MEDIA REPORTS AND INFLATION EXPECTATIONS

Dissertation
zur Erlangung des wissenschaftlichen Doktorgrades (Dr. rer. pol.)
an der Fakultät Wirtschafts- und Sozialwissenschaften der
Universität Hamburg

vorgelegt von

JAN-OLIVER MENZ

17. Juli 2014

Erstgutachter: Prof. Dr. Ulrich Fritsche
Zweitgutachter: Prof. Dr. Jan-Egbert Sturm

Preface

This dissertation was written while I was a research assistant in the DFG-Project "Inflation Expectation Formation and Information Transmission on Households' Expectations: Stickiness, Agenda-Setting and Uncertainty" of Prof. Dr. Ulrich Fritsche at the University of Hamburg. I thank Ulrich Fritsche for providing excellent research advise and constructive critique of earlier drafts of my research. I enjoyed the opportunity to work in an environment which has been marked by a highly inspiring and pleasant atmosphere. I thank Ulrich Fritsche for giving me the opportunity to attend various research conferences and summer schools, as well as for being able to gain experience in teaching economics. Finally, I appreciate his continuing patience regarding the longer than expected finalization of the dissertation.

A number of colleagues and friends have contributed a great deal to this dissertation. Above all, I thank Lena Dräger for excellent research collaboration, helpful comments and discussions, and for being the kind of office-mate a Ph.D. student has to have in order to be able to write his dissertation. I am also extremely grateful to various comments and support by Jana Görner, Artur Tarassow and Martin Sauber. Furthermore, I am indebted to Eva Arnold, Roberta Colavecchio, Nadja König, Katharina Glass, Marcel Garz, Ingrid Größl and Bernd Lucke. Torsten Osigus, Ulrich Brandt and Phillip Poppitz have provided very excellent research assistance throughout the years. Finally, Claudia Ranft has done a great job in organizing many administrative issues.

I am indebted to the German Research Foundation (DFG) for financial support and to Media Tenor for providing the data on news coverage.

Finally, this dissertation would not have been finished without the love and support of Cindy Hayes.

Frankfurt am Main, December 2013 *Jan-Oliver Menz*

Contents

List of Figures

List of Tables

Summary

This dissertation explores the various links between news media coverage of inflation and the inflation expectations of households. Since the beginning of 2000, a number of alternative models of expectation formation have been proposed seeking to overcome "the limits of rational expectations" (Pesaran, 1987). A common feature of these new approaches consists in relaxing an important assumption of the rational expectations paradigm: that households use the latest available information set when forming beliefs about the future. Throughout this dissertation, we will thus test which kind of information households rely on when forecasting inflation, focusing in particular on the role of the news media.

In the first chapter, we provide a brief overview of some models of expectation formation that are relevant for our analysis. In addition, we present the concept of agenda setting used by researchers in communication studies to analyze the news media. In Chapter (3), we then test the epidemiology model of expectations proposed by Carroll (2003) in great detail. Since this model suggests a direct impact of news media coverage on household inflation expectations, it seems to be well suited for answering our research questions. Using survey data on inflation expectations in the U.S. over the period 1980-2011, and news coverage of inflation in *The New York Times*, we provide empirical evidence supporting the epidemiology model. Households are found to adjust their beliefs to the average inflation forecast of experts, whereas the speed of adjustment rises in line with the number of news reports on inflation. The speed of updating varies significantly over time: households rely more on experts in periods of low inflation and during economic crises. Applying our analysis using both macro and micro survey data on expectations, we find that the news media effect is larger on the micro level. Looking at households with different news perceptions, we find that those who claim to have heard news on inflation commit larger forecast errors than other households while at the same time being more receptive to media reports. Finally, our results suggest that the media effect is nonlinear: An increasing number of news reports increases the impact from expert expectations, whereas the adjustment takes place only gradually and depends on a threshold level of news reports.

The next chapter applies the framework of the epidemiology model to different household groups and news media sources. Using German data from 1999-2010, we try to explain the stylized fact that households disagree considerably in their beliefs on future prices depending on their socioeconomic background. For example, low-income or unemployed households are often found to commit larger forecaster errors than high-income households. We test the hypothesis that these differences emerge from socioeconomic news exposure, meaning that households belonging to different socioeconomic groups read different newspapers. And since the media differ in the extent and the way they cover economic topics such as inflation, the information set of their corresponding readers will differ. Constructing an index

of newspaper coverage and TV coverage, we indeed observe considerable heterogeneity in news consumption across income, age and occupation groups. Furthermore, we find that constructing an index of news reports by aggregating all available newspaper and TV reports can be misleading. Coverage of inflation in *Tagesschau*, Germany's most influential TV evening news show, is found to increase the gap between households and professional forecasters, while a rising number of articles published in *BILD*, Germany's most prominent tabloid, brings households closer to the best available forecast. Finally, it is important to distinguish between the effects of a rise in the number of news reports and a change in the journalists' judgment of inflation. Whereas households' expectation gaps increase if *BILD* presents inflation in a negative way thereby possibly inducing a media bias, more negative coverage in *Tagesschau* narrows the gap between households and professional forecasters.

In the final chapter, we extend the framework of the epidemiology model by including the number of Google search requests of inflation. This measure can be understood as a proxy for the demand of information in the sense that households will search for inflation on the web if they need do know more about the current or future price environment. Internet search data could also serve as a complement to inflation expectations measured by surveys. Whereas surveys suffer from the "cheap talk"-problem arising from the fact that respondents do not have an incentive to provide their best forecast, households will only search for inflation if they really want to use this information. Using U.S. data from 2005-2011, we find that the number of Google search requests reacts in a meaningful way to fundamental economic data. Google users distinguish between headline and core inflation and they react asymmetrically: the demand for information increases if core inflation falls whereas in periods of historically high inflation rates, the number of search requests is significantly larger. Estimating various Vector Autoregressive Models, we find that households' inflation forecasts are driven by TV reports, newspaper articles, and Google search requests, while the feedback effect from expectations on web searches is rather small and estimated less precisely. About 20% of the forecast error variance decomposition of households' inflation expectations can be explained by Google search requests.

Zusammenfassung

Die vorliegende Dissertation untersucht die verschiedenen Wechselwirkungen zwischen Medienberichterstattung über Inflation und den Inflationserwartungen von Haushalten. Seit dem Beginn der 2000er Jahre sind einige alternative Erwartungsbildungsmodelle vorgeschlagen worden die das Ziel haben, "die Grenzen von rationalen Erwartungen" (Pesaran, 1987) zu überwinden. Ein gemeinsames Merkmal dieser neuern Modelle besteht darin, eine wichtige Annahme des rationalen Erwartungsbildungsparadigmas aufzugeben, wonach Haushalte immer alle aktuell verfügbaren Informationen verwenden um Einschätzungen über die Zukunft vorzunehmen. In den einzelnen Kapiteln dieser Dissertation werden wir daher testen, auf welche Informationen sich Haushalte beziehen wenn sie Erwartungen über die zukünftige Inflation bilden, wobei ein besonderes Augenmerk auf die Rolle der Medien gelegt wird.

Im ersten Kapitel geben wir zunächst einen kurzen Überblick über die verschiedenen Erwartungsbildungsmodelle die für unsere Untersuchung relevant sind. Außerdem beschreiben wir das "Agenda-Setting"-Konzept, das häufig in kommunikationswissenschaftlichen Studien verwendet wird, um die Rolle der Medienberichterstattung zu untersuchen. In Kapitel (3) untersuchen wir anschließend detailliert das "Epidemiologie"-Modell von Carroll (2003). Dieses Modell scheint deshalb besonders für die Analyse unseres Untersuchungsgegenstandes geeignet, da es einen direkten Einfluss der Medienberichterstattung auf die Inflationserwartungen von Haushalten ableitet. Mit Hilfe von Umfragedaten zu Inflationserwartungen in den USA über den Zeitraum 1980-2011, sowie Daten zur Medienberichterstattung über Inflation in der *New York Times* zeigen wir, dass das "Epidemiologie"-Modell durchaus von den Daten gestützt wird. Haushalte passen ihre Erwartungen an die Meinungen von Experten an, wobei die Anpassungsgeschwindigkeit mit der Anzahl der Medienberichte über Inflation ansteigt. Außerdem zeigt sich, dass die Anpassungsgeschwindigkeit nicht immer gleich ist: Haushalte beziehen sich stärker auf Experten in Zeiten niedriger Inflation sowie während der Finanzkrise. Indem wir unsere Analyse sowohl mittels Makroals auch mittels Mikroumfragedaten durchführen, können wir zeigen, dass sich auf der Mikroebene stärkere Medieneffekte finden lassen. Unterscheidet man Haushalte nach ihrer individuellen Informationswahrnehmungen, so lässt sich feststellen, dass Individuen die angeben, zuletzt Neuigkeiten über Inflation gehört zu haben, einem größeren Prognosefehler unterliegen und außerdem stärker auf Medienberichte reagieren. Außerdem deuten unsere Ergebnisse daraufhin, dass der Medieneffekt nichtlinear wirkt: Mit steigender Anzahl an Medienberichten über Inflation erhöht sich der Einfluss der Experten auf die Erwartungsbildung der Haushalte, wobei die Anpassung nur langsam von statten geht und außerdem vom durchschnittlichen Niveau der Berichterstattung abhängt.

Im nächsten Kapitel wird das "Epidemiologie-Modell" auf verschiedene Haushaltsgruppen

und Medien angewandt. Unter Verwendung von Umfragedaten in Deutschland im Zeitraum 1999-2010 versuchen wir ein wiederkehrendes Muster in Umfragen zu erklären, das darin besteht, dass sich die Inflationserwartungen je nach sozioökonomischem Hintergrund der Befragten stark unterscheiden. Zum Beispiel ist oft zu beobachten, dass Niedrigeinkommensbezieher oder Arbeitslose größere Prognosefehler begehen als Haushalte die zu höheren Einkommensgruppen zählen. Wir testen die Hypothese dass sich die beobachteten sozioökonomischen Unterschiede in den Inflationserwartungen dadurch erklären lass, dass sich der Medienkonsum verschiedener Haushaltsgruppen unterscheidet. Und da sich die Medien im Ausmaß sowie der Art und Weise der Berichterstattung über Inflation unterscheiden, führt dies dazu, dass Haushalte ihre Erwartungen auf Basis unterschiedlicher Informationen bilden. Anhand der Berechnung eines Index der Berichterstattung über Inflation in Zeitungen und Fernsehen können wir zeigen, dass sich der Medienkonsum in der Tat zwischen Einkommens-, Alters- und Berufsgruppen unterscheidet. Außerdem belegen unsere Ergebnisse, dass die Verwendung eines aus mehreren Einzelmedien aggregierten Medienindex irreführend sein kann. Berichterstattung über Inflation in der *Tagesschau*, Deutschlands wichtigster Nachrichtensendung, führt dazu, dass Haushalte in ihren Erwartungen stärker von Experten abweichen, während eine Ausweitung der Berichterstattung in *BILD*, Deutschlands meistgelesenem Boulevardmedium, die Haushaltserwartungen den Expertenprognosen annähert. Schließlich ist es wichtig, zwischen den Auswirkungen einer größeren Anzahl von Medienberichten und einer Veränderung in der Einschätzung der verantwortlichen Journalisten zu unterscheiden. Während sich die Erwartungslücke der Haushalte erhöht wenn *BILD* die Inflationsentwicklung stark negativ darstellt, so führt eine negativere Berichterstattung in der *Tagesschau* dazu, dass sich die Haushaltserwartungen den Expertenmeinungen annähern.

Im letzten Kapitel erweitern wir den Ansatz des "Epidemiologie"-Modells indem wir die Anzahl der Googlesuchanfragen nach Inflation einbeziehen. Googlesuchanfragen können als Proxy für die Informationsnachfrage von Nutzern interpretiert werden, unter der Annahme dass Haushalte dann im Internet nach Informationen über Inflation suchen wenn sie mehr über die derzeitige oder zukünftige Preisentwicklung wissen müssen. Internetsuchdaten lassen sich daneben auch als Ergänzung zu durch Umfragen gemessenen Inflationswertungen verstehen. Während die Qualität von Umfragen unter dem „cheap talk"-Problem leiden, das dadurch entsteht, dass Umfrageteilnehmer keinen Anreiz haben, ihre bestmögliche Inflationsschätzung anzugeben, so werden Haushalte nur nach Informationen im Internet suchen, wenn sie diese auch wirklich nutzen wollen. Mittels U.S.-Daten von 2005 bis 2011 zeigen wir, dass die Anzahl der Googlesuchanfragen in sinnvoller Art und Weise auf ökonomische Fundamentaldaten reagiert. Googlenutzer unterscheiden zwischen Gesamtinflations- und Kerninflationsrate wobei ihre Reaktion asymmetrisch ist: Die Informationsnachfrage geht zurück wenn die Kerninflationsrate fällt, während in Zeiten historisch hoher Inflationsraten die Informationsnachfrage ansteigt. Anhand der Schätzung mehrerer Vektorautoregressiver Modelle finden wir, dass die Inflationserwartungen von Haushalten sowohl von TV-Nachrichten, Zeitungsartikeln als auch von der Zahl der Googlesuchanfragen auch Inflation abhängen, während der Feedbackeffekt von Erwartungen auf die Informationsnachfrage eher gering ist. Ungefähr 20% der prognostizierten Fehlerdekomposition (FEVD) der Inflationserwartungen lassen sich durch die Googlesuchanfragen erklären.

Chapter 1

Introduction

In his seminal paper introducing the concept of "rational expectations" into economic theory, Muth (1961) suggests that "in order to explain fairly simply how expectations are formed, we advance the hypothesis that they are essentially the same as the predictions of the relevant economic theory." Later in his paper, he rephrases this statement: in his point of view, expectations are rational if "the subjective probability distribution of outcomes tends to be distributed, *for the same information set*, about the prediction of the theory (italics added)".

It is the purpose of this dissertation to analyze the condition "for the same information set" in some detail. Again expressed in the words of Muth (1961): "We shall examine the effect (...) of differences in the information possessed by various firms in the industry. Whether such biases in expectations are empirically important remains to be seen." As it has turned out in recent studies, and as we will also emphasize throughout the dissertation, these biases are indeed empirically important.

Policy Implications Modeling the process of expectation formation in an adequate way is important for a number of reasons. The workhorse of modern macroeconomics, the New Keynesian Dynamic Stochastic General Equilibrium (DSGE) Model[1], places expectations at center stage. Consumers form beliefs about the future path of their life-time income and about future price changes. Via the Euler equation, these predictions feed directly into today's consumption and saving decisions. Similarly, firms hold beliefs about future costs, profits and price changes, and set their profit-maximizing price according to these expectations. Following the financial crisis in 2008, the role of expectations has gained further importance. Eggertsson and Woodford (2003) have analyzed the consequences of the zero lower bound on interest rates within the framework of the standard DSGE model concluding that in such a case, the management of expectations ("forward guidance") becomes the key instrument of monetary policy. Schmitt-Grohe and Uribe (2013) try to explain the jobless

[1] See for textbook expositions Walsh (2003), Woodford (2003) and Galí (2008); as well as the seminal paper by Clarida et al. (1999).

growth recovery in Japan and the U.S. arguing that this scenario results from a downward shift in agents' inflation expectations.

The baseline DSGE model assumes that expectations are formed rationally in the sense of Muth (1961). However, as it has been shown by Mankiw and Reis (2002), among others, the policy conclusions drawn from the DSGE model can change substantially if the assumption of rational expectations is relaxed. As an example, they show that disinflation induced by monetary policy always leads to a contraction in output. And Wiederholt (2013) suggests that the policy conclusions derived from the DSGE model including the zero lower bound depend heavily on the assumed process of expectation formation. Therefore, it is of great importance to assess whether rational expectations or alternative theories capture the expectation formation of economic agents in the most sensible way.

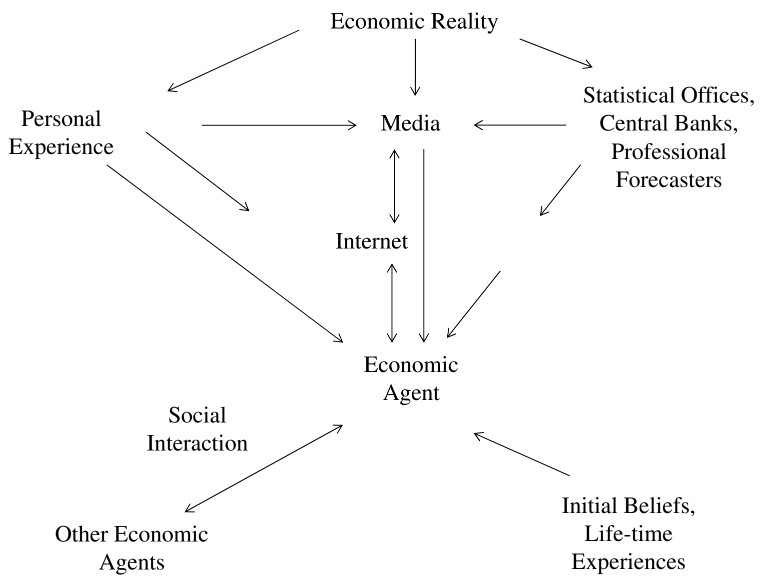
Survey Data One way to test competing models of expectation formation is by use of survey data. As Coibion and Gorodnichenko (2012) have put it: "What can survey forecasts tell us about informational rigidities? - A lot.". Using survey data to test models of expectation formation dates back at least to Turnovsky (1970) but research has intensified only recently due to advances in hardware capacity and the availability of surveys that cover a sufficient period of time. Besides of survey data, expectations could also be measured with the help of financial market data.[2] This has the advantage that the expectations of agents are directly reflected in economic decisions whereas opinions expressed in surveys can suffer from the "cheap talk"-problem meaning that agents do not face any consequences if their responses to a questionnaire are far a away from reality. However, since only a very small fraction of economic agents actively trades on financial markets[3], using such data does not necessarily capture the beliefs of the general public. Therefore, throughout this dissertation, we measure inflation expectations by means of survey data that are representative of the entire population.

[2]The most prominent approach consists of computing "break-even inflation rates". In 1997, the U.S. government introduced "Treasury inflation protected securities" (TIPS) which pay investors an extra dividend if the general price index changes. Hence, the difference between the yields on nominal bonds and the yields on TIPS can be interpreted as investors' inflation expectations even if the difference is also affect by investors' risk assessment and liquidity premium. Still, in a short note, Groen and Middeldorp (2013) show that inflation expectations derived from break-even inflation rates show at least some comovement with expectations derived from household survey data. Note however, that the forecast horizon of financial market expectations typically refers to the long horizon, while survey participants are often asked about their expectations for the following year and the DSGE framework uses expectations for the next period, mostly a quarter. See Schulz and Stapf (2009) and Ejsing et al. (2007) for studies computing market-based inflation expectations for the Euro Area.

[3]Stock market participation rates are typically found to range from 40% in Australia, over 30% in the UK, 25% in the U.S. to below 10% in Germany (Giannetti and Koskinen, 2010).

Information Flows and Media Reports There are various reasons why agents use different information sets when forming expectations about future outcomes of economic variables.[4] Figure (1.1) summarizes the various information flows that determine the information set of economic agents.

Figure 1.1: Information Flows in the Economy



Above all, there is the *economic reality* expressed for example in the development of prices. However, this reality does not have to be the same for each agent. While the inflation rate is designed such that it captures price changes of a consumption bundle that is representative of an average consumer, this rate does not reflect the price development of agents whose consumption decisions deviate considerably from the average.[5] Furthermore, economic reality as measured in official numbers by *statistical institutes* can be different from the reality perceived by agents. Differences can arise from statistical issues: new products enter the representative consumption basket only with some delay, hence, price changes of goods that are already much in demand do not yet enter the official inflation rate.[6] Similarly, agents might weight price changes of goods they encounter in everyday life much more compared

[4] As we will discuss below, individuals could also have similar information sets but differ in the way they proceed this information. This hypothesis, advanced by theories of "rational inattention" does lead to similar outcomes.

[5] As an example, compare the typical consumption bundles of old and young individuals.

[6] For example, in the German consumer price index (CPI), the cost of living only enters via rental contracts that have been signed in the past ("Bestandsmieten"). By contrast, new contracts ("Neuvertragsmieten") that are subject to much larger rent increases are not yet part of the CPI. As a result, households' "real" cost of living is underestimated in the official price data.

to their corresponding weight in the official inflation rate (Dräger et al., 2009). As a result, the information set of agents is also determined by *personal experience*. Furthermore, agents collect information and form believes about the future by talking to friends, family members and colleagues implying that *social interaction* plays an important role in determining households' (inflation) expectations. Next, households' beliefs can be influenced by memorable events that have occurred in the past. Malmendier and Nagel (2013) have shown that American households who have grown up in the 1970s and thus have experienced high inflation rates still expect higher future inflation compared to households whose *life-time experience* does not include this period.

In this dissertation, we test whether relaxing the Muthian "for-the-same-information-set"-condition affects agents' expectation formation. More precisely, we analyze the link between news coverage of inflation in the *media* and the inflation expectations of households. Starting with Carroll (2003), economists have documented that the news media play an important role in shaping households' beliefs about future prices. In Chapters (2) and (3) of this dissertation, we add to this literature by analyzing in detail some open issues about the link between media coverage and inflation expectations. In Chapter (4) we then explore whether households also rely on information from the *internet* when forming beliefs about future prices. Note that the questions and explanations we raise throughout this dissertation do not only apply to the formation of *inflation* expectations, but are also relevant for expectations on future income, interest rates or job security. However, we focus entirely on inflation expectations given their prominent role in macroeconomic models and policy debates.[7]

In the reminder of this introduction, we first provide a brief summary of different expectation formation models in Section (1.1) before describing research in communication studies on the role of the news media for determining the general public's beliefs on economic issues in Section (1.2). Section (1.3) then presents the objectives and the outline of the dissertation.

1.1 Models of Expectations Formation

The importance of expectations and the effects of different models of expectation formation can be illustrated within the famous cobweb model analyzed by Kaldor (1934).[8] In the model, demand D_t is determined negatively by the rate of change of the market price π_t: [9]

[7] Garz (2013) provides a detailed treatment of the links between news coverage of unemployment and households' income and job security expectations. Dräger et al. (2013) explore whether households' expectations of various economic variables taken together fit to the predictions of key economic equations such as the Phillips curve or the Taylor rule.
[8] The following exposition builds on Evans and Honkapohja (2001).
[9] Note that the original model is expressed in price levels P_t. We have chosen, for illustrative purposes, to use price changes instead. The results hold for both formulations.

$$D_t = \alpha_0 - \alpha_1 \pi_t + \varepsilon_{1,t}, \tag{1.1}$$

where α_0, α_1 are fixed parameters and $\varepsilon_{1,t}$ denotes a stochastic process. Next, supply S_t depends positively on the expected rate of price changes π_t^{exp}:

$$S_t = \beta_0 + \beta_1 \pi_t^{exp} + \varepsilon_{2,t}, \tag{1.2}$$

again with fixed parameters β_0, β_1 and a stochastic process $\varepsilon_{2,t}$ different from $\varepsilon_{1,t}$. Since the inflation rate π_t is not yet known at the time when the supply of period t is determined, producers have to rely on expected inflation when deciding on the amount of output they want to supply. Solving the model under the market clearing condition $D_t = S_t$ yields the reduced form

$$\pi_t = \gamma_0 + \gamma_1 \pi_t^{exp} + \nu_t; \text{ where } \gamma_0 = \frac{\alpha_0 - \beta_0}{\alpha_1}; \; \gamma_1 = \frac{-\beta_1}{\alpha_1}; \; \nu_t = \frac{\varepsilon_{1,t} - \varepsilon_{2,t}}{\alpha_1} \tag{1.3}$$

Hence, the rate of inflation π_t depends on the rate of expected inflation π_t^{exp} plus some random error. We will now illustrate the impact of different expectation rules using the reduced form of the Cobweb model.

1.1.1 Adaptive Expectation Formation

The hypothesis of *adaptive expectations* suggests that agents rely on past data to forecast future price changes. In its simplest form, agents take the latest observable inflation rate as their prediction of the future rate of price changes:

$$\pi_t^{exp} = \pi_{t-1} \tag{1.4}$$

Therefore, since agents are assumed to expect that past inflation will also hold at present, the formation of expectations is entirely backward-looking.[10] Using this price formula in the reduced form (1.3), we get the result that inflation is given by an AR(1) process:[11]

$$\pi_t = \gamma_0 + \gamma_1 \pi_{t-1} + \nu_t \tag{1.5}$$

Nerlove (1958) proposes a refinement of the simple formula in equation (1.4):

[10] Since we focus on inflation expectations, we can ignore complications arising from data revisions. In case of GDP where much larger revisions occur, Arnold (2013) has shown that some forecasters focus on the initial release whereas others try to target the latest revision.

[11] Apart from the lack of output, this formulation gives the traditional Phillips Curve.

$$\pi_t^{exp} = \pi_{t-1}^{exp} + \eta\left(\pi_{t-1} - \pi_{t-1}^{exp}\right) \tag{1.6}$$

According to this rule, agents will adjust their forecast from the previous period π_{t-1}^{exp} if their past forecast error $\left(\pi_{t-1} - \pi_{t-1}^{exp}\right)$ has been different from zero. Since equation (1.6) also applies to previous periods, one can express the adaptive expectation formula as:

$$\pi_t^{exp} = \eta \sum_{j=0}^{\infty} (1-\eta)^j \pi_{t-j} \tag{1.7}$$

Hence, expected inflation is given by a weighted average of all past inflation rates where more weight is attached to data points in the recent past. Using adaptive expectations in the Cobweb model yields:

$$\pi_t = \gamma_0 + \gamma_1 \eta \sum_{j=0}^{\infty} (1-\eta)^j \pi_{t-j} + \nu_t \tag{1.8}$$

Hence, under this rule, inflation depends *inter alia* on the full history of price changes.

Muth (1961) and his heirs have criticized the use of adaptive expectation rules for three reasons. First, it has to be answered which of the various rules should actually be applied. Since in the end it is left to the researcher whether to use (1.2), (1.3) or some other variant, the adaptive expectation hypothesis suffers from a great degree of ambiguity. Muth (1961) has pointed out that one might even have to chose a different rule for different markets. Second, if it is really true that agents only use past information to predict future price changes, economists could easily produce better forecasts already by taking into account announced tax changes that will affect future inflation. Third, assuming that economic agents use some sort of dynamic optimization while at the same only employing past data to form expectations has been criticized for its inconsistency.[12] Furthermore, Lucas (1976) has pointed out that predicting the effects of economic policy in models with adaptive expectation formation can be highly misleading since it assumes that agents do not adjust their behavior in response to policy changes.

1.1.2 Rational Expectations

Seeking to solve the problems of adaptive expectations, Muth (1961) has proposed to use the concept of *rational expectations*. According to this hypothesis, agents know exactly how

[12]For example, in modern consumption theory following Modigliani and Brumberg (1954) and Friedman (1957), households are assumed to choose their optimal consumption path depending on their life-time income. Since this assumes a considerable amount of computation capacity and forward-lookingness, it seems odd to assume at the same time that consumers will base their forecast of their life-time income only on past data.

the economy works, i.e. referring to the Cobweb model, they know equations (1.1) - (1.3), the parameter values α_i, β_i and the behavior of the stochastic processes $\varepsilon_{i,t}$. If these conditions are satisfied, one can apply the mathematical expectations operator to the expectation formula:

$$\pi_t^{exp} = E_{t-1}\pi_t \text{ and } \pi_{t+1}^{exp} = E_t\pi_{t+1} \tag{1.9}$$

Note that the mathematical expectation $E_{t-1}\pi_t$ is conditional on the information set available at time $t-1$. Using rational expectations in the Cobweb model yields

$$\begin{aligned} \pi_t &= \gamma_0 + \gamma_1 E_{t-1}\pi_t + \nu_t \\ \Leftrightarrow E_{t-1}\pi_t &= \gamma_0 + \gamma_1 E_{t-1}\pi_t \\ \Leftrightarrow E_{t-1}\pi_t &= (1-\gamma_1)^{-1}\gamma_0, \end{aligned} \tag{1.10}$$

where the second line is computed by taking conditional expectations on both sides and assuming that $\nu_t \sim iid\,(0, \sigma_\nu^2)$. Using (1.10) in (1.3) leads to

$$\pi_t = (1-\gamma_1)^{-1}\gamma_0 + \nu_t \tag{1.11}$$

Hence, under rational expectations, inflation depends on a constant term plus a random process. There is no impact from past inflation rates. Due to its internal consistency, rational expectations have become the benchmark approach among the different models of expectation formation.

1.1.3 Learning Models

In the aftermath of the "rational expectations revolution" , economists generally lost interest in developing and testing different models of expectation formation. As Manski (2004) has put it: "Rather than speculate on how expectations actually are formed, they follow convention and assume rational expectations." However, the assumption of rational expectations has already been criticized in the 1980s, mainly because it requires that agents possess a deep knowledge of the economy and are both able and willing to conduct the necessary computations leading to the rational expectations forecast. In response to this critique, Evans and Honkapohja (2001) have proposed a learning approach to the formation of expectations which is based on weaker assumptions than the rational expectations formula. In the learning approach, agents are supposed to act like econometricians. This is motivated by the fact that economists themselves do not know the true model of the economy and the exact parameter values of single equations such as the reduced form of the

Cobweb model. Instead, they use data and econometric techniques to find the best estimate. Therefore, according to Evans and Honkapohja (2001), the hypothesis of rational expectations implicitly assumes that economic agents have more information on how the economy works than trained economists. In the learning approach, by contrast, agents also have to estimate the true parameters of the model.

In the literature, various learning rules have been proposed. If agents know that inflation behaves roughly as in equation (1.11), i.e., that it depends on a constant plus a random error, but if they do not know exactly the values of γ_0 and γ_1, they have to estimate it. The most natural way to estimate the constant is by use of the sample mean gained from a series of past observations on prices. Thus, expected inflation will be given by

$$\pi_t^{exp} = \frac{1}{t}\sum_{i=0}^{t-1} \pi_i \tag{1.12}$$

Under this rule, the solution of the Cobweb model becomes

$$\pi_t = (1 - \hat{\gamma}_1)^{-1}\,\hat{\gamma}_0 + \nu_t, \tag{1.13}$$

where $\hat{\gamma}_0, \hat{\gamma}_1$ denote the parameter values estimated by "agents acting like econometricians." Note that the solution of the learning approach converges to the rational expectation solution if $\gamma_1 < 1$. In addition to the simple mean estimate, Evans and Honkapohja (2001) introduce a least squares learning rule that applies if inflation also depends on exogenous variables such as, for example, the interest rate. In this case, the reduced form of the Cobweb model in (1.3) is transformed into

$$\pi_t = \gamma_0 + \gamma_1 \pi_t^{exp} + \gamma_2 i_{t-1} + \nu_t \tag{1.14}$$

In order to find the true parameter values of this equations, agents will run a least squares regression of π_t on the interest rate i_{t-1} and a constant. Despite its milder assumptions, learning approaches still demand a great deal of computational ability of agents especially if one thinks through the various complications that can arise when estimating equations similar to (1.14).

1.1.4 Sticky Information and the Epidemiology of Expectations

About a decade ago, Mankiw and Reis (2002) proposed the *sticky information* approach to the process of expectation formation. As they put it: "The essence of the model is that information about macroeconomic conditions diffuses slowly through the population." As a result, agents do not always act on the latest available information set but ignore new data, either because searching and acquiring information is costly or because processing and using this information to produce a good forecast is costly. Costs can arise in terms of time, effort or money. According to the sticky information approach, in each period, only a fraction λ of agents receives the latest observable information set and computes the rational expectation forecast. The remaining fraction $(1-\lambda)$ does not update its information set but sticks to its forecast made in previous periods.[13] Under the assumption of sticky information, agents' expectations can be expressed as

$$\pi_t^{exp} = \lambda \sum_{i=0}^{\infty} (1-\lambda)^i E_{t-1-i}\pi_t \tag{1.15}$$

Note that if all agents use the latest available information, i.e. if $\lambda = 1$, we get the rational expectations formula in equation (1.9). Interestingly, the possibility of sticky information has already been mentioned by Muth (1961). On page 321, he briefly analyzes the case "that some of the firms have access to later information than the others", yielding the expression

$$\pi_t^{exp} = \lambda \varepsilon_{t-1} + \sum_{i=2}^{\infty} \varepsilon_{t-i}, \tag{1.16}$$

where λ again is the fraction of firms that has access to the latest available information, whereas the remaining firms can only use information up to period $t-2$.

