPS measure right.

[45]Being completely rigorous it is theoretically problematic to use this estimate, since it is exactly this estimate agents, under the theory, are trying to learn.

[46]Observe that the standard deviation bound is not the subjective uncertainty, rather the value that constitutes (maximum) dispersion.

Chapter 5

Conclusion

5.1 Discussion and Conclusions

Traditional models for how financial analysts construct their forecasts predict that dispersion in forecasts is driven by asymmetric information. When dispersion in forecasts increases as a result of asymmetric (private) information, it simultaneously implies that subjective uncertainty decreases, i.e. agents become more certain. In contrast, empirically dispersion in forecasts tends to be large in environments, such as firms experiencing losses, that heuristically indicate uncertainty. The hypothesis that losses are linked to uncertainty originates from evidence in analysts written reports, where analysts describe negative earnings environments as having low, or weak visibility. Theory and evidence thus indicate contradictory explanations for (increases in the) dispersion in forecasts. Theoretical models indicate that increased dispersion implies increased subjective certainty, whereas anecdotal evidence from analysts reports suggests that increased dispersion implies increased subjective uncertainty.

Studies in Accounting Research typically take the private information assumption as given, and consequently invoke the prediction of private information increasing dispersion as a more or less untested explanatory factor for empirically observed increases in forecast dispersion. This study shows that private information cannot explain the observed magnitude of forecast dispersion.

The study is able to challenge the private information acquisition hypothesis for increasing dispersion through the following. The study develops a representation/ application of the models of Barron et al. (1998) and Barry and Jennings (1992) that explicitly takes into account the amount of common information that becomes available through earnings announcements over time. When companies release earnings to the market on an annual basis at minimum, (public) information about earnings must successively add more information to analysts, causing convergence of individual beliefs. The study tracks the convergence of subjective beliefs that results from common information and subsequently derives maximum bounds for dispersion in forecasts, conditional on the amount of received common information. The study then empirically tests the compatibility of the maximum bounds with observed levels of forecast dispersion, conditional on the amount of common information that has become available.

The main theoretical results are that not only does there exist a maximum amount for dispersion[1], but this maximum amount must decline monotonically over time at a predetermined rate. The convergence of beliefs, and subsequently the convergence of the maximum amount of dispersion, is an outcome of Bayesian learning where agents must become more certain upon receiving more information. Each annual earnings release contributes more information about the parameter that is being forecasted and leads to a decrease in subjective uncertainty.

The simple Bayesian model in which each earnings release represents a draw from a commonly observed signal distribution and leads to declining maximum amounts of dispersion, is a single parameter model. The single parameter setup ensures that the theoretically correct conditioning set is that of the signal distribution variance. In this single parameter setting this conditioning variable, the informative distribution or Data Generating Process (DGP), is assumed to equal firm level earnings. Through assuming that the observed historical sample variance of earnings has converged sufficiently to represent population values, the empirical mapping replaces the theoretical variance with observed historical earnings variance. In doing so, the study is able to present the maximum level of dispersion as a ratio of the dispersion in forecasts to earnings variability ($\sqrt{V}/\sigma$). When the true variance of signals is time invariant, this ratio depends exclusively on the number of periods. Having established the theoretical benchmark, the study empirically evaluates the maximum bounds for dispersion in forecasts $(\sqrt{V}/\sigma)_{max}$ implied by the above model.

The main empirical result is that the *magnitude* of the dispersion in forecasts is not compatible with the standard Bayesian learning model if agents treat the informative distribution as fixed. The theoretically implied maximum bound $(\sqrt{V}/\sigma)_{max}$, is exceeded in the data both in general, and further, especially around negative earnings. This implies that the increases in dispersion cannot solely be driven by an acquisition of private information.

[1] A result derived by Barry and Jennings (1992), where the maximum is attained for private information being equal to common information.

A second empirical result derives from the theoretical predication that if there have been sufficient amounts of dispersion in the past, increases in private information should yield *decreases* in dispersion. A straightforward implication of this theoretical result is that one would expect to find numerous companies for which e.g. negative earnings yield lower dispersion. One would consequently also expect the average effect of losses on forecast dispersion to be muted or even negative. Such an effect cannot be identified in the data.

The empirical results, where observed values for forecast dispersion breach theoretically implied bounds, show that asymmetric information alone cannot explain the magnitudes of disagreement, given the underlying model. The implication of these results is that when dispersion increases, the explanation that rests on increased private information and nesting increased subjective *certainty*, is rejected. The study therefore challenges interpretations made in studies such as Lang and Lundholm (1996), Adut et al. (2008), Ali et al. (2009), Barron et al. (2009), where (increases in) forecast dispersion is thought to signify an increase in information asymmetry. In light of the combined evidence of this study, information asymmetry alone cannot cause levels of dispersion that are observed in the data and this study suggest that conclusions in previous studies resting on the private information acquisition hypothesis could benefit from being re-evaluated[2].

In order to be able to explain magnitudes of forecast dispersion observed in the data, this study offers an alternative mechanism capable of generating observed levels of forecast dispersion. This explanation, that maintains the rational learning model, but relaxes the "fixedness" of the informative distribution, is that agents, when faced with uncertain forecasting environments, such as those associated with firms experiencing losses, actually treat the parameter as coming from a new distribution. Agents must then restart learning, implying that they start out with new priors. This encompasses the auxiliary implication that agents actually become *more uncertain* about the parameter (earnings) they are estimating. In these environments the interpretation on large increases in dispersion now changes dramatically; instead of being the result of increased private information (with the implication of increased certainty), increases in dispersion instead indicate real increases in uncertainty about firm earnings. This explanation finds support in empirical estimations.

Apart from the empirical results of this study, the threshold level for the maximum amount of dispersion can easily be applied by other researchers wishing to determine

[2]An earlier version (2007) of the Adut et al. (2008) paper does hypothesize loosely that bad news environments could be associated with higher levels of uncertainty, but the Adut et al. (2008) version of the paper subsequently moves the stance in favor of the private information acquisition hypothesis.

whether asymmetric information alone can be determined to be the driver behind dispersion in analysts' forecasts. The threshold level, $(\sqrt{V}/\sigma)_{max} = \sqrt{1/4n}$, can be used in a backward sense to assess what level of volatility/uncertainty for public information, σ, observed levels of dispersion are compatible with, when one accepts the assumption of signals being fixed over time. This results in $\sigma_{(min)} = 2\sqrt{nV}$, which is now the *minimum* implied amount of volatility that the observed levels of dispersion, in conjunction with the assumed length of learning, are compatible with. If earnings volatility is below this level, it is equivalent with dispersion exceeding its maximum value. The resulting value thus obtained for σ can be compared against corresponding observed values in one's data set, and can thus be used in evaluation of the reasonability of the private information assumption being the single driver behind disagreement in the data[3].

5.2 Discussion on alternative theoretical explanations

The combined theoretical and empirical results of this study indicate that the dispersion in forecasts (beliefs) cannot be explained by asymmetric information alone in a rational Bayesian learning setup where learning about the parameter is fixed. The study presents a possible mechanism to explain the observed magnitude of forecast dispersion but alternative explanations are also possible. For example, the derived maximums withhold the assumption of honest forecasting. E.g. Ottaviani and Sorensen (2010) note in relation to the Keane and Runkle (1990) finding that asymmetric (private) information drives differences in forecasts, that such conclusions rest on the maintained assumption of honest forecasting and could also be due to strategic forecasting. Since this study does not attempt to derive bounds for dispersion in a setting similar to that of Ottaviani and Sorensen (2010), it is not assessed directly whether strategic forecasting affects the results of this study.

In some sense the same caveat applies to Harris and Raviv (1993) and Kandel and

[3]As an example, Figure 12 shows that dispersion in forecasts in 2009 is 0.25. Since in light of the current data set $n = 14$, σ_{min} above becomes: $\sigma_{min} = 2\sqrt{14} \times 0.25 = 1.87$. Because the average coefficient of variation for all earnings in the sample was determined to be 1.76 in Chapter 3, observed levels of dispersion in 2009 are either too high to be supported by private information acquisition, or the volatility of earnings in the aggregate is measured incorrectly. These results become even stronger when the restrictions on December fiscal year ends, as well as the full availability of earnings histories are enforced. Dispersion in 2009 remains close to the value above, and is 0.245, whereas the coefficient of variation for earnings drops to 1.35, implying greater exceedance. Note however that this is just an example - as discussed at length previously, the aggregate data used in Figure 12 is not necessarily directly comparable with theory, and neither does it necessarily correctly identify n.

Pearson (1995); since this study does not attempt to derive maximum bounds for dispersion in a setup exactly equal to either one of the studies, it is difficult to make explicit statements about whether their models can support the observed *magnitude* of forecast dispersion. Nevertheless, it is possible that heterogeneous *processing* of public information happens in reality, but whether it affects a model similar to the one in this study is left open.

The results of this study also warrant a discussion on how they relate to Johnson (2004). First, the conclusion that dispersion is related to uncertainty fits well with Johnson (2004). The likelihood of the candidate explanation by Johnson (2004) in driving the results here however, is somewhat difficult to assess. First, the continuous time framework in Johnson (2004) makes a direct mapping difficult, and neither is that the objective of this study. Secondly, Johnson (2004) is explicit about not modeling the expectation formation process. As such, it is difficult to assess to what bounds, if any, dispersion in his model has to conform. Johnson (2004) does mention that "Conversely, they [the forecasters] could all place enormous confidence in their own estimates while differing wildly from each other", but this is mentioned in a more general discussion, not in context of the model. Obviously, such behavior is not possible in the setup of this study, and the mere existence of a maximum for the dispersion in forecasts, rules out such effects, depending on the definition "wildly".

On a more qualitative level there are similarities between Johnson (2004) and the sketch of an explanation presented in this study. The theory in Johnson (2004) rests on the idea of (unpriced) parameter risk, with a guiding notion that asset values are unobservable. The explanation provided here, where agents must restart their learning procedures qualitatively rests on a similar idea, since this is hypothesized to occur when firms earnings processes undergo pre-unknowable shifts. In this sense one could think of asset values or firm fundamentals as being unobservable, or at least perhaps, unlearnable. Thus, while potential mechanisms might differ, the main take away is the same in this study as in that of Johnson (2004) - dispersion is related to uncertainty.

The Johnson (2004) model, however, has some further implications that are more difficult to reconcile with the empirical evidence that dispersion is related to losses, evidence strongly supported both in this study and to various extents in Adut et al. (2008), Ali et al. (2009) and Xu and Zhao (2010). A particular implication derives from the following, found in Johnson (2004): "[T]o the extent that dispersion of earnings expectations is under the control of firms themselves, they might actually benefit, via a lower cost of equity capital, by increasing disagreement." In a simplistic case where negative earnings are the only driver behind disagreement, firms knowing the relation between losses and dispersion, could at least in this sense benefit from reporting losses in order to increase the equity value of the firm. Such an effect seems unlikely.

5.3 Discussion and implications for pricing

If increased dispersion implies increased uncertainty and the pathway for the uncertainty increase is through an actual increase in the variance of the DGP, it is of interest to consider potentially resulting asset pricing implications. First, however, it is important to re-emphasize that the learning model analyzed in this study is contained to the expectation formation process, and as such does not generate any formal or direct asset pricing implications. Neither are the empirical estimations asset pricing tests.

If uncertainty instead of asymmetric information drives dispersion, it fits intuitively with Guntay and Hackbarth (2010) where dispersion is associated with higher credit spreads, and Avramov et al. (2009), where dispersion is related to worse credit ratings. Worse credit ratings and higher credit spreads could be interpreted as resulting from increases in uncertainty[4]. A higher credit spread can be interpreted as a higher required return, where the higher return is achieved through a lower price. A similar pattern can be envisioned for stocks, where higher uncertainty would yield price discounts. Avramov et al. (2009) find that in periods of deteriorating credit conditions, firms with low credit ratings experience increases in dispersion and large price drops. Concerning both stocks and bonds, investors would be willing to pay a lower price for the uncertainty involved, since if the expectations of analysts proxy for average investor expectations, the predictions could be assumed to carry over in the sense that where analysts are more uncertain, the average investor becomes more uncertain. Here the implication would be that investors are uncertainty-averse, or that risk and uncertainty have a one to one correspondence. The above pricing implications are, however, problematic, since results derived from evidence in this

[4]It is worth emphasizing that neither of the papers actually use the asymmetric information explanation and instead both can be seen to derive the uncertainty premise from Johnson (2004). Guntay and Hackbarth (2010) in fact suggest that *uncertainty* about future cash flows is the culprit behind their result. Avramov et al. (2009) also lean towards a similar explanation. Avramov et al. (2009) argue that financial distress drives the dispersion effect, and that periods of financial distress are associated with rising uncertainty about firm fundamentals. The point here is to consider which explanation, certainty or uncertainty fits the overall picture better. Also, the fact that the uncertainty explanation seems the likelier choice, does not in itself refute an asymmetric information explanation, but since the results of this study reject the asymmetric information explanation, the fit with an uncertainty explanation warrants consideration.

study imply all effects are idiosyncratic. The notion above where investors pay a lower price for uncertainty is thus not a straightforward effect if uncertainty or risk in this setting is not systematic.

The standard asset pricing prediction is that priced risk should lead to lower prices (higher expected return). If dispersion proxies for risk, risk averse investors would be willing to pay less for assets with higher dispersion. If the risk is priced, price discounts arise due to prices being at levels such that the correctly priced assets offer compensation for the risks in the form of higher expected returns. This is the rationale behind the initial hypothesis in Diether et al. (2002) - if dispersion proxies for priced risk, dispersion should be related to higher returns. The Diether et al. (2002) finding is the exact opposite - dispersion is related to lower returns.

The question that arises is whether dispersion is a manifestation of risk, uncertainty or both and if there are different idiosyncratic and systematic elements at play. The theoretical results of this study are derived purely on a firm level, and thus models the uncertainties of agents in an idiosyncratic setting. The evidence of this study shows that stocks with high dispersion are associated with future losses, as evidenced also in Adut et al. (2008). The conclusion of this study is also that (increases in) dispersion indicate increased uncertainty about the future of the firms' earnings, occurring through a proposed pathway of increased variance or uncertainty for the earnings distribution. This is qualitatively very similar to Guntay and Hackbarth (2010), who advocate that "dispersion appears to proxy largely for future cash flow uncertainty in corporate bond markets". If negative earnings indicate true increases in *idiosyncratic* levels of uncertainty, then it is possible to have lower associated returns. In fact, Johnson (2004) notes that an opposite risk return dynamic is indeed possible when the risk is idiosyncratic - expected returns decrease with the level of idiosyncratic risk. Furthermore, the evidence in Ali et al. (2008) suggests that the Diether et al. (2002) dispersion lower return finding goes away after controlling for the relation of future earnings.

A potential explanation that unifies the evidence of this study with that of the evidence in the literature, is one where sudden increases in uncertainty[5] force idiosyncratic elements to affect price adjustments. The reasoning is as follows. A temporary increase in idiosyncratic uncertainty about a firm's fundamentals first

[5]If increases in uncertainty are modeled as changes in the DGP, it poses problems for standard asset pricing models since these usually require time series to be stationary. The stationarity of time series is also a requirement for standard rational asset pricing to hold, since investors price assets on the basis of their risks and expected returns. In order to be able to compute correct risks and returns, distributions will generally have to be stable so that information can be summarized by the first moments of the distribution.

leads to higher dispersion as per the explanation proposed in this study. This increase in uncertainty occurs because the firm itself undergoes an unexpected change. If the uncertainty, indicated by dispersion, leads to subsequent future losses (and permanent lower earnings), the value of the firm decreases. The dispersion - negative future earnings association is confirmed by the results of this study, as well as those in Ali et al. (2009) and Adut et al. (2008). Because of the increased uncertainty, lower earnings and overall worsened situation of the firm, the stock price declines. The declining prices in the worsened environment is confirmed by Avramov et al. (2009). Sudden price declines should also by construction materialize as a lower return. Consequently, the dispersion-return relation first documented by Diether et al. (2002) is not surprising. A simultaneous worsened credit spread as in Guntay and Hackbarth (2010), and credit downgrades as in Avramov et al. (2009) are in some ways expected effects for abruptly worsened firm specific conditions. Finally, once the pricing adjustment from the increase in idiosyncratic uncertainty is concluded, learning and uncertainty stabilize and result in standard pricing predictions applying.