Using the sticky information formula (1.15) in the reduced form of the Cobweb model, we get[14]:

$$\pi_t = \gamma_0 + \gamma_1 \lambda \sum_{i=0}^{\infty} (1-\lambda)^i E_{t-1-i}\pi_t + \nu_t \tag{1.17}$$

Note that in contrast to the solution with rational expectations, where inflation is determined by current expectations about future prices, the sticky information variant models current inflation as a function of past expectations.

Next, it is important to note that the fraction of rational consumers λ does not have to be fixed over time. Modeling the microeconomic foundation of the sticky information model, Reis (2006) shows that the optimal length of inattentiveness d_t^* is given by

$$d_t^* = \frac{1}{r} \ln \left(1 + \sqrt{\frac{4K}{\alpha \sigma_Y^2}}\right) \tag{1.18}$$

Hence, agents are more rational the higher the real interest rate r, the higher the volatility of

[13] This is the same mechanism introduced by Calvo (1983) in the context of staggered price setting behavior of firms.

[14] This expression is similar to the sticky information Phillips curve derived by Mankiw and Reis (2002).

the income shock σ_Y^2, the higher the coefficient of risk aversion α, and the lower the costs of processing information K.

In an alternative version of the sticky information approach, Carroll (2003) suggests that the costs of processing information depend negatively on the amount of news coverage. In his *epidemiology model of expectation formation*, the inflation expectations of agents are partly determined by the best available forecast and partly by agents' past expectations:

$$\pi_t^{exp} = \lambda(MEDIA_t)\pi_t^{exp,prof} + (1 - \lambda(MEDIA_{t-1})\pi_{t-1}^{exp} \tag{1.19}$$

As a proxy for the best available forecast, he uses the average forecast of professional forecasters $\pi_t^{exp,prof}$ since this is the inflation forecast agents typically read about in the news media. Agents get closer to the best available forecast if the media increases the amount of news coverage about inflation $MEDIA_t$: the more news reports on inflation published by newspapers, the higher the likelihood that agents will read about the best available forecast, and thus, the lower the costs of processing information. Hence, according to this view, agents do not spend time collecting data and trying to estimate the unknown parameters of the true model, but simply rely on the news media to get the latest available inflation forecast. It is this epidemiology model of inflation expectations that we are going to test throughout the dissertation.

1.1.5 Further Models

Before proceeding with a detailed literature overview on how the news media typically cover economic topics such as inflation, it is important to note that our list of expectation formation hypotheses is by no means exhaustive.

Sims (2003) has proposed a model of *rational inattention*. In contrast to the sticky information approach, agents are allowed to update their information set each period, however, they face constraints in processing this information. Since the rational inattention approach generally leads to the same conclusions as the sticky information model, we forgo a further exposition.[15] A key difference is worth highlighting, however. In models of sticky information, a fraction of agents forms expectations rationally, whereas in the rational inattention framework, agents can never compute the rational expectation forecast.

Akerlof et al. (1996, 2000) provide a model assuming *near-rationality*. In this approach, the degree of rationality depends on the level of inflation. If inflation is close to normal, agents tend to ignore new information given that small deviations from the best available forecast do not matter that much. However, in times of high or very low inflation, agents face a growing incentive of getting the latest data and thereby avoiding costs from falsely predict-

[15] See Dräger and Lamla (2013b) for an empirical comparison of sticky information and rational inattention models.

ing future price changes. As a result, the Phillips curve becomes nonlinear.
Finally, some approaches stress that expectations are *heterogeneous*. In our description of the most prominent models of expectation formation, we have implicitly retained a core assumption of many DSGE models, namely the use of a representative agent.[16] According to this modeling approach, it is either assumed that all agents behave in the same way, or that their individual decisions can be described by the behavior of one agent. Therefore, it does not have to be the case that every agents forms rational expectations, as long as the average computed from a number of different forecasts can be taken as rational. The sticky information model, by contrast, gives rise to heterogeneous agents. In each period, a fraction of agents computes rational forecasts, while the remaining fraction uses forecasts made in previous periods. Branch (2004, 2007) has offered evidence that agents switch between different models of expectation formation and Coibion and Gorodnichenko (2012) suggest that one should model the forecasts of consumers, firms and workers in a different way. Until present, it seems an open question whether it is more important to include some form of backward-lookingness or a considerable degree of heterogeneity into macroeconomic models. Wiederholt (2013) suggests that the policy conclusions drawn from a DSGE model including the zero lower bound are more sensitive to the assumption that inflation expectations are build in a purely forward-looking manner. It is the goal of this dissertation to test models of sticky information and the dependence of inflation expectations on news coverage, while also providing evidence in favor of heterogeneous expectations.

1.2 News Coverage of Inflation and Agenda Setting

In the next section, we provide a detailed overview of research in communication studies dealing with the question of how the media covers economic issues such as inflation and how this affects the opinions of readers. Among different theories modeling the impact of the news media, we focus on the theory of *agenda setting*. Furthermore, note that in the various studies the media agenda is measured by counting the number of articles or television reports that contain a certain term such as “inflation”. Further details on collecting data on media reports will be given throughout the dissertation.

1.2.1 The Idea of Agenda Setting and its Historical Development

Agenda setting[17] can be understood as the media’s influence on what people think is the most important event at present and/or in the future. In the words of McCombs (2004),

[16] See Hartley (1997) for a critical overview.
[17] McCombs (2004) offers an excellent overview on the concept of agenda setting. Quiring (2004) analyzes the impact of news coverage on political elections in Germany, and Hagen (2005) and Bachl (2008) investigate the news effects on economic sentiment in Germany.

p.37: agenda setting is defined as the "successful transfer of salience from the media agenda to the public agenda", where agenda means a "ranking of the relative importance of public issues" (Dearing, 1989, p.310).

The origin of agenda setting theory dates back to the "Chapel Hill study", implemented during the U.S. presidential election of 1968 (McCombs and Shaw, 1972). However, already some years earlier, Cohen (1963), p.13 in a frequently quoted phrase, has stated that the press "may not be successful in telling people what to think, but it is stunningly successful in telling its readers what to think about". Later, McCombs (2004) has extended this view claiming that recent advances in theory suggest that the media also influence *how* people think about certain topics. Empirically, agenda setting theorists have mostly used the "most-import-problem (MIP)-question" as the dependent variable capturing public opinion. More precisely, survey respondents are asked "What is the most import problem facing this country today?" (Soroka, 2002). This specific measure has to be kept in mind when assessing the results put forward by agenda setting theory concerning inflation. It is *not* shown that the media affect the exact number or tendency of individuals' beliefs about future prices, but instead, whether the media increase agents' attention towards inflation relative to other economic or even broadly political topics in general.

Psychologically, the existence of agenda setting effects is explained by the concept of *the need for orientation*.[18] This means that human beings have a general desire to understand the environment in which they live, and that they try to satisfy this need by using personal experience, personal conversation, and information obtained from the media. Two factors determine this need for orientation, first the relevance of a given topic for people's lives, and second people's uncertainty about this topic. If both relevance and uncertainty are high, people feel a strong need for orientation and seek for orientation in the media; hence, the agenda setting effects become stronger.

McCombs (2004), referring to the so-called "Acapulco-typology", lists four different perspectives of agenda setting. In the broadest perspective, called *competition*, agenda setting means the transfer of an entire agenda from the media to the aggregate of the public opinion. With regard to inflation expectations, the perspectives three and four are particularly important. The former, called *natural history*, takes a closer look at the evolution of the link between a single item on the media agenda with the aggregate public agenda, whereas the latter, denoted as *cognitive portrait*, investigates the effect of a single issue on the media agenda on the agenda of single individuals.[19] Historically, this typology is a result of four phases of agenda setting theory (See McCombs, 2004). In a first step, researchers focused on

[18] See McCombs (2004), and Matthes (2006). Ju (2008) quotes studies emphasizing the role of *accessibility* as the underlying reason for media effects.

[19] Perspective two, called *automaton*, evaluates the link between the entire media agenda and the agenda of a single individual. McCombs (2004) denies the relevance of this perspective since it rarely happens that indeed the whole ranking is transferred from the media to the individual.

issue salience, i.e. how the media manages to transfer its agenda to the public. In a second step, conditions were explored determining the strength of these agenda setting effects, i.e. whether some topics are obtrusive or unobtrusive, hence being more or less open for media influence or personal experience. Next, *attribute salience* was added to the picture, exploring how the media shape the way people perceive a certain topic, and not only whether they pay attention to an issue at all. Finally, research has been directed to the sources of the media agenda, exploring whether it depends on the policy agenda, interest groups, or real word conditions.

Taking a closer look at the transfer of the media agenda to the public agenda, one can thus distinguish two dimensions, issue salience and attribute salience.[20] Moreover, not every public issue is easily transferable via the media. Following Zucker (1978), the literature distinguishes "obtrusive issues", i.e. topics people encounter in their daily lives, and "unobtrusive issues", i.e. more complex and abstract issues that people mainly experience and understand by following the news media. Unobtrusive issues can both be complicated topics such as the working of monetary policy, or simple issues being relevant on the national level in contrast to the micro perspective of individuals' daily lives. This is the reason that agenda setting theorists mostly argue that inflation might be rather obtrusive, since individuals experience price developments every day through personal experience. However, it might also be justified to argue that agents might experience only certain prices such as food or fuel prices by direct experience, whereas the general inflation rate can only be learned about by following the news media.[21]

1.2.2 Empirical Evidence with Regard to Inflation

Whether the media affect the public opinion with regard to inflation has been investigated in a number of papers in the agenda setting literature. Mostly, inflation is found to be an unobtrusive topic, suggesting at most a very modest media effect. Soroka (2002) links public and policy agenda setting research by investigating the mutual effects of media reports, public opinion, and policy agenda in Canada between 1985 and 1995. Running Granger causality tests and calculating impulse response functions in a three equation SUR-system, he finds no significant effect from the media agenda on the public agenda with regard to inflation. Instead, the public agenda is found to be highly autocorrelated and affected by the monthly inflation rate. In a subsequent paper using aggregate times series for the United Kingdom from 1986 to 2000, Soroka (2006) shows that the media report more news about rising in-

[20] This is similar to the volume channel and tone channel coined by Lamla and Lein (2010).

[21] Ju (2008), investigating the effect of media coverage and real economic conditions on public opinion in South Korea, refers to the literature on "macroeconomic illiteracy". He expects that individuals gain a sense for the general direction of economic variables, such as inflation, from the media, rather than a particular number of the inflation rate.

flation than on decreasing inflation.[22] The same finding has been reported by Harrington (1989) in a study about three major U.S. television networks and Goidel and Langley (1995) in an analysis of articles in the New York Times. Zhu et al. (1993) in an article testing the relative importance of agenda setting effects and social interaction find that for inflation, social interaction is more important than media reports. However, the authors also find a significant recruitment effect of the media, i.e. more media reports induce people to name inflation as the most important topic. Besides their model's extension to include the effects of social interaction, it also belongs to a couple of papers investigating the potential nonlinear effects of agenda setting. Analyzing television broadcasts in the United States between 1974 and 1980, Behr and Iyengar (1985) find that only above average inflation rates lead to higher news coverage of inflation, explaining this modest effect by the continuing high inflation rate during this period. Furthermore, their results suggest a clear agenda setting effect with regard to inflation: rising inflation leads to more media reports which subsequently affect the public opinion. There is no direct link between inflation or food prices on public opinion, but a feedback effect from public concern about inflation on media coverage. Slightly extending the sample period used by Behr and Iyengar (1985), Harrington (1989) adds to their results by showing that television broadcasts report more on rising than on decreasing inflation. In another early study, Winter et al. (1982) calculated cross correlations and find a positive media setting effect for the first month, but insignificant lags thereafter. Criticizing Behr and Iyengar (1985) for not having controlled for autocorrelation in the news series of inflation, Demers et al. (1989) find no agenda setting effect for inflation. More recently, agenda setting theory has also been applied in economics. Larcinese et al. (2011) explore the possibility that left-wing newspapers in the U.S. write more about bad economic news if the incumbent president is a Republican, and vice versa. They find that such a partisan bias exist with regard to news on unemployment, but not in case of inflation.

1.2.3 The Paradox of Agenda Setting

Agenda setting theory would suggest that the effect of news reports should be stronger if one uses micro data instead of macro data due to the higher precision of micro data:[23] the researcher can control for the media use of a specific individual, whereas macro data implicitly assumes that all participants in a survey follow the new media. Given that only a fraction of the population follows the news,[24] using the average of a survey of a number of different individuals should result in lower agenda setting effects. However, summarizing the research conducted in communication theory, the opposite seems to be true. Agenda

[22] He also shows that the public reacts asymmetrically to bad news on unemployment, without testing for an asymmetric impact of the inflation rate.

[23] See Rössler (1999) for a distinction between micro and macro data in agenda setting theory.

[24] In a survey conducted on U.S. households, Blinder and Krueger (2004) show that only 46% of respondents mention TV broadcasts as the most important source of information on economic policy.

setting effects are found to be relatively strong on the macro level whereas studies using panel or micro data find much lower effects. This finding contrary to initial expectations is called *the paradox of agenda setting*.

Maurer (2004) suggests two explanations for the paradox of agenda setting. First, social interaction or interpersonal communication might lower the effects of agenda setting. Since the use of micro data captures the fact that some individuals follow the news while others tend to ignore it, this increases the probability that individuals with different beliefs meet each other inducing some individuals to drop the ideas/information they got from the media. This effect might even be reinforced if the information from the media is noisy, i.e., if some newspapers write about falling inflation, while others mention increasing prices or do not write about inflation at all. These effects might get lost by aggregating individual data resulting in strong media effects on the aggregate level.
Second, social interaction could also amplify the effects of news coverage. This is the case if some "news-followers" transmit the information originally received from the news media to others who do not follow the news (Price, 1988). This can be especially true, if the media news is relatively uniform. Hence, one would only find low direct media effects by using micro data, since a group of individuals does not get the information from the media but from other individuals, whereas aggregating over all individuals incorporates the news-triggering effect of social interaction. Calculating correlation coefficients in a short panel on individuals' political assessment of politicians in Germany, Maurer (2004) finds some empirical evidence for this distinction.

1.3 Research Questions and Outline of the Dissertation

In this dissertation, we explore the links between news coverage of inflation and the inflation expectations of economic agents. Our main goal consists of testing whether the predictions of the epidemiology model are supported by the data.

1.3.1 The Epidemiology Model of Expectation Formation

In Chapter (2), the basic equation of the epidemiology model given in (1.19) is analyzed in great detail. If expectations are indeed sticky, rather than rational or adaptive, a number of policy implications emerge. First, allowing for sticky information in macroeconomic models leads to a sluggish behavior of output and inflation in response to monetary policy shocks, which is a robust stylized fact documented in the empirical literature (Mankiw and Reis, 2006). Second, if agents can deliberately choose not to pay attention to all available information, this can affect the optimal monetary policy strategy (Paciello and Wiederholt, forthcoming). Third, policy makers in general and central bankers in particular are concerned with the question whether expectations are well anchored i.e. whether inflation expectations are close to the target rate. Most of the empirical literature on anchoring, however, uses data for professional forecasters, assuming implicitly that these are fully matched by the general public. By contrast, the epidemiology model predicts that the expectations of households and experts can differ substantially, depending on the amount of news coverage. And if agents disagree persistently with respect to future outcomes of economic variables, this might call for an adapted communication strategy of central banks (Sims, 2009).

Testing the epidemiology model, we follow much of the literature and proxy the best available forecast with survey expectations of professional forecasters. More precisely, we use U.S. data on household expectations from the Michigan Survey and the Survey of Professional Forecasters (SPF) from January 1980 - November 2011. We focus on three dimensions of the epidemiology model. First, we analyze whether the expectation formation process changes over time, i.e. whether households build different forecasts in times of high or low inflation, or in times of economic crisis. Moreover we test whether the degree of updating varies over time in line with the amount of news coverage on inflation. Second, we use both aggregate and micro survey data in our analysis thereby studying whether the results depend on the aggregation level of the data employed. In order to test for the "paradox of agenda setting", we separate the full sample of survey respondents into households who have heard news on economic issues, on inflation, and on good or bad news on inflation. We then test if these groups are more receptive to news media coverage compared to others, and if their forecast error is lower. Finally, research in psychology (Batchelor, 1986) suggests that individuals only pay attention to news if the stimulus passes a certain threshold. On the other hand, there can be a satiation level: If the media treat a certain topic extensively over some period, readers loose interest and are thus less willing to react to new incoming information. We test for the possible non-linearity of news media effects by fitting Smooth Transition Autoregressive Models. For the best of our knowledge, this is first time that non-linear news effects on the inflation expectation of households are investigated in the literature.

Our empirical analysis yields the following results. Overall, we find that the epidemiology model is supported by the data. Households partly use the best available forecast and their own past forecast when forming beliefs about future inflation. In addition, households adjust more to experts in times of low and stable inflation and during economic crisis. More news coverage of inflation generally lowers the gap between households' and professional forecasters' predictions, however, the effect is not stable over time. In times of falling inflation, the news media lower the expectation gap, whereas in times of economic crisis, more articles on inflation increase the gap. Comparing the results using macro and micro data, we find that the speed of updating is lower if we use micro level data. In contrast to the degree of updating, the media effect is found to be larger on the micro level. Looking at households with different news perceptions, we find that those who claim to have heard news on inflation commit larger forecast errors than other households while at the same time being more receptive to media reports. Finally, our analysis suggests that the media effect is non-linear. An increasing number of news reports strengthens the impact from expert expectations. For all households, the adjustment takes place only gradually, whereas those who have heard news about inflation are much quicker in reacting to rising amounts of news coverage.

1.3.2 Socioeconomic Expectation Formation and News Media Exposure

In Chapter (3), we apply the epidemiology model to different groups of households. As it is well known from international studies, households with low income and low education, females, unemployed, and young and old individuals have higher inflation expectations and forecast errors compared to other households. Whereas the reasons for these expectation differentials are still up to debate in the economic literature, economic policy will be affected through various channels.

First, allowing for heterogeneity of expectations has found to to be important to explain stylized facts such as the hump-shaped response of output and inflation to monetary policy shocks. Second, anchoring agents' inflation expectations might call for different communication strategies of central banks if households persistently form expectations in different ways. Third, rising disagreement on the future path of prices might be a sign of uncertainty with possible effects on economic risk-taking. And fourth, if some demographic groups tend to have forecast errors that are persistently above average, this might call for economic policies mitigating the resulting effects on the distribution of wealth and income.

Our results show that in Germany, expectation differentials of households with regard to income, age, and occupation can be explained by different group-specific inflation rates and socioeconomic media consumption. From 1999-2010, we analyze the links between households' inflation expectations and inflation rates, as well as the news coverage of inflation in 10 different news sources.

We find that inflation expectations are higher for households with low income, for young households and for the unemployed. Moreover, the same types of households show larger deviations from the best available forecast, which we proxy with professional forecasters' expectations. We find that the higher expectation gaps of young and old households as well as the rising deviation with lower income levels can be explained by higher inflation rates of these groups, while no such effect can be observed for occupation groups. With regard to the news media, we document considerable heterogeneity in news consumption of different newspapers and TV news shows for income, age and occupation groups. It thus seems that media coverage offers some explanation on why households with a different socioeconomic background disagree on the future path of prices. Depending on whether different news media report negatively or positively about inflation, this will narrow or widen the gap between experts' inflation forecasts and households' inflation expectations

1.3.3 Internet Search Data as Alternative Measure of Inflation Expectations

Whereas our previous analysis has been built on measuring inflation expectations with the help of survey data, Chapter (4) extends this analysis. Whereas survey data have proven to be useful in forecasting inflation (Ang et al., 2007) and in predicting individual investment decisions (Armantier et al., 2012), they also face a number of difficulties. Results can strongly depend on the exact question wording, which is particularly relevant with regard to inflation forecasts since respondents easily confuse price level and inflation rate depending on how they are asked (Bruine de Bruin et al., 2012, Dräger and Fritsche, 2013). Moreover, designing and implementing a questionnaire consumes time and money, hence, existing surveys often face a small sample problem, both across time and respondents. Third, survey respondents might lack an incentive to state their best possible expectations due to the absence of financial consequences or peer pressure. Moreover, if the same individuals participate repeatedly a survey, learning effects might result in much better predictions compared to individuals that do not take part in the survey. Finally, many countries still lack surveys that ask respondents to express their expectations in terms of a precise number or within predefined ranges. Instead, qualitative answers are provided making it necessary to apply data transformations that depend on various, often restrictive assumptions (Nardo, 2003).

In this chapter, we propose the use of Google search requests as a supplementary measure for inflation expectations. People increasingly turn to the internet if they feel the need to get more information on a certain topic. Compared to surveys, internet search intensity does not depend on framing effects stemming form question wording. Moreover, the number of searches comes as a by-product of users' internet activities: if individuals use the Google web page in order to find information on a certain topic, they do so because they already feel the need to get informed, either because they are reluctant to seem uninformed in daily talks, or because they have a specific economic transaction in mind which makes it necessary to possess the latest news on inflation. Finally, since Google search data is available on a weekly basis, this means that internet search requests could serve as a supplement to the existing survey data which is often compiled on a monthly basis and only released with some time lag. This is of particular interest for monetary policy that seeks to monitor price developments as timely as possible.

We analyze U.S. data from January 2005 to May 2011 on households' and professional forecasters' expectations measured via survey data, newspaper articles and television reports on inflation, and Google search requests for inflation. The contribution of this chapter is twofold: First, we explore the news content of web searches on inflation. More precisely, we want to know whether search intensity evolves in a systematic way that can be attributed to real economic data. Note that Google searches might simply mirror the news coverage of inflation in the media, hence there might be no additional gain of using web searches in addition to the number of newspaper articles. To test whether Google searches are different, we compare the reaction of Google searches, TV reports and newspaper articles to changes in prices, variables describing the stance of monetary policy and lagged values of households' and professional forecasters' expectations. In a second part, we take into account the various feedback effects among the news media, Google search requests and the inflation expectations of households and professional forecasters by estimating Vector Autoregressive models.

Our results show that users' demand for information can indeed be linked to economic fundamentals: Google search requests can be explained by price changes much better than media reports. Google users distinguish between headline and core inflation and they react asymmetrically: the demand for information increases if core inflation falls. In periods of historically high inflation rates, the number of search requests is significantly larger. Moreover, internet users pay attention to central bank behavior: unscheduled conference calls as well as issued statements increase search intensity. In addition, we find a positive effect from households' inflation expectations in the previous period on search requests: Google users seek for additional information if they predict prices to rise in the future. Higher inflation forecasts of experts only marginally increase Google search requests, but if professional forecasters disagree a lot on future prices, the resulting uncertainty leads to a large increase in Google users' demand for information.

With regard to the results of the VAR models, we find that television news coverage is driving newspaper coverage, in addition to a feedback effect. Building on this result, we show that Google search requests for inflation are mainly determined by TV reports and only to a lesser degree by newspaper articles. Again, we find considerable feedback effects, suggesting that journalists consider the interests of their readers when deciding on the newspaper's agenda. Finally, taking into account households' and professional forecasters' inflation expectations, we show that households' forecasts are driven by TV reports, newspaper articles, and Google searches, while the feedback effect from expectations on web searches is rather small and estimated less precisely. Furthermore, the impulse response function from shocks on web searches to expectations is estimated more efficiently for weekly data, which indicates that the demand for new information has a rather short-run impact on peoples' expectations. About 20% of the forecast error variance decomposition of households' inflation expectations can be explained by Google search requests.

Chapter 2

Unfinished Business in the Epidemiology of Inflation Expectations

2.1 Introduction

The debate on how economic agents form expectations about the future centered for a long time around two competing approaches, the rational expectations hypothesis and the proposition of adaptive expectations (See e.g. Gertchev (2007) for a critical discussion). More recently, theories of sticky information (Mankiw and Reis, 2002) and rational inattention (Sims, 2003) have proposed a compromise between these contradicting approaches. In these models, agents are assumed to be forward-looking, but do not always adjust their expectations to the best available forecast. Instead, based on the assumption that gathering and processing information is costly, agents are supposed to only slowly update their expectations. Therefore, the predictions of agents consist partly of the best available forecast (which might be rational) and the forecast that has been made in earlier periods. According to the epidemiology model proposed by Carroll (2003), the relative weight of these two ingredients is determined by the ease of availability of incoming information which are disseminated by the news media. Hence, in times where information can be accessed easily, households' expectations should be closer to the best available forecast compared to periods of below average information flows.

If expectations are indeed sticky, rather than rational or adaptive, a number of policy implications emerge. First, allowing for sticky information in macroeconomic models leads to a sluggish behavior of output and inflation in response to monetary policy shocks, which is a robust stylized fact documented in the empirical literature (Mankiw and Reis, 2006, 2007). Second, if agents can deliberately choose not to pay attention to all available information, this can affect the optimal monetary policy strategy (Paciello and Wiederholt, forthcoming). Third, policy makers in general and central bankers in particular are concerned with the question whether expectations are well anchored i.e. whether inflation expectations are close to the target rate. Most of the empirical literature on anchoring, however, uses data for professional forecasters, assuming implicitly that these are fully matched by the general public (see Dräger and Lamla (2013a) for an exception). But if expectations are formed in line with the hypotheses of sticky information or the epidemiology model, it might well be the case that the predictions of households deviate from the expectations of experts.[1] Finally, in addition to adjusting their forecasts only gradually, households are often found to disagree about the future (Armantier et al., 2012). And if agents disagree persistently with respect to future outcomes of economic variables, this might call for an adapted communication strategy of central banks (Sims, 2009).

In this chapter, we test the sticky information expectation hypothesis in some detail. More precisely, we focus on the epidemiology model proposed by Carroll (2003) since it provides a direct way to test the "ease-of-information hypothesis" by incorporating a prominent role of the news media.[2] In line with most of the literature, we refer to inflation expectations and use survey data for household expectations and proxy the best available forecast with survey expectations of professional forecasters. More precisely, we use U.S. data from the Michigan Survey and the Survey of Professional Forecasters (SPF) from January 1980 - November 2011. We focus on three dimensions of the epidemiology model. First, we analyze whether the expectation formation process changes over time, i.e. whether households build different forecasts in times of high or low inflation, or in times of economic crisis. Moreover, related to Lamla and Sarferaz (2012), we test whether the degree of updating varies over time in line with the amount of news coverage on inflation. Second, we use both aggregate and micro survey data in our analysis thereby studying whether the results depend on the aggregation level of the data employed. As it has been stressed by Dovern et al. (2013), theories of expectation formation are mostly formulated on the individual level and are subsequently applied to aggregate survey data using for example the cross-sectional mean forecast. However, aggregating individual survey responses might be problematic if it masks important heterogeneity on the micro level. This is particularly relevant with respect to measuring the strength of the news media effect. Empirical research in communication studies has found much stronger media effects on the aggregate level compared to the micro level (Krause and Gehrau, 2007). A possible explanation for this "paradox of agenda setting" might be that only a part of the population follows the news media and subsequently circulates the information to non-users. In order to test this explanation, we separate the full sample of survey respondents into households who have heard news on economic issues, on inflation, and on good or bad news on inflation. We then test if these groups are more receptive to news me-

[1] In this context, Coibion and Gorodnichenko (2012) have recently emphasized that it is important to allow for different expectation formation processes of different agents.

[2] The baseline equation of the sticky information model and the epidemiology model is essentially the same. For differences with respect to the rational inattention variant, which are beyond the scope of this paper, see for example Dräger and Lamla (2013b) and Dovern et al. (2013).

dia coverage compared to others, and if their forecast error is lower. Finally, we test whether the news media effect is non-linear. Research in psychology suggests that individuals only pay attention to news if the stimulus passes a certain threshold.[3] On the other hand, there can be a satiation level: If the media treat a certain topic extensively over some period, readers loose interest and are thus less willing to react to new incoming information. We test for the possible non-linearity of news media effects by fitting Smooth Transition Autoregressive Models. For the best of our knowledge, this is first time that non-linear news effects on the inflation expectation of households are investigated in the literature.

Our empirical analysis yields the following results. Overall, we find that the epidemiology model is supported by the data. Households partly use the best available forecast and their own past forecast when forming beliefs about future inflation. In addition, households adjust more to experts in times of low and stable inflation and during economic crisis. More news coverage of inflation generally lowers the gap between households' and professional forecasters' predictions, however, the effect is not stable over time. In times of falling inflation, the news media lower the expectation gap, whereas in times of economic crisis, more articles on inflation increase the gap. Comparing the results using macro and micro data, we find that the speed of updating is lower if we use micro level data. While this result is in contrast to other findings (Dräger and Lamla, 2013b), it might stem from the fact that we also consider households who only take part in the survey once, and that we allow for time variation. In contrast to the degree of updating, the media effect is found to be larger on the micro level. Looking at households with different news perceptions, we find that those who claim to have heard news on inflation commit larger forecast errors than other households while at the same time being more receptive to media reports. Finally, our analysis suggests that the media effect is non-linear. An increasing number of news reports increases the impact from expert expectations. For all households, the adjustment takes place only gradually, whereas those who have heard news about inflation are much quicker in reacting to rising amounts of news coverage.

Our paper is closely related to two recent studies of the micro data of the Michigan Survey. Dräger and Lamla (2013b) show that the updating frequency of households is much higher if the analysis is conducted with micro data. Coverage of inflation in the news media is found to have no effect on the updating frequency and the precision of forecasts. By contrast, if participants claim to have heard news on inflation, they are more likely to adjust their expectations resulting, however, in a larger forecast error. Pfajfar and Santoro (2013) show that a rising amount of news coverage *increases* the gap between households' and experts expectations which is in marked contrast to the prediction of the epidemiology model. Both of these studies use the short rotating panel dimension of the Michigan survey. In each

[3]This point relates the to so called "Weber-Fechner law" stating that the effect of a stimulus is not constant but depends on the initial level of the stimulus. See for an application to the perceptions of inflation Thaler (1980) and Batchelor (1986).

month, about 40% of all participants are reinterviewed a second time six months after the first interview. While using the rotating panel is appealing due to the fact that it allows studying whether and how the same individual changes her expectations over time, it also has some disadvantages. Individuals might pay (more) attention to the news simply because they participate in a survey. Since they know that they will interviewed a second time, they try their best to look good when being faced with the interviewer.[4] Moreover, the individual updating period is fixed by assumption. The second interview will take place six months after the first one, and if individuals have changed their forecasts several times in between the survey rounds, this will not appear in the responses. Finally, it remains unknown whether participants in the second interview will be reminded of the forecasts they have made in the first interview. Therefore, expectation updating might arise simply because participants do not remember their previous forecast.

Besides of the rotating panel dimension, both Dräger and Lamla (2013b) and Pfajfar and Santoro (2013) use the fraction of households who have heard news on inflation interchangeably with the amount of news coverage in the media. Whereas self-reported news might be preferable to the number of newspaper articles because it measures the actual information set of households more closely, it also suffers from severe overreporting (Prior, 2009). Therefore, we take a slightly different perspective and test whether households who claim to have heard news about inflation are also affected more by the news media compared to other households.

We start our analysis with a brief exposition of the epidemiology model and a discussion of its particular features that we are going to analyze in detail (Section 2.2). In Section (2.3), we describe the data set and provide summary statistics of the micro data of households' inflation expectations which already provides important insights about the expectation formation. The empirical analysis is divided into three parts. We start with estimating the epidemiology model without news media in Section (2.4), before including news coverage in a linear framework in Section (2.5) and allowing for non-linear effects in Section (2.6). Section (2.7) summarizes the results.