The explanation for the observed magnitudes of dispersion in forecasts that this study provides, where breaks in the earnings series cause increases in (idiosyncratic[6] [7]) uncertainty and the restarting of estimation procedures, is in some sense only a somewhat crude sketch of a potential mechanism. The fact that the simple Bayesian model taking into account the convergence on common information cannot yield observed levels of dispersion that are high enough, might be due to simply an over simplified modelling setup. Perhaps more intricate ways of handling expectations, or even more importantly models that encompass richer probability dynamics for the DGP in models such as rational beliefs (Kurz, 1994), could provide a formal solution to the problem, encompassing predictions an all phenomena relating to dispersion

[6]Note that the reason for discussing increases in uncertainty (changes in the DGP) as an idiosyncratic effect is that the model used in the study is idiosyncratic (in terms of information. Recall that the model is *not* an asset pricing model). Potential structural changes might well have a systematic element, in the sense that firms to different extents load on what would be interpreted as a risk factor. The problem with this is however twofold. First, from a theoretical point of view, having a systematic uncertainty increase across firms would potentially require a different information structure, since agents could be learning about a firm's earnings process through information about other firms. Secondly, the Diether et al. (2002) result very much suggests that there is indeed a systematic (cross sectional) pricing *effect*, but this effect is difficult to interpret as a *risk factor*, precisely because the relation has the wrong sign.

[7]However, since the main regressions in Diether et al. (2002) utilize short return horizons, it is not impossible that the negative dispersion return relation being picked up is temporary uncertainty shocks. Increases in dispersion are followed by large price declines (Avramov et al., 2009) and at least on a firm level, such an effect will show a short term negative return by construction. Whether the downward adjusted prices following uncertainty and dispersion increases are rewarded with a higher return over subsequent longer horizons, possibly supporting a risk factor explanation, is left open.

and uncertainty.

The above does however *not* alter the fact that standard conclusions of private information yielding increases in dispersion, which are a direct consequence of the underlying model, are shown not to hold in this study. The results of this study *do* show that private information cannot explain observed levels of forecast dispersion. Consequently, the private information acquisition assumption that implies increased certainty when dispersion increases is rejected by the joint evidence of this study. While there is no attempt made at horse racing the alternative mechanism presented in this study against potential explanations from other theories, the alternative explanation does find support in the data.

Generally, in what is correctly pointed out by Johnson (2004), we do not have data on how confident forecasters are (subjectively) about their estimates, and data on forecast dispersion cannot *in itself* come to the rescue. Conclusions from dispersion data are thus always made in conjunction with a theory of how beliefs are formed. This is as true in this study, as in others using data on forecast dispersion and prediction on the direction of subjective uncertainty (increases/decreases) is assessed indirectly, in conjunction with an underlying model. Interestingly however, there exists unique data linking dispersion to subjective uncertainty in a study using the NBER-ASA survey of economic forecasters by Zarnowitz and Lambros (1987) and in Bomberger (1996) using Livingston survey data. Johnson (2004) summarizes the link between dispersion and uncertainty resulting from the studies of Zarnowitz and Lambros (1987) and Bomberger (1996) in that while caution in warranted in interpreting the results of the two above studies, the intuition linking dispersion to uncertainty is on solid ground.

In conclusion, the survey evidence by Zarnowitz and Lambros (1987) and Bomberger (1996), the anecdotal evidence from analysts' reports indicating increased uncertainty, the explanations in Guntay and Hackbarth (2010), and Avramov et al. (2009), and the joint evidence of this study where the increased certainty hypothesis is clearly rejected, all strongly support the case for dispersion in forecasts in reality being linked to future uncertainty.

5.4 Contributions

This final section of the study summarizes and makes explicit the various contributions that this study makes. These contributions are the following:

1) On a very general level, this study points out the limitations of employing the asymmetric information Bayesian Gaussian learning model as an explanation for disagreement in practice. The main result is that given an assumed signal distribution (DGP) with a quantifiable fixed variance, only a handful of observations (signal realizations) leads to the maximum magnitude of disagreement being constrained to only a fraction of the variance of the DGP.

The result of the monotonically decreasing maximum amount of disagreement, along with a notably small magnitude of disagreement in relation to the variance of the DGP, can be seen as a combination of two elements. One element, the rate of convergence on multiple realizations from an informative signal distribution, can be traced back to a representation of Bayesian updating in standard textbook expositions, such as in Gelman et al. (2004, pp. 49). The second element is that theoretically disagreement attains its maximum when private information equals common information, as shown by Barry and Jennings (1992). The contribution lies in combining these two elements, yielding a model that accounts for both effects simultaneously. Such a combination is justified by an aspiration for realism and the notion of Brown (1993), who implicitly suggests that learning occurring around earnings announcements is connected over multiple periods. The implications of the model are monotonically declining maximum levels of disagreement and disagreement levels that at maximum can only be a fraction of the variance of the assumed signal distribution. The model and its implications are contributing in the following ways:

-Constructing a model explicitly considering both elements, the rate of convergence of learning and the maximum levels of disagreement, yields implications and restrictions on disagreement that are previously unexplored.

-A thorough analysis of the model implications, monotonically declining maximum levels of disagreement and theoretically supported levels of disagreement that are small in relation to the assumed DGP, illustrate the limitations of using the asymmetric information Gaussian Bayesian learning model as model for attaining realistic levels of disagreement.

-The implications from the model not only offer a contribution to the study of dispersion in financial analysts' forecasts, but offer a general theoretical contribution to the use of Bayesian models of asymmetric information in economic modelling - disagreement resulting from asymmetric information is small, and can at most be a fraction[8] of the uncertainty or variance of commonly observed information.

[8]The size of the fraction being guided by the number of realizations form the (fixed) informative

2) The results of the study contribute directly to the literature in Accounting Research, particularly to that concerned with dispersion in analysts' forecasts. By taking the purely theoretical results in 1) and proxying the variance of the DGP with observed firm level historical variance, empirical estimations show that observed levels of dispersion do not find theoretical support. In showing that observed levels of forecast dispersion are too high to be supported by a standard Bayesian learning model under asymmetric information augmented by an explicit consideration of public information dissemination, the study challenges the use of increases in forecast dispersion as evidence of increased information asymmetry alone. The combined theoretical and empirical evidence of the study instead suggest alternative explanations are needed to explain both increases and levels of forecast dispersion. Where the standard interpretation of increased dispersion resulting from increases in private information by construction must imply increases in certainty, the study instead suggests that increases in dispersion may well be driven by increases in subjective *uncertainty*.

While there exists studies that invoke a notion of uncertainty as an explanation for dispersion in forecasts such as Guntay and Hackbarth (2010), Avramov et al. (2009) and Johnson (2004), this study is the first to explicitly *prove* that the opposite explanation for forecast dispersion, asymmetric information, is in reality not possible.

3) Finally, the results of this study can be easily employed by researchers wishing to assess whether asymmetric information alone can drive disagreement. By assuming that the DGP remains fixed, and the variance of the DGP can be reasonably proxied, the study contributes by offering a simple formula for assessing whether observed levels of dispersion are compatible with a theoretical explanation that encompasses asymmetric information as the driver behind increases in forecast dispersion.

---

distribution.

References

Abarbanell, J., Lanen, W., Verrecchia, R., 1995. Analysts' forecasts as proxies for investor beliefs in empirical research. Journal of Accounting and Economics, 20, 31-60.

Adut, D., Sen, P. K., Sinha, P., 2008. Properties of the Variance of Analysts' Forecast of Earnings in Good-News and Bad-News Environments: Theory, Evidence and Usefulness. Working paper series.

Ali, A., Liu, M.H., Xu, D., Yao, T., 2009. Corporate Disclosure, Analyst Forecast Dispersion, and Stock Returns. Working paper series.

Ang, A., R. J. Hodrick, Y. Xing, and X. Zhang, 2006, The Cross-Section of Volatility and Expected Returns, Journal of Finance, 61(1), 259-299.

Avramov, D., T. Chordia, G. Jostova, and A. Philipov, 2009. Dispersion in Analysts Earnings Forecasts and Credit Rating. Journal of Financial Economics 91, 83-101.

Bali, T., Cakici, N., 2008. Idiosyncratic Volatility and the Cross Section of Expected Returns. Journal of Financial and Quantitative Analysis, Vol. 43, No. 1, 29-58.

Barron, O., Kim, O., Lim, S., Stevens, D., 1998. Using analysts' forecasts to measure properties of analysts' information environment. The Accounting Review, 73, 421-433.

Barron, O., Stanford, M., Yu, Y., 2009. Further evidence on the relation between analysts forecast dispersion and stock returns, Contemporary Accounting Research 26, 329 - 357.

Barry, C., Jennings, R., 1992. Information and diversity of analyst opinion, Journal of Financial and Quantitative Analysis 27, 169-183.

Berger, J., 1985. Statistical decision theory and Bayesian Analysis (2nd ed.). Springer-Verlag, New York.

Bomberger, W., 1996. Disagreement as a measure of uncertainty. Journal of Money, Credit, and Banking 28, 381392.

Bradshaw, M.T., Drake, M.S., Myers, J.N., Myers, L.A., 2012. A re-examination of analysts superiority over time-series forecasts of annual earnings. Review of Accounting Studies 17, 944968.

Brown, L., Han, J., 1992. The Impact of Annual Earnings Announcements on Convergence of Beliefs. The Accounting Review, Vol. 67, No. 4, 862-875

Brown, L. D., 1993. Earnings forecasting research: Its implications for capital markets research. International Journal of Forecasting 9: 295-320.

Brown, L., 2000. I/B/E/S research bibliography, 6th edition. New York: I/B/E/S International.

Chamley, C., 2004. Rational Herds. Cambridge University Press.

Cochrane, John H., 2005. Asset Pricing (Revised Edition), Princeton University Press.

Cragg, J., Malkiel, B., 1968. The Consensus and Accuracy of Some Predictions of the Growth of Corporate Earnings. Journal of Finance, Vol. 23, iss. 1, p.67-84

DeGroot, M., 1970. Optimal Statistical Decisions. McGraw Hill, New York.

Diether, K., Malloy, C., Scherbina, A., 2002. Differences of opinion and the cross-section of stock returns. Journal of Finance 57, 2113-2141

Duru, Augustine, and David M. Reeb, 2002, International diversification and analysts' forecast accuracy and bias, The Accounting Review 77, 415-433.

Gelman, A., Carlin, J. B., Stern, H., S., Rubin, D., B., 2004. Bayesian Data Analysis, Second Edition. Chapman and Hall/CRC Texts in Statistical Science.

Givoly, D., Lakonishok, J., 1984. Properties of analysts' forecasts of earnings: A review and analysis of the research. Journal of Accounting Literature, 3, 117-152.

Graham, B., Dodd, D. L., 1934. Security Analysis, McGraw Hill.

Guntay, L., Hackbarth, D., 2010. Corporate bond credit spreads and forecast dispersion. Journal of Banking and Finance, 34, 23282345

Harris, M., Raviv, A.,1993. Differences of opinion make a horse race, Review of

Financial Studies 6, 473-506.

Hwang, L., Jan, C., Basu, S., 1996. Loss firms and analysts' earnings forecast errors. Journal of Financial Statement Analysis 1, 18-30.

Johnson, T. C., 2004. Forecast Dispersion and the Cross Section of Expected Returns. The Journal of Finance 59 (5), 1957-1978.

Kandel, E., Pearson, N.D., 1995. Differential interpretation of public signals and trade in speculative markets. Journal of Political Economy 103, 831-872.

Keane, M.P., Runkle, D.E., 1990. Testing the rationality of price forecasts: new evidence from panel data. American Economic Review 80, 714-735.

Keane, M.P., Runkle, D.E., 1998. Are financial analysts forecasts of corporate profits rational? Journal of Political Economy 106, 768-805.

Kurz, M., 1994. On the structure and diversity of rational beliefs. Economic Theory Volume 4, Issue 6, pp 877-900

Lang, M., Lundholm, R., 1996. Corporate disclosure policy and analyst behavior. The Accounting Review 71 (October): 467-492.

Miller, E., 1977. Risk, uncertainty and divergence of opinion. Journal of Finance 32, 1151-1168.

Minton, B. A., Schrand, C. M., Walther, B. R., 2002. The Role of Volatility in Forecasting. Review of Accounting Studies, 7, 195215.

Morse, D., Stephan, J., Stice, E. K. 1991. Earnings announcements and the convergence (or divergence) of beliefs. The Accounting Review 66, 376-88.

Muth, J.F., 1961. Rational expectations and the theory of stock price movements. Econometrica 29, 315-335.

Ottaviani, M., Sorensen, P. N., 2006. The strategy of professional forecasting. Journal of Financial Economics 81, 441-466.

Park, C., 2005. Stock return predictability and the dispersion in earnings forecasts. Journal of Business 78 (6), 2351-2376.

Prendergast, C., Stole, L., 1996. Impetuous youngsters and jaded oldtimers: acquiring a reputation for learning. Journal of Political Economy, 104, 1105-1134.

Raiffa, H., and R. Schlaifer., 1961. Applied Statistical Decision Theory. Cambridge, MA. The M.I.T. Press

Ramnath, S., Rock, S., Shane, P., 2008. The financial analyst forecasting literature: A taxonomy with suggestions for further research. International Journal of Forecasting, 24, 3475.

Särkkä, S., 2013. Bayesian Filtering and Smoothing. Cambridge University Press.

Satterthwaite, F. E. 1946. An approximate distribution of estimates of variance components. Biometrics Bulletin, 2, 110-114.

Verrecchia, R. 1983. Discretionary disclosure. Journal of Accounting and Economics 5 (3): 179-194.

White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica, 48, 817838.

Xu, Y., Zhao, Y., 2010. Re-assess the Analyst Forecast Dispersion Puzzle. Working paper series.

Yu, J., 2011. Disagreement and return predictability of stock portfolios, Journal of Financial Economics 99, 162-183.

Zarnowitz, V., Lambros, L.A., 1987. Consensus and uncertainty in economic prediction, Journal of Political Economy 95, 591-621.

Appendix A

Consider a 2 period case. There are N agents (indexed by i:j,k,l,...,N) who forecast an unknown earnings variable θ. At $t = t_0$, an arbitrary point in time, there exists only common information. Prior beliefs at $t = t_0$, about θ are summarized by $\theta \sim N(\mu_0, \tau_0^2)$. At $t = t_1$, new information (common) arrives in the form of an (annual) earnings release. This is modeled as a signal, observed by all agents, informative on θ, and is parameterized as $y \sim N(\theta, \sigma^2)$.[1] Agents now update their beliefs in accordance with equation (1) as:

$$\mu_{t_1} = E[\theta|y] = \frac{\frac{1}{\tau_0^2}\mu_0 + \frac{1}{\sigma^2}y_{t_1}}{\frac{1}{\tau_0^2} + \frac{1}{\sigma^2}}.$$

Notice the use of μ_{t_1} directly, instead of the individual forecasts $u_{t_1 i}$: since the information in the *realization*, y_{t_1}, is observed by all, everyone updates their expectation equally.

At $t = t_2$, agents gain access to private information. Private information is introduced in the standard fashion; that is, as a signal, z_i, informative on θ. In particular, $z_i \sim N(\theta, \nu_i^2)$. The agent now observes its *realization*, $z_{t_2 i}$ and updates his expectation of θ according to:

$$u_{t_2 j}(\neq u_{t_2 k}) = E[\theta|z_i] = \frac{\frac{1}{\tau_1^2}\mu_{t_1} + \frac{1}{\nu_i^2}z_{t_2 i}}{\frac{1}{\tau_1^2} + \frac{1}{\nu_i^2}}.$$

Here, the posterior expectation, after having observed the signal at $t = t_2$, is expressed in terms of the prior at $t = t_1$.[2] In terms of the initial prior from $t = t_0$, the expectation takes the form:

$$u_{t_2 i} = E[\theta|z_i] = \frac{(\frac{1}{\tau_0^2} + \frac{1}{\sigma^2})\left(\frac{\frac{1}{\tau_0^2}\mu_0 + \frac{1}{\sigma^2}y_{t_1}}{\frac{1}{\tau_0^2} + \frac{1}{\sigma^2}}\right) + \frac{1}{\nu_i^2}z_{t_2 i}}{\frac{1}{\tau_0^2} + \frac{1}{\sigma^2} + \frac{1}{\nu_i^2}}$$

$$u_{t_2 i} = \frac{\frac{1}{\tau_0^2}\mu_0 + \frac{1}{\sigma^2}y_{t_1} + \frac{1}{\nu_i^2}z_{t_2 i}}{\frac{1}{\tau_0^2} + \frac{1}{\sigma^2} + \frac{1}{\nu_i^2}} \tag{28}$$

[1]Here, the analysis does not take a stand on how σ^2 and τ_0^2 are related.