2.2 The Epidemiology Model

According to the epidemiology model of inflation expectations proposed by Carroll (2001, 2003, 2005), in each period t, only a fraction λ of households adjusts its expectations to the best available forecast whereas the remaining fraction $1 - \lambda$ sticks to the forecasts made in the previous period.[5] Thus, denoting households' expectations with $\pi_t^{exp,hh}$ and the best available forecast with $\pi_t^{exp,prof}$, households' one-year-ahead forecast is given as a weighted

[4]Since the Michigan Survey is a telephone interview, there is less anonymity compared to a written survey.
[5]The detailed derivation of the model is given in Section (A.6.1) in the Appendix.

average such that

$$\pi_t^{exp,hh} = \lambda \pi_t^{exp,prof} + (1-\lambda)\pi_{t-1}^{exp,hh} \tag{2.1}$$

As a proxy for the best available forecast, Carroll suggests the use of the mean forecast computed from a number of professional forecasters, which is justified by the fact that the news media regularly report the inflation forecast published by research institutes or centrals banks. Furthermore, Carroll suggests that households should get closer to professional forecasters' expectations the more the news media report about inflation since this increases the likelihood that an expert is quoted stating his outlook on future inflation.[6] Thus, one can rewrite equation (2.1) as

$$\pi_t^{exp,hh} = \lambda\left(MEDIA_t\right)\pi_t^{exp,prof} + (1-\lambda\left(MEDIA_t\right))\pi_{t-1}^{exp,hh}, \tag{2.2}$$

where $MEDIA_t$ denotes the number of newspaper articles about inflation. Typically, this second equation is estimated by using the transformed version

$$GAPSQ_t = \alpha_1 + \alpha_2 MEDIA_t + \varepsilon_t, \tag{2.3}$$

where $GAPSQ_t = \left(\pi_t^{exp,hh} - \pi_t^{exp,prof}\right)^2$ denotes the squared gap between households' and professional forecasters' expectations. According to the model, one would expect that $\alpha_2 < 0$: the more news coverage about inflation, the lower the gap between households and experts.

The "squared gap-equation", however, is only a reduced form and cannot be fully derived from the structural form of the epidemiology model.[7] Rearranging (2.2) yields:

[6] Expressed in the words of Carroll (2003), p.275: "We will assume that households believe that experts have some ability to directly estimate the past and present values of inflation (...) (through deeper knowledge of how the economy works, or perhaps some private information); thus, households can rationally believe that a forecast from a professional forecaster is more accurate than a simple adaptively rational forecast that they could construct themselves." This argument, however, relies on the assumption that professional forecasters are indeed better in predicting inflation than households. While this is true in general, as it has been documented by Thomas (1999) comparing the SPF and the Michigan survey over the time span 1980-1997, it is by no means obvious that this should always be the case. While households' financial well-being directly depend on the accuracy of their expectations, experts are paid on their fame and reputation instead of the precision of their forecasts. As it has been argued by Ottaviani and Sorensen (2006), some forecasters can have a strong incentive to deviate from the consensus forecast in order to gain reputation, which can subsequently result in larger forecast errors compared to the prediction of households.

[7] I thank Ulrich Fritsche for raising this point.

$$\begin{aligned}
\pi_t^{exp,hh} &= \lambda\left(MEDIA_t\right)\pi_t^{exp,prof} + (1-\lambda\left(MEDIA_t\right))\pi_{t-1}^{exp,hh} \\
\pi_t^{exp,hh} &= +\lambda MEDIA_t\pi_t^{exp,prof} + \pi_{t-1}^{exp,hh} - \lambda MEDIA_t\pi_{t-1}^{exp,hh} \mid -\pi_t^{exp,prof} + \pi_t^{exp,prof} \\
\pi_t^{exp,hh} - \pi_t^{exp,prof} &= \lambda MEDIA_t\pi_t^{exp,prof} + \pi^{exp,prof} + \pi_{t-1}^{exp,hh} - \lambda MEDIA_t\pi_{t-1} \\
GAP_t &= \underbrace{\lambda(\pi_t^{exp,prof} - \pi_{t-1}^{exp,hh})}_{=\alpha_2} MEDIA_t + \underbrace{(\pi_t^{exp,prof} + \pi_{t-1}^{exp,hh})}_{=\alpha_1}
\end{aligned} \tag{2.4}$$

Hence, we can find a structural form relating the gap to the level of media reports, whereas the squared gap would have to be related to the squared number of news reports.
From the expression in (2.4), note that we should get a negative news effect α_2 if $\pi_t^{exp,prof} < \pi_{t-1}^{exp,hh}$, i.e., if the latest available forecast is lower than the forecast of households from the previous period. Hence, if households observe that their past prediction was above the latest available forecast, they will lower their prediction if the media write about the experts forecast. By contrast, we should get a positive news effect if $\pi_t^{exp,prof} > \pi_{t-1}^{exp,hh}$ meaning that households will raise their prediction if they observe that their previous forecast was below the recent forecast of experts. If there was no deviation, there is no media effect. In both cases, however, media reports will lower the gap, since households adjust towards the best available forecast. Expressed in terms of the squared gap, we should find that media reports (expressed in squared terms) lower the gap if households' past forecast was above the forecast of experts, but increase the gap if households were below experts.

In the reminder of this chapter, we will follow much of the literature and estimate the reduced form (2.3) leaving a more detailed treatment of the structural form for further research. For the time being, we take the rather loose derivation of the gap equation as additional motivation for directly estimating the nonlinear formulation of the epidemiology model given in equation (2.2).

Summing up, the epidemiology model comes in two versions. Equation (2.1) is the version without news, and it simply states that households only partially adjust to the best available forecast while another fraction keeps its inflation forecast from the past. Until present, the empirical evidence of the epidemiology model without news has been mixed. While some authors find empirical support in the data (e.g. Carroll, 2003 for the U.S., Döpke et al., 2008 for European countries, and Lamla and Lein, 2010 for Germany), others come to the contrary conclusion (Lanne et al., 2009, Luoma and Luoto, 2009), emphasizing the superior role of the adaptive expectation hypothesis.

The second variant of the epidemiology model stated in equations (2.2) and (2.3) highlights the distinct role of the media in the updating process. An increase in the amount of news about inflation should lower the distance between the inflation forecasts of households and experts. As regards the second version, the empirical evidence has been rather weak. Whereas Carroll (2003) presents supportive evidence, others have generally rejected the negative news effect derived from the epidemiology model (Lamla and Lein, 2010, Pfajfar and Santoro, 2013). In addition, as is has been shown by Menz and Poppitz (2013) and Dräger and Lamla (2013b), the result depends on the type of news used in the empirical analysis.

We add to this literature by analyzing in detail three features of the epidemiology framework.

Time-Dependent Updating The baseline epidemiology model suggests that the adjustment parameter to the best available forecast is fixed over time. However, using a Bayesian state-space approach, Lamla and Sarferaz (2012) have shown that λ varies substantially over time and that the time-dependence is at least weakly determined by the amount of news coverage. In addition to news coverage, the amount of trust people have in economists and people working in the financial sector might be another crucial determinant of the updating mechanism. Since households do not have the time or the ability to check whether experts make precise forecasts, they will only rely on these predictions if they belief that experts in general make a good job. However, as it has been documented by survey data, this trust has eroded dramatically during the financial crisis. Finally, the degree of updating can depend on the level and the variance of the actual inflation rate. As it has been suggested by Akerlof et al. (1996, 2000), in times of low inflation, households are less willing to spend time on getting the best forecast. However, if inflation moves above or below its "normal" level, households put more effort in forecasting prices. We test for possible state-dependent updating by performing QLR tests to identify structural breaks in equation (2.1) and (2.3). Furthermore, we test whether the expectation formation is different during recessions.

Macro and Micro Data The epidemiology model is formulated on the aggregate level, however, Dräger and Lamla (2013b) and Pfajfar and Santoro (2013) have shown that the degree of updating is typically higher if the analysis is conducted with micro survey data. Therefore, we estimate the epidemiology model using both macro and micro data, and distinguish households according to whether they have heard news on inflation or not. We compute the forecast errors of different households groups and test whether households who have perceived some news are better in forecasting inflation than other households. As regards the empirical analysis, in order to avoid the potential pitfalls of the rotating panel dimension of the Michigan survey, we employ the full sample of the survey. Facing a repeated cross section, this leaves us with the problem of how to proxy households' lagged inflation expectations. In the analysis using micro data, we start with using the cross-sectional mean expectation assuming that households do not refer to their own forecast from the past but to the average prediction of the general public. This can be motivated with the presence of a learning process in which households share their beliefs and thus converge to the cross-sectional average (Malmendier and Nagel, 2013). As an alternative approach, we fit a pseudo panel defining cohorts using the age of survey participants.

Non-Linear Media Effects The original version of the epidemiology model in equation (2.2) including media coverage is formulated such that the amount of news coverage affects the degree of updating in a non-linear way. With the exception of Lamla and Sarferaz (2012), the literature has bypassed the non-linear framework by focusing on the linear transformation in equation (2.3). This formulation, however, can be too restrictive. Households might not always react to the news media in the same way, instead, they might either miss a single article about inflation if the general interest in inflation is low. Likewise, news coverage might reach a satiation level beyond which readers ignore additional articles on inflation. Finally, the degree of non-linearity and the attention and satiation level might differ across households, depending on whether households generally follow the news or not. In order to test for non-linear media effects, we estimate equation (2.2) with the Smooth Transition Autoregressive model.

2.3 The Data Set and Preliminary Analysis

2.3.1 Inflation Expectations and Media Reports

In order to test the epidemiology model in detail, we use survey data for households' inflation expectations stemming from the *Michigan Survey of Consumers*.[8] Survey participants are asked to answer the questions "During the next 12 months, do you think that prices in general will go up, or go down, or stay where they are now?", and "By about what percent do you expect prices to go up, on the average, during the next 12 months?". In our analysis, we use the answers to the second question since it gives us a direct quantitative estimate of households' one-year-ahead inflation expectations. Each month, about 500 individuals take part in the survey. In line with the theoretical model, we use the cross-sectional mean for the estimation with aggregate data. Typically, the mean is corrected for outliers by truncating extreme answers such as an expected rate of inflation of 50% to more meaningful numbers. As it is shown by Curtin (1996), the cross-sectional mean is affected by the applied truncation rule albeit the impact is relatively small. Since we want to keep as much information as possible while at the same time avoiding distorted estimates due to outliers, we adopt

[8]The data can be downloaded at Michigan. Further details on the surveys used in this paper can be found in Thomas (1999) and Croushore (1993).

a rather conservative truncation rule of +/- 30% throughout the analysis.[9] Data from the Michigan Survey is available on a monthly basis from January 1978 onwards.

As regards professional forecasters' expectations, we employ data from the *Survey of Professional Forecasters (SPF)* which is available since the third quarter of 1981.[10] Each quarter, about fifty economists mainly working in nonfinancial and financial firms are asked to state their quarter-by-quarter forecast for the CPI inflation rate over the next year. These quarterly forecasts are transformed into a one-year-ahead prediction via a geometric average:

$$\pi_{t,t+4}^{exp,prof} = 100\left\{\left[\left(1+\frac{\pi_{t,t+1}^{exp}}{100}\right)\left(1+\frac{\pi_{t+1,t+2}^{exp}}{100}\right)\left(1+\frac{\pi_{t+2,t+3}^{exp}}{100}\right)\left(1+\frac{\pi_{t+3,t+4}^{exp}}{100}\right)\right]^{1/4}-1\right\} \quad (2.5)$$

In what follows, the quarterly SPF one-year-ahead forecast is transformed into monthly forecasts by linearly interpolating the missing months. We choose to conduct our analysis on monthly data instead of computing quarterly averages of the Michigan data for the following reasons. First, we want to keep as many observations as possible. Second, since our focus is on explaining households' expectations, we did not want to impose too many a priori restrictions on our dependent variable. Finally, the news media are relatively fast in emphasizing certain topics, hence, the actual impact of news media coverage on the expectation formation of households might be downplayed by an analysis using quarterly data. As a cross-check of the interpolation, we compare the SPF series with data from *Consensus Economics*, a survey conducted on a monthly basis.[11] Figure (A.1) in the Appendix plots our interpolated SPF series together with the Consensus forecast. With the exception of the financial crisis in 2008/2009 where the volatility of the Consensus forecast is much higher compared to the SPF, the two series move quite closely together. Hence, we are confident that our results are not affected by interpolating the quarterly SPF forecast.

Due to the availability of the micro data of the Michigan survey, our analysis covers the time span January 1980 - November 2011.[12] Figure (2.1) shows households' inflation expectations measured by the Michigan survey, professional forecasters' price predication, and the annual change of the seasonally adjusted CPI index. They gray shaded areas denote recession periods as dated by the NBER.[13] We can roughly distinguish four periods. From the beginning of the sample until the mid 1990s, the two series of expected inflation

[9] Due to the dependence of the mean on different truncation rules, the median might be a more robust measure of the general public's inflation forecast. However, since the theoretical model derives predictions only for the mean, we also stick to this measure in the empirical analysis. Furthermore, whereas the mean is typically 1 percentage point higher than the median, the two series move very closely together.

[10] For data download and further information, see SPF.

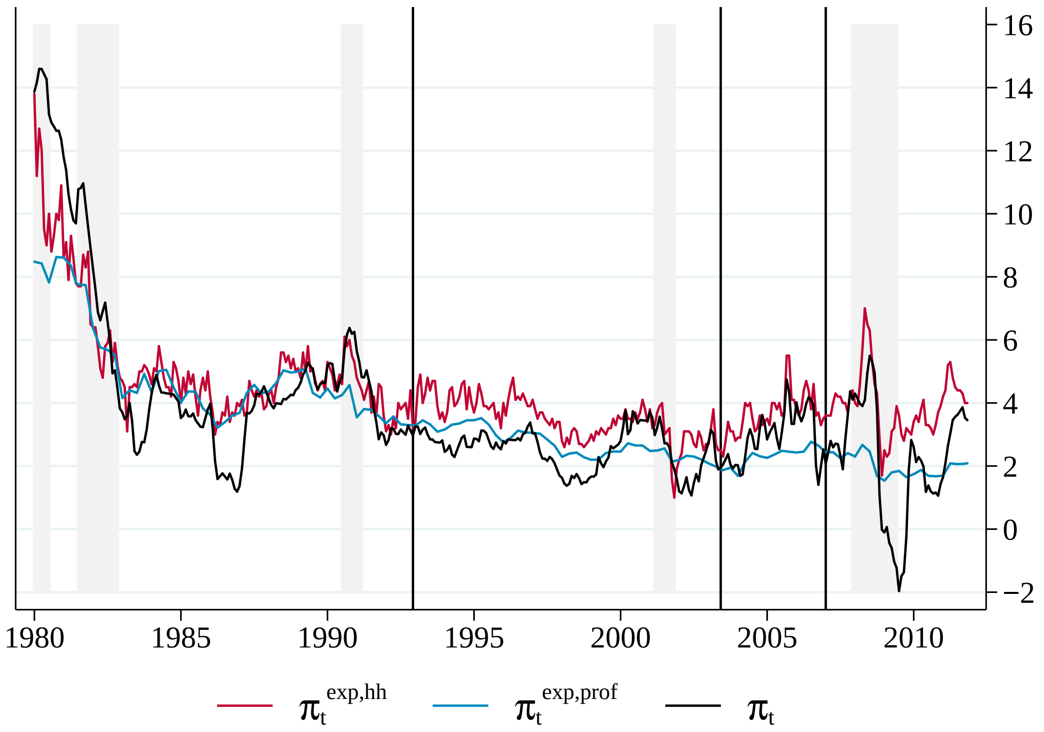
[11] See Consensus for details. We did not use the Consensus survey in the analysis since it is subject to a fee and covers a shorter time span than the SPF.

[12] In order to include the missing data from the SPF in 1980, we follow Luoma and Luoto (2009) and proxy the CPI forecast with the prediction for the GDP deflator.

[13] See NBER for details.

moved fairly closely together. Afterwards, households' expectations shifted upwards and constantly stayed above the forecast of experts but the two series still behaved rather similarly. Since 2003, households' forecast trended upwards in line with the rising inflation rate, whereas experts continued to expect an inflation rate of about 2%. Finally, households' expectation fluctuated a lot since the beginning of the financial crisis, while the prediction of experts remained rather constant. As we will discuss in more detail below, these different sub-periods correspond to those found by structural break tests of the epidemiology model which are highlighted by the vertical lines in the graph.

Figure 2.1: Households' and Professional Forecasters' Inflation Expectations



Note: The graph shows the mean inflation expectations of households ($\pi_t^{exp,hh}$), and of professional forecasters ($\pi_t^{exp,prof}$), together with the annual inflation rate π_t. The gray shaded areas denote NBER recessions, and the vertical lines indicate the structural breaks in 1992:12, 2003:06 and 2007:01 found in applying the QLR-test to equation (2.6).

In order to measure media coverage of inflation, we follow Carroll (2003) and Pfajfar and Santoro (2013) and count all articles published in *The New York Times* and *The Washington Post* that contain words with the root "inflation". The corresponding articles can be accessed in the database *Lexis Nexis*[14] and are available on a monthly basis since January 1980. This way of measuring news coverage of inflation has the advantage that the data is costless and readily available, but also suffers from some limitations. First, the automatic search procedure of *Lexis Nexis* does not allow us to detect whether an article containing the word

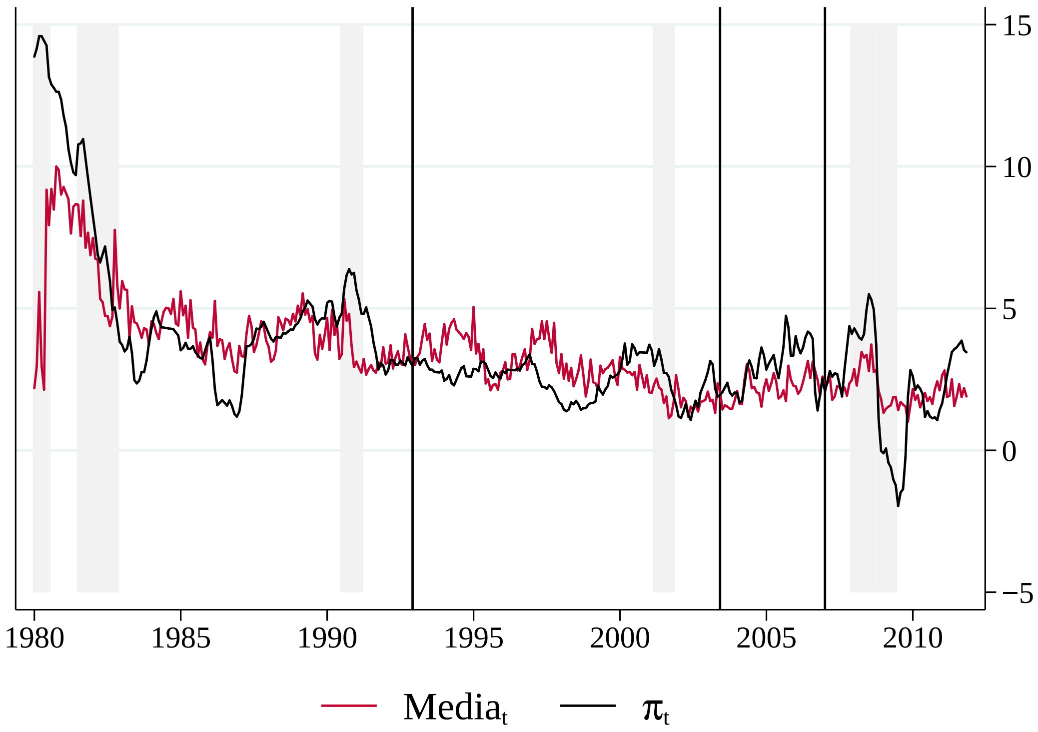
[14]See Lexis Nexis for details.

"inflation" is actually about current or future price developments or refers to an historical episode unrelated to the present situation. Second, we cannot separate press reports from opinions, or capture whether an article describes the current inflationary (or deflationary) environment as problematic. However, we can compare our media measure with a more sophisticated news series compiled by the media research institute *Media Tenor*[15]. This data is manually collected and adjusted and thus does not suffer from the problems of the Lexis Nexis series. Comparing the two news series in Figure (A.2) in the Appendix reveals that the differences are not too large, overall, we find a correlation of 0.7. We do not use the Media Tenor series in this chapter since it is only available since January 1998, however, due to the close connection of the series with the Lexis Nexis measure, we are confident that our results are not affected too much by the measurement problems of our news series.

Next, we have to scale the media data in order to rule out that a decreasing number of articles on inflation is simply due to the fact that the size of the newspaper is shrinking. Following Carroll (2003), we divide the number of news reports by the maximum number of articles on inflation published in any quarter of the sample. While this might not fully take into the effect of a shrinking newspaper size, our series is quite close to an alternative scaling procedure used by Pfajfar and Santoro (2013). As it is shown in Figure (A.2) in the Appendix, dividing the number of articles on inflation by the total number of articles in each quarter does lead to virtually unchanged results. Plotting our news series in Figure (2.2) shows that in general media coverage of inflation moves together with the inflation rate, albeit the correlation varies slightly over time.

[15]See Media Tenor for details.

Figure 2.2: News Coverage of Inflation



Note: The graph shows the number of news reports on inflation published in *The New York Times* and *The Washington Post*, together with the annual inflation rate π_t. The news series is scaled by its maximum value. The gray shaded areas denote NBER recessions, and the vertical lines indicate the structural breaks in 1992:12, 2003:06 and 2007:01 found in applying the QLR-test to equation (2.3).

2.3.2 Micro Level Data of Households' Inflation Expectations

We now take a closer look at the micro level data of households' inflation expectations provided by the *Michigan Survey*, notably because the survey contains a question on the economic news households have perceived in the period before answering the survey. Survey participants are asked: "During the last few months, have you heard of any favorable or unfavorable changes in business conditions?". If respondents answer in the affirmative, they are asked "What did you hear?". In Table (2.1), we provide summary statistics of the micro data of the Michigan Survey keeping in mind that expected inflation rates are truncated at $+/-30\%$.

Table 2.1: Summary Statistics - Micro Data - Michigan Survey

	Total	News Heard	News Heard: Inflation	Bad News: Infaltion	Good News: Inflation
N	233361	137480	15498	3126	12456
%N / N total	100.0	58.9	6.6	1.3	5.3
N missing	22031	10399	1291	128	1166
% missing / N total	9.4	4.5	0.6	0.1	0.5
% missing / N	9.4	7.6	8.3	4.1	9.4
% extreme / N nonmissing I	5.3	5.7	23.0	80.7	28.6
% extreme / N nonmissing II	1.1	1.9	16.7	79.2	21.0
% females	0.5	0.5	0.5	0.4	0.5
∅ age	45.4	45.5	44.5	42.0	45.1
∅ income	45120.6	50173.1	49255.5	49378.4	49090.0
∅ education	13.5	14.0	14.1	14.6	14.0
$\varnothing\pi^{exp}$	4.6	4.5	6.1	3.4	6.8
$\varnothing\pi^{exp,sd}$	5.8	5.7	6.4	4.8	6.6
$\varnothing GAPSQ^{SPF}$	30.9	29.0	37.0	22.0	41.0

Note: The table provides summary statistics for all survey participants ("Total"), and those who have stated to have heard news about changes in the economy ("News Heard"), about inflation ("News Heard: Inflation"), as well as bad and goods news on inflation. Expected inflation rates are truncated at $+/-30\%$. N denotes the total number of responses, $N/N\ total$ gives the fraction of answers with respect to the total number of survey participants. $N\ missing$ sums the number of missing responses, and $\%\ missing/N\ total$, and $\%\ missing/N$ shows the fraction of missing answers relative to the total number of responses and the number of answers in each category (news heard, news heard: inflation, ...). $\%\ extreme/N\ nonmissing\ I/II$ computes the percentage of extreme answers relative to the number of nonmissing responses, where in I, extreme answers are defined as expected inflation rates $> 15\%$ and $< -5\%$, whereas II applies +30% as the upper limit. The rows $\%\ females$, $\varnothing age$, $\varnothing income$, $\varnothing education$ show the average number of females, as well as the average age, income and education of each category, where education is defined as the number of years in school. Finally, the last three columns show the average expected inflation rate, the average standard deviation of expected inflation, and the squared expectation gap with respect to professional forecasters surveyed in the SPF. Sample: 1980:01-2011:11.

The summary statistics reveal a number of interesting features. First, we note that out of a total of 233,361 survey responses over the time period 1980:01-2011:11, only a fraction of 59% claims to have heard about changes in economic conditions during the previous months. Considering only news about inflation, this fraction drops to 6.6%. Second, looking at the number of missing responses reveals that in total, about 10% do not answer the question on inflation expectations. The fraction of non-responses slightly drops if participants claim to have heard news about changes in the economy and about inflation and declines further for those who have heard bad news about inflation. While this finding might be taken as informal evidence that people who follow the news are better able to give a precise estimate of future inflation, having heard good news increases the fraction of non-responses. Third, we compute the percentage of extreme answers provided by survey participants relative to the number of non-missing values, defining both an estimate of future inflation of below

-5% and above $+15\%$, as well as below -5% and above $+30\%$ as an extreme answer. The results are fairly surprising. Whereas across all households, about 5% of those who have given an estimate of future inflation end up choosing extreme values, this number does not fall if we select only those individuals who have heard about economic news. Taking only those who have heard news about inflation, 23% give extreme answers, whereas for negative news about inflation, even a total of 80% (!) provide estimates for future inflation beyond -5% and $+15\%$. This latter result is in contrast to the general hypothesis of the epidemiology model or the sticky information model, namely that better informed households should give better expectations. While we seek to explore this issue further below, at the moment, we can only come up with a suggestive explanation. Instead of expecting that individuals seek to improve their forecast in response to having received bad news, it might rather be the case that bad news frighten individuals leading to extreme forecasts.[16] Fourth, we check whether a different sociodemographic background influences households' perception of economic news. As it turns out, this is not the case: For each news category, we find an equal amount of male and female respondents, an average age of about 45 years, and an income level of about $50,000. Moreover, the the level of education does not seem to increase households' attention to economic news. Finally, for each news category, we compute the average expected inflation rate over time. Compared to all households, the mean expected inflation rate is higher if participants have heard news about inflation, lower for bad news on inflation and highest for good news on inflation.[17] The same holds true for the standard deviation, and for the expectation gap defined as the squared difference between households' and experts' inflation forecast.

In order to analyze the differences in inflation expectations between households with different news perceptions in more detail, we compute the root mean squared forecast error (RMSE) using $e_t^{exp,prof} = \pi_t^{exp,prof} - \pi_{t+12}$.[18] This allows us to assess whether experts are indeed better in forecasting inflation than households, and whether survey participants who claim to have heard news about inflation are better in predicting future price changes compared to all households. Given that we do not know a priori whether survey participants try to forecast headline or core inflation, we compute the forecast errors for both series. Moreover, we calculate the forecast error for different subsamples classified with the help of structural break tests to be discussed below.

[16] Using the pseudo panel dimension of the Michigan survey, Dräger and Lamla (2013b) and Pfajfar and Santoro (2013) have indeed found that having heard bad news on inflation increases households' forecast error. In our paper, however, we did not want to use households' news perception as explanatory variable, but instead investigate whether households who have heard about inflation show different updating behavior and reaction to media reports. As we will show below, households who have heard bad news on inflation adjust faster to the best available forecast and are also more receptive to media reports. Since the forecast error of these households is higher compared to other survey participants, our results suggest that this can be explained by a false adjustment to media reports.

[17] Remember that these averages are computed with truncated data.

[18] Results are virtually the same if the mean absolute error is used. Detailed tables are available on request.

Overall, we find that experts are indeed better than households in forecasting inflation.[19] This holds true for forecast errors with respect to headline inflation given in Table (2.2) and with respect to core inflation in Table (2.3). Only between 1980:01 and 1992:12, experts are slightly worse in predicting headline inflation compared to all survey participants. More surprisingly, however, we find that households who have perceived news about inflation are worse in predicting future price changes compared to all households. And the largest errors are made by those who have heard bad news on inflation, which corresponds to our earlier finding that this group of households also has the highest fraction of extreme responses. This result might partly been driven by the low number of participants stating to have heard news about inflation. However, we also do not observe a significant improvement of expectations for households having heard news about economic issues in general, where the number of responses is considerably larger. Next, we observe that the forecasting ability varies over time. For both experts and the full sample of households, the forecast error reaches its lowest value between 1993:01 and 2003:06, a time period where the level and in particular the volatility of the inflation rate have been low. Finally, the pattern of forecast errors is unaffected by the choice of the inflation rate, with the exception of the fact that both households and experts make lower errors if the less volatile core inflation is used.

Table 2.2: Forecast Precision: RMSE - Headline Inflation

	80/1-11/11	80/1-92/12	93/1-03/6	03/7-07/1	07/2-11/11
$e_t^{exp,spf}$	1.32	1.43	0.80	1.18	2.04
$e_t^{exp,hh\,agg}$	1.78	1.59	1.39	1.31	3.19
$e_t^{exp,hh\,all}$	1.64	1.34	1.25	1.30	3.15
$e_t^{exp,hh\,nh}$	1.62	1.33	1.16	1.27	3.17
$e_t^{exp,hh\,ninfl}$	2.29	1.97	2.30	1.52	3.54
$e_t^{exp,hh\,ngood}$	2.44	2.16	2.07	1.94	4.08
$e_t^{exp,hh\,nbad}$	3.03	2.73	3.33	1.61	4.02
$mean(\pi_t)$	3.63	5.25	2.52	2.93	2.22
$var(\pi_t)$	6.69	10.18	0.43	0.68	3.38

Note: The RMSE is defined as $e_t = \sqrt{\frac{1}{T}\sum\left(\pi_{t,t+12}^{exp} - \pi_{t+12}\right)^2}$. $e_t^{exp,hh\,agg}$ denotes the forecast error computed with the cross-sectional mean of households' inflation expectations as provided by the Michigan University. By contrast, $e_t^{exp,hh\,all}$ uses the cross-sectional mean computed from the micro data using the truncation rule $+/-30\%$. $mean(\pi_t)$ gives the average inflation rate, and $var(\pi_t)$ the variance.

[19]Note that we use two series for the average inflation expectation of all households. $e_t^{exp,hh\,agg}$ denotes the forecast error computed with the cross-sectional mean of households' inflation expectations as provided by the Michigan University. By contrast, $e_t^{exp,hh\,all}$ uses the cross-sectional mean computed from the micro data using the truncation rule $+/-30\%$.

Table 2.3: Forecast Precision: RMSE - Core Inflation

	80/1-11/11	80/1-92/12	93/1-03/6	03/7-07/1	07/2-11/11
$e_t^{exp,spf}$	0.75	1.01	0.49	0.31	0.59
$e_t^{exp,hh\ agg}$	1.40	1.05	1.24	1.50	2.42
$e_t^{exp,hh\ all}$	1.31	1.01	1.06	1.44	2.32
$e_t^{exp,hh\ nh}$	1.29	1.05	0.95	1.40	2.31
$e_t^{exp,hh\ ninfl}$	2.02	1.56	2.17	1.77	2.99
$e_t^{exp,hh\ ngood}$	2.20	2.19	1.92	1.52	3.20
$e_t^{exp,hh\ nbad}$	2.79	2.19	3.24	2.00	3.75
$mean(\pi_t)$	3.66	5.71	2.56	2.03	1.77
$var(\pi_t)$	6.09	7.36	0.18	0.24	0.32

Note: The RMSE is defined as $e_t = \sqrt{\frac{1}{T}\sum\left(\pi_{t,t+12}^{exp} - \pi_{t+12}\right)^2}$. $e_t^{exp,hh\ agg}$ denotes the forecast error computed with the cross-sectional mean of households' inflation expectations as provided by the Michigan University. By contrast, $e_t^{exp,hh\ all}$ uses the cross-sectional mean computed from the micro data using the truncation rule $+/-30\%$. $mean(\pi_t)$ gives the average inflation rate, and $var(\pi_t)$ the variance.

2.3.3 Testing for Unit Roots and Cointegration

Before estimating the epidemiology model, we have to test the order of integration of the time series. We first test for unit roots with the standard Dickey-Fuller-GLS test (Elliott et al., 1996), the KPSS test (Kwiatkowski et al., 1992), and the Phillips-Perron test (Phillips and Perron, 1988). Second, we use the Zivot-Andrews test (Zivot and Andrews, 1992), and the Clemente-Montanes-Reyes test (Clemente et al., 1998) that allow for the possibility of structural breaks. As it has been shown by Perron (1990), if a series is stationary but has a structural break, conventional unit root tests often fail to reject the wrong null hypothesis of a unit root.[20]

The results of the various unit root tests are given in Table (A.1). Overall, experts' inflation expectations and the annual inflation rate have a unit root. Only in two cases, the t-statistic of the Phillips-Perron test and the innovative outlier version of the Clemte-Montanes-Reyes test, we reject the null hypothesis of non-stationarity. By contrast, the news media series is found to be stationary by most of the tests. Finally, results are rather mixed with regard to households' inflation expectations. The Dickey-Fuller test, the KPSS test, and the additive outlier version of the Clemente-Montanes-Reyes test suggest non-stationarity, whereas the Phillips-Perron test, the Zivot-Andrews test and the innovative outlier version of the

[20]Further details on the employed unit root tests can be found in Section (A.6.2) in the Appendix.

Clemente-Montanes-Reyes test reject the null of a unit root. Overall, we continue to assume that expected inflation of households and experts and inflation itself have a unit root, while we take the media series as stationary. Furthermore, the testing procedure has emphasized the importance of accounting for structural breaks during our sample period.

We move on by testing for cointegration. According to the epidemiology model, we expect a cointegrating relationship between households' inflation expectations and professional forecasters' predictions. In addition, if households' do not fully adjust to experts in the long-run, but show some adaptive expectation formation, we could also find a cointegration relationship between inflation expectations of households and experts, together with the inflation rate. We apply two test procedures. First, following Engle and Granger (1987), we check whether the residuals of the cointegrating regressions $\pi_t^{exp,hh} = \beta_1 \pi_t^{exp,prof} + \varepsilon_t$ and $\pi_t^{exp,hh} = \beta_1 \pi_t^{exp,prof} + \beta_2 \pi_t + \varepsilon_t$ are stationary thereby implying that the original series are cointegrated. Moreover, we include a constant in the cointegrating equations allowing for the possibility that households do not fully adjust to experts' forecasts even in the long-run but have a constant bias. Second, as in case of the unit root tests, the cointegration test is affected by structural breaks. Gregory et al. (1996) have shown that if the cointegrating vector shifts over time, the standard ADF test often fails to identify a true cointegration relationship. Gregory and Hansen (1996a,b) have thus proposed modified versions of the ADF test that can account for the presence of one unknown structural break.[21] With regard to the interpretation of the results, Gregory and Hansen (1996a) suggest the following approach. If the standard ADF test does not reject the null hypothesis of no cointegration, but the modified ADF does, one can conclude that there exists a cointegrating relationship with a structural break. By contrast, if both tests reject the null hypothesis, this should not lead to the conclusion that a structural break exists, since the standard ADF test is more powerful in testing against cointegration without time shifts. In this second case, further tests are needed to identify possible breaks in the long-run relationship.

Table (A.2) in the Appendix shows the results of the relationship between the expected inflation series only. If we allow for a constant in the long-run equation, the standard ADF test rejects the null hypothesis of no cointegration, whereas adding a constant only suggests cointegration if the updated critical values of MacKinnon (2010) are used. Moving on to the Gregory-Hansen test, we find robust evidence of cointegration. As it is shown in Table (A.3), the same result is found if we include headline inflation in the long run relationship. Hence, we take these results as robust evidence for cointegration, while again noting that the relationship between expected inflation of households and experts and actual inflation is likely to vary over time.

[21] Again, details on the test procedures are given in Section (A.6.2).

2.4 The Epidemiology Model Without News

2.4.1 Aggregate Data

We start our empirical analysis by briefly replicating and extending Carroll (2003)'s original results, i.e. we estimate the epidemiology model given in equation (2.1) with aggregate survey data:

$$\pi_t^{exp,hh} = \alpha_0 + \alpha_1 \pi_t^{exp,prof} + \alpha_2 \pi_{t-1}^{exp,hh} + \varepsilon_t \tag{2.6}$$

The epidemiology model holds if $\alpha_1 + \alpha_2 = 1$, and $\alpha_0 = 0$. We then augment this baseline regression by adding the annual inflation rate of the previous period, thereby allowing for adaptive expectation formation:

$$\pi_t^{exp,hh} = \alpha_0 + \alpha_1 \pi_t^{exp,prof} + \alpha_2 \pi_{t-1}^{exp,hh} + \alpha_5 \pi_{t-1} + \varepsilon_t, \tag{2.6a}$$

where $\pi_{t-1} \in \{\pi_{t-1}^{CPI}, \pi_{t-1}^{CORE}\}$. We include both headline inflation CPI and core inflation $CORE$, given that the latter receives a lot of attention in policy discussions and public debates in the US. Finally, we allow for different updating coefficients in times of recessions by including interaction dummies that take the value 1 in recessions periods according to the NBER dating procedure. Thus, we have[22]

$$\pi_t^{exp,hh} = \alpha_0 + \alpha_1 \pi_t^{exp,prof} + \alpha_2 \pi_{t-1}^{exp,hh} + \alpha_3 rec \pi_t^{exp,prof} + \alpha_4 rec \pi_{t-1}^{exp,hh} + \varepsilon_t \tag{2.6b}$$

Before estimating the baseline equation of the epidemiology model together with its two extensions, we test for unknown structural breaks using the Quandt-Likelihood-Ratio test developed by Quandt (1960) and Andrews (1993).[23] Following Hansen (2001), we apply the OLR test to the entire sample, split the sample if we find evidence for a structural break and apply the test again on the longest of the two resulting time spans. The epidemiology model is then estimated for the implied subperiods. As we have already mentioned earlier, the structural breaks found by the QLR test fit quite well to the four different regimes that we have identified in the graphical analysis.[24] Across the different model specifications, we find a break at the end of the 1980s or the beginning of the 1990s, which corresponds to

[22]Since we have found robust evidence for cointegration, we estimate the different equations in levels following the super consistency argument of Engle and Granger (1987) whereby parameter estimates converge faster to their true values than consistent estimates of stationary variables. On the other hand, estimating the epidemiology model in an error correction framework would be more efficient if the residuals are serially correlated (Kirchgässner and Wolters, 2007). In our approach, however, for sake of comparison with much of the literature, we have chosen to apply robust standard errors.

[23]We use the critical values from Stock and Watson (2007).

[24]The estimated break tests can be found in Table (A.4) in the Appendix. We ignore structural breaks if splitting the sample at this date leads to a time period with too few observations.

the period of falling inflation and a close comovement between households' and experts' inflation forecasts. Afterwards, until the second break typically found in the mid of the 2000s, households' forecasts have been higher than the prediction of experts. Finally, we identify a break at the financial crisis around the end of 2007, splitting the remaining sample into a period marked by rising household expectations and stable expert predictions, and a fourth period characterized by volatile household expectations whereas experts did not react that much to the pronounced up and down of the inflation rate.[25]

The corresponding estimation results of the epidemiology model can be found in Table (2.4), where the upper panel reports the results for the baseline version, whereas the lower panel adds the lagged headline inflation. We estimate each of the models for the full sample 1980:01-2011:11 and for 1980:01-2009:06 when adding the recession dummies, as well as for the subsamples implied by the QLR-tests. In general, we find a relatively low degree of updating between 0.14 and 0.20 for the different versions of the epidemiology model estimated over the full sample period. This is in line with the literature using quarterly data for the US (Carroll, 2003, Pfajfar and Santoro, 2013) and monthly data for European countries (Döpke et al., 2008). However, if we split the sample at the structural breaks suggested by the QLR test, we get a considerably larger degree of updating for all periods and model variants. In addition, in the financial crisis, households are found to react even stronger to experts compared to the previous periods, i.e. it seems that households look for external advice the more uncertain they are about future prices.

Next, if we judge the performance of the baseline version of the epidemiology model according to whether the data rejects the restrictions $\alpha_1 + \alpha_2 = 1$ and $\alpha_0 = 0$, the results also depend on the time span. While for the full sample, the restrictions are generally rejected, thus providing evidence against the epidemiology model, we get the contrary result for some subperiods. Since there is no clear pattern of rejection and nonrejection, we conclude that this test is not be a robust tool to judge the performance of the epidemiology model.

As regards the effect of past inflation, we generally find evidence for some degree of adaptive expectation formation. Rising headline inflation affects households' expectation positively over the full sample, albeit the effect has vanished since the middle of the 2000s.[26] Finally, we observe that the updating process does not change much during recessions. Only during the disinflation period in the 1980s, households' paid more attention to experts during recessions.

[25]We have also applied the multiple breakpoint tests proposed by Bai (1997) and Bai and Perron (1998). Overall, the results confirm our previous analysis (results are available upon request). We can identify the break at the beginning of the 1990s and at the financial crisis, whereas the break at the beginning of 2000 is dated some months earlier.

[26]The results are similar using core inflation instead of headline inflation, apart from the fact that rising core inflation affects households' expectations *negatively*. The results are show in Table (A.5) in the Appendix.

Table 2.4: Results: Aggregate Data

(a) Baseline Model

	80/1-11/11	80/1-89/11	89/12-03/6	03/7-08/1	08/2-11/11	80/1-09/6	80/1-81/11	81/12-92/12	93/1-08/1	08/1-09/6
$\pi_t^{exp,prof}$	0.20***	0.58***	0.51***	0.54*	1.54***	0.20***	2.30**	0.49***	0.32***	1.56***
	(0.04)	(0.11)	(0.08)	(0.30)	(0.38)	(0.04)	(1.01)	(0.11)	(0.10)	(0.35)
$\pi_{t-1}^{exp,hh}$	0.76***	0.53***	0.52***	0.54***	0.58***	0.74***	0.32	0.33***	0.64***	0.52***
	(0.03)	(0.08)	(0.07)	(0.13)	(0.10)	(0.04)	(0.22)	(0.09)	(0.06)	(0.15)
$\pi_t^{exp,prof} rec$						0.11***	0.24***	0.02	-0.11	0.15
						(0.03)	(0.08)	(0.06)	(0.14)	(0.17)
$\pi_{t-1}^{exp,hh} rec$						-0.07***	-0.06	0.02	-0.00	0.05
						(0.03)	(0.06)	(0.05)	(0.11)	(0.10)
$cons$	0.36***	-0.40*	0.27*	0.49	-1.29**	0.38***	-13.09*	0.91**	0.44**	-1.75***
	(0.08)	(0.24)	(0.16)	(0.42)	(0.53)	(0.11)	(6.84)	(0.40)	(0.22)	(0.50)
R^2	0.89	0.91	0.76	0.52	0.85	0.90	0.65	0.68	0.66	0.86
N	383	119	163	55	46	354	23	133	169	29
Wald	0.016	0.026	0.604	0.702	0.000	0.430	0.046	0.088	0.157	0.000

(b) Including CPI Inflation

	80/1-11/11	80/1-92/12	93/1-03/6	03/7-07/1	07/2-11/11	80/1-09/6	80/1-92/12	93/1-04/8	04/9-09/6
$\pi_t^{exp,prof}$	0.14***	0.41***	0.61***	0.49	0.89***	0.15***	0.31***	0.43***	1.41***
	(0.04)	(0.10)	(0.11)	(0.34)	(0.29)	(0.05)	(0.11)	(0.10)	(0.33)
$\pi_{t-1}^{exp,hh}$	0.66***	0.44***	0.42***	0.35*	0.75***	0.58***	0.50***	0.45***	0.41***
	(0.04)	(0.08)	(0.09)	(0.19)	(0.11)	(0.06)	(0.09)	(0.09)	(0.16)
π_{t-1}^{CPI}	0.10***	0.13**	0.13**	0.19	-0.04	0.16***	0.14**	0.17**	0.16
	(0.03)	(0.05)	(0.06)	(0.13)	(0.06)	(0.04)	(0.06)	(0.07)	(0.12)
$\pi_t^{exp,prof} rec$						0.10***	0.11***	-0.17	0.11
						(0.03)	(0.04)	(0.12)	(0.17)
$\pi_{t-1}^{exp,hh} rec$						0.05	-0.13	0.22	0.30*
						(0.06)	(0.11)	(0.30)	(0.16)
$\pi_{t-1}^{CPI} rec$						-0.12**	0.03	-0.23	-0.28*
						(0.05)	(0.09)	(0.32)	(0.15)
$cons$	0.62***	0.28	0.01	0.71	-0.69	0.65***	0.37	0.35*	-1.63***
	(0.12)	(0.27)	(0.17)	(0.46)	(0.55)	(0.12)	(0.27)	(0.18)	(0.62)
R^2	0.90	0.91	0.75	0.54	0.80	0.91	0.91	0.73	0.76
N	383	156	126	43	58	354	156	140	58
Wald	0.000	0.661	0.009	0.880	0.018	0.010	0.476	0.229	0.000

Note: HAC standard error in parentheses. Wald denotes p-value of test $H_0 : \alpha_1 + \alpha_2 = 1$, $H_0 : \alpha_1 + \alpha_2 + \alpha_5 = 1$, $H_0 : \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 1$, and $H_0 : \alpha_1 + ... + \alpha_6 = 1$. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

2.4.2 Micro Data

We now reestimate the baseline epidemiology model, but use micro data instead of the cross-sectional aggregate mean of households' forecasts. By doing this, we can check whether our previous results also hold if we analyze the expectations of individuals directly. Furthermore, we are able to test whether households that claim to have heard news about inflation have a higher updating coefficient compared to others. Applied to micro data, the epidemiology model including expectations of each household i and the lagged inflation rate is given as[27]:

$$\pi_{i,t}^{exp,hh} = \alpha_0 + \alpha_1 \pi_t^{exp,prof} + \alpha_2 \pi_{i,t-1}^{exp,hh} + \alpha_5 \pi_{t-1} + \varepsilon_t \tag{2.7}$$

Since the micro data set of the Michigan survey is designed as a repeated cross section, we have to find a proxy for households' inflation expectations in the previous period. For that purpose, we fit both a pooled OLS regression using the cross-sectional average of inflation expectations in the previous period, as well as a pseudo-panel approach using cohort expectations. Both approaches have their advantages and disadvantages. Pooling the data keeps the estimation simple and allows us to evaluate the different updating coefficients of households who have followed the news. By contrast, the pseudo panel approach keeps some of the cross-section variation included in the data, but requires further assumptions for carrying out the estimation.

Pooled OLS In the pooled regression approach, we follow an idea that has also been used by Malmendier and Nagel (2013). If households with different expectations share their opinion, they might eventually adjust to the overall inflation forecast hold by all households in the economy. Therefore, we use the cross-sectional average of households' expectations in the previous period as a proxy for individuals' past inflation forecast. Furthermore, we add a set of dummy variables to account for the impact of demographic factors on expectations that has been often found in the literature (Menz and Poppitz, 2013). More precisely, we include the categories sex, race, family status, number of children, place of residence, education and age.[28] Using micro data allows us to separate individuals who have heard news about inflation from others who did not. For that purpose, we estimate the baseline model (2.7) using only the expectations of households who claim to have heard news about inflation where we assume that those individuals adjust to the average expected inflation in the previous period of the corresponding news heard group. Assuming instead that individu-

[27]We drop the recession dummy from our regressions given its small and mostly insignificant impact in the previous analysis using aggregate data.

[28]Due to a break in the series, we did not include income. However, since income is highly correlated with education, we are confident that this will not affect our results too much. Furthermore, we are not able to include housing status and stock ownership, because these categories are only added to the survey in 1990.

als, no matter whether they have heard news about inflation, adjust to the average expected inflation of all households leads to fairly similar results.[29] Denoting the average expected inflation rate of households that have heard news about inflation with $\bar{\pi}_t^{exp,NEWS}$, we have

$$\pi_{i,t}^{exp,hh} = \alpha_0 + \alpha_1 \pi_t^{exp,prof} + \alpha_2 \bar{\pi}_{t-1}^{exp,NEWS} + \pi_{t-1} + \varepsilon_t,$$
$$\text{where } NEWS \in \{NH, NINFL, NGOOD, NBAD\}$$
$$\text{and } \bar{\pi}_{t-1} = \sum_{i=1}^{N} \pi_{i,t-1} \tag{2.8}$$

We expect individuals who claim to have followed the news to form expectations that are closer to the best available forecast compared to other households.

The results of the pooled OLS regressions including lagged headline inflation are shown in Tables (2.5) and (2.6).[30] Note that we split the sample according to the structural breaks found in our estimates using aggregate data, both for sake of comparison and for including the time variation in expectation updating. For lack of space, we do not show the results of the demographic variables, however, the results are fairly similar to those previously found in the literature.

Comparing first the baseline model in Table (2.5) with its counterpart using aggregate data in Table (2.4(b)), we observe a fairly similar impact of households' lagged inflation expectations. By contrast, the degree of updating to the best available forecast differs significantly. For the whole sample, rising inflation expectations of professional forecasters actually *lower* households' forecasts. Furthermore, while the effect becomes positive if we take into account the structural breaks, the degree of updating is found to be lower compared to the estimation using aggregate expectations. This is true for all periods with the exception of the financial crisis.

[29]Results are not shown but are available upon request.

[30]Results are very similar if we drop the inflation rate from the estimation. See Tables (A.6) and (A.7) in the Appendix.

Table 2.5: Results: Micro Data Including CPI Inflation I

	80/1-11/11	80/1-92/12	93/1-03/6	03/7-07/1	07/2-11/11	80/1-11/11	80/1-92/12	93/1-03/6	03/7-07/1	07/2-11/11
$\pi_t^{exp,prof}$	-0.06***	0.28***	0.34***	0.36***	1.02***	0.01	0.32***	0.41***	0.28	1.20***
	(0.02)	(0.04)	(0.07)	(0.14)	(0.12)	(0.02)	(0.05)	(0.07)	(0.18)	(0.15)
$\pi_{t-1}^{exp,hh}$	0.64***	0.46***	0.42***	0.44***	0.76***					
	(0.02)	(0.03)	(0.05)	(0.07)	(0.04)					
$\bar{\pi}_{t-1}^{exp,NH}$						0.73***	0.56***	0.50***	0.45***	0.70***
						(0.02)	(0.04)	(0.05)	(0.08)	(0.05)
π_{t-1}^{CPI}	0.08***	0.04*	0.20***	0.18***	-0.01	0.08***	0.05**	0.12***	0.19***	-0.03
	(0.01)	(0.02)	(0.04)	(0.05)	(0.02)	(0.01)	(0.02)	(0.05)	(0.06)	(0.03)
$cons$	2.24***	1.54***	1.44***	0.81**	-1.06***	1.70***	0.78***	0.92***	0.88*	-0.87*
	(0.13)	(0.22)	(0.21)	(0.36)	(0.40)	(0.17)	(0.29)	(0.28)	(0.47)	(0.50)
R^2	0.08	0.08	0.04	0.03	0.06	0.09	0.09	0.04	0.04	0.06
N	184886	85621	55063	18742	25460	111494	51898	30693	11162	17741
Wald	0.000	0.000	0.345	0.835	0.000	0.000	0.017	0.616	0.542	0.000

Note: Standard error in parentheses. Wald denotes p-value of the test $H_0 : \alpha_1+\alpha_2+\alpha_3 = 1$. Demographic control variables such as gender, education, race, and age are included in each regression. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Separating households according to their self-reported news perceptions, we observe some interesting results. First, our estimates in Table (2.5) suggest that households who have perceived some news on economic issues in the previous months tend to have a slightly higher degree of updating and also stick more strongly to average household expectations formed in the past. Interestingly, this effect changes over time: both for the full sample and for the period 2003:07-2007:01, the updating coefficient is not significantly different from zero and the point estimate of the subperiod is lower compared to the estimation using all households.

As regards households who have perceived news about inflation, the results in Table (2.6) support the hypothesis that following the news lowers the gap between households' and experts' expectations and the degree of expectation stickiness. For all periods except the 2003:07-2007:01 period, we get higher updating coefficients together with a lower impact of households' predictions from the previous period. Finally, if we look at the inflation news in more detail, we find important differences with regard to the updating process. Households who have heard good news about inflation only adjust to the expectations of experts over the full sample and between 1980:01 and 1992:12. Whereas the degree of updating is lower over the full sample compared to all households, we find a much higher impact for the fist subperiod when inflation came down by 14 percentage points. Finally, we find the largest updating coefficients if we consider only households who have heard bad news about inflation.

Table 2.6: Results: Micro Data Including CPI Inflation II

	80/1-11/11	80/1-92/12	93/1-03/6	03/7-07/1	07/2-11/11	80/1-11/11	80/1-92/12	93/1-03/6	03/7-07/1	07/2-11/11	80/1-11/11	80/1-92/12	93/1-03/6	03/7-07/1	07/2-11/11
$\pi_t^{exp,prof}$	-0.09	0.33**	0.67**	-0.93*	1.59***	0.23**	0.60***	0.73	-0.15	0.92	0.08	0.47**	1.19***	-0.53	1.70***
	(0.06)	(0.17)	(0.29)	(0.52)	(0.40)	(0.12)	(0.23)	(0.51)	(0.93)	(1.29)	(0.08)	(0.22)	(0.38)	(0.62)	(0.41)
$\bar{\pi}_{t-1}^{exp,NINFL}$	0.37***	0.20***	0.21***	0.11	0.63***										
	(0.04)	(0.07)	(0.07)	(0.12)	(0.11)										
$\bar{\pi}_{t-1}^{exp,NGOOD}$						0.17***	0.17**	0.23**	0.00	-0.03					
						(0.05)	(0.07)	(0.10)	(0.15)	(0.17)					
$\bar{\pi}_{t-1}^{exp,NBAD}$											0.24***	0.09	0.10	0.02	0.51***
											(0.04)	(0.08)	(0.07)	(0.13)	(0.10)
π_{t-1}^{CPI}	0.34***	0.30***	0.51***	0.39***	-0.05	0.19***	0.04	0.52	-0.38**	0.26	0.30***	0.26***	0.16	0.33**	-0.06
	(0.04)	(0.08)	(0.19)	(0.13)	(0.08)	(0.07)	(0.11)	(0.38)	(0.19)	(0.20)	(0.05)	(0.10)	(0.23)	(0.15)	(0.08)
$cons$	2.68***	0.29	1.80	5.04***	-0.73	1.81*	-0.65	0.40	6.45**	5.17	3.26***	0.84	2.71	4.65***	-0.69
	(0.57)	(1.00)	(1.41)	(1.54)	(1.47)	(0.93)	(1.29)	(2.13)	(3.14)	(5.09)	(0.69)	(1.34)	(1.78)	(1.76)	(1.55)
R^2	0.11	0.14	0.08	0.03	0.06	0.07	0.09	0.08	0.09	0.14	0.11	0.11	0.08	0.03	0.05
N	10820	5086	1580	2013	2141	2670	1756	442	264	208	8165	3364	1125	1742	1934
Wald	0.000	0.071	0.191	0.004	0.001	0.000	0.148	0.256	0.113	0.900	0.000	0.160	0.258	0.040	0.002

Note: Standard error in parentheses. Wald denotes p-value of the test $H_0 : \alpha_1 + \alpha_2 + \alpha_3 = 1$. Demographic control variables such as gender, education, race, and age are included in each regression. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Pseudo Panel We next turn to our second approach to proxy lagged household inflation expectations by fitting a pseudo-panel. Building on the literature (McKenzie, 2004, Verbeek, 2012), we compute cohort averages using the year of birth of survey participants as the cohort variable, where we follow Malmendier and Nagel (2013) and compute the cohort averages using sample weights. Denoting the mean inflation expectation of individuals belonging to cohort c in period t as $\bar{\pi}_{c,t}^{exp,hh}$, we estimate the following equation:

$$\bar{\pi}_{c,t}^{exp,hh} = \alpha_0 + \alpha_1 \pi_t^{exp,prof} + \alpha_2 \bar{\pi}_{c,t-1}^{exp,hh} + \varepsilon_t \tag{2.9}$$

In addition, we include cohort specific fixed effects α_c which are defined as dummy variables that take the value 1 if individuals belong to the corresponding cohort:

$$\bar{\pi}_{c,t}^{exp,hh} = \alpha_c + \alpha_1 \pi_t^{exp,prof} + \alpha_2 \bar{\pi}_{c,t-1}^{exp,hh} + \varepsilon_t \tag{2.10}$$

In the literature on pseudo panels (Deaton, 1985, Moffitt, 1993, Collado, 1997), it has been emphasized that getting consistent estimates of equations (2.9) and (2.10) depends on the specification of the cohorts. A cohort is defined by Deaton (1985) as "a group with fixed membership". This rules out the use of the "news heard" variable, since individuals do not always hear news on inflation over time. Therefore, we follow much of the literature and define individual membership to cohorts by year of birth. In the empirical application, one has to determine the number of cohort as well as the size of cohorts, where it has been shown that both choices affect the consistency of the estimator (McKenzie, 2004). As it has been documented by Glocker and Steiner (2007), the error terms in equations (2.9) and (2.10) are correlated with lagged household expectations, thus OLS will be biased. This bias does not arise from the typical correlation of fixed effects with the error term in dynamic panels, but rather stems from the measurement error given that a cohort in period t does not contain the same individuals in period $t-1$. This bias will vanish if either the number of cohorts tends to infinity, or if the size of the cohorts gets large. However, the cross-section of our data set is relatively small, since only about 500 individuals are interviewed each periods. We choose to construct 10 cohorts by separating households into age groups 20 to 25, 25 to 30, ..., 65 to 70. This yields an average cohort size of 43 if we consider all households.[31] Hence, albeit we expect biased estimates, we still estimate our pseudo panel with OLS instead of fitting an IV regression. As it has been shown by McKenzie (2004), the IV estimator does not suffer from the downward bias of OLS, but its results vary a lot in simulation studies in case of small cohort sizes which is true for our data set. Therefore, we estimate our pseudo panel with OLS, keeping in mind that the estimators are expected to be biased downwards. Finally, note that we estimate the model separately for each of the four news heard answer

[31] For some months at the end of the sample, we do not have observations for the two youngest age cohort groups. We interpolate the missing data.

categories. In this case, we make sure that we have a cohort size of at least 10 in each month, whereas over time, we still get average cohort sizes of about 44.
Overall, the results confirm those of the OLS estimates. Looking at the expectations of the full set of households, and of those who have heard news about changes in business conditions in Table (2.7), we find much larger updating coefficients for all sample periods. In addition, the degree of expectation stickiness turns out to be lower, however, the difference is not that large if we take into account the downward bias of the pseudo panel estimation. Next, looking at households who have heard news about inflation in Table (2.8), the results are generally similar compared to the pooled OLS approach, whereas again, we find a slightly lower impact of households' lagged inflation expectations. As regards the differences between households' information sets, we do not observe higher degrees of updating for households who have heard news about changes in business conditions, or about inflation. Only for bad news about inflation, experts' forecast exert a higher impact on households over the full sample period.[32] Overall, the pseudo panel yields more reasonable results, given that we do not get the *negative* updating coefficient that has been found in some cases of the pooled OLS approach.

Table 2.7: Results: Pseudo Panel I

	80/2-11/11	80/2-92/12	93/1-03/6	03/7-07/1	07/2-11/11	80/2-11/11	80/2-92/12	93/1-03/6	03/7-07/1	07/2-11/11
$\pi_t^{exp,prof}$	0.14***	0.37***	0.81***	0.76***	1.50***	0.14***	0.48***	0.40***	0.71***	1.60***
	(0.02)	(0.06)	(0.07)	(0.20)	(0.22)	(0.03)	(0.07)	(0.04)	(0.25)	(0.27)
$\bar{\pi}_{t-1}^{expw,hh}$	0.30***	0.25***	0.07**	0.12**	0.22***					
	(0.02)	(0.02)	(0.03)	(0.05)	(0.04)					
$\bar{\pi}_{t-1}^{expw,hh} NH$						0.23***	0.16***	0.15***	0.14***	0.24***
						(0.02)	(0.02)	(0.02)	(0.05)	(0.04)
π_{t-1}^{CPI}	0.21***	0.14***	0.29***	0.29***	0.15***	0.26***	0.15***	0.18***	0.28***	0.13***
	(0.02)	(0.03)	(0.04)	(0.07)	(0.04)	(0.02)	(0.03)	(0.02)	(0.08)	(0.05)
$cons$	1.85***	1.49***	0.34**	1.10***	-0.17	1.87***	1.15***	1.33***	1.18**	-0.42
	(0.08)	(0.19)	(0.17)	(0.39)	(0.43)	(0.10)	(0.22)	(0.12)	(0.50)	(0.53)
R^2	0.52	0.59	0.29	0.30	0.33	0.42	0.52	0.51	0.22	0.26
N	3820	1550	1260	430	580	3820	1550	2810	430	580
Wald	0.000	0.000	0.002	0.311	0.000	0.000	0.000	0.000	0.550	0.000

Note: Standard error in parentheses. Wald denotes p-value of test $H_0 : \alpha_1 + \alpha_2 + \alpha_3 = 1$. Regressions include cohort dummy variables. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

[32] Note that this result might stem from the fact that the cohort size becomes relatively small if we consider only households who have heard news about inflation.

Table 2.8: Results: Pseudo Panel II

	80/2-11/1	80/2-92/12	93/1-03/6	03/7-07/1	07/2-11/1	80/6-09/8	80/6-92/12	93/1-03/6	03/7-07/1	07/2-09/8	80/2-11/1	80/2-92/12	93/1-03/6	03/7-07/1	07/2-11/1
$\pi_t^{exp,prof}$	0.14*	0.38**	0.61***	-0.26	1.11**	0.05	0.11	0.31***	0.48	-0.57	0.36***	0.52**	0.77***	0.23	1.16***
	(0.08)	(0.19)	(0.23)	(0.54)	(0.48)	(0.05)	(0.14)	(0.11)	(0.30)	(0.55)	(0.08)	(0.22)	(0.23)	(0.59)	(0.44)
$\bar{\pi}_{t-1}^{expw,hh} NINFL$	0.33***	0.21***	0.46***	0.31***	0.33***										
	(0.02)	(0.02)	(0.03)	(0.05)	(0.04)										
$\bar{\pi}_{t-1}^{expw,hh} NGOOD$						0.72***	0.59***	0.85***	0.67***	0.70***					
						(0.01)	(0.02)	(0.01)	(0.04)	(0.04)					
$\bar{\pi}_{t-1}^{expw,hh} NBAD$											0.45***	0.35***	0.58***	0.35***	0.60***
											(0.01)	(0.02)	(0.02)	(0.05)	(0.04)
π_{t-1}^{CPI}	0.23***	0.24***	0.34**	-0.02	-0.03	0.07**	0.05	-0.04	-0.09	0.23**	0.03	0.08	0.09	-0.23	-0.26***
	(0.05)	(0.09)	(0.16)	(0.18)	(0.09)	(0.03)	(0.07)	(0.07)	(0.10)	(0.09)	(0.05)	(0.10)	(0.16)	(0.19)	(0.08)
$cons$	1.92***	0.92	-0.17	3.76***	1.69*	0.42***	1.18***	-0.60**	-0.36	1.54	1.76***	0.77	0.17	3.38***	0.66
	(0.24)	(0.60)	(0.63)	(1.10)	(0.96)	(0.16)	(0.43)	(0.29)	(0.60)	(1.19)	(0.26)	(0.71)	(0.62)	(1.20)	(0.88)
R^2	0.20	0.18	0.27	0.12	0.15	0.55	0.42	0.75	0.60	0.65	0.27	0.20	0.40	0.15	0.40
N	3710	1540	1260	430	480	3504	1504	1260	430	310	3718	1548	1260	430	480
Wald	0.000	0.119	0.044	0.033	0.347	0.000	0.002	0.215	0.819	0.210	0.001	0.653	0.031	0.190	0.215

Note: Standard error in parentheses. Wald denotes p-value of test $H_0 : \alpha_1 + \alpha_2 + \alpha_3 = 1$. Regressions include cohort dummy variables. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

2.5 Including News I: Expectation Gap

2.5.1 Aggregate Data

We now extend the analysis by adding the news media to the baseline framework. As we have stated above, this can be done by assuming linear media effects and regressing the squared gap between households' and experts inflation forecast on the number of media reports, or by allowing for nonlinear media effects and estimating the epidemiology model (2.2) directly. We start with the gap equation (2.3) using aggregate data:

$$GAPSQ_t = \alpha_1 + \alpha_2 MEDIA_t + \alpha_3 \pi_{t-1} + \varepsilon_t$$
$$\text{where } GAPSQ_t = \left(\pi_t^{exp,hh} - \pi_t^{exp,prof}\right)^2 \tag{3}$$

According to Carroll (2003), a rising number of news reports $MEDIA_t$ should lower the squared difference between households' and experts' expectations. We also include the lagged inflation rate in order to rule the possibility that the media simply report the most recent price changes without adding additional information. As in our previous analysis, we run QLR tests to check for structural breaks.