[2]Notice that generally, the variables in the conditioning set are indeed variables and therefore not indexed by time. This makes the proposition that agents have to have knowledge of the distributional properties of their signals explicit.

It is informative to use the formulation above, which shows the dependence of agents i's time $t = t_2$ expectation, $u_{t_2 i}$, on the realization of y at time $t = t_1$.

At both $t = t_0$ and $t = t_1$, μ_0 and μ_{t_1} respectively do not depend on i, as all agents make the same forecast. Consequently, dispersion is given by $d_{t_0} = \frac{1}{N-1}\sum_{i=1}^{N}(u_{t_0 i} - \bar{u}_{t_0})^2 = 0$ and $d_{t_1} = \frac{1}{N-1}\sum_{i=1}^{N}(u_{t_1 i} - \bar{u}_{t_1})^2 = 0$. At $t = t_2$ however, dispersion will be present since realizations of private signals will lead to $u_{t_2 j} \neq u_{t_2 k}$. It is therefore the case that: $d_{t_2} = \frac{1}{N-1}\sum_{i=1}^{N}(u_{t_2 i} - \bar{u}_{t_2})^2 \geq 0$.

To see why diversity cannot be directly predicted by information that is public ,consider a particular time $t = t_1$ realization of y, $\underline{y}_{t_1}$:

$$u_{2i} = E[\theta|z_i] = \frac{\frac{1}{\tau_0^2}\mu_0 + \frac{1}{\sigma^2}\underline{y}_{t_1} + \frac{1}{\nu_i^2}z_{t_2 i}}{\frac{1}{\tau_0^2} + \frac{1}{\sigma^2} + \frac{1}{\nu_i^2}}$$

The realization $\underline{y}_{t_1}$ has already occurred at time $t = t_1$, and is part of the prior of agent i, who at time $t = t_2$ observes a realization $z_{t_2 i}$ from $z_i \sim N(\theta, \nu_i^2)$. Recall also that y (or in this case $\underline{y}_{t_1}$) is a realization from $y \sim N(\theta, \sigma^2)$. Thus, $\underline{y}_{t_1}$ and $z_{t_2 i}$ are *realizations* from two *independent* probability distributions, and any realization from one distribution cannot predict realizations from the other.

Another way of seeing this is by noticing that d_{t_2} is a function of $u_{t_2 i}$ (the differences between the i's being the driver behind the variation) and that the differences in $u_{t_2 i}$'s are determined by realizations of z_i's . Realizations from z_i, are however unpredictable, as they are driven by the white noise process $\varepsilon_i \sim N(0, \nu_i^2)$. As the single driver behind variations in d_{t_2} is the white noise process ε_i, it implies that d_{t_2} is itself unpredictable. Therefore, it cannot be the case that any realizations, or sets of realizations from y could predict d_{t_2}.*Q.E.D.*

Appendix B

This section shows how to derive the expression for the dispersion in forecasts, through aggregating over all agents' individual expectations.

The general form for the variance of u_i is:

$$var(u_i) = \frac{1}{N-1}\Bigg[\left(\frac{h\mu_0 + s_j z_j}{h+s_j} - \frac{1}{N}\left(\frac{h\mu_0 + s_j z_j}{h+s_j} + \frac{h\mu_0 + s_k z_k}{h+s_k} + ... + \frac{h\mu_0 + s_N z_N}{h+s_N}\right)\right)^2$$

$$+\left(\frac{h\mu_0 + s_k z_k}{h+s_k} - \frac{1}{N}\left(\frac{h\mu_0 + s_j z_j}{h+s_j} + \frac{h\mu_0 + s_k z_k}{h+s_k} + ... + \frac{h\mu_0 + s_N z_N}{h+s_N}\right)\right)^2 + ...$$

$$... + \left(\frac{h\mu_0 + s_N z_N}{h+s_N} - \frac{1}{N}\left(\frac{h\mu_0 + s_j z_j}{h+s_j} + \frac{h\mu_0 + s_k z_k}{h+s_k} + ... + \frac{h\mu_0 + s_N z_N}{h+s_N}\right)\right)^2\Bigg].$$

The dispersion in forecasts is the variance of individual forecasts. By assuming just one private signal distribution (but different realizations), it is possible to derive an expression for the dispersion in forecasts:

$$var(u_i) = \frac{1}{N-1}\Bigg[\left(\underbrace{\frac{h\mu_0 + s z_j}{h+s} - \underbrace{\frac{1}{N}\underbrace{\left(\frac{h\mu_0 + s z_j}{h+s} + \frac{h\mu_0 + s z_k}{h+s} + ... + \frac{h\mu_0 + s z_N}{h+s}\right)}_{(iv)}}_{(v)}}_{(vi)}\right)^2 +$$

$$+\left(\frac{h\mu_0 + s z_k}{h+s} - \frac{1}{N}\left(\frac{h\mu_0 + s z_j}{h+s} + \frac{h\mu_0 + s z_k}{h+s} + ... + \frac{h\mu_0 + s z_N}{h+s}\right)\right)^2 + ...$$

$$... + \left(\frac{h\mu_0 + s z_N}{h+s} - \frac{1}{N}\left(\frac{h\mu_0 + s z_j}{h+s} + \frac{h\mu_0 + s z_k}{h+s} + ... + \frac{h\mu_0 + s z_N}{h+s}\right)\right)^2\Bigg]$$

Combining terms in (iv) gives (iv):

$$\frac{h\mu_0 + sz_j + h\mu_0 + sz_k + ... + h\mu_0 + sz_N}{h+s}$$

$$\frac{s(z_j + z_k + ... + z_N) + N \times h\mu_0}{h+s}$$

$$\frac{s(\sum_{i=1}^{N} z_i) + N \times h\mu_0}{h+s}$$

Multiplying in $1/N$ gives (v):

$$\frac{s(\frac{1}{N}\sum_{i=1}^{N} z_i) + h\mu_0}{h+s}$$

Where $\frac{1}{N}\sum_{i=1}^{N} z_i$ is the mean of the signal realizations. Now consider (vi):

$$\frac{h\mu_0 + sz_j}{h+s} - \left(\frac{s(\frac{1}{N}\sum_{i=1}^{N} z_i) + h\mu_0}{h+s}\right)$$

$$\frac{h\mu_0 + sz_j - s(\frac{1}{N}\sum_{i=1}^{N} z_i) - h\mu_0}{h+s}$$

This is an important step since it can now be seen that dependence on (all) prior information drops out from the expression for the variance of u_i, through the elimination of $h\mu_0$. (vi) thus becomes:

$$\frac{sz_j - s(\frac{1}{N}\sum_{i=1}^{N} z_i)}{h+s}$$

The variance expression now looks like:

$$var(u_i) = \frac{1}{N-1}\Bigg[\left(\frac{sz_j - s(\frac{1}{N}\sum_{i=1}^{N} z_i)}{h+s}\right)^2 + \left(\frac{sz_k - s(\frac{1}{N}\sum_{i=1}^{N} z_i)}{h+s}\right)^2 + ...$$

$$... + \left(\frac{sz_N - s(\frac{1}{N}\sum_{i=1}^{N} z_i)}{h+s}\right)^2\Bigg]$$

$$var(u_i) = \frac{1}{N-1}\left[\frac{s^2(z_j - \frac{1}{N}\sum_{i=1}^{N} z_i)^2}{(h+s)^2} + \frac{s^2(z_k - \frac{1}{N}\sum_{i=1}^{N} z_i)^2}{(h+s)^2} + ...\right.$$

$$\left.... + \frac{s^2(z_N - \frac{1}{N}\sum_{i=1}^{N} z_i)^2}{(h+s)^2}\right]$$