Starting with the results of these tests, the procedure identifies break dates that are fairly similar to those found earlier.[33] We find breaks at the beginning of the 1990s, at the beginning of the 2000s, and at the aftermath of the financial crisis. As in our previous analysis of the model without news coverage, we test whether the news media affect inflation expectations differently in recessions by adding a dummy variable such that equation (2.3) becomes

$$GAPSQ_t = \alpha_1 + \alpha_2 MEDIA_t + \alpha_3 \pi_{t-1} + \alpha_4 MEDIA_t rec + \alpha_5 \pi_{t-1} rec + \varepsilon_t \tag{2.3a}$$

We expect that in times of bad economic conditions, consumers will turn to the news media to keep up-to-date to the present and the future state of the economy. Hence, changes in the number of news reports as well as in the inflation rate should exert a stronger impact on the expectation gap during recessions.

The results shown in Table (2.9) reveal some interesting insights. First, estimating the baseline equation (2.3) for the full sample, we find the expected negative news effect, together with a positive impact of the inflation rate. This result is in contrast to Pfajfar and Santoro (2013) who have found a positive news effect using quarterly data and neglecting possible structural breaks. However, it turns out that a rising news coverage of inflation does not al-

[33]See Table (A.8) for details on the QLR tests. Again, our results are broadly confirmed by applying the multiple breakpoint test producers of Bai (1997) and Bai and Perron (1998).

ways lower the difference between households' and experts' beliefs on future prices. While this is true for the period 1980:01-1992:12, the coefficient is much lower and not significantly different from zero between 1993:01-2003:07, and even turns positive thereafter. In addition, more news reports increase the expectation gap the strongest during the financial crisis. Second, we observe that in general, the news effect is not different during recessions, with the exception of the period 2004:02-2009:06 where the positive news effect becomes even stronger.[34]

Table 2.9: Results: Expectation Gaps - Aggregate Data

	80/1-11/11	80/1-92/12	93/1-03/7	03/8-06/12	07/1-11/11	80/1-09/6	80/1-90/6	90/7-92/12	93/1-04/2	04/3-09/6
$MEDIA_t$	-8.54***	-5.99***	-0.48	4.33	37.18***	-6.92***	-5.53***	3.88**	-0.97	22.40**
	(2.09)	(1.09)	(0.59)	(7.44)	(12.07)	(1.71)	(1.46)	(1.95)	(0.69)	(11.31)
$MEDIA_t rec$						1.31	-1.51	1.40	0.39	15.82**
						(1.54)	(1.13)	(1.27)	(2.28)	(6.76)
π_{t-1}^{CPI}	0.54***	0.53***	0.45***	0.87**	0.26	0.40***	0.45***	0.09	0.44***	0.50
	(0.13)	(0.06)	(0.09)	(0.43)	(0.40)	(0.14)	(0.09)	(0.13)	(0.10)	(0.47)
$\pi_{t-1}^{CPI} rec$						0.11	0.15*	0.06	-0.02	-0.17
						(0.12)	(0.09)	(0.06)	(0.16)	(0.55)
$cons$	2.38***	0.67	-0.24	-1.32	-4.16*	1.96***	0.84*	-1.26	-0.02	-4.35**
	(0.65)	(0.41)	(0.25)	(1.35)	(2.16)	(0.46)	(0.47)	(0.93)	(0.27)	(2.10)
R^2	0.18	0.40	0.20	0.20	0.41	0.21	0.44	0.68	0.17	0.51
N	382	155	127	41	59	353	125	30	134	64

Note: HAC standard error in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

2.5.2 Micro Data

Before thinking of possible explanations for the positive news effect, we reestimate the model using micro data, checking whether the result depends on the level of aggregation. Hence, we fit equation (2.3) with pooled OLS, again adding dummy variables to control for the effects of individuals' demographic and socioeconomic characteristics. Furthermore, we test whether the media have a different impact on the expectation gap of households who have recently heard news about inflation by computing gaps using only the different household groups who have heard news about economic issues in general, inflation, or good or bad news about inflation.

Table (2.10) shows the results. In contrast to the model without news media, we find larger media effects if we use micro data instead of macro data. Regarding the direction of the media effect, we now find *positive* media effects for all model specifications, i.e., the more

[34]If we estimate equation (2.3) without the inflation rate, the recession dummy is found to be significantly positive for the entire sample as well. See Table (A.9).

articles about inflation written in the media, the larger the difference between households' and experts' expectations. As regards the time variation of the media effect, we report the strongest impact during the financial crisis. However, the results do not depend on whether or not the financial crisis is included in our sample, since the media effect is significantly positive in all subsamples for households having heard news about inflation. Overall, we find that an increasing news coverage of inflation has a lower impact on the expectation gap of individuals who have heard news on inflation. By contrast, the effects are larger for households who have heard good or bad news on inflation. It is important to emphasize that we observe considerable time variation. Households having heard news on inflation in general or bad news are not affected by news coverage during the period of disinflation 1980:01-1992:12, and during the period of rising inflation 2003:08-2006:12. By contrast, households having heard good news on inflation only react to the news media during the period of disinflation. Finally, as in the estimation using aggregate data, a rising inflation rate increases the expectation gap.[35]

[35]The results with regard to the media effects remain the same if we drop the inflation rate. See Table (A.10).

Table 2.10: Results: Expectation Gaps - Micro Data

	all households					news heard				
	80/1-11/11	80/1-92/12	93/1-03/7	03/8-06/12	07/1-11/11	80/1-11/11	80/1-92/12	83/1-03/7	03/8-06/12	07/1-11/11
$MEDIA_t$	17.27***	5.39**	12.59***	5.52	47.50***	12.27***	3.45	-0.91	19.67**	54.09***
	(1.61)	(2.55)	(3.41)	(7.34)	(9.18)	(1.90)	(3.00)	(3.97)	(8.97)	(11.12)
π_{t-1}^{CPI}	2.95***	3.13***	4.98***	3.45***	-1.18***	2.91***	3.16***	4.17***	2.87***	-1.25***
	(0.10)	(0.13)	(0.47)	(0.42)	(0.28)	(0.11)	(0.16)	(0.55)	(0.51)	(0.34)
$cons$	33.30***	45.62***	23.81***	18.39***	30.15***	37.98***	51.08***	33.55***	17.40***	35.24***
	(1.92)	(3.34)	(3.14)	(3.36)	(4.56)	(2.45)	(4.25)	(3.93)	(4.43)	(5.89)
R^2	0.04	0.05	0.03	0.02	0.01	0.04	0.05	0.03	0.03	0.01
N	184886	85621	55482	17872	25911	111494	51898	30936	10660	18000

	news inflation					good news					bad news				
	80/1-11/11	80/1-92/12	93/1-03/7	03/8-06/12	07/2-11/11	80/1-11/11	80/1-92/12	93/1-03/7	03/8-06/12	07/1-11/11	80/1-11/11	80/1-92/12	93/1-03/7	03/8-06/12	07/1-11/11
$MEDIA_t$	14.13**	0.17	70.44***	-0.89	113.42***	19.49*	24.76*	17.37	-6.74	-98.14	20.23***	-1.47	117.95***	5.30	132.63***
	(5.52)	(8.15)	(23.64)	(24.54)	(31.60)	(10.55)	(13.99)	(38.87)	(46.35)	(111.49)	(6.68)	(10.26)	(30.97)	(26.92)	(33.46)
π_{t-1}^{CPI}	3.26***	3.34***	3.44	3.50***	-1.35	2.18***	2.10***	1.01	1.42	4.25	3.01***	3.00***	2.33	2.57	-2.40
	(0.31)	(0.41)	(3.01)	(1.32)	(1.30)	(0.66)	(0.75)	(5.78)	(1.85)	(3.05)	(0.37)	(0.54)	(3.74)	(1.58)	(1.48)
$cons$	51.55***	81.64***	46.41**	26.91**	22.06	43.99***	45.40***	129.09***	2.65	52.99	55.92***	97.95***	12.10	32.57**	19.65
	(8.58)	(14.70)	(22.05)	(12.76)	(17.67)	(13.07)	(17.08)	(36.32)	(21.96)	(50.34)	(10.73)	(20.64)	(27.22)	(14.20)	(18.97)
R^2	0.07	0.08	0.04	0.04	0.03	0.06	0.07	0.07	0.12	0.05	0.06	0.06	0.06	0.04	0.04
N	10820	5086	1591	1953	2190	2711	1769	456	250	236	8171	3364	1141	1708	1958

Note: Standard error in parentheses. Demographic control variables such as gender, income, education, race, and age are included in each regression. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

2.6 Including News II: STAR

In the final section of this chapter, we fit the epidemiology model with media coverage by allowing for nonlinear news effects. Instead of using the linear gap equation (2.3), we estimate equation (2.2) directly by applying the smooth transition (STAR) model proposed by Teräsvirta (2004) which allows for threshold effects and time variation in the estimated coefficients.

$$\pi_t^{exp,hh} = \lambda\left(MEDIA_t\right)\pi_t^{exp,prof} + (1-\lambda\left(MEDIA_t\right))\pi_{t-1}^{exp,hh} \tag{2}$$

Albeit the STAR model has also been extended to the panel framework (Gonzalez et al., 2005), we refrain from estimating the panel smooth transition model since it is unclear how the model behaves when applied to a pseudo panel. Hence, we estimate the nonlinear equation of the epidemiology model using aggregate data. As before, we compute the cross-sectional mean expectation of all households, and separately for those survey participants who have perceived some news on economic conditions and inflation.
We adopt the theoretical expression of the epidemiology model in (2.2) to the STAR framework as follows:[36]

$$\begin{aligned}\pi_t^{exp,hh} &= \alpha_0 + \alpha_1\pi_t^{exp,prof} + \alpha_2\pi_{t-1}^{exp,hh} + \alpha_3\pi_{t-1}\\ &\quad + G(q_t;\gamma,c)\left[\beta_0 + \beta_1\pi_t^{exp,prof} + \beta_2\pi_{t-1}^{exp,hh} + \beta_3\pi_{t-1}\right] + \varepsilon_t\end{aligned} \tag{2.2a}$$

where

$$G(MEDIA_t;\gamma,c) = \left(1 + exp\{-\gamma\prod_{j=1}^{m}(MEDIA_t - c)\}\right)^{-1} \tag{2.11}$$

The first line of equation (2.2a) gives the linear version of the epidemiology model without news media effect that we have estimated above. The second line incorporates possible nonlinear news effects: The coefficients in the square bracket describe the impact of professional forecasters expectations, households' lagged forecast from the previous period and the inflation rate if news coverage is above a certain threshold. The threshold effect is captured with the transition function G which is defined by the transition variable $MEDIA_t$, the degree of nonlinearity given by the shape of the transition function γ, and the threshold c. If γ is large, the model tends to a regime-switching framework where the coefficients α_i (β_i) describe the behavior of the model if the number of news reports is below (above) c. By contrast, if γ is

[36] As in the epidemiology model without media reports, including an error correction term in the STAR equation as in Lütkepohl et al. (1999) could yield more efficient estimates. We leave this for further research.

low, we would interpret the result as such that households slowly change their reaction to professional forecasters' expectations if the media increase the amount of news coverage on inflation. In the extreme case $\gamma = 0$, we are back to the linear model. Note that both the shape and the threshold are determined endogenously during the estimation process.

Applied to the epidemiology model, we expect that households react more strongly to an upward revision of expert predictions in periods of high news coverage of inflation, while at the same time relying less on their own belief formed in the previous period. Therefore, we should get $\beta_1 > 0$, and $\beta_2 < 0$. In addition, households might also react differently to the inflation rate depending on how prominently prices are discussed in the media. Therefore, we also include the inflation rate (and a constant) in the nonlinear part of the model and test whether this increases the fit of the model.

Before estimating the model, we check whether the relationship between households' and experts' expectations is indeed nonlinear, and whether the nonlinearity depends on the number of news reports. For that purpose, we apply the nonlinearity test proposed by Teräsvirta (2004), which approximates the nonlinear function G with its first-order Taylor expansion around $\gamma = 0$.[37] The test regression is given by

$$y_t = \theta_0' \mathbf{x}_t + \sum_{j=1}^{3} \theta_j' \tilde{\mathbf{x}}_t q_t^j + u_t^*, \tag{2.12}$$

where y_t is the dependent variable, $\mathbf{x}_t$ a vector of explanatory variables, q_t the transition variable and the error term $u_t^* = u_t + R$ with R as the remainder of the Taylor expansion. In this regression, the null hypothesis of linearity is $H_0 : \theta_1 = \theta_2 = \theta_3 = 0$, since the coefficients θ_j are a function of β_j and c. We test both for nonlinearity stemming from the number of news reports $MEDIA_t$ and from variation over time including a time trend. Finally, the smooth transition model is estimated with conditional maximum likelihood after performing a grid search on various values of γ and c.[38] As it is emphasized by Teräsvirta (2004), if γ is found to be large meaning that the model is close to a regime-switching model, numerical problems can often affect the estimated standard deviation of γ, which becomes equally large. However, given the nonexistence of the null hypothesis of the standard t-test of statistical significance of γ, these numerical problems should not lead to the conclusion that the model is in fact linear.

Table (2.11) shows the results for the full sample, separating households according to whether they have heard news about inflation. Starting with the linearity tests, we clearly reject the null hypothesis of linearity for all households and for households who have heard some news, only marginally missing the critical value for households who have heard news about inflation. By contrast, we cannot reject the null hypothesis for households who state to have

[37] Using the Taylor approximation is necessary given the fact that the linearity test is not defined under the null hypothesis $H_0 : \gamma = 0$. See for details Teräsvirta (2004).

[38] Note that it is also possible to allow for more than one threshold or, likewise, more than two regimes.

heard good or bad news on inflation.

Table 2.11: STAR - Full Sample - Different Information Sets

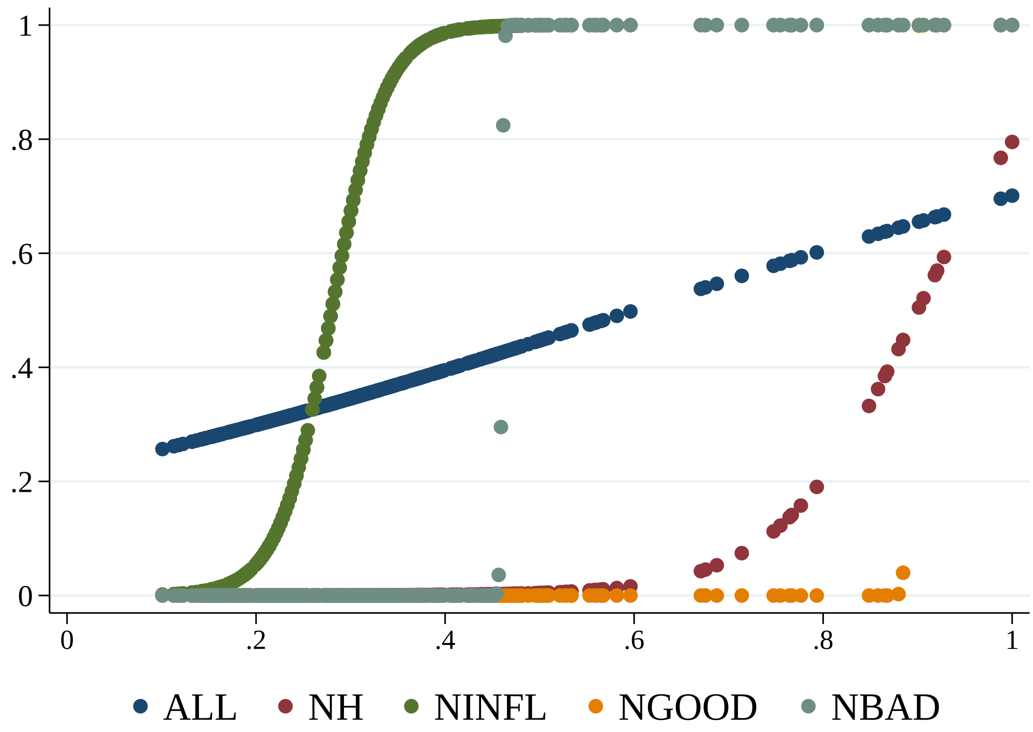
	all households	news heard	news inflation	news good	news bad
			linear part		
$cons$	0.62***	0.57***	1.71***	0.57**	2.09***
	(0.11)	(0.10)	(0.31)	(0.25)	(0.36)
$\pi_t^{exp,prof}$	0.14***	0.08**	-0.21	0.52***	0.42***
	(0.04)	(0.03)	(0.22)	(0.08)	(0.15)
$\pi_{t-1}^{exp,NEWS}$	0.71***	0.74***	0.47***	0.20***	0.21***
	(0.05)	(0.04)	(0.09)	(0.05)	(0.05)
π_{t-1}	0.04	0.04	0.29***		0.14*
	(0.03)	(0.03)	(0.07)		(0.08)
			nonlinear part		
$consG$	-	-	-	75.29***	-
				(24.72)	
$\pi_t^{exp,prof}G$	0.52**	2.61****	0.40**	-8.86***	0.35*
	(0.22)	(0.91)	(0.17)	(2.91)	(0.20)
$\pi_{t-1}^{exp,NEWS}G$	-0.89***	-2.60***	-0.27***	0.24	-0.34**
	(0.26)	(0.91)	(0.10)	(0.38)	(0.16)
$\pi_{t-1}G$	0.35***	-	-	-	-
	(0.11)				
γ	2.13*	13.56*	35.60	590.96	1004.50
	(1.24)	(8.18)	(42.91)	(3.10E+08)	(2.63E+05)
c_1	0.60***	0.90***	0.28***	0.89	0.46***
	(0.09)	(0.01)	(0.01)	(4.46E+03)	(0.03)
AIC	-1.42	-1.39	0.76	1.31	1.38
SBIC	-1.33	-1.31	0.85	1.39	1.46
HQ	-1.39	-1.36	0.80	1.35	1.41
$R^2 adj.$	0.91	0.88	0.46	0.32	0.29
Linearity Test					
w.r.t. $NEWS_t$	0.000	0.000	0.177	0.924	0.618
w.r.t. $TIME$	0.000	0.000	0.000	0.329	0.000

Note: Standard errors in brackets. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Values of linearity tests are p-values derived from the F-statistic.

Looking at the linear part of the STAR model which corresponds to the epidemiology model without news coverage, we get positive and significant coefficients of both experts' and households' expectations. Moreover, for all households except those who have heard good or bad news about inflation, the degree of updating is found to be relatively low. Next, looking at the nonlinear part of the model, we find evidence supporting the epidemiology model. The coefficient of experts' expectations is positive and statistically significant, whereas lagged household expectations have a negative impact. Hence, it seems that more news coverage increases the impact of professional forecasters' predictions on households' expectation formation, and lowers the degree of expectation stickiness. This holds true for all households except those who have heard good news on inflation. Interestingly, we find that the media also increases the degree of adaptive expectation formation. A rising number of articles increases the impact of the lagged inflation rate.

The STAR model allows us to endogenously determine how exactly the nonlinear media effect influences households' expectation formation. For that purpose, we plot the estimated transition functions in Figure (2.3) which have been calculated according to equation (2.11).

Figure 2.3: Estimated Transition Functions



Note: The graph shows the estimated transition functions that are computed using equation (2.11). ALL is the transition function of all households with threshold 0.6, NH the function for households who have heard about economic news in general with threshold 0.9, $NINFL$ is the function for households who have heard news about inflation with threshold 0.3, and $NGOOD$ and $NBAD$ show the functions of those who have heard good and bad news on inflation, with thresholds 0.9 and 0.5. Note that there are also some observations in the nonlinear regime of $NGOOD$, but these are covered by the other graphs.

Looking at the estimated thresholds c_1 of the transition functions, we observe that the strength of the media effect differs according to whether households follow the news or not. For all households, the threshold is found to be at 0.6, i.e. at an average amount of coverage. Households who claim to follow news on economic issues have a much larger threshold at 0.9, whereas households who have heard news about inflation already react to media reports if news coverage exceeds a value of 0.3. Moreover, the transition from the linear to the nonlinear regime differs remarkably. The full sample of survey participants adjusts very smoothly to rising amounts of newspaper articles on inflation, whereas having heard news in general and in particular news on inflation leads to a much quicker adjustment. Having heard any news on inflation, or bad news on inflation does not make a large difference. The threshold is lower compared to the full household sample, and the adjustment is much quicker. Finally, the STAR model does not perform well for households who have heard good news on inflation. The null hypothesis of linearity cannot be rejected and for the good news group, the estimated transition function while being very steep results in the fact that most of the estimated reactions belong to the linear regime. And as regards the estimated coefficients, the model suggests that households react *less* strongly to experts if newspapers write more about inflation.

Note that the linearity test with respect to the time trend suggests a structural break in the full sample for all households groups except for those who have heard good news on inflation. Hence, in the appendix, we also show the estimated STAR models for the different subperiods that have been identified earlier keeping in mind that the relatively low number of observations in some periods might affect the results. Overall, our general results also hold if we split the sample into different periods. Taking the full set of households (Table (A.11)), the media thresholds are considerably lower since 1993:01 and the transition functions get steeper throughout. If we estimate the STAR model for different households groups, no matter if households have heard some news, good, or bad news on inflation, the results are fairly similar with the exception of the time span 2003:08-2006:12 (Tables (A.12)-(A.15)). In this period, we find that more news coverage makes households react *less* to experts and stick more to their own forecast from the previous period. Finally, during the financial crisis period 2007:01-2011:11, households who have heard bad news on inflation adjust less to experts in response to rising news coverage. This latter result might be due to the fact that households have a different interpretation of "bad news on inflation" than professional forecasters. While the actual inflation rate dropped below zero in 2009, households seemed to think that the financial turmoil would result in higher inflation. By contrast, experts' expectations remained rather constant at an inflation rate of about 2% (see Figure 2.1)

2.7 Conclusion

In this chapter, we have analyzed three particular features of the epidemiology model: the time-dependence of the updating and expectation formation process, differences arising from the use of macro or micro level data, and possible non-linear news media effects.

Overall, we find that there is substantial time-variation in expectation formation process:

Households make different forecasts in periods of disinflation, rising inflation, low and stable inflation and during economic crisis. More precisely, households react more to experts if inflation is low and stable and during economic crisis, no matter if we use micro or macro data. This partly supports the idea of a state-dependent Phillips Curve suggested by Akerlof et al. (1996, 2000). The news effect also varies over time. Whereas in general, a rising number of news reports lowers the difference between households and experts, this effect mainly arises in times of disinflation. By contrast, more news coverage widens the gap during economic crisis.

As regards the use of macro or micro data, we find that our results do not depend on the exact specification made about households' lagged inflation expectations. Proxying these with the cross-sectional mean forecast as well as using a pseudo panel set up leads to similar conclusions. Overall, and in contrast to the results provided by Dräger and Lamla (2013b), we find lower updating frequencies if we use micro data. This difference might stem from the fact that we do not restrict the sample to include only households who are interviewed twice. In addition, we account for structural breaks in the data. As regards the news media, we get larger news effects if we use micro data, which is in contrast to the "paradox-of-agenda-setting" suggested by Krause and Gehrau (2007). Distinguishing households according to their self-reported information set results in the surprising observation that those who have perceived news on inflation are worse in forecasting inflation compared to other households. Moreover, a larger fraction of these households expect inflation rates that are above 15% or below -5%. The largest media effects are found for households who have heard some news or bad news about inflation.

Finally, we find robust evidence for non-linear news effects. Households in general adjust more to the best available forecast if news coverage moves beyond its normal level where the adjustment is generally slow. Looking at different types of households, we find that those who have heard some news or bad news on inflation are much quicker in adjusting to rising news coverage. Moreover, these types of households seem to have higher attention levels as they already react more to experts if the general level of news coverage is still low.

Summing up, our analysis generally supports the epidemiology model while at the same time raising some new questions. Households adjust gradually to experts, but the strength of this link varies over time. The degree of updating is generally lower on the individual level, but the news effect is found to be stronger. Households who have heard news on inflation react more strongly to the best available forecast, however, they are worse in forecasting inflation compared to other households.

Chapter 3

Households' Disagreement on Inflation Expectations and Socioeconomic Media Exposure in Germany

3.1 Introduction

The reasons why households with low income and low education, females, unemployed, and young and old individuals have higher inflation expectations and forecast errors compared to other households are still unclear. Some studies propose that these expectation differentials arise from different consumption baskets, while others suggest that they simply reflect differences in financial literacy. In this chapter, we explore another driving force of the demographic heterogeneity of inflation expectations, namely the impact of news media coverage. Models of sticky information (Mankiw and Reis, 2002) and rational inattention (Sims, 2003) propose that households' inflation expectations in the long run move in line with the best available forecast in the economy. In the short run, however, consumers' expectations may deviate considerably from the best available forecast, since the costs of gathering and processing this forecast might be too high. Carroll (2003) has argued that the news media can strengthen the link between households' and professional forecasters' expectations: the more articles published about inflation, the higher the likelihood that consumers get to know the best available forecast.

Carroll's epidemiology model of expectation formation relies on three crucial assumptions. First, households possess equal capacity of understanding and processing the media articles. Second, all agents have the same reading propensity, and third, all media sources report on inflation in a similar vein. Each of these assumptions can be questioned, and relaxing them might help explain demographic differences in inflation expectations. Regarding households' processing capacities, studies on financial literacy (Lusardi and Mitchell, 2008, Bruine de Bruin et al., 2010) show that the accuracy of inflation expectations depends on demographic characteristics of individuals. Hence, even in times of high news coverage, some households might still deviate from the best available forecast, if they have difficulties to understand media reports and thus do not incorporate the latest available information. Second, reading propensities differ considerably across households (Schoenbach et al., 1999), a feature that Carroll (2003) himself has already tried to take into account. Third, the various news media cover inflation in a different way. Anecdotal evidence indicates that the yellow press as well as TV channels with a focus on entertainment devote less space to inflation in ordinary times, but increase their coverage significantly and in an often exaggerated way if something unusual happens. By contrast, state-funded TV channels seem to report on a more regular and accurate basis on inflation. It is the aim of this chapter to relax these three assumptions and to test whether allowing for socioeconomic news coverage can help explain the demographic differences in inflation expectations often found in the literature. Besides the news media and professional forecasters' expectations, households rely on further sources of information to build their expectations. According to the "availability hypothesis" (Tversky and Kahneman, 1973), households tend to have a better memory for prices they pay more frequently. Hence, if people are asked for their expectations about future price developments, it is not clear whether they refer to CPI inflation reported in the media or to prices they encounter in their everyday life. We take this into account by computing household-specific inflation rates that closely match typical spending patterns of the demographic groups in our data set. Furthermore, at the moment people state their expectations, they might not remember exactly the entire price changes of their household-specific goods basket, but only prices that have risen a lot. We account for this selective perception by including households' nowcast of the current inflation rate, the so-called inflation perceptions. Overall, we thus simultaneously explore three sources of expectation differentials: media effects, inflation rates, and inflation perceptions. For reasons of data availability, we use monthly survey data for German households' inflation expectations distinguishing between age, income and occupation groups together with 10 different news media sources over the time span January 1999 - March 2010.

Accounting for the determinants of the heterogeneity of inflation expectations is important for a number of reasons. As it has been nicely summarized by Gnan et al. (2011), if expectations differ among agents, this will affect economic policy through various channels. First, heterogeneity of expectations has found to be important to explain stylized facts such as the hump-shaped response of output and inflation to monetary policy shocks (Mankiw and Reis, 2006). Second, anchoring agents' inflation expectations might call for different communication strategies of central banks if households persistently form expectations in different ways (Sims, 2009). Third, as it is argued by Bomberger (1996), rising disagreement on the future path of prices might be a sign of uncertainty with possible effects on economic risk-taking. Fourth, if expectations affect current inflation as it is the case in the forward-looking New Keynesian Phillips Curve, does this relationship change if there is considerable heterogeneity in expectations? Finally, if some demographic groups tend to have forecast errors that are persistently above average, this might call for economic policies mitigating the resulting effects on the distribution of wealth and income (Doepke and Schneider, 2006).

Our analysis makes the following contributions. First, in line with previous findings in the literature, we observe that inflation expectations depend on demographics also in Germany, albeit differences are not that large. Inflation expectations are higher for households with low income, for young households and for the unemployed. Moreover, the same types of households show larger deviations from the best available forecast, which we proxy with professional forecasters' expectations. Besides of deviating more in absolute terms, these household-groups also show larger fluctuations with regard to experts' expectations.
Second, we try to explain these demographic differences with household-specific inflation rates, inflation perceptions and news coverage. We find that the higher expectation gaps of young and old households as well as the rising deviation with lower income levels can be explained by higher inflation rates of these groups, while no such effect can be observed for occupation groups. Across all household groups, inflation perceptions do not play a role in determining inflation expectations. With regard to the news media, we observe considerable heterogeneity in news consumption of different newspapers and TV news shows for income, age and occupation groups. It thus seems that media coverage offers some explanation on why households with a different socioeconomic background disagree on the future path of prices. Furthermore, we find that constructing an index of news reports by aggregating all available newspaper and TV reports can be misleading. Coverage of inflation in *Tagesschau*, Germany's most influential TV evening news show, is found to increase the gap between households and professional forecasters, while a rising number of articles published in *BILD*, Germany's most prominent tabloid, brings households closer to the best available forecast. Finally, it is important to distinguish between the effects of a rise in the number of news reports (volume channel) and a change in the journalists' judgment of inflation (tone channel). Whereas households' expectation gaps increase if *BILD* presents inflation in a negative way thereby possibly inducing a media bias, more negative coverage in *Tagesschau* narrows the gap between households and professional forecasters.

We start this chapter with a short description of Carroll (2003)'s epidemiology model and its application to the demographic dependence of households' inflation expectations. We then describe the data set and our estimation strategy, before presenting our results and discussing directions for further research. A detailed literature summary of the different sources of households' disagreement on inflation expectations that have been proposed in the literature is provided in the Appendix.

3.2 The Dependence of Inflation Expectations on Socioeconomic Characteristics

It is a robust finding in the empirical literature that inflation expectations depend on households' socioeconomic background. Among other characteristics, high-income households and better educated individuals tend to report lower expectations, the unemployed generally state higher expectations, and young and old households expect inflation to be higher compared to middle age households. Expressed formally, for different households groups j, we observe:

$$\pi_{j,t+1}^{exp,hh} = f\left(\underset{(-)}{income}, \underset{(-)}{education}, \underset{(+)}{unemployed}, \underset{(+/-)}{age}\right) \tag{3.1}$$

This pattern is found in various studies for different countries, different time periods and for both qualitative and quantitative surveys (Bryan and Venkatu, 2001b,a, Blanchflower and MacCoille, 2009, Bruine de Bruin et al., 2010). We offer a detailed survey of the evidence in the Appendix (B.1).

Besides expecting higher inflation in absolute terms, the same groups of households also make larger forecast errors:

$$e_{j,t+1} = f\left(\underset{(-)}{income}, \underset{(-)}{education}, \underset{(+)}{unemployed}, \underset{(+/-)}{age}\right), \text{ where } e_{j,t+1} = \pi_{j,t+1}^{exp,hh} - \pi_{t+1} \tag{3.2}$$

Evidence has been provided for example by Souleles (2004) for the US, Blanchflower and MacCoille (2009) for the UK, and Leung (2009) for New Zealand. Since no such study has been conducted for Germany, it is the first goal of this chapter to establish comparable evidence using German data.

A number of different explanations have been proposed in order to explain this pattern, such as different degrees of financial literacy across households (Burke and Manz, 2011, Bruine de Bruin et al., 2010), household-specific inflation rates (Jonung, 1981, Bryan and Venkatu, 2001a) or household-specific inflation perceptions (Blanchflower and MacCoille, 2009). However, a systematic summary of the literature, which is provided in Appendix (B.1), reveals that most studies only test one explanation at a time, without assessing the possible impact of alternative reasons of why households' inflation expectations systematically depend on their socioeconomic background. For this reason, we try to test simultaneously as many of the proposed explanations as possible, in order to assess their relative importance. Furthermore, we add to the literature by suggesting that household-specific news consumption is responsible for the socioeconomic differences in inflation expectations.

The role of news reports in shaping households' belief about future inflation has originally been emphasized by Carroll (2003). According to his epidemiology model, only a fraction λ of households forms expectations in line with the best available forecast $E_t[\pi_{t+1}]$, whereas the remaining part $1-\lambda$ sticks to their beliefs built in the previous period. Thus, the mean expectations computed across all households is given as a weighted average:

$$\pi_{t,t+1}^{exp,hh} = \lambda E_t[\pi_{t+1}] + (1-\lambda)\pi_{t-1,t}^{exp,hh} \tag{3.3}$$

Next, Carroll (2003) assumes that households think that experts are better in forecasting inflation than themselves. Thus, one can use the average of the inflation expectations provided by professional forecasters, $\pi_{t,t+1}^{exp,prof}$, as a proxy for the best available forecast in the economy. And, since households get to know experts' expectations via reading newspapers or watching television, this suggests that news coverage is an important driver of households' inflation expectations.[1]. If the media report a lot about inflation, this increases the probability that households receive this information and subsequently update their expectations to expert forecasts that are often quoted in the news. Note that models of sticky information (Mankiw and Reis, 2002) and rational inattention (Sims, 2003) imply a similar role of the news media. According to these models, households do not form expectations rationally if the costs of gathering and processing information are too high. Instead, they receive the most recent inflation forecast from following the news media, whereas in times of large media coverage of inflation, households face lower search costs and are thus quicker to adjust to expert forecasts. Expressed formally, the epidemiology model allowing for an effect from news coverage is given as:

$$GAPSQ_t = \alpha_0 + \alpha_1 News_t, \tag{3.4}$$

where $GAPSQ_t\left(\pi_{j,t}^{exp,hh} - \pi_t^{exp,prof}\right)^2$ is the squared difference of households' expectations and the expectations of professional forecasters.[2] Following the epidemiology model or models of sticky information, one would expect a negative news effect, i.e. more newspaper articles or television reports should lower the gap between experts and households.

This model can be related to the question on demographic differences in inflation expectations by assuming that households have different reading propensities resulting in household-specific news effects:

$$GAPSQ_{j,t} = \alpha_{j,0} + \alpha_{j,1} News_t \tag{3.5}$$

In the working paper version of his paper, Carroll (2001) argues in favor of such heteroge-

[1]Supportive evidence for the role of news in explaining inflation expectations is provided by Carroll (2003), Dräger (2011) and Lamla and Lein (2010), whereas Pfajfar and Santoro (2013) do not find significant news effects.

[2]Using the absolute gap instead of the squared gap does not change the results qualitatively.

neous news effects. If, for example, low-income households have a lower reading propensity, a rise in news coverage of inflation would have a lower effect on this group compared to the remaining income groups. According to Schoenbach et al. (1999), in Germany, males, older households, better educated and households with higher income read newspapers more frequently compared to others. As a result, the expectation gap of low income households will be larger, since they are less likely to update to the best available forecast in the economy. We thus take the epidemiology model allowing for different news effects across households as the starting point for our analysis of demographic differences in inflation expectations. Note that arguing in terms of "expectation gaps" instead of "forecast errors" or "absolute values of inflation expectations" does not affect our general conclusions: As we will show below, those household groups that express the highest inflation expectations are generally the same that make the largest forecast errors and also show the largest expectation gaps. Moreover, we will take the perspective of households throughout the analysis. While it has been shown that experts occasionally also adjust to households, the expectation gap of households and experts is mainly driven by households adjusting to experts (Menz, 2013). Keeping this in mind, we state a first testable hypothesis:

Hypothesis 1 *The extent to which households adjust to experts when forecasting inflation depends negatively on the amount of news coverage on inflation. The larger expectation gaps of some household groups result from lower news effects due to different reading propensities.*

In what follows, we relax and test a number of assumptions of the epidemiology model expressed in terms of group-specific expectation gaps. So far, the baseline version in equation (3.5) assumes that the effect of news coverage is the same for all different newspapers and television shows. For the purpose of explaining socioeconomic news consumption, this assumption is too restrictive, given that households of different age, income, or occupation prefer different news sources. Thus, distinguishing between various print and TV media, our second hypothesis is given as

Hypothesis 2 *Households react differently to different news sources, depending on their socioeconomic characteristics.*

Next, it is important not only to account for the amount of news coverage, but also for its tone. Gentzkow and Shapiro (2010), among others, show that the media "slant" the news, i.e. certain news are discussed more prominently and in a different light than others, depending *inter alia* on readers' initial beliefs. In the context of inflation expectations, Lamla and Lein (2010) and Dräger (2011) report evidence that households react strongly to news on inflation if articles are written in a negative tone, i.e. if journalists argue that current or future inflation is a serious problem for the economy. Again, we expect households to react differently to media slant, depending on their socioeconomic background. For example, better educated households could be less receptive for overly negative newspaper articles, whereas younger households with less personal experience might react more strongly to negative news reports. Thus, we state our third hypothesis as

Hypothesis 3 *Households do not only react to the amount of news coverage but also to its tone. Depending on the demographic background, negative news on inflation are perceived differently than positive news.*

Finally, the epidemiology model excludes some factors that possibly affect households' inflation expectations. Since we ultimately want to explain the demographic differences in expectation gaps, we have to account for at least three more variables that have been proposed in the literature as determinants of socioeconomic disagreement in inflation expectations.
First, as it is argued by Akerlof et al. (1996, 2000), the heterogeneity of households' inflation expectations depends negatively on the level of the overall inflation rate. Mankiw et al. (2003) for the US and Gnan et al. (2011) for Euro Area countries present supportive evidence for the near-rationality hypothesis of Akerlof. Furthermore, the epidemiology model has been criticized for excluding adaptive expectation formation. Instead of sticking to their own past expectations, non-updating households could simply adjust to the most recent inflation rate (Luoma and Luoto, 2009) . However, we expect that the inflation rate does not have the same effect on all households. If high-income households are more forward-looking than low-income households, a positive increase of inflation should have a lower impact on households at the top of the income distribution. Therefore, we test a fourth hypothesis:

Hypothesis 4 *Households' do not only adjust to the best available forecast or stick to their own past expectations, but they also react positively on the actual inflation rate. The effect varies with households' socioeconomic background: The larger expectation gaps of some households might be due to a larger degree of adaptive expectation formation.*

However, it is not obvious that households have the official inflation rate in mind when forming expectations about future prices. Instead, they might refer to price changes of a consumption bundle which is more closely linked to their own spending behavior. And as it has been argued by various authors beginning at least with Michael (1979), households with low income, low education, and the elderly face above average inflation rates. Thus, our next hypothesis is given as

Hypothesis 5 *Households mainly react to their group-specific inflation rates instead of overall inflation. Since households with different demographic characteristics face systematically different inflation rates, the effect of price changes on expectation gaps will vary as well.*

Finally, research in psychology shows that households have difficulties in recalling prices they have paid, even of goods they have bought only recently (Ranyard et al., 2008). If this is true, households would not base their expectations on actual group specific inflation rates, but instead use an own estimate of past prices, the so-called perceived inflation rate. Since the ability to remember past prices varies with the age of households, or since low income households will face a greater need to remember prices, we would also expect group-specific effects from perceived inflation. Hence, we test a final hypothesis:

Hypothesis 6 *Instead of overall inflation or group-specific inflation, households use an own estimate of past price changes, the perceived inflation rate, to form expectations. Since the ability and necessity to remember past prices can be related to demographics, we expect that the impact of perceived inflation varies across households.*

Summing up, we test an extended version of the baseline epidemiology model:

$$GAPSQ_{j,t} = f_i\left(News_{i,t}, \pi_t, \pi_{j,t}, perc_{j,t}\right) \tag{3.6}$$

Here, $News$ captures either the total amount of media coverage about inflation or its tone, for different media sources i, π_t is the actual inflation rate, $\pi_{j,t}$ gives the inflation rate corresponding to household j, and $perc_{j,t}$ denotes household-specific inflation perceptions.

3.3 Data

This section describes the data on household-specific inflation expectations and perceptions, group-specific inflation rates, professional forecasters' expectations and news coverage. Overall, our sample covers the period 1999M1-2010M3. All data sources can be found in Table (B.2) in the Appendix.

The household-specific inflation expectations and perceptions are taken from the Consumer Survey conducted by the European Commission (EC), whereas household-specific inflation rates are derived using data from Eurostat. Unfortunately, the demographic categories of the EC survey do not match entirely with the categories used to compute household-specific inflation rates. In Table (3.1), we show the categories that are possible to merge, namely *age*, *income*, and *occupation*. Even if the classifications are slightly different, we think that this should not affect the results too much. It is not possible to include education, since no data is available for household-specific inflation rates.

Table 3.1: Match of Demographic Groups

HH-Expectations (EC)	HH-Inflation (Eurostat)	Variable Label
total	total inflation	macro
Age Groups		
16-29	0-30	ylt30
30-49	30-44	y3044
50-64	45-59	y4559
65+	60+	yge60
Income Groups		
1st quartile	1st income quintile	inc1
2nd quartile	2nd income quintile	inc2
3rd quartile	4th income quintile	inc3
4th quartile	5th income quintile	inc4
Occupation Groups		
skilled manual workers	manual workers in industry and services	wman
self employed and professional	self-employed	wfree
unemployed	unemployed	wune

3.3.1 Household-specific Inflation Expectations

The Consumer Survey of the European Commission consists of qualitative data. Each month, a random sample of households in different European countries is asked the following question: "By comparison with the past 12 months, how do you expect that consumer prices will develop in the next 12 months?". Respondents can choose between six answer categories: "rise a lot", "rise moderately", "rise slightly", "stay about the same", "fall", "don't know". The EC publishes the resulting response fractions, both on the aggregate household level and for different demographic groups. Unfortunately, the underlying micro data is not available.

For the purpose of explaining the expectation gaps of different households, we need to quantify the qualitative survey responses using the probability method proposed by Carlson and Parkin (1975). The use of this method has been sometimes criticized in the literature, as for example recently by Breitung and Schmeling (2013). However, since we only have qualitative data at hand, we have no choice but to accept the disadvantages of the probability method. Since a detailed discussion of the quantification procedure is beyond the scope of this chapter, we propose a brief description in the Appendix (B.2). At the moment, it suffices to stress that the probability method has to assume a probability distribution and a scaling parameter. For the former, we use the normal distribution, whereas for the latter, we could either use the aggregate inflation rate, as it is usually done in the literature, or household-specific inflation rates.[3] Using the official inflation rate assumes that survey participants refer to the overall price development at the time they answer the questionnaire. However, if individuals base their inflation expectations on past price changes of those goods categories they are more familiar with, it might be more appropriate to employ household-specific inflation rates in the quantification process. Since the EC survey only refers to "consumer prices" instead of "prices in general" or "inflation rate", both versions are possible. Hence, the choice of the appropriate inflation rate used to scale households' qualitative expectations is an empirical question. We thus calculate the recursive HP-filter over 20 months prior to each survey date, using both aggregate inflation and household-specific inflation.[4]

In Table (3.2), we compare the mean, the standard deviation, and the root mean squared error of households' quantified inflation expectation. The results suggest that households tend to base their expectations on group-specific inflation: for all households, the RMSE's are lower if we quantify the qualitative answers with household-specific inflation (columns (3) and (4)) compared to aggregate inflation (columns (7) and (8)). Furthermore, households are better in predicting changes in the aggregate price level rather than changes of their group-specific inflation rate. Thus, it seems that households participating in the survey refer to overall inflation but evaluate the expected changes against their group-specific inflation rate. Hence, in the remaining part of the chapter, we use group-level inflation rates to quantify inflation expectations.[5]

[3]The construction of household-specific inflation rates is described in the next section.

[4]The results do not change much if we use different lags to calculate the HP-filter.

[5]Results are qualitatively similar if we employ overall inflation.

Table 3.2: Results: Forecast Errors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	hh-inflation				aggregate inflation				GAPSQ			
	mean	sd	$RMSE$ $\pi_{j,t}$	$RMSE$ π_t	mean	sd	$RMSE$ $\pi_{j,t}$	$RMSE$ π_t	$mean$ $\pi_{j,t}$	sd $\pi_{j,t}$	$mean$ π_t	sd π_t
prof	1.497	0.471	0.944	.	.	.	.	.	.	.	.	.
all	1.118	0.442	1.122	1.122	1.118	0.442	1.122	1.122	0.309	0.303	0.309	0.303
ylt30	1.144	0.449	1.171	1.094	1.074	0.418	1.198	1.125	0.273	0.258	0.345	0.316
y3044	1.203	0.478	1.187	1.089	1.106	0.437	1.218	1.124	0.231	0.233	0.310	0.298
y4559	1.253	0.500	1.166	1.066	1.144	0.458	1.208	1.116	0.213	0.232	0.293	0.299
yge60	1.283	0.509	1.177	1.051	1.152	0.464	1.238	1.129	0.213	0.246	0.301	0.312
inc1	1.264	0.548	1.255	1.104	1.168	0.471	1.270	1.121	0.272	0.329	0.291	0.304
inc2	1.226	0.514	1.192	1.100	1.148	0.467	1.216	1.128	0.253	0.283	0.292	0.299
inc3	1.237	0.482	1.169	1.075	1.132	0.445	1.213	1.126	0.219	0.240	0.301	0.306
inc4	1.240	0.471	1.151	1.035	1.102	0.435	1.214	1.116	0.181	0.177	0.310	0.302
wman	1.221	0.460	1.152	1.064	1.123	0.426	1.190	1.108	0.218	0.231	0.302	0.298
wfree	1.209	0.481	1.164	1.073	1.100	0.441	1.207	1.123	0.224	0.222	0.316	0.305
wune	1.296	0.540	1.267	1.101	1.179	0.465	1.288	1.125	0.227	0.268	0.270	0.276

Note: Sample: 1999M1-2010M3. RMSE is the root mean squared error of inflation expectations and actual inflation 12 months ahead, π_t denotes aggregate inflation and $\pi_{j,t}$ is the representative inflation rate of household-group j. $GAPSQ$ is the squared difference between households' and professional forecasters' inflation expectations.

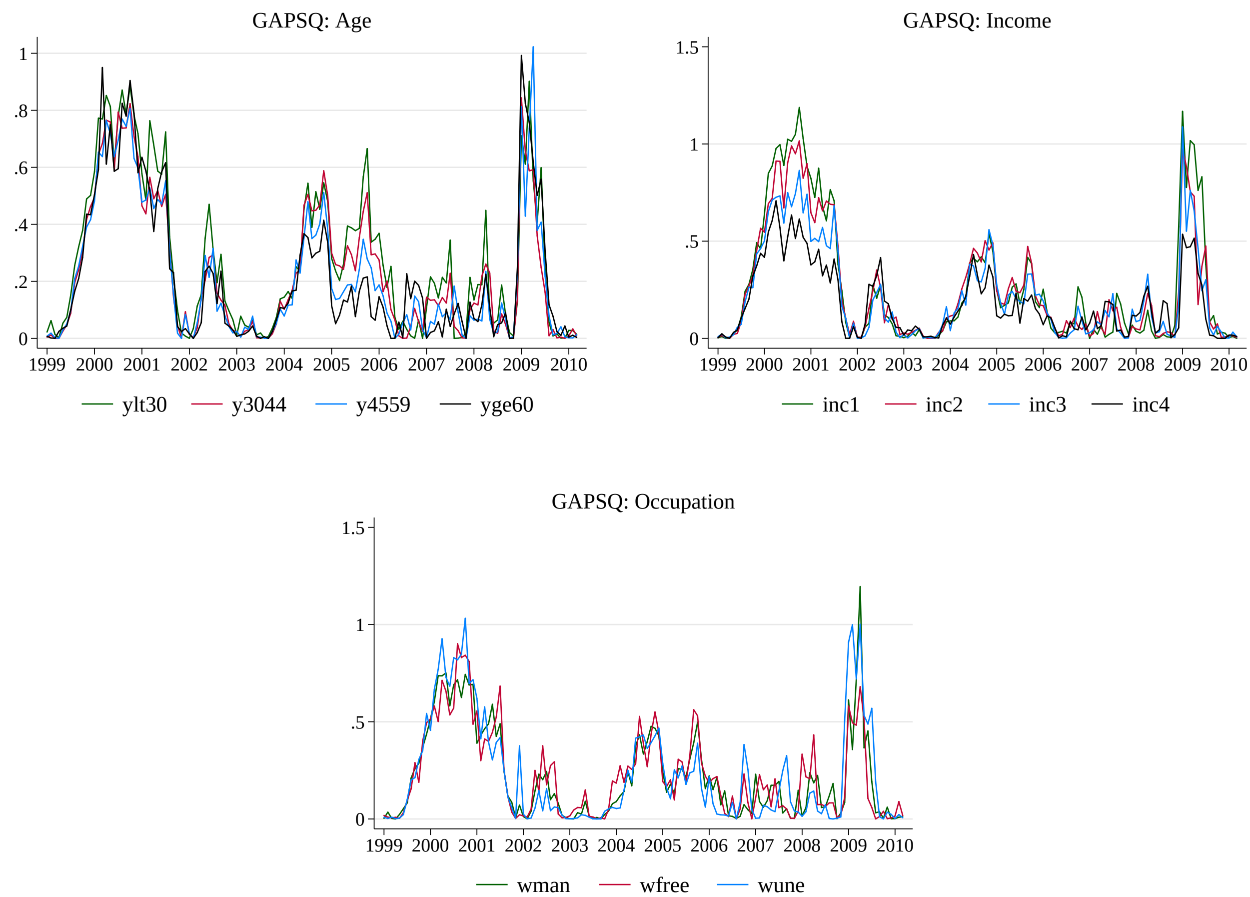
Next, we check whether the general findings with regard to the demographic expectation differentials also hold in Germany.[6] Overall, the differences of quantified inflation expectations are relatively minor across demographic groups, which might be due to fact that we can only use group level data instead of micro data.[7] Still, the summary statistics in Table (3.2) reveal pattern in households' inflation expectations that are similar to those reported in the literature. The older the households, the higher their expectations. Unemployed people have higher expectations than manual workers and self-employed. With regard to the income differentials, the results are less clear-cut. In accordance with the literature, the poorest households have the highest inflation expectations. However, moving from the second income quartile to the fourth quartile, we observe rising inflation expectations, but, turning to the RMSE, households' forecast error constantly falls with rising income. Whereas the unemployed are considerably worse in forecasting their group-specific inflation compared to manual workers and self-employed, no clear pattern emerges for age groups.

[6]We plot households' quantified inflation expectations in Figure (B.2) in the Appendix.

[7]Moreover, Gnan et al. (2011) report marked differences between European countries: Whereas the within-group disagreement does not differ much between household-groups in France, Germany, and Slovakia, the remaining Euro Area countries exhibit much larger deviations.

Comparing household expectations with expert expectations, the demographic pattern becomes more explicit. First, we get lower expectation gaps if we quantify households' expectations using group specific inflation (column (9)) compared to aggregate inflation (column (11)). Second, the expectation gaps are larger if households are unemployed, belong to low-income groups, or to the youngest age group. Plotting the expectation gaps for each household groups in Figure (3.1) also shows some variation over time, with the largest gaps in 2000/2001 and 2009.

Figure 3.1: The Expectation Gaps of Households



3.3.2 Household-specific Inflation Rates and Perceptions

The household-specific inflation rates are taken from Colavecchio et al. (2011). The authors compute fictitious group-specific inflation rates by combining household expenditure patterns from the Household Budget Surveys (HBS) of the European Commission with the harmonized inflation rates for different goods categories according to the "Classification of Individual Consumption by Purpose (COICOP)". We refer to their paper for a detailed description.

As we have mentioned above, we use these household-specific inflation rates for the quantification of inflation expectations on the group level. Moreover, we can test whether households react to changes in overall inflation or to price changes that are closer related to their group-specific spending patterns. However, when forming their expectations, households could also use their estimates of current inflation as a benchmark. This perceived inflation rate can be computed from the EU Consumer Survey as well. In addition to asking households to state their beliefs on future prices, the survey includes a question on perceived inflation: "How do you think that consumer prices have developed over the last 12 months?", offering respondents the same answer categories as for the expectation series. Again, we apply the probability method as described in Appendix (B.2) to quantify the qualitative perception series.

3.3.3 Media Data

The media data is compiled by the media research institute *Media Tenor*[8]. Newspaper articles and television reports are searched for the keywords "inflation", "deflation", "price increase", "price cut", "price stability" and "oil price", followed by a human-based content analysis of the news reports that have been picked up. This detailed coding allows us, for example, to distinguish reports with a main focus on Germany from reports that mention inflation in other countries. In total, ten different media sources are included, ranging from one national daily newspaper (*BILD*), over two national weekly magazines (*Der Spiegel*, *Focus*) to seven evening news shows on TV (*Tagesschau*, *Heute*, *Heute Journal*, *Tagesthemen*, *SAT1 18:30*, *RTL Aktuell*, and *Pro7 Nachrichten*).

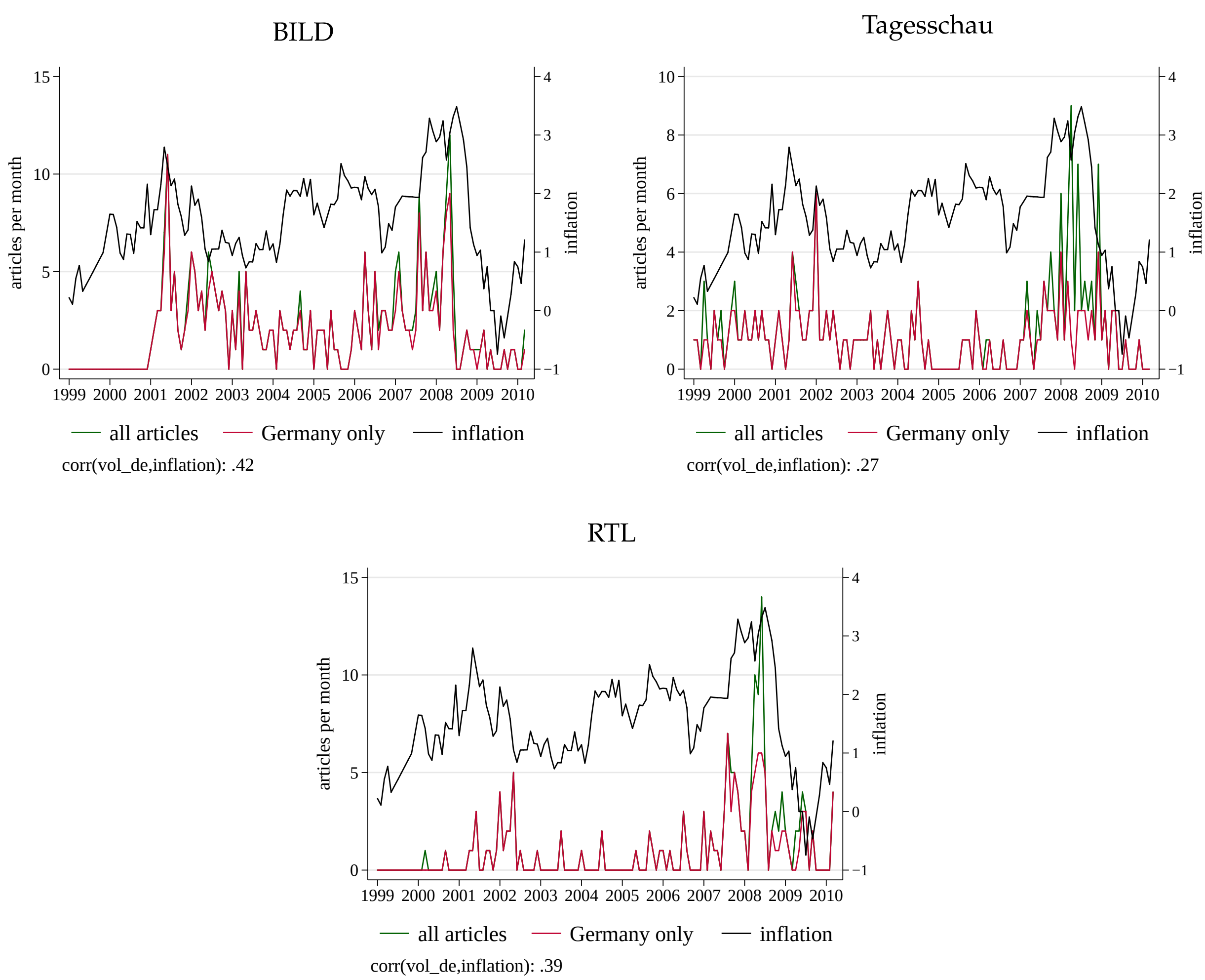
In what follows, we mainly focus on the daily newspaper *BILD*, the most important public news broadcast *Tagesschau* and the most influential private channel *RTL*, in order to keep the exposition tractable. The monthly sum of newspaper articles and TV reports of these news sources are shown in Figure (3.2), together with the annual inflation rate and distinguished between all articles and news reports that deal only with Germany.[9] Overall, the media follow a similar trend: news coverage tends to peak in 2002M1 and 2008M1 across

[8] http://www.mediatenor.com/

[9] The graphs for the remaining news media can be found in Figure (B.4) in the Appendix.

all media. In addition, most of the articles and TV reports deal with inflation in Germany, the only exception being the period of the financial crisis. Still, there are differences between media sources. The daily tabloid *BILD* covers inflation in nearly every month, whereas the public evening news show *Tagesschau* covers inflation on a more regular basis than the private TV channel *RTL*. Accordingly, the correlation of news coverage with annual inflation varies between single media sources. Whereas news coverage in *Tagesschau* has a correlation coefficient of .27, *BILD* and *RTL* react slightly stronger to inflation.

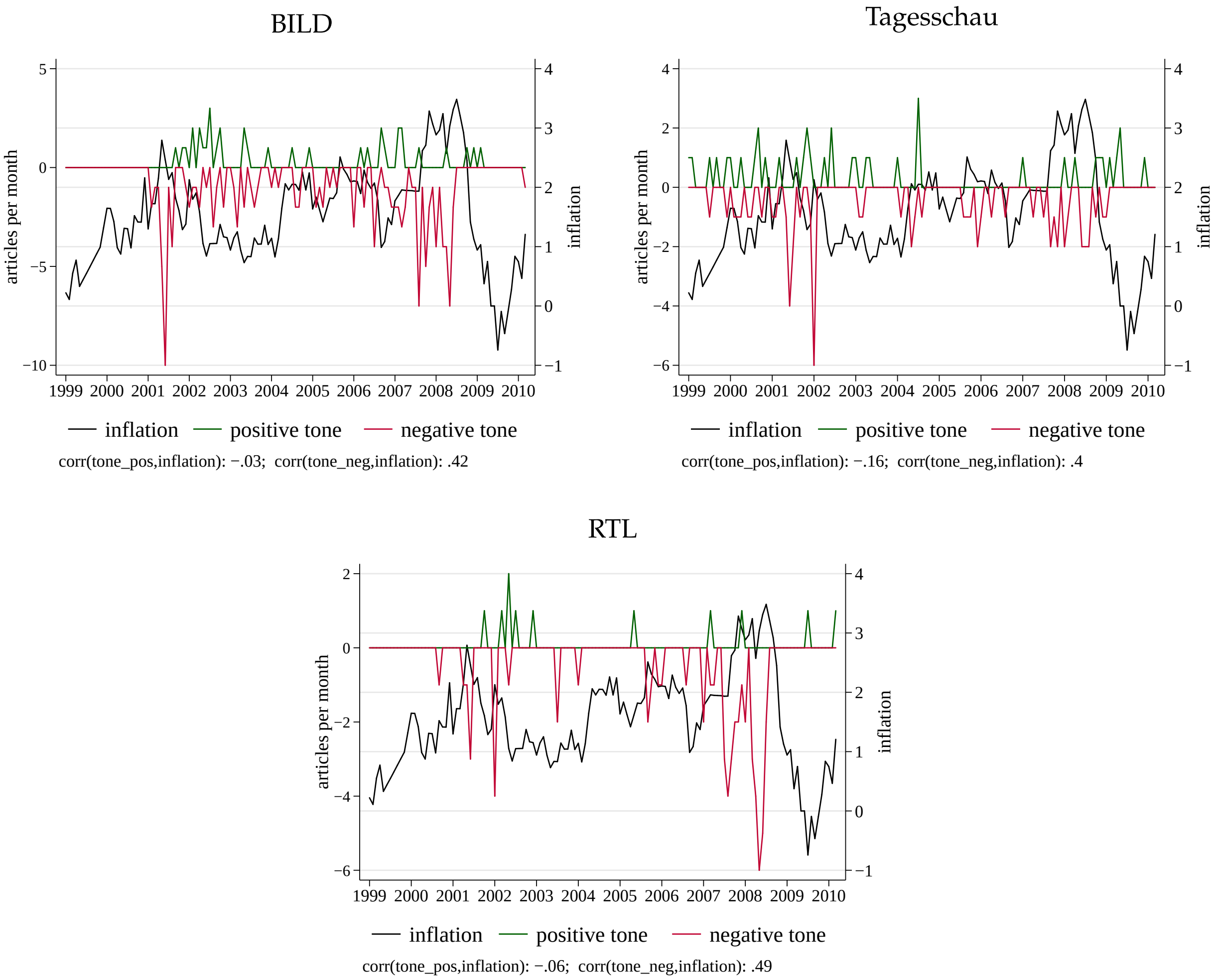
Figure 3.2: Media Coverage I: Number of News Reports About Inflation per Month



Besides the total amount of news coverage, our media data set also allows us to include a tone variable, which can be captured via the *valuation* and the *context* of an article. The valuation of an article is more narrowly defined. As an example, a statement such as "hyperinflation destroys the savings of citizens" would be coded as negative valuation. In addition, the context of an article takes into account a broader judgment. For example, the sentence "inflation has been consistently higher than in other OECD countries" receives a negative context in the coding. These classifications can depend on the interpretation of the individual coder, however, *Media Tenor* reports to have a high intercoder reliability.

In the following, we only plot the number of positive and negative articles using the context variable since the single news media only show very low numbers of news reports with a narrowly defined judgment (valuation). As it is shown in Figure (3.3), we generally observe a rising number of negative reports and a drop in the number of positive articles if inflation rises.[10] With regard to the heterogeneity of news coverage, on average, *Tagesschau* has the most balanced coverage about inflation topics in terms of valuation as well as context. The tabloid *BILD*, by contrast, mostly covers inflation with a negative tone.

Figure 3.3: Media Coverage II: Number of Negative and Positive News About Inflation per Month



[10]This picture also holds for the remaining news media, see Figure (B.5).

3.4 Estimation Strategy

As regards the estimation, we start with specifying a baseline version of the epidemiology model in equation (3.6), i.e. for different household groups, we explain the squared gap between households' inflation expectations and experts' forecast, with overall and household-specific inflation rates, inflation perceptions and news media variables.

In a first set of equations, we test the **Hypothesis 1**, i.e. we evaluate whether the impact of the overall number of newspaper articles $News_t^{print}$ and the number of TV reports on inflation $News_t^{tv}$ differs across household groups. Furthermore, we simultaneously test **Hypothesis 4 - 6** by including overall and household-specific inflation as well as household-specific inflation perceptions. Thus, for each age group, income group and occupation group j, we estimate

$$\begin{aligned} GAPSQ_{j,t} = \alpha_{j,1} + \alpha_{j,2}\pi_{t-1} + \alpha_{j,3}News^{print} + \alpha_{j,4}News^{tv} + \alpha_{j,5}\left(\pi_{j,t} - \pi_t\right) \\ + \alpha_{j,6}\left(perc_{j,t} - perc_t\right) + \varepsilon_{j,t} \end{aligned} \tag{3.7}$$

Three points have to be mentioned. First, we follow Anderson et al. (2010) and include the overall inflation rate π_t with its first lag to take into account that the official price statistic is only released with a delay of one month. Second, we do not use the raw series of household-specific inflation rates and perceptions, but calculate the deviations of group-specific inflation rates from aggregate inflation rate, $\pi_{j,t} - \pi_t$, as well as the difference between group-specific perceptions and aggregate perceptions, $\pi_{j,t}^{perc} - \pi_t^{perc}$.[11] By using price differentials, we belief to be closer to the underlying information processing of households: these might either increase their inflation expectations in response to rising aggregate inflation, or if their group-specific inflation deviates considerably from overall inflation. We include the contemporaneous value of inflation differentials assuming that households immediately realize price changes of their group-specific consumption bundle. Third, the news variables are computed as follows. For each month, we sum all articles that mention inflation in each of the 10 different news sources. Then, in order to account for the fact that the size of newspapers has been changing over time, we divide the monthly sums by their maximum value over the entire sample. Finally, for computing the overall number of newspaper articles $News_t^{print}$ and TV reports $News_t^{tv}$ we weight the single newspapers by their print run and the TV reports by the number of daily viewers.[12,13]

Next, we disaggregate the news variables, and include the volume of inflation reports in

[11] The resulting series are shown in Figure (B.6) in the Appendix.