Sum over the squared differences from the mean of the realizations. The realizations are indexed by $i = j, k, l, ..., N$, so we thus have $\sum_{i=1}^{N}(z_i - \frac{1}{N}\sum_{i=1}^{N} z_i)^2$, yielding the following expression for the variance:

$$var(u_i) = \frac{1}{N-1}\left[\frac{s^2[(z_j - \frac{1}{N}\sum_{i=1}^{N} z_i)^2 + (z_k - \frac{1}{N}\sum_{i=1}^{N} z_i)^2 + ... + (z_N - \frac{1}{N}\sum_{i=1}^{N} z_i)^2]}{(h+s)^2}\right]$$

$$var(u_i) = \frac{1}{N-1}\left[\frac{s^2[\sum_{i=1}^{N}(z_i - \frac{1}{N}\sum_{i=1}^{N} z_i)^2]}{(h+s)^2}\right]$$

Now multiply in $\frac{1}{N-1}$ and observe that now the right-hand tem in the numerator is the expression for the variance of z_i:

$$var(u_i) = \frac{s^2 \overbrace{\left[\frac{1}{N-1}\sum_{i=1}^{N}(z_i - \frac{1}{N}\sum_{i=1} N z_i)^2\right]}^{var(z_i)}}{(h+s)^2}$$

Recalling that the variance of z_i was given by ν_i^2 and furthermore that $\nu_i^2 = 1/s_i$ (or in this case $\nu^2 = 1/s$) , it is the case that:

$$var(u_i) = \frac{s^2\nu^2}{(h+s)^2} = \frac{s^2\frac{1}{s}}{(h+s)^2}$$

Which finally yields:

$$var(u_i) = \frac{s}{(h+s)^2}$$

This final result is the same as the one given in Eq. 19 in Barron et al. (1998).

Appendix C

This section expands the analysis of forecast dispersion from Appendix B, and derives an expression for the variance of forecasts in a case where agents are not assumed to have equal priors and receive both private and commonly observed signals. This corresponds to a period following a period where agents already received (one, or a sequence of identical) private signals.

Again, only one signal distribution is assumed (but different realizations across all agents). Before proceeding, notice that the first term in (vi), $(hu_j + sz_j + ry)/(h + s + r)$, actually implies[1]:

$$\frac{h_{(n)}u_{(n-1)j} + s_{(n)}z_{(n)j} + ry}{h_{(n)} + s_{(n)} + r} \tag{29}$$

but time subscripts are suppressed for notational clarity throughout the derivations. The rationale for $u_{(n-1)j}$ referring to the previous period, $n-1$, is that the prior that agents use are their posteriors from the previous periods. At this point, the priors, $u_{(n-1)i}$, are not specified in any way, and thus the analysis is general. Thus, the sub-indices, j, refer to the same agent but are temporally separated, and by assumption, uncorrelated (This obviously applies analogously to the other terms as well).

The expression for the variance is now given by:

$$var(u_n i) =$$

[1]In regards to the model in "The Model", where the $n = 2$ forecast is analyzed, this would imply: $u_{2i} = (h_{(2)}u_{(1)i} + s_{(2)}z_{(2)i} + ry)/(h_{(2)} + s_{(2)} + r)$

$$= \frac{1}{N-1}\Bigg[\left(\underbrace{\frac{hu_j+sz_j+ry}{h+s+r} - \underbrace{\frac{1}{N}\underbrace{\left(\frac{hu_j+sz_j+ry}{h+s+r} + \frac{hu_k+sz_k+ry}{h+s+r} + ... + \frac{hu_N+sz_N+ry}{h+s+r}\right)}_{(iv)}}_{(v)}}_{(vi)}\right)^2$$

$$+\left(\frac{hu_k+sz_k+ry}{h+s+r} - \frac{1}{N}\left(\frac{hu_j+sz_j+ry}{h+s+r} + \frac{hu_k+sz_k+ry}{h+s+r} + ... + \frac{hu_N+sz_N+ry}{h+s+r}\right)\right)^2$$

$$...+\left(\frac{hu_N+sz_N+ry}{h+s+r} - \frac{1}{N}\left(\frac{hu_j+sz_j+ry}{h+s+r} + \frac{hu_k+sz_k+ry}{h+s+r} + ... + \frac{hu_N+sz_N+ry}{h+s+r}\right)\right)^2\Bigg]$$

Combining terms in (iv) gives (iv):

$$\frac{hu_j + sz_j + ry + hu_k + sz_k + ry + ... + hu_N + sz_N + ry}{h+s+r}$$

$$\frac{s(z_j + z_k + ... + z_N) + h(u_j + u_k + ... + u_N) + N \times ry}{h+s+r}$$

$$\frac{s(\sum_{i=1}^{N} z_i) + h(\sum_{i=1}^{N} u_i) + N \times ry}{h+s+r}$$

Multiplying in $1/N$ gives (v):

$$\frac{s(\frac{1}{N}\sum_{i=1}^{N} z_i) + h(\frac{1}{N}\sum_{i=1}^{N} u_i) + ry}{h+s+r}$$

Where $\frac{1}{N}\sum_{i=1}^{N} z_i$ is the mean of the signal realizations and $\frac{1}{N}\sum_{i=1}^{N} u_i$ is the mean of the priors, containing past signals . Now consider (vi):

$$\frac{hu_j + sz_j + ry}{h+s+r} - \left(\frac{s(\frac{1}{N}\sum_{i=1}^{N} z_i) + h(\frac{1}{N}\sum_{i=1}^{N} u_i) + ry}{h+s+r}\right)$$

$$\frac{hu_j + sz_j - s(\frac{1}{N}\sum_{i=1}^{N} z_i) - h(\frac{1}{N}\sum_{i=1}^{N} u_i)}{h+s+r}$$

Whereas in Appendix B, all dependence on the prior dropped out of the numerator, this is not the case here[2] and (vi) looks like :

$$\frac{sz_j - s(\frac{1}{N}\sum_{i=1}^{N} z_i) + hu_j - h(\frac{1}{N}\sum_{i=1}^{N} u_i)}{h+s+r}$$

The variance expression now looks like:

$$var(u_n i) = \frac{1}{N-1}\Bigg[\left(\frac{sz_j - s(\frac{1}{N}\sum_{i=1}^{N} z_i) + hu_j - h(\frac{1}{N}\sum_{i=1}^{N} u_i)}{h+s+r}\right)^2 + \left(\frac{sz_k - s(\frac{1}{N}\sum_{i=1}^{N} z_i) + hu_k - h(\frac{1}{N}\sum_{i=1}^{N} u_i)}{h+s+r}\right)^2 + ... + \left(\frac{sz_N - s(\frac{1}{N}\sum_{i=1}^{N} z_i) + hu_N - h(\frac{1}{N}\sum_{i=1}^{N} u_i)}{h+s+r}\right)^2\Bigg]$$

[2] notice however that all dependence of the last common signal, y, drops out of the numerator. The variance remains affected by the common information however, through the scaling by r.

$$var(u_n i) = \frac{1}{N-1}\Bigg[\underbrace{\left(\frac{[sz_j - s(\frac{1}{N}\sum_{i=1}^{N} z_i)] + [hu_j - h(\frac{1}{N}\sum_{i=1}^{N} u_i)]}{h+s+r}\right)^2}_{(vii)}$$

$$+\left(\frac{[sz_k - s(\frac{1}{N}\sum_{i=1}^{N} z_i)] + [hu_k - h(\frac{1}{N}\sum_{i=1}^{N} u_i)]}{h+s+r}\right)^2 + ...$$

$$...+\left(\frac{[sz_N - s(\frac{1}{N}\sum_{i=1}^{N} z_i)] + [hu_N - h(\frac{1}{N}\sum_{i=1}^{N} u_i)]}{h+s+r}\right)^2\Bigg]$$