[12] In Figure (B.3) in the Appendix, we plot the average number of readers per newspaper issue and the average number of daily viewers of TV news shows.

[13] Correlation of the two news indexes only reaches .4, so there should be no multicollinearity problem. The same is true for the correlation between household-specific inflation rates and inflation perceptions.

BILD, *Tagesschau*, and *RTL* separately, thereby testing the **Hypothesis 2** stating that households of different socioeconomic background choose different news sources to get information about inflation. We choose to only include the three most important news sources in order to keep the estimation and interpretation tractable. The results remain the same if we use the entire media data set. Hence, equation (3.7) is modified such that

$$\begin{aligned} GAPSQ_{j,t} &= \alpha_{j,1} + \alpha_{j,2}\pi_{t-1} + \alpha_{j,3}News_t^{Bild} + \alpha_{j,4}News_t^{Tag} + \alpha_{j,5}News_t^{RTL} \\ &+ \alpha_{j,6}\left(\pi_{j,t} - \pi_t\right) + \alpha_{j,7}\left(perc_{j,t} - perc_t\right) + \varepsilon_{j,t} \end{aligned} \tag{3.8}$$

Note that since we do not have data on the relative amount of time households spend watching television or reading the newspapers, we cannot weight the single media indexes.
Next, we replace the volume of news media coverage with the tone of media reports thereby testing **Hypothesis 3**. We distinguish between the number of negative news $News^{neg}$ and positive news $News^{pos}$, and employ the two different codings used by *Media Tenor*, context con and valuation val. The news variables with a negative tone are highly correlated (.8), however, this hardly affects the results. The third equation is given as:

$$\begin{aligned} GAPSQ_{j,t} &= \alpha_{j,1} + \alpha_{j,2}\pi_{t-1} + \alpha_{j,3}News^{pos\_con} + \alpha_{j,4}News^{neg\_con} + \alpha_{j,5}News^{pos\_val} \\ &+ \alpha_{j,6}News^{neg\_val} + \alpha_{j,7}\left(\pi_{j,t} - \pi_t\right) + \alpha_{j,8}\left(perc_{j,t} - perc_t\right) + \varepsilon_{j,t} \end{aligned} \tag{3.9}$$

Finally, we also use the disaggregated tone variables, regressing the expectation gaps on the number of news reports with a positive tone in *BILD*, *Tagesschau*, and *RTL* on the one hand, and on the media reports with a negative judgment on the other hand. Since single news media only show very low numbers of news reports if we classify the journalists' judgment in a narrow sense, we only employ the broader definition included in *context* in the estimation. Our final equations are thus given by:

$$\begin{aligned} GAPSQ_{j,t} &= \alpha_{j,1} + \alpha_{j,2}\pi_{t-1} + \alpha_{j,3}News^{Bild\,con\,pos} + \alpha_{j,4}News^{Tag\,con\,pos} \\ &+ \alpha_{j,5}News^{RTL\,con\,pos} + \alpha_{j,6}\left(\pi_{j,t} - \pi_t\right) + \alpha_{j,7}\left(perc_{j,t} - perc_t\right) + \varepsilon_{j,t} \end{aligned} \tag{3.10}$$

$$\begin{aligned} GAPSQ_{j,t} &= \alpha_{j,1} + \alpha_{j,2}\pi_{t-1} + \alpha_{j,3}News^{Bild\,con\,neg} + \alpha_{j,4}News^{Tag\,con\,neg} \\ &+ \alpha_{j,5}News^{RTL\,con\,neg} + \alpha_{j,6}\left(\pi_{j,t} - \pi_t\right) + \alpha_{j,7}\left(perc_{j,t} - perc_t\right) + \varepsilon_{j,t} \end{aligned} \tag{3.11}$$

It is worth noting, at this point, that there are probably a number of feedback effects between the variables under investigation. Of particular importance, it might be fairly restrictive to treat media coverage as an exogenous variable for explaining households' expectations. Mullainathan and Shleifer (2005) and Gentzkow and Shapiro (2010) have argued that under certain conditions, newspapers slant their news coverage in the direction of the initial beliefs of their readers. Therefore, we take a systems approach and model news coverage in each of the estimated equations as an endogenous variable. More precisely, we relate media coverage to economic developments and agents' thoughts about the future:

$$NEWS_{i,t} = \beta_1 + \beta_2 NEWS_{i,t-1} + ... + \beta_6 NEWS_{i,t-5} + \beta_7 \pi_t + \beta_8 \pi_t^{exp,hh} + \beta_9 \pi_t^{exp,prof} + \varepsilon_t \quad (3.12)$$

Hence, we explain the news coverage of different media sources with aggregate inflation π_t, the mean inflation expectations of all households $\pi_t^{exp,hh}$, and the mean price projection of professional forecasters $\pi_t^{exp,hh}$. While it stands to reason that news media relate their coverage to actual inflation and to the best available forecasts, it might be less obvious why this should also be the case for households' expectations. However, Mullainathan and Shleifer (2005) and Gentzkow and Shapiro (2010) have illustrated that consumer preferences are an important driver of newspaper coverage.

We estimate the resulting system of equations via Three-Stage-Least-Squares (3SLS). Allowing for endogeneity of news coverage, we expect the error terms of the equations explaining the expectation differentials to be correlated with the news variables. Furthermore, this endogeneity is also a potential source of correlation of the error terms across the different equations of the system, albeit not the only one. If inflation expectations are affected in a similar way by common shocks such as monetary policy decisions, this will also violate the assumption of independent errors across equations. In the latter case, we could use seemingly unrelated regressions (SUR) to account for this problem, but SUR will not give us consistent estimates if some of the explanatory variables are endogenous. We thus present estimates using system 3SLS, also discussing the differences compared to an equation-by-equation SUR approach. For the implementation of 3SLS, all variables other than the endogenous variables of our system are taken as instruments. Using these instruments, in a first stage, the predicted variables of the dependent variables are estimated, which are then used in a second step to consistently estimate the error terms of the different equations in the system. Finally, the estimated covariance matrix is used together with the predicted values of the right-hand-side endogenous variables computed in the first stage, to estimate the structural equations (3.7) - (3.12) of the system. For the estimation of the news equations (3.12), we allow for up to six lags of the media variables in order to account for the persistence of news coverage, and choose those lag length which yields the best overall fit. Overall, the results do not depend on the exact number of lags. In what follows, for sake of brevity, we do not report the results of the media equations. These are available upon request.

3.5 Results

We now present the results of our empirical analysis. In the following section we describe in detail the results of the 3SLS-estimation, and discuss differences with equation-by-equation SUR regressions. Furthermore, we have also tested whether the reported differences in the estimated coefficients are significantly different across household groups. While we cannot reject the hypothesis of coefficient equality in some cases, we choose to report results of unconstrained regressions throughout. Generally, our conclusions do not change if we estimate restricted regressions. Second, one could question the way we quantify the qualitative survey responses on inflation expectations. We have shown in Table (3.2) that households' forecast errors and expectation gaps are considerably lower if we use household-specific inflation as the reference level which makes us confident that this is the appropriate quantification variable. Still, we also repeat our empirical analysis using aggregate inflation in the quantification process. Overall, the results are fairly similar for both specifications. [14]

3.5.1 The Volume of News Coverage

We start with explaining the expectation gaps with the weighted number of newspaper articles and television reports, the results are summarized in Table (3.3).
Beginning with the inflation rates, across all household groups, we observe stronger effects from household-specific price indexes compared to the overall inflation rate. Aggregate inflation raises the expectation gap of younger households, and of manual workers and the self-employed. By contrast, the coefficients of household-specific inflation are generally larger, and also help explain part of the observed demographic heterogeneity in expectations. Compared to middle-age households, younger and older survey participants deviate more from the best available forecast in response to an increase in their corresponding inflation rate. Moreover, we observe slightly larger coefficients the poorer the households, which helps explain the larger expectation gap of low-income households. However, group-specific inflation cannot explain the larger expectation gap of the unemployed. With regard to inflation perceptions, we do not find any impact for the different household groups. These findings support the hypothesis that households focus more on price changes of goods that they encounter in everyday life than on headline inflation. In addition, the memory of consumption decisions is more important than the perception of a general price trend.

With respect to the news media, we generally observe that a rising number of articles or

[14] Detailed results of restricted 3SLS and SUR regressions and of models using aggregate inflation to quantify households' expectations are not shown but are available upon request.

Table 3.3: Results: Aggregate Volume - Endogenous News Coverage

	ylt30	y3044	y4559	yge60	inc1	inc2	inc3	inc4	wman	wfree	wune
π_{t-1}	0.12***	0.09***	0.02	-0.01	-0.01	0.07*	0.04	0.02	0.07*	0.10***	0.03
	(0.04)	(0.03)	(0.03)	(0.03)	(0.05)	(0.04)	(0.03)	(0.02)	(0.04)	(0.04)	(0.04)
$News_t^{pr\ index}$	-0.65***	-0.64***	-0.36*	-0.60***	-1.94***	-1.60***	-0.93***	-0.49***	-0.65***	-0.34	-1.40***
	(0.22)	(0.20)	(0.19)	(0.19)	(0.29)	(0.26)	(0.22)	(0.16)	(0.24)	(0.23)	(0.26)
$News_t^{tv\ index}$	-0.61**	-0.45**	-0.25	0.17	0.47*	0.08	0.04	0.09	-0.42*	-0.65***	-0.05
	(0.24)	(0.21)	(0.21)	(0.21)	(0.28)	(0.25)	(0.20)	(0.15)	(0.22)	(0.21)	(0.26)
$\pi_{j,t} - \pi_t$	0.13**	0.09	0.16**	0.21***	0.24***	0.21***	0.19***	0.18***	0.29***	0.31***	0.19***
	(0.07)	(0.06)	(0.07)	(0.05)	(0.06)	(0.07)	(0.06)	(0.05)	(0.10)	(0.10)	(0.06)
$perc_{j,t} - perc_t$	-0.01	0.05	0.02	-0.10	-0.03	-0.11*	0.03	-0.03	-0.04	-0.07	-0.01
	(0.07)	(0.09)	(0.07)	(0.07)	(0.07)	(0.06)	(0.06)	(0.05)	(0.06)	(0.05)	(0.05)
cons	0.34***	0.30***	0.30***	0.31***	0.57***	0.45***	0.33***	0.21***	0.31***	0.25***	0.46***
	(0.05)	(0.04)	(0.04)	(0.04)	(0.07)	(0.06)	(0.04)	(0.03)	(0.05)	(0.04)	(0.06)
R^2	0.233	0.252	0.265	0.368	0.233	0.209	0.275	0.327	0.279	0.260	0.285
N	130				130				130		

Note: Unconstrained 3SLS regressions using equations (3.5) and (3.10). Equation (3.10) is estimated using 5 lags of the dependent variables. *<0.1, **<0.05, *** p<0.01. Numbers in brackets denote standard errors. Sample 1999M1-2010M3. R^2 is calculated as correlation coefficient from actual values and predicted values from 2nd stage regression.

television stories *lowers* the gap between households' and professional forecasters' expectations. This is an important result, since this is the first time that the negative news effect originally put forward by Carroll (2003) has been confirmed in the literature.[15] Furthermore, we observe that the strength of the news effect differs both across households and across print media and television. In general, newspaper coverage is found to have a larger effect than television reports. Across household groups, however, aggregate print media coverage does not help explain the heterogeneity of households' expectation gaps. While we observe significantly larger coefficients for low income households, since the effect is negative, we would conclude that more newspaper articles lower the expectation gap of the poor more strongly as it is the case for rich households. The same result holds true for the unemployed. By contrast, aggregate television news do give rise to larger expectation gaps of poor, unemployed, and older households. While we do not find an effect from TV news that is significantly different from zero for households older than 44 and for the unemployed, more television reports significantly *increase* the expectation gap of households in the lowest income category without affecting the remaining quartiles.

Finally, we compare the 3SLS regressions with SUR estimates, the detailed results are found in Table (B.3) in the Appendix. While the general picture remains unchanged, the SUR results are different in two respects. First, and as a general feature of all regressions applying SUR to the set of equations (3.7) - (3.11), the coefficients of the news variables are much lower. Second, we do not find an impact from Television news and slightly less evidence of heterogeneity in the effects of newspaper articles.

[15]By contrast, Pfajfar and Santoro (2009, 2013) either find no news effect at all or a positive sign.

Next, we disaggregate the news indexes but use only the number of media reports in the three most important news sources *BILD*, *Tagesschau*, and *RTL*.[16] Compared to the previous estimates, the results shown in Table (3.4) confirm our conclusions with regard to the impact of aggregate and group-specific inflation, as well as inflation perceptions. Overall, group-specific inflation is more important than headline inflation, the effects of household-specific inflation are heterogeneous and help to some degree explain the expectation gap of the poor, the young and the old, and perceptions are generally not significant.

Disaggregating the news media, however, yields some interesting results. First, we find opposite media effects from *Tagesschau* on the one hand, and *BILD* and *RTL* on the other hand. An increase in news coverage in the latter lowers the gap between households and professional forecasters, as we would expect: following the idea of Carroll (2003), more news reports should increase the probability that households read about the best available forecast and subsequently update their beliefs on future prices. However, more news coverage in *Tagesschau* widens the expectation gap. This seems puzzling since the *Tagesschau* is associated with reputable quality journalism, while *BILD* and *RTL* are Germany's leading tabloid and private channel often marked by sensation reporting. We think that part of this surprising result stems from the fact that public TV channels such *Tagesschau*, due to its educational mandate, reports about inflation on a rather regular and neutral basis without overemphasizing unusual price changes. We further investigate this result in the next section.

Table 3.4: Results: Disaggregate Volume - Endogenous News Coverage

	ylt30	y3044	y4559	yge60	inc1	inc2	inc3	inc4	wman	wfree	wune
π_{t-1}	0.07**	0.06**	-0.00	-0.01	-0.03	0.03	0.00	-0.00	-0.00	0.02	-0.04
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)	(0.02)	(0.03)	(0.03)	(0.04)
$News_t^{Bild}$	-0.66***	-0.72***	-0.49***	-0.65***	-1.29***	-0.97***	-0.58***	-0.25*	-0.31	-0.08	-0.95***
	(0.21)	(0.19)	(0.19)	(0.19)	(0.27)	(0.24)	(0.19)	(0.15)	(0.20)	(0.19)	(0.22)
$News_t^{Tag}$	1.06***	0.91***	0.87***	0.79***	0.97***	0.77***	0.73***	0.47***	0.98***	0.90***	1.15***
	(0.25)	(0.22)	(0.22)	(0.22)	(0.29)	(0.25)	(0.20)	(0.15)	(0.23)	(0.21)	(0.26)
$News_t^{RTL}$	-0.82***	-0.62***	-0.40**	-0.13	-0.10	-0.26	-0.23	-0.10	-0.57***	-0.77***	-0.25
	(0.19)	(0.17)	(0.16)	(0.17)	(0.23)	(0.20)	(0.16)	(0.12)	(0.17)	(0.15)	(0.19)
$\pi_{j,t} - \pi_t$	0.14**	0.08	0.13*	0.18***	0.23***	0.20***	0.19***	0.21***	0.28***	0.29***	0.20***
	(0.07)	(0.06)	(0.07)	(0.05)	(0.06)	(0.07)	(0.06)	(0.05)	(0.09)	(0.09)	(0.06)
$perc_{j,t} - perc_t$	0.03	0.06	-0.00	-0.12*	0.01	-0.08	0.04	-0.04	-0.02	-0.05	-0.01
	(0.09)	(0.10)	(0.07)	(0.06)	(0.07)	(0.05)	(0.06)	(0.05)	(0.06)	(0.05)	(0.05)
cons	0.21***	0.19***	0.21***	0.23***	0.38***	0.29***	0.21***	0.14***	0.17***	0.13***	0.29***
	(0.06)	(0.05)	(0.05)	(0.05)	(0.07)	(0.06)	(0.05)	(0.03)	(0.05)	(0.04)	(0.06)
R^2	0.306	0.321	0.3478	0.392	0.336	0.312	0.355	0.428	0.303	0.398	0.372
N	130				130				130		

Note: Unconstrained 3SLS regressions using equations (3.5) and (3.10). Equation (3.10) is estimated using 5 lags of the dependent variables. *<0.1, **<0.05, *** p<0.01. Numbers in brackets denote standard errors. Sample 1999M1-2010M3. R^2 is calculated as correlation coefficient from actual values and predicted values from 2nd stage regression.

Second, we observe considerable heterogeneity of news effects across different household groups. Regarding age, we get significantly larger effects of coverage in *RTL* the younger

[16]The results using the entire media data set are qualitatively similar and are available upon request.

the survey participants. This result matches a well-known pattern in German media consumption, namely that the viewers of *RTL* tend on average to be younger than those of other channels. Similarly, news coverage in *Tagesschau* has a lager effect on younger households, whereas the impact of *BILD* is rather homogeneous across age groups. Separating households according to income, while no effect is found for *RTL*, news coverage of *BILD* and *Tagesschau* affect households the more the lower their income. However, given that the *BILD* lowers the expectation gap, we should get lower expectation gaps of the poor compared to the rich, which is in contrast to what we observe in the data. This result, puzzling at first glance, could also be understood in a different way. Households with the worse expectations react more to any news about inflation than other households which are less prone to media effects in general. Finally, with regard to occupation groups, we observe that *Tagesschau* increases the expectation gap of the unemployed by more than the gaps of manual workers and self-employed. However, *BILD* strongly reduces the difference between the expectations of unemployed and professional forecasters, without affect the remaining occupation groups.

Again, applying SUR instead of system 3SLS yields slightly different results (see Table B.4). Most importantly, we do not find an effect of news coverage in *Tagesschau* on young households, while by contrast, media coverage in *RTL* is estimated to be significantly negative for income groups.

Summing up, we find that the pure volume of news coverage indeed helps explain the heterogeneity of households' expectation gaps, and that summing across all media sources masks important effects. Next, we move from the volume to the tone of media reports in order to shed more light on our previous, sometimes striking results.

3.5.2 The Tone of News Coverage

As before, we first present results of media indexes with a positive and a negative tone, before distinguishing the effects between single media sources. The results using aggregate tone variables are shown in Table (3.5), and again replicate the effects of inflation and perceptions. Low-income households even deviate more strongly from experts compared to what we found before.

Next, moving from the volume to the tone of media reports leads to the following conclusions. First, we find that the results are surprisingly sensitive to the underlying coding of the tone of news reports. Defining the tone of an article in a very narrow sense ($News_t^{pos\ val}$ and $News_t^{con\ val}$), we get positive news effects on expectation gaps, no matter if journalists judge the inflation environment positively or negatively. By contrast, if we classify the tone in a broader sense, we get negative coefficients for both positive and negative news cover-

Table 3.5: Results: Aggregate Tone - Endogenous News Coverage

	ylt30	y3044	y4559	yge60	inc1	inc2	inc3	inc4	wman	wfree	wune
π_{t-1}	0.06	0.03	-0.01	-0.06*	-0.11**	-0.05	-0.02	0.01	0.02	0.03	-0.04
	(0.04)	(0.03)	(0.03)	(0.03)	(0.05)	(0.04)	(0.03)	(0.02)	(0.04)	(0.03)	(0.04)
$News_t^{pos\ con}$	-0.52**	-0.67***	-0.36*	-0.20	-0.20	-0.59**	-0.23	0.17	-0.57**	-0.53**	-0.42
	(0.25)	(0.22)	(0.21)	(0.23)	(0.28)	(0.24)	(0.19)	(0.15)	(0.24)	(0.23)	(0.28)
$News_t^{neg\ con}$	-0.66	-0.70	-0.54	-1.09**	-2.27***	-1.45***	-0.98**	-0.71**	-0.58	-0.52	-1.91***
	(0.56)	(0.49)	(0.48)	(0.51)	(0.65)	(0.54)	(0.44)	(0.34)	(0.48)	(0.45)	(0.57)
$News_t^{pos\ val}$	0.81***	0.81***	0.59**	0.36	0.66*	0.85**	0.61**	0.20	0.23	0.54*	-0.10
	(0.28)	(0.26)	(0.25)	(0.25)	(0.40)	(0.34)	(0.27)	(0.20)	(0.31)	(0.29)	(0.37)
$News_t^{neg\ val}$	1.56***	1.51***	0.99*	1.27**	2.90***	2.08***	1.35***	0.92***	1.39***	1.17**	2.88***
	(0.59)	(0.52)	(0.52)	(0.54)	(0.66)	(0.55)	(0.45)	(0.35)	(0.49)	(0.46)	(0.58)
$\pi_{j,t} - \pi_t$	0.18***	0.13**	0.15**	0.22***	0.34***	0.31***	0.23***	0.20***	0.27***	0.28***	0.24***
	(0.06)	(0.06)	(0.06)	(0.05)	(0.08)	(0.08)	(0.06)	(0.05)	(0.09)	(0.09)	(0.07)
$perc_{j,t} - perc_t$	0.05	0.05	-0.04	-0.05	0.01	-0.08	0.04	-0.02	-0.05	-0.05	-0.04
	(0.08)	(0.10)	(0.06)	(0.08)	(0.09)	(0.06)	(0.05)	(0.05)	(0.07)	(0.06)	(0.06)
cons	0.27***	0.27***	0.27***	0.28***	0.40***	0.36***	0.23***	0.10**	0.29***	0.24***	0.39***
	(0.06)	(0.06)	(0.05)	(0.06)	(0.08)	(0.07)	(0.05)	(0.04)	(0.06)	(0.06)	(0.07)
R^2	0.255	0.278	0.294	0.379	0.292	0.297	0.307	0.404	0.267	0.303	0.272
N	132				132				132		

Note: Unconstrained 3SLS regressions using equations (3.5) and (3.10). Equation (3.10) is estimated using 3 lags of the dependent variables. *<0.1, **<0.05, *** p<0.01. Numbers in brackets denote standard errors. Sample 1999M1-2010M3. R^2 is calculated as correlation coefficient from actual values and predicted values from 2nd stage regression.

age.[17] While we do not have an obvious explanation for this result, as we will show below, disaggregating the media indexes changes this result.

As regards heterogeneity, we find larger media effects for old and young households, for low income households and for the unemployed. Looking at the SUR estimates in Table (B.5), we do not find media effects of positive articles and TV reports. Still, we observe that reports with a negative tone broadly defined closes the expectation gap whereas the narrow definition leads to the opposite conclusion.

Finally, we turn to the effects of the single news media and show the results using the number of articles with a positive tone and with a negative judgment in *BILD*, *Tagesschau*, and *RTL* in Tables (3.6) and (3.7). Remember that we restrict ourselves to the use of the context variable since the more narrowly defined valuation concept only delivers a very small number of articles with an explicit tone.

Starting with the number of positive reports, we generally find less evidence of media effects. More positive news coverage in *BILD* lowers the expectation gap for all households, while we find a significant impact of positive news in *Tagesschau* only for the youngest households and for *RTL* only for the highest income quartile. The effect of positive coverage in *BILD* is larger for low income households and for the unemployed. Applying SUR estimates results in significantly positive coefficients for positive news coverage in *Tagess-*

[17] Lamla and Lein (2010) find that a negative tone increases the gap between professional forecasters and households in the aggregate. Their result might, *inter alia*, stem from the fact that they only apply the narrow coding of the news reports in their data set.

chau for nearly all household groups. The remaining results are unchanged (see Table B.6).

Table 3.6: Results: Disaggregate Positive Tone - Endogenous News Coverage

	ylt30	y3044	y4559	yge60	inc1	inc2	inc3	inc4	wman	wfree	wune
π_{t-1}	0.01	-0.01	-0.03	-0.04	-0.11***	-0.06*	-0.03	-0.00	-0.03	-0.01	-0.10***
	(0.03)	(0.03)	(0.02)	(0.02)	(0.04)	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)
$News_t^{Bild\ con\ pos}$	-0.37*	-0.66***	-0.40**	-0.43**	-0.76***	-0.85***	-0.51***	-0.08	-0.38*	-0.21	-0.64***
	(0.21)	(0.19)	(0.18)	(0.18)	(0.25)	(0.22)	(0.17)	(0.13)	(0.20)	(0.18)	(0.23)
$News_t^{Tag\ con\ pos}$	0.47**	0.20	0.16	0.06	-0.03	-0.26	-0.10	0.09	0.08	0.11	0.17
	(0.23)	(0.20)	(0.19)	(0.20)	(0.30)	(0.26)	(0.21)	(0.15)	(0.23)	(0.22)	(0.27)
$News_t^{RTL\ con\ pos}$	-0.10	0.08	0.14	0.15	0.30	0.22	0.31	0.24*	-0.13	-0.22	0.06
	(0.26)	(0.23)	(0.22)	(0.23)	(0.28)	(0.24)	(0.19)	(0.14)	(0.24)	(0.22)	(0.29)
$\pi_{j,t} - \pi_t$	0.16**	0.10	0.14**	0.20***	0.27***	0.24***	0.21***	0.20***	0.26***	0.26***	0.24***
	(0.07)	(0.07)	(0.07)	(0.05)	(0.06)	(0.07)	(0.06)	(0.04)	(0.09)	(0.08)	(0.06)
$perc_{j,t} - perc_t$	0.16*	0.12	-0.01	-0.11	0.01	-0.10*	0.02	-0.01	-0.02	0.00	0.03
	(0.09)	(0.11)	(0.07)	(0.07)	(0.07)	(0.06)	(0.06)	(0.04)	(0.07)	(0.06)	(0.06)
cons	0.25***	0.27***	0.28***	0.30***	0.49***	0.44***	0.30***	0.15***	0.29***	0.24***	0.39***
	(0.06)	(0.05)	(0.05)	(0.05)	(0.08)	(0.07)	(0.05)	(0.04)	(0.05)	(0.05)	(0.06)
R^2	0.302	0.217	0.274	0.368	0.228	0.156	0.228	0.416	0.318	0.396	0.276
N	129				129				129		

Note: Unconstrained 3SLS regressions using equations (3.5) and (3.10). Equation (3.10) is estimated using 6 lags of the dependent variables. *<0.1, **<0.05, *** p<0.01. Numbers in brackets denote standard errors. Sample 1999M1-2010M3. R^2 is calculated as correlation coefficient from actual values and predicted values from 2nd stage regression.

Turning to the effects of negative news coverage, the results in Table (3.7) suggest that households deviate more from experts if *BILD* and *RTL* increase the number of news reports presenting inflation as a problem. Since the effects are significantly larger for young households, the poor, and the unemployed, negative news coverage indeed makes an important contribution to explaining why households' inflation expectations differ with respect to their socioeconomic background. By contrast, more negative news coverage in *Tagesschau* lowers the gap between households and professional forecasters, while the effect is larger for the young and the old, low-income households, and not significantly different from zero for occupation groups. Assuming exogeneity of news coverage and using SUR delivers a fairly different picture. According to the results in Table (B.7), *BILD* has no significant impact, *Tagesschau* affects the poor and the unemployed negatively, and negative news coverage in *RTL* seem to raise the expectation gap of low-income households.

Summing up, we find a number of interesting results if we split the aggregate tone variable into the three most important single news media. Remember that we were surprised to find that news coverage in *Tagesschau* widens the gap between households' and experts' inflation expectations. Distinguishing positive from negative media reports, this result does not hold anymore. Instead, a more negative judgment of price developments in *Tagesschau* moves households closer to the best available forecast. The contrary results arise for the media effects of private TV news and tabloid newspapers: In this case, a more positive news coverage makes people to be more in line with experts, while more negative news raises the expectation gap.

Table 3.7: Results: Disaggregate Negative Tone - Endogenous News Coverage

	ylt30	y3044	y4559	yge60	inc1	inc2	inc3	inc4	wman	wfree	wune
π_{t-1}	0.06*	0.05	-0.01	-0.04	0.01	0.05	0.04	0.01	0.08**	0.10***	0.08*
	(0.04)	(0.03)	(0.03)	(0.03)	(0.05)	(0.04)	(0.03)	(0.02)	(0.04)	(0.04)	(0.05)
$News_t^{Bild\ con\ neg}$	0.80*	0.82**	0.26	0.36	2.34***	1.63***	0.91**	-0.41	1.40***	0.90**	3.05***
	(0.41)	(0.37)	(0.38)	(0.37)	(0.54)	(0.48)	(0.40)	(0.35)	(0.47)	(0.46)	(0.51)
$News_t^{Tag\ con\ neg}$	-1.26***	-1.14***	-1.11***	-1.45***	-1.52***	-1.31***	-0.86***	-0.57**	-0.41	-0.05	-0.20
	(0.43)	(0.39)	(0.38)	(0.38)	(0.45)	(0.39)	(0.32)	(0.24)	(0.44)	(0.40)	(0.54)
$News_t^{RTL\ con\ neg}$	0.73**	0.58*	0.61**	0.47*	0.23	0.44	0.42	0.70***	0.32	0.34	-0.42
	(0.33)	(0.30)	(0.29)	(0.29)	(0.39)	(0.34)	(0.28)	(0.23)	(0.32)	(0.30)	(0.36)
$\pi_{j,t} - \pi_t$	0.13**	0.08	0.11	0.20***	0.22***	0.18**	0.14**	0.21***	0.30***	0.28***	0.21***
	(0.07)	(0.06)	(0.07)	(0.05)	(0.07)	(0.07)	(0.06)	(0.06)	(0.10)	(0.10)	(0.06)
$perc_{j,t} - perc_t$	0.04	0.05	-0.04	-0.13**	0.08	-0.04	0.03	-0.11**	-0.09	-0.07	0.02
	(0.07)	(0.09)	(0.07)	(0.06)	(0.07)	(0.06)	(0.06)	(0.06)	(0.07)	(0.06)	(0.06)
cons	0.21***	0.18***	0.21***	0.23***	0.36***	0.27***	0.20***	0.11***	0.20***	0.16***	0.32***
	(0.05)	(0.04)	(0.04)	(0.04)	(0.07)	(0.06)	(0.04)	(0.03)	(0.05)	(0.04)	(0.06)
R^2	0.227	0.201	0.237	0.310	0.189	0.167	0.246	0.294	0.214	0.230	0.172
N	133				133				133		

Note: Unconstrained 3SLS regressions using equations (3.5) and (3.10). Equation (3.10) is estimated using 2 lags of the dependent variables. *<0.1, **<0.05, *** p<0.01. Numbers in brackets denote standard errors. Sample 1999M1-2010M3. R^2 is calculated as correlation coefficient from actual values and predicted values from 2nd stage regression.

Cautiously speaking, these conflicting results might be interpreted as follows. *BILD* and *RTL* might overemphasize negative price developments, even if professional forecasters do not judge the situation as badly as the media. As a result, households following these news sources deviate from experts when forming beliefs about future inflation. By contrast, if *BILD* and *RTL* exceptionally present inflation as unproblematic, households' expectations will come back to professional forecasters' beliefs. For news coverage in *Tagesschau*, a different story could be told. As we have argued before, *Tagesschau* reports on inflation in a very regular manner. Moreover, the tone of its TV reports are much more balanced compared to *BILD* and *RTL* whose coverage of inflation is mainly dominated by negative news. Thus, it is likely that a negative judgment of inflation in *Tagesschau* describes the situation in a much more adequate way which is more in line with the opinions of professional forecasters. As a result, more negative news coverage in *Tagesschau* lowers the expectation gap of households. In addition, our results could also be understood from a different perspective. For nearly all of the estimated models, we found larger media effects for the young, the old, the poor, and the unemployed, however, since the signs of the estimated coefficients are sometimes negative, this would suggest that the expectation gaps and forecast errors of these groups are lower than they actually are. However, it could be the case that those groups that are better in forecasting inflation - high income, middle age and employed households - are simply not as prone to change their expectations as soon as they hear about information in the media. By contrast, households that are worse in predicting prices seem to react strongly to any piece of news, and thus change their beliefs more frequently. The fact that those households with the largest expectation gap and forecast error are the same whose expectation are the most volatile in terms of the standard deviation (see Table 3.2), gives some evidence for this interpretation.