Expanding (vii), we have:

$$\frac{[sz_j - s(\frac{1}{N}\sum_{i=1}^{N} z_i)]^2 + 2[sz_j - s(\frac{1}{N}\sum_{i=1}^{N} z_i)][hu_j - h(\frac{1}{N}\sum_{i=1}^{N} u_i)] + [hu_j - h(\frac{1}{N}\sum_{i=1}^{N} u_i)]^2}{(h+s+r)^2}$$

$$\frac{s^2[z_j - (\frac{1}{N}\sum_{i=1}^{N} z_i)]^2 + 2sh[z_j - (\frac{1}{N}\sum_{i=1}^{N} z_i)][u_j - (\frac{1}{N}\sum_{i=1}^{N} u_i)] + h^2[u_j - (\frac{1}{N}\sum_{i=1}^{N} u_i)]^2}{(h+s+r)^2}$$

Using (vii) in the expression for the variance we have:

$$var(u_n i) = \frac{1}{N-1}\Bigg[$$

$$\frac{s^2[z_j - (\frac{1}{N}\sum_{i=1}^{N} z_i)]^2 + 2sh[z_j - (\frac{1}{N}\sum_{i=1}^{N} z_i)][u_j - (\frac{1}{N}\sum_{i=1}^{N} u_i)] + h^2[u_j - (\frac{1}{N}\sum_{i=1}^{N} u_i)]^2}{(h+s+r)^2}$$

$$+\frac{s^2[z_k - (\frac{1}{N}\sum_{i=1}^{N} z_i)]^2 + 2sh[z_k - (\frac{1}{N}\sum_{i=1}^{N} z_i)][u_k - (\frac{1}{N}\sum_{i=1}^{N} u_i)] + h^2[u_k - (\frac{1}{N}\sum_{i=1}^{N} u_i)]^2}{(h+s+r)^2}+..$$

$$...+\frac{s^2[z_N - (\frac{1}{N}\sum_{i=1}^{N} z_i)]^2 + 2sh[z_N - (\frac{1}{N}\sum_{i=1}^{N} z_i)][u_N - (\frac{1}{N}\sum_{i=1}^{N} u_i)] + h^2[u_N - (\frac{1}{N}\sum_{i=1}^{N} u_i)]^2}{(h+s+r)^2}\Bigg]$$

Since we are summing over all agents (realizations) $i = j, k, l, ..., N$, the variance can be expressed as:

$$var(u_i) = \frac{1}{N-1}\Bigg[\frac{s^2\sum_{i=1}^{N}[z_i - (\frac{1}{N}\sum_{i=1}^{N} z_i)]^2}{(h+s+r)^2}$$

$$+\frac{2sh\sum_{i=1}^{N}[z_i - (\frac{1}{N}\sum_{i=1}^{N} z_i)][u_i - (\frac{1}{N}\sum_{i=1}^{N} u_i)]}{(h+s+r)^2}$$

$$+\frac{h^2\sum_{i=1}^{N}[u_i - (\frac{1}{N}\sum_{i=1}^{N} u_i)]^2}{(h+s+r)^2}\Bigg]$$

Multiplying in $\frac{1}{N-1}$, and using the fact that u_i is actually u_{n-1i}:

$$var(u_n i) = \Bigg[\frac{s^2 \overbrace{[\frac{1}{N-1}\sum_{i=1}^{N}(z_i - \frac{1}{N}\sum_{i=1}^{N} z_i)]^2}^{var(z_i)}}{(h+s+r)^2} + \frac{2sh \overbrace{[\frac{1}{N-1}\sum_{i=1}^{N}(z_i - \frac{1}{N}\sum_{i=1}^{N} z_i)(u_i - \frac{1}{N}\sum_{i=1}^{N} u_i)]}^{covar(z_i, u_{n-1i})}}{(h+s+r)^2} + \frac{h^2 \overbrace{[\frac{1}{N-1}\sum_{i=1}^{N}(u_i - \frac{1}{N}\sum_{i=1}^{N} u_i)^2}^{var(u_{n-1i})}}{(h+s+r)^2} \Bigg]$$

Which for simplicity can be written:

$$var(u_n i) = \frac{s^2[var(z_i)]}{(h+s+r)^2} + \frac{2sh[covar(z_i, u_{n-1i})]}{(h+s+r)^2} + \frac{h^2[var(u_{n-1i})]}{(h+s+r)^2}$$

First note that the covariance term is equal to zero as long as signals are drawn independently[3].

Thus, for this analysis, where independence is assumed, we have:

$$var(u_n i) = \frac{s^2[var(z_i)] + h^2[var(u_{n-1i})]}{(h+s+r)^2}$$

,putting back all relevant time subscripts, we have the final result:

$$var(u_n i) = \frac{s_n^2[var(z_{ni})] + h_n^2[var(u_{n-1i})]}{(h_n+s_n+r)^2} \tag{30}$$

,where the subscripts on s result from the fact that this is the latest signal, received at time n.

[3]This is not the same as having correlated draws at each period, rather here the assumption is only that signals are independent across time.

Appendix D

This section derives an expression for the evolution of forecast dispersion *over time*. Time subscripts on the private signal are omitted since the signal is assumed to be constant over time and thus $s_1 = s_2 = ... = s_n = s$. The indices on the variables are the ones referring to the period, labeled n-indexing, described in "The Model" in detail.

Agents make their initial forecasts from information contained in their common prior and individual signal realizations (this is labelled n=1). The study now considers what happens at n=2 and onwards.

Starting with the result preceding the final result for the variance of the forecasts from Appendix C (and writing $u_{1i} = u_i$) , for $n = 2$ we have:

$$var(u_2 i) = \frac{s^2[var(z_i)] + h^2[var(u_i)]}{(h + s + r)^2}$$

Recalling that the variance of z_i was given by ν^2 and furthermore that $\nu^2 = 1/s$, the variance expression becomes:

$$var(u_2 i) = \frac{s^2 \frac{1}{s} + h^2[var(u_i)]}{(h + s + r)^2} = \frac{s + h^2[var(u_i)]}{(h + s + r)^2}$$

The question is: what is the variance of the prior (u_i)? One might be tempted to plug in the variance for the private signal, $1/s$, as the variance of the prior is driven by the variance of the signal. This is not correct however, since the extent to which the variance of the prior is affected by the variance of the private signal in the previous (first) period, will depend on the weights that were placed on the signal and the prior in the previous (first) period.

The answer lies in observing that the variance of the prior in the expression above, or expressed in another way: the *uncertainty* of the prior at $n = 2$, is given by the variance of the *posteriors* from the previous (first) period (where the prior was a constant). The cross sectional variance of the posteriors (forecasts) is given exactly by the expression for the variance of the forecasts from Appendix C: $var(u_i) = \frac{s}{(h_1+s)^2}$, as agents will, in a "closed" context, use their posteriors from the previous period as their prior in the current period (Gelman et al., 2004). Note that h_1 denotes the precision on commonly available information at the outset/first/beginning period.

Now, put time-subscripts on the precision of the prior in the expression for the variance of the forecasts at $n = 2$ ($var(u_2 i)$):

$$var(u_2 i) = \frac{s + h_2^2[var(u_i)]}{(h_2 + s + r)^2}$$

Next, substitute $var(u_i)$ with the expression for the variance, or uncertainty, from the first period,

$$var(u_i) = \frac{s}{(h_1 + s)^2}$$

Yielding:

$$var(u_2 i) = \frac{s + h_2^2(\frac{s}{(h_1+s)^2})}{(h_2 + s + r)^2}$$

Since precisions add linearly[1], assuming that the previous period was the first, consisting of a common prior and a private signal part, we can express the precision at $n = 2$ in terms of the precision and the signal at $n = 1$: $h_2 = h_1 + s$. Substituting:

$$var(u_2 i) = \frac{s + (h_1 + s)^2(\frac{s}{(h_1+s)^2})}{(h_1 + s + s + r)^2}$$

$$var(u_2 i) = \frac{2s}{(h_1 + 2s + r)^2}$$

The final step is noticing that that the weight placed on the common signal y, is r and is in this case equal to h_1, since y is in this case a (commonly observed) signal, and agents will weight it by its perceived precision h_1, obtained from the historical record. Thus $r = h_1$, so we have that:

$$var(u_2 i) = \frac{2s}{(h_1 + 2s + h_1)^2} = \frac{2s}{(2h_1 + 2s)^2}$$

[1]Look at e.g. Eq. (2), or Gelman et al. (2004). $\tau_1^2 = \left(\frac{1}{\tau_0^2} + \frac{1}{\sigma^2}\right)^{-1} \Leftrightarrow \frac{1}{\tau_1^2} = \frac{1}{\tau_0^2} + \frac{1}{\sigma^2}$. Using precisions: $h_1 = h_0 + s$. Note however that this is in terms of t-indexing, which in terms of n-indexing is $h_2 = h_1 + s$.

Analogously, for n=3, we have:

$$var(u_3 i) = \frac{s + h_3^2[var(u_2 i)]}{(h_3 + s + r)^2}$$

$$var(u_3 i) = \frac{s + h_3^2(\frac{2s}{(2h_1+2s)^2})}{(h_3 + s + r)^2}$$

and using $h_3 = h_2 + s + r = h_1 + s + s + r = h_1 + s + s + h_1 = 2h_1 + 2s$, we have:

$$var(u_3 i) = \frac{s + (2h_1 + 2s)^2(\frac{2s}{(2h_1+2s)^2})}{(2h_1 + 2s + s + h_1)^2}$$
$$var(u_3 i) = \frac{3s}{(3h_1 + 3s)^2}$$

Simple forward substitution then gives the main result:

$$var(u_n i) = \frac{ns}{(nh_1 + ns)^2}$$

Using variances instead of precisions:

$$var(u_n i) = \frac{n\frac{1}{\nu^2}}{(n\frac{1}{\tau^2} + n\frac{1}{\nu^2})^2}$$

Re-arranging finally yields:

$$var(u_n i) = \frac{(\tau^2)^2\nu^2}{n(\tau^2 + \nu^2)^2}$$

, where τ^2 refers to the initial starting prior.

It was noted earlier that $r = h_1$, so we can also express the main result in terms of r:

$$var(u_n i) = \frac{ns}{(nr + ns)^2} \tag{31}$$

, implying that that the variance distribution used is the observable historical variance σ^2, so we have:

$$var(u_n i) = \frac{(\sigma^2)^2\nu^2}{n(\sigma^2 + \nu^2)^2}$$

Appendix E

This section attempts to illustrate more intuitively how the variable DTSdirect ($\sqrt{V}/\sigma$) performs under different assumptions for σ. In the example, there is a negative earnings realization at $t = T$. As usual, this increases both dispersion in forecasts at $t = T$, as well as $\sigma_{t>T}$.

The measures in the figure below are described in event time. There is a negative earnings realization occurring at $t = T$.

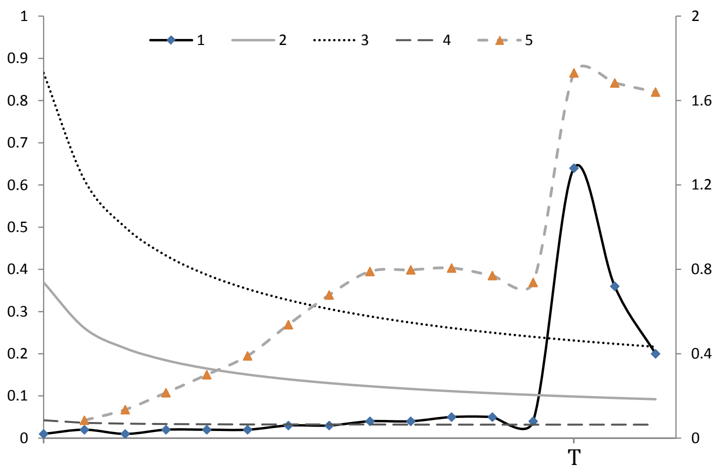

Figure 17

1. $\sqrt{V}$, is the observed standard deviation of forecasts.

2. $(\sqrt{V}/\sigma_{(T-1)})_{max}$. In the empirical estimations, this is the *threshold value* for the maximum bound, as a function of n, corresponding to the measuring of DTSdirect in the study. It compares the standard deviation in forecasts with the lagged value of the (recursive) standard deviation of earnings, which in turn is assumed to constitute σ. Here however, since the analysis only pertains to one company, the bound $(\sqrt{V}/\sigma_{(T-1)})_{max}$ is actually multiplied through by $\sigma_{(T-1)}$, so that in the Figure, the bound is directly in units of the standard deviation of forecasts (earnings). The maximum (threshold) value for *dispersion*, $\sqrt{V}_{max}$, (not the bound) at T is approximately 0.10. The calculation resulting in the value is described below.

Using Eq. (18) , where V_{max} is shorthand for $var_{max}(u_{ni})$, $V_{max} = 1/4nr \Leftrightarrow \sqrt{V}_{max} = \sqrt{1/4nr} \Leftrightarrow \sqrt{V}_{max}/\sqrt{1/r} = \sqrt{1/4n}$. Since $(1/r) = \sigma^2$ it follows that $\sqrt{V}_{max}/\sigma = \sqrt{1/4n}$, or in the notation of the study: $(\sqrt{V}/\sigma)_{max} = \sqrt{1/4n}$.

At $t = T$, real n is equal to 13 and consequently the maximum value for DTSdirect (maximum threshold value for $n = 13$) is thus 0.1387. Through multiplying by $\sigma_{(T-1)}$, which here equals 0.74, we get 0.10 for the absolute maximum magnitude for the dispersion in forecasts, $\sqrt{V}$, given that $\sigma_{(T-1)}$ corresponds to agents' information sets.

Obviously, the standard deviation of forecasts, $\sqrt{V}$ (line 1), heavily exceeds this maximum value at $t = T$. The value of $\sqrt{V}$ at $t = T$ equals 0.64.

3. $(\sqrt{V}/\sigma_{(t>T)})_{max}$. This is the corresponding implied threshold using forward looking information. However, due to the break in the earnings series occurring at T, conditioning on $\sigma_{(t>T)}$ for $t < T$ is suspect. Again, in the Figure, the bound $(\sqrt{V}/\sigma_{(t>T)})_{max}$ is multiplied by $\sigma_{(t>T)}$ so that it is in the same units as the standard deviation of forecasts (earnings) for the company.

4. $(\sqrt{V}/\hat{\sigma})_{max}$. The notation here is somewhat off. This is actually the maximum level in absolute dispersion, calculated through assuming that the estimated σ at each step is the noise of the signal. A better notation would be $(\sqrt{V})_{max|\hat{\sigma}}$, or $(\sqrt{V} \mid \hat{\sigma})_{max}$, where $\hat{\sigma}$ is understood to refer to the standard deviation being estimated separately at each step.

5. σ estimated on a rolling, lagged basis, as in the study.

Fredrik Le Bell

The Time Series Convergence of Dispersion in Financial Analysts' Forecasts

Financial analysts play an important role in capital markets as information intermediaries. In filtrating information, resulting in earnings forecasts, analysts generally tend to disagree. This thesis focuses on the disagreement between financial analysts.

Previous research and data indicate that when a company reports losses, analysts start disagreeing more about the future earnings of that company. Intuition would suggest that this is due to more uncertainty. Anecdotal evidence from analysts earnings reports corroborates this intuition, finding analysts more uncertain around negative earnings, precisely where disagreement tends to increase.

However, the theoretical models for belief formation that lay the mathematical foundations for this thesis, incorporate a somewhat strange implication - If analysts start disagreeing more, it can only mean they become more certain. In the theoretical setup, one that is used extensively in the literature, it is only asymmetric information that can give rise to increased disagreement.

In order to resolve the certainty/uncertainty contradiction, this thesis shows that a model taking into account the public information flow in earnings announcements over time, can produce only small levels of disagreement between analysts, levels of disagreement that are too small to encompass observed levels of disagreement.

As a result, this thesis concludes that the theoretical models used in the literature for explaining analyst disagreement, as such seem insufficient, and increases in disagreement could instead be interpreted as increased uncertainty, in accordance with evidence from analysts' reports.

The evidence in this thesis contributes to the Accounting literature, since many studies employ these models the other way around, in that when an increase in disagreement in empirical data is observed, the observed disagreement is thought to signify an increase in asymmetric information. Extensive reliance on the underlying models obscures our understanding of the uncertainty dynamics around e.g. earnings announcements. Earnings announcements are paramount for price discovery in practice, and are also extensively studied in the Accounting literature. The results in this thesis indicate that conclusions in other studies regarding increases in disagreement resulting from information asymmetry might be premature.

9 789517 657624 >

Åbo Akademi University Press | ISBN 978-951-765-762-4