3.6 Conclusion

Recently, economic research has intensified in modeling heterogeneity and exploring the implications of heterogeneous agents in macroeconomic models (Hommes, 2006). In this chapter, we have analyzed the heterogeneity of inflation expectations in Germany, and, more precisely, the dependence of inflation forecasts on the demographic characteristics of households. In line with similar studies in the literature, we have found higher inflation expectations and forecast errors of households with lower income, younger households, and unemployed individuals. Furthermore, the same household groups show the largest deviations from expert expectations. We have tested the relative explanatory power of three sources that might drive these demographic expectation differentials. While we did not find an impact of aggregate inflation and household-specific inflation perceptions, we were able to identify household-specific inflation rates and heterogeneous news media consumption as main determinants of expectation differentials. Poorer and younger households deviate much more from expert forecasts in response to a change in their group-specific inflation rates, and households in lower income categories, unemployed, and younger and older households also react more strongly to news reports. Furthermore, we have shown that it is important to distinguish between different media sources, and to take into account the tone of news reports.

Our findings suggest important implications for communication strategies of central banks. If some household groups show systematic biases in inflation expectations and forecast errors, and if these differences are related to specific newspaper consumption, "the ideal communication strategy might then be multi-tiered" (Sims, 2009). Central bankers rarely appear on television, but if it is TV reports that systematically raise the forecasts of some household groups, this might be problematic. Furthermore, if some households rely more on their group-specific inflation rate instead of overall inflation, the credibility of the central bank might be undermined.

We think that several directions of further research seem to be worth following. Until now, possible differences in inflation expectations between creditors and borrowers have not yet been explored. This might be an important issue, due to the implications for redistribution effects and risk-taking on financial markets. A further question that we have left aside in this chapter is whether the reported differences in expectations are short-run or long-run phenomena. Anderson et al. (2010) have shown that the differences become minor because households learn over time. However, an impulse is needed to make this learning mechanism work, such as participating in a survey or individually-adapted communication policies. Also, as we have mentioned above, expectation differentials in Germany are found to be minor. Since we have chosen Germany mainly because of the availability of a large media data set, it would be interesting to see whether our results hold also in other countries, where demographic differences are more pronounced. Finally, it could be worth exploring one possible interpretation of our results, namely that those households with the worse expectations seem to react to any news, whereas households with better forecast capacities appear more confident with respect to their own beliefs about future prices and thus more reluctant to change these beliefs in response to news media information.

Chapter 4

Google Search Requests, the News Media and Inflation Expectations

4.1 Introduction

Already early in 1970, economists have used survey data to test different hypotheses of (inflation) expectation formation (Turnovsky, 1970). Paradoxically, the Lucas critique (Lucas, 1976) while putting expectations at center stage in macroeconomics, lead economists to abandon the analysis of the growing survey data sets. With the words of Manski (2004): "Rather than speculate on how expectations actually are formed, they follow convention and assume rational expectations." However, reaching "the limits to rational expectations" (Pesaran, 1987), the question of how to explain individuals' expectations was back on the agenda. Ironically, already Turnovsky's study had difficulties in confirming the hypothesis of rational expectations. Since then, economists have made increasing use of survey data for various purposes and set up new surveys to cover topics that had previously been neglected. To pick only two examples, Ang et al. (2007) show that inflation expectations derived from surveys possess much better forecaster performance than alternative measures gained from financial instruments or macroeconomic variables. Recently, Armantier et al. (2012) document that individuals after having revealed their inflation expectations in a survey subsequently react on these stated beliefs in an experiment modeling investment decisions.[1]

Besides these merits, survey expectations suffer from a number of weaknesses. Results can strongly depend on the exact question wording, which is particularly relevant with regard to inflation forecasts since respondents easily confuse price level and inflation rate depending on how they are asked (Bruine de Bruin et al., 2012, Dräger and Fritsche, 2013). Moreover, designing and implementing a questionnaire consumes time and money, whereas existing

[1] Inoue et al. (2009) argue that inflation can be predicted better using expectations derived from an Euler equation modeling intertemporal consumption decisions. However, this does not question the superior forecast performance of survey data compared to other methods.

surveys often face a small sample problem, both across time and respondents. Third, survey respondents might lack an incentive to state their best possible expectations due to the absence of financial consequences and peer pressure. In addition, Kahneman (2011) has recently claimed that economic decisions are taken by the use of two mental "systems": while filling in a survey puts the respondent into a situation that activates his cognitive reasoning, consumption decisions by contrast might primarily be governed by intuition. In this regard, if the same individuals participate repeatedly in the same survey, learning effects might result in much better predictions compared to individuals that do not take part in the survey. Finally, many countries still lack surveys that ask respondents to express their expectations in terms of a precise number or within predefined ranges. Instead, qualitative answers are provided making it necessary to apply data transformations that depend on various, often restrictive assumptions (Nardo, 2003).

In this chapter, we propose the use of a supplementary measure for inflation expectations and explore its usefulness: the number of Google search requests for inflation. People increasingly turn to the internet if they feel the need for more information on a certain topic. And among the various search engines, Google currently has a market share of about 66% in the U.S., hence representing the majority of all search queries.[2] While Google search data have successfully been used in forecasting (see our survey below), we propose its application in the context of inflation expectations since it does not suffer from the disadvantages of survey data. Internet search intensity does not depend on framing effects stemming form question wording, nor do the raw series have to be quantified. The data can easily be downloaded without charge, and since the number of downloads is virtually unlimited, the small sample problem is avoided. Moreover, the number of searches comes as a by-product of users' internet activities, hence search intensity is not affected by the particular circumstances of a survey or a telephone interview. If individuals use the Google web page in order to find information on a certain topic, they do so because they already feel the need to get informed, either because they are reluctant to seem uninformed in daily talks, or because they have a specific economic transaction in mind which makes it necessary to possess the latest news on inflation. This alleviates the "cheap talk" problem encountered in survey data. Finally, since Google search data is available on a weekly basis, this means that internet search requests could serve as a supplement to the existing survey data which is often compiled on a monthly basis and only released with some time lag. This is of particular interest for monetary policy that seeks to monitor price developments as timely as possible. It is against this background that the Bank of England states: "The Bank will continue to monitor these data (...). As further developments are made in this area (...), these data are likely to become an increasingly useful source of information about economic behavior" (McLaren and Shanbhogue, 2011).

[2]Web-statistics should be read carefully due to varying methods of calculation. However, the two leading web analyzers comScore (2012) and Experian Hitwise (2012) both report a market share of 66%.

In this chapter, we thus aim at exploring whether Google search requests can deliver new insights to the process of inflation expectation formation. For that purpose, we analyze U.S. data from January 2005 to May 2011 on households' and professional forecasters' expectations measured via survey data, newspaper articles and television reports on inflation, and Google search requests for inflation. In line with Carroll (2003)'s epidemiology model, we think of households adjusting both their demand for information and their expectations to the opinions of experts via the news media.

It is important to note at this stage that we do not consider Google search requests as an alternative measure for inflation expectations, but rather as a supplementary variable that can shed more light on the direction of future price expectations. Individuals can search for inflation in the web because they have heard about inflation in the news media and want to get more information on the topic which might subsequently result in an update of their expectations. Alternatively, according to the expectancy confirmation hypothesis (Traut-Mattausch et al., 2004, 2007), individuals might already expect higher prices in the future and seek to confirm their initial beliefs. While we aim at exploring these links in our empirical analysis, we broadly consider Google search request as a measure of attention (Da et al., 2011) and of the demand for information.

The contribution of this chapter is twofold: First, we analyze the news content of web searches on inflation. More precisely, we want to know whether search intensity evolves in a systematic way that can be attributed to real economic data. Note that Google searches might simply mirror the news coverage of inflation in the media, hence there might be no additional gain of using web searches in addition to the number of newspaper articles. To test whether Google searches are different, we compare the reaction of Google searches, TV reports and newspaper articles to changes in prices, variables describing the monetary policy and lagged values of households' and professional forecasters' expectations. In a second part, we take into account the various feedback effects among the news media, Google search requests and the inflation expectations of households and professional forecasters by estimating Vector Autoregressive models.

Our results show that users' demand for information can indeed be linked to economic fundamentals: Google search requests can be explained by price changes much better than media reports. Google users distinguish between headline and core inflation and they react asymmetrically: the demand for information increases if core inflation falls. Furthermore, internet users understand the difference between relative and overall price changes: they search less for inflation if the relative price variability increases. In periods of historically high inflation rates, the number of search requests is significantly larger. Also, and in contrast to media coverage, stock prices do not affect internet searches for inflation, but rising oil prices are found to reduce users' demand for information on inflation. Moreover, internet users pay attention to central bank behavior: unscheduled conference calls as well as issued statements increase search intensity. In addition, we find a positive effect from house-

holds' inflation expectations in the previous period on search requests: Google users seek for additional information if they belief prices to rise in the future. Higher inflation forecasts of experts only marginally increase Google search requests, but if professional forecasters disagree a lot on future prices, the resulting uncertainty leads to a large increase in Google users' demand for information.
With regard to the results of the VAR models, we find that television news coverage is driving newspaper coverage, in addition to a feedback effect. Building on this result, we show that Google search requests for inflation are mainly determined by TV reports and only to a lesser degree by newspaper articles. Again, we find considerable feedback effects, suggesting that journalists consider the interests of their readers when deciding on the newspaper's agenda. Finally, taking into account households' and professional forecasters' inflation expectations, we show that households' forecasts are driven by TV reports, newspaper articles, and Google searches, while the feedback effect from expectations on web searches is rather small and estimated less precisely. Furthermore, the impulse response from shocks on web searches to expectations is estimated more efficiently for weekly data, which indicates that the demand for new information has a rather short-run impact on peoples' expectations. About 20% of the forecast error variance decomposition of households' inflation expectations can be explained by Google search requests.

We start the chapter with a brief description of studies that use Google search requests in economics, with a special focus on how web query data can fit into the expectation formation process (Section 4.2). We then explain our estimation approach in Section (4.3) before describing the compilation of the media and Google data in Section (4.4). Subsequently, Section (4.5) presents the results and Section (4.6) concludes and discusses various directions for further research.

4.2 Google Econometrics: A Literature Review

In recent years, the internet has become an additional source of data on economic behavior. Edelman (2012) lists various studies that make use of internet data. Topics covered range from "labor and demographic economics" over "macroeconomics and monetary economics" to "economic history". Furthermore, he summarizes the various types of data that are available online, and generally have the advantage of being free of charge, timely availability, and being collected as a by-product of real user behavior. In a similar vein, Varian (2010) describes how the internet changed the nature of economic transactions. In the context of inflation, a project has recently been launched at the MIT designed to collect billions of retail prices from online sources in order to compute a daily inflation rate for the U.S. and other countries.[3] Cavallo (2013) uses this data to construct an online-based consumer price

[3]See Daily Price Index.

index for five Latin American countries. With the exception of Argentina, where officials are suspected to manipulate official price data, the online-based index captures the official fairly well.
This section provides an overview of using Google search requests in economic research. We summarize both the work that has been conducted with respect to nowcasting and forecasting economic variables with the help of internet search data, and discuss how Google search data might be related to expectation formation.

The literature on using Google data in forecasting models can be summarized as follows. As we describe in detail in the literature survey in Section (C.2) in the appendix, overall, the work conducted so far suggests very good now- and forecasting performance of Google search data. However, two technical questions are still up to debate. The first concerns the choice of the appropriate keyword searches: some authors simply use the variable of interest as the keyword ("job", "car sales", ...), while others start from the entire list of search categories provided by Google and subsequently reduce the number of queries applying statistical techniques such as principal component analysis. The second question is related to the time aggregation necessary for forecasting. Since the Google series is compiled on a weekly basis, whereas macroeconomic variables are mostly available on a monthly or quarterly frequency, the Google series has to be aggregated.[4] This is far from trivial: The week used by Google always ranges from Sunday to Sunday, hence one has to avoid overlapping data. However, as will become clear below, neither the keyword choice problem nor the time aggregation issue is relevant for our analysis.

In the context of inflation expectations, there is by now only one paper that employs internet search data. Guzmán (2011) uses a full set of 38 measures of inflation expectations for the U.S., including Google search requests for the word "inflation", and compares their forecast performance with regard to future inflation.[5] Importantly, Guzmán (2011) interprets Google search requests for inflation as a measure of revealed expectations: In her point of view, people only devote time for internet searches of inflation if they feel concerned of the future price development. Her analysis provides a number of interesting findings. Starting with long-run Granger causality tests of inflation expectations and actual future inflation, most of the expectation series are found to Granger-cause future prices changes, however, the Google series is the only variable that does not exhibit a feedback from actual inflation to expected inflation. Next, following the standard rationality tests conducted by Thomas (1999), Guzmán (2011) shows that the Google search data is biased but efficient if past inflation, oil prices, unemployment and money growth are tested individually. The most remarkable result, however, concerns the out-of-sample forecast performance: For the time span Jan-

[4] So far, mixed data sampling regression models (MIDAS) suggested by Ghysels et al. (2005, 2006) have not yet been used in the context of Google search data.
[5] The list of expectation measures consists of survey data of households, firms and professional forecasters, as well as expectations derived from a yield difference using the Treasury Inflation Protected Securities.

uary 2006 to October 2008, the root-mean-squared forecast error of Google search requests is considerably lower compared to all other expectation measures. Hence, it seems that using internet search data works fairly well in the context of inflation expectations. However, Guzmán (2011)'s analysis should be treated with care. She only uses one keyword and neglects the potential problems caused by Google's random sampling procedure (see below). Also, she aggregates the originally weekly Google data to monthly series which drops a lot of information.

In a more general perspective, as we have outlined in the introduction, treating Google search requests as an alternative expectation measure is not the only possible interpretation. Da et al. (2011) suggest to use Google search requests as a measure of revealed *attention*: If financial investors do not fully pay attention to news, they do not incorporate all available information in their investment decisions. One way of measuring investor attention consist of using the number of articles published in the news media, a route that has also been suggested by Carroll (2003) in the context of inflation expectations. If news coverage of a particular stock or inflation is high, it is assumed that this piece of information will soon or later reach all economic agents. Da et al. (2011), by contrast, argue that using Google search data to capture individuals' attention is a much more direct and timely measure. In this context, Granka (2010) is the only paper so far that analyzes empirically the links between television and print media news coverage on the one hand, and Google search requests on the other. Comparing the decay of interest in the different media following political as well as sensational news, her results indicate that Google searches are more closely aligned to TV broadcast than to newspaper articles. In addition, she finds that Google users are rather quick to loose interest in political events.

Hence, following the literature, one might interpret Google data as a measure of the demand for information whereas the news media provide the supply of information. However, the link between Google searches on the one hand and expectations and behavior on the other hand is less clear-cut. In line with models of rational inattention (Sims, 2003) and sticky-information (Mankiw and Reis, 2003), one might expect that households' inflation expectations are more rational, i.e. closer to the best-available forecast in an economy, if search intensity is high. To this effect, Google search requests might serve as a link between households' and professionals' expectations, instead of a proxy of expectations themselves. This poses the question of what determines search intensity. While the individual's demand for information might be driven by news coverage of future events in the media, individuals could also increase their search intensity for reasons that are entirely independent of their expectations of the future. To provide an example: Anvik and Gjelstad (2010) use a search-and-matching model in the labor market following Mortensen and Pissarides (1994) to motivate their use of Google search data to predict the unemployment rate. In this context, the number of internet searches related to unemployment would capture the search intensity of workers. While people might increase their search intensity for job vacancies if they expect higher unemployment in the future, there could also be various other reasons to look for a new job in the internet, such as dissatisfaction with one's current job or salary which is entirely unrelated to the individual's expectations of the future. A similar reasoning might be at work in the context of inflation expectations. People might search for inflation in the internet if they want to buy a house, an IPod, make a financial investment, or feel a general need to get information on the state of the economy as a whole. Hence, using Google search data as a one-to-one equivalent for expectations might be too simplistic. It is the purpose of this chapter to explore the various links between newspaper reports, Google search data, and inflation expectations in more detail.

4.3 The Information Content of Google Search Requests

4.3.1 The Information Content of Web Searches and Different News Media

In our analysis of the usefulness of Google search data with regard to household expectations, we first check whether users' search behavior can be explained by the economic environment. We also compare the information content of Google searches with newspaper articles and television reports. Modeling the "market for news", Mullainathan and Shleifer (2005) show that the degree to which the news media emphasize a certain topic depends not only on the preferences of journalists, but more importantly on the initial beliefs of their readers. Gentzkow and Shapiro (2010) provide empirical support for this hypothesis with respect to various political topics. Hence, it is crucial to know whether Google users react differently to prices and related economics variable than the news media, i.e. whether the reaction of the demand for information can be separated from the supply of information.
For this purpose, we regress the number of news stories in *The New York Times*, NYT_t, and the number of TV reports, TV_t, on the one hand, and the number of Google searches on the other hand on economic variables related to prices. More precisely, we estimate two equations:

$$NEWS_{i,t} = \beta_{i,0} + \mathbf{X}\beta_{i,1} + \beta_{i,2}\pi_{t-1}^{exp,hh} + \beta_{i,3}\pi_{t-1}^{expdis,hh} + \varepsilon_{i,t} \tag{4.1}$$

$$NEWS_{i,t} = \beta_{i,0} + \mathbf{X}\beta_{i,1} + \beta_{i,2}\pi_{t-1}^{exp,hh} + \beta_{i,3}\pi_{t-1}^{expdis,hh} + \beta_{i,4}\pi_{t-1}^{exp,prof} + \beta_{i,5}\pi_{t-1}^{expdis,prof} + \varepsilon_{i,t}, \tag{4.2}$$

with $NEWS_{i,t} = \{NYT_t, TV_t, Google_t\}$, and the matrix $\mathbf{X}$ containing a list of explanatory variables. We use OLS and adjust our estimates for serial correlation employing Newey-

West standard errors throughout.[6] Since the time span that a topic is dealt with in the media as well as users' attention span is fairly short, we estimate these two equations both on monthly and on weekly data. For the latter case, we interpolate all variables that are available only on a monthly basis by using the spline method.

Four sorts of economic variables might lead journalists to increase news coverage and readers to search for more information in the web: changes in consumer prices, price developments in related markets such as oil prices and stock prices, variables capturing the decisions of the central bank, and agents' beliefs and disagreement about future price changes. Table (4.1) shows the definitions and data sources of all variables.

Whereas the literature on media effects on inflation expectations following Carroll (2003) treats the number of media reports as exogenous variable, scholars in communication science have argued in favor of a mutual causality: the public's concern of economic issues depends on media coverage, but at the same time, the media also react to prevailing beliefs of their readers (Behr and Iyengar, 1985). Similarly, Mullainathan and Shleifer (2005) show that the degree of the news media's bias in reporting depends on the opinions of their readers. Hence, we expect that the news media increase coverage if households expect higher prices in the future. In the estimations, we use lagged households expectations $\pi_{t-1}^{exp,hh}$ to rule out endogeneity due to the mutual causality problem. With regard to Google search requests, internet users might seek for additional information if they expect higher inflation in the future, a link which is referred to as expectancy confirmation hypothesis (Traut-Mattausch et al., 2004). Furthermore, Mullainathan and Shleifer (2005) derive an impact from reader heterogeneity: the more diverge the beliefs of readers, the stronger the bias introduced by the media. To this effect, the new media might have an interest to report on controversial topics since this could attract more readers. We capture this effect by adding households' forecast disagreement $\pi_{t-1}^{expdis,hh}$, again lagged by one period. Finally, we also include the inflation expectations and disagreement of professional forecasters, $\pi_{t-1}^{exp,prof}$ and $\pi_{t-1}^{expdis,prof}$, which serve as a proxy for the best available forecast in the economy. This follows the idea proposed by Carroll (2003) according to which the media tend to quote experts in their articles about inflation and that households' inflation expectations are more in line with those of professionals' if the media increase the amount of coverage. For the same reasons, professional forecasters' expectations are used to explain Google searches.

Besides the prevailing mood of agents in an economy, journalists and internet users are expected to react to hard facts. Starting with prices, we use both headline inflation, i.e. changes in the overall price index, and core inflation, calculated as CPI without food and energy. Exploring the effect of core inflation is especially important in the U.S., where both

[6] Generally, it might be preferable to estimate seemingly unrelated regressions in order to account for the existence of unobserved shocks that affect each of the three news variables contemporaneously. Attention of readers and Google users is easily directed to special events such as 9/11, the death of Michael Jackson and the like. However, since our explanatory variables do not differ across the media series, SUR delivers the same results as single equation OLS, see Greene (2003), p.343.

the general public and the central bank are concerned more with core inflation (Blinder and Reis, 2005).[7] Furthermore, we use the annualized monthly inflation rate throughout, for both theoretical and econometric reasons. With regard to the former, note that it is typically the period-by-period change in the price level that is used in macroeconomic models, hence, it is of general interest whether the media react more to monthly than to annual inflation. Branch (2004) argues that in the U.S., it is indeed monthly changes in the CPI that are primarily reported in the media. Finally, using annualized monthly inflation instead of annual inflation avoids possible negative effects of moving-average terms in the residuals. We include the level of core and headline inflation, together with positive and negative changes, i.e. $\Delta^{+}\pi_t^{ALL}$ contains all positive changes of headline inflation and is zero if inflation falls, whereas accordingly, $\Delta^{-}\pi_t^{ALL}$ consists of the negative changes. By looking at positive and negative changes separately, we can allow for possible asymmetric effects in both media reports and Google search requests in case if the coefficients of positive and negative price changes are significantly different. A large body of work stemming from communication science argues that the news media emphasize bad news over positive events (see Soroka, 2006 for references). Possible explanations for this asymmetric media reporting range from supply-driven factors - e.g., ideological preferences of journalists, political watchdog function of the media - to demand-driven news slanting - journalists reacting to the negativity bias of their readers (Mullainathan and Shleifer, 2005; Gentzkow and Shapiro, 2010). Similarly, Google users might also care more about negative news than about positive news, as it is predicted by prospect theory and its concept of loss aversion (Kahneman and Tversky, 1979).[8]

With regard to the effects from inflation on media reports, we expect the coefficients of positive and negative changes of inflation to be positive and of similar size if news coverage is symmetric. In case of asymmetric reporting, a rise in inflation should lead to more articles or Google search requests than falling price changes of the same magnitude. Note, however, that is is by no means obvious that rising inflation is coded as negative news thus leading to a more pronounced media coverage. While this might be true in general, it might also happen that journalists increase the amount of articles if inflation is falling, especially in times of deflation or in periods where inflation is close to zero. Figure (4.1) shows that the CPI inflation in the U.S. was indeed negative in 2009, hence, we could also expect a negative sign at least for the CPI inflation rate. In the literature, Soroka (2006) finds that *The Times* published significantly more articles about inflation if inflation is increasing whereas falling price changes hardly affect coverage at all. However, looking only at articles dealing with decreasing inflation, he finds the adverse effect: falling inflation increases the amount of news coverage on falling inflation.

[7]The correlation coefficient between headline inflation and core inflation is only .48, hence, multicollinearity should be a minor problem.

[8]Applying the concept of loss aversion in a panel study on inflation *perceptions*, Dräger et al. (2014) find that the inflation nowcast of households indeed depends asymmetrically on price changes.

Two further variables capturing the inflation environment are added. First, we use a measure for relative price variability computed as the interquartile range between 33 main components of the overall CPI following Bryan and Cecchetti (1994). As it is shown by Mankiw et al. (2003), rising variability between different prices raises the disagreement of households' inflation expectations. The effect on media coverage is less clear-cut. If some goods show considerable price increases compared to other goods, journalists might consider this as an interesting piece of news and raise coverage. By contrast, larger price variability could also lead to more disagreement among journalists resulting in a decrease in the amount of coverage if the media do not know whether the changes of single prices will affect the overall price level. The same mechanisms could also be at work for Google search requests. Next, we include a dummy variable that is equal to one if inflation is above a long-run average. Malmendier and Nagel (2013) show that individuals tend to collect a memory of historical inflation rates throughout their lives which will feed into their current inflation expectations. Hence, we would expect that individuals increase their demand for information if inflation is above a long-run average. With regard to media coverage, journalists frequently report on extraordinary events thereby capturing the attention of their readers, i.e. they increase coverage if inflation is on historically high or low levels. We calculate the long-run average using inflation rates from 1946 until each month of our sample.

Next, we explore whether oil prices capturing supply-driven price pressures affect the amount of inflation coverage. If journalists and Google users think that rising oil prices will spur increases in goods prices in the future, the number of articles and search requests will rise. Alternatively, rising oil prices could also decrease news coverage of inflation if consumers associate oil prices with other topics such as the economic situation in general or a looming recession. In this case, Google users devote more attention and time to the search for different topics. A similar effect could be at work for stock prices. If a rise in share prices is put in context to an economic recovery, it will work as a proxy for demand-driven inflation thus increasing coverage and web searches. By contrast, rising stock prices might drive off journalists' and users' attention from inflation.

A final set of variables is composed of actions of the central bank. If the Fed raises the level of the Federal Funds rate, this might signal fear of future price pressures thereby increasing the number of articles and Google searches. However, the effect can also work the other way round: if the Fed increases the interest rate, this might take inflation off the journalists' and the general public's agenda. In order to take into account the change at the head of the Fed, we interact the Federal Funds rate with a chairman dummy set to zero for the Greenspan presidency and one since February 2006 when Bernanke took over. Whereas the Fed holds eight regular meetings per year, in the event of extraordinary circumstances, policy decisions are taken via unscheduled conference calls. The possible effect of such policy surprises is included by another dummy variable set to one for those periods in which a conference call takes place. Finally, we investigate whether the communication strategy of the Fed has an impact on the amount of news coverage on inflation. In January 2000, the Fed decided to issue a statement after each meeting, independently of whether the interest rate has been changed. The dummy variable $statement$ will thus test whether the media's and Google users' interest in inflation is higher in periods of central bank meetings.

Table (4.1) summarizes the full set of the explanatory variables.

Table 4.1: Explanatory Variables

Variable	**Definition**	**Source**
π_t^{ALL}	Headline inflation, seasonally adjusted	Bureau of
$\Delta^+\pi_t^{ALL}$	Positive changes of headline inflation, zero otherwise	Labor Statistics
$\Delta^-\pi_t^{ALL}$	Negative changes of headline inflation, zero otherwise	
π_t^{CORE}	Core inflation, all items less food and energy, sa	
$\Delta^+\pi_t^{CORE}$	Positive changes of core inflation, 0 otherwise	
$\Delta^-\pi_t^{CORE}$	Negative changes of core inflation, 0 otherwise	
$price\ variability$	Measure of relative price variability, interquartile range between 33 components of overall CPI	
π_t $above\ average$	Dummy variable, set to 1 if inflation above a long-run average calculated using data from 1946.	
$log(oil\ price)$	Log of oil price, measured in US-Dollar per barrel	Fed of St.Louis[a]
$log(S\&P500)$	Log of S&P500 stock index	Robert Shiller[b]
$Fed\ Funds\ Rate$	Level of Federal Funds Rate	FOMC
$\Delta\ Fed\ Funds\ Rate$	Change of Federal Funds Rate	
$conference\ call$	Dummy, set to 1 if Fed holds an unscheduled conference call	
$statement$	Dummy, set to 1 if Fed issues a state after policy meeting	
$chair$	Dummy, set to 1 for Bernanke presidency, 0 for Greenspan	
$\pi_{t-1}^{exp,hh}$ $median$	Median of households' inflation expectations	Michigan Survey,
$\pi_{t-1}^{exp,hh}$ $mean$	Mean of households' inflation expectations	Consensus
$\pi_{t-1}^{expdis,hh}$ iqr	Interquartile range of households' inflation expectations	Economics and
$\pi_{t-1}^{expdis,hh}$ var	Variance of households' inflation expectations	Survey of
$\pi_t^{exp,prof}$ $median$	Median of professional forecasters' inflation expectations	Professional
$\pi_t^{expdis,prof}$ iqr	Interquartile range of professional forecasters' expectations	Forecasters (SPF)
$\pi_t^{exp,prof}$ $mean$	Mean of professional forecasters' inflation expectations	
$\pi_t^{expdis,,prof}$ var	Variance of professional forecasters' expectations	

(a) See Fed Oil Price Data.

(b) See Shiller Stock Price Data.

4.3.2 Interactions Between Google Searches and the News Media

Having analyzed whether Google search requests react to economic variables and how these reactions compare to news media coverage, we now explore the various feedback effects between the news media, Google searches and inflation expectations by estimating Vector Autoregressive (VAR) models.

First, we expect television and newspapers to react to each other, with causality running from TV to newspapers since the release of the CPI figures is reported in the evening news on the same day, while it is only dealt with in the newspapers during the following days. In order to check for this expected causality, we start with estimating a bivariate VAR model including television reports and newspaper articles on a daily basis.[9] Looking a TV news and print media separately instead of aggregating them into one single media index is important since Americans spend more time watching television compared to reading newspapers. Blinder and Krueger (2004) surveyed households in the U.S.: 46% of respondents picked television as most import source of information about economic policy, compared to 18% choosing newspapers and 10% the internet. Furthermore, the results of Granka (2010) suggest a much tighter link between Google searches and television compared to Google searches and newspaper articles.

Second, we proceed fitting a VAR using the news variables together with the Google search requests which allows us to check whether internet users react differently to television and newspapers while taking into account feedback effects from the users' interests on the news media agenda. It is important to note that the mode of journalism in a TV channel is fairly distinct from the one in a newspaper. Television has the advantage of being able to react much faster to incoming news than newspapers, which can only reach their readers on the next day. By contrast, television broadcasts are much more limited by space and by their attention span of their viewers. Hence, one might expect that inflation receives much larger public attention if it is part of TV broadcast, but that it takes much more extraordinary changes in the inflation rate until it is considered newsworthy. At the same time, we expect that newspapers offer a much more pronounced and detailed treatment of inflation to their readers, resulting in longer-lasting effects on Google search requests. Since the Google data is collected on a weekly basis, we estimate the VAR both using weekly data counting the number of articles and TV reports published each week, and using monthly data computing the monthly average of the Google series. In order to rule out spurious results that might arise if each of the series does not react to each other but on actual inflation, we add the annualized monthly inflation rate as an exogenous variable. For the VAR using weekly data, we use the interpolated inflation rate.

[9] Research in communication science often concludes that newspaper coverage drives television news rather than the other way round (Vliegenthart and Walgrave, 2008). Most of this research, however, lacks a sound empirical methodology.

Third, we estimate a large VAR adding households' and professional forecasters' inflation expectations to the news media and the Google data. This setting mirrors Carroll (2003)'s epidemiology model: households's inflation expectations are expected to be driven by professional forecasters' expectations via the news media. Our VAR framework allows for feedback effects between these variables, while at the same time considering TV and newspaper reports differently and including households' demand for information on prices in addition to their expectations. Once again, we include the inflation rate as exogenous variable and apply the VAR to both monthly and weekly data, the latter using the interpolated inflation expectations. We expect professional forecasters' expectations to Granger cause the other series, such that the release of the best available forecast is transmitted via the news variable to households that subsequently increase their demand for additional information on inflation and adjust their inflation expectations.

The analysis of the VAR models proceeds in three steps. We first select the lag order and run Granger causality tests to determine the Cholesky ordering. In a second step, impulse response functions are analyzed before we finally compute the Forecast Error Variance Decompositions (FEVD).

4.4 The Data

This section presents the main variables used in the subsequent analysis, namely the news coverage of inflation in the print media and on television, Google search requests for inflation, and inflation expectations of households' and professionals.

Starting with the expectation series, we take the monthly one-year-ahead inflation expectations from the Michigan survey[10]. Since 1978, the University of Michigan asks a randomly selected sample of 500 American households to state their expected rate of inflation within predefined ranges. The resulting individual responses are then transformed into an aggregate mean and median series, together with the corresponding variance and interquartile range (IQR) to measure households' disagreement on future price developments.[11] In what follows, we use both the mean and the median series in the estimations.

The inflation expectations of experts are taken from *Consensus Economics*[12], an economic survey firm situated in London which covers the forecasts of a number of experts from public research institutes and private enterprises for variety of variables and different countries. We decided to use this data set since it is compiled on a monthly basis, in contrast to the widely used, but only quarterly implemented *Survey of Professional Forecasters (SPF)* which is now conducted by the Federal Reserve Bank of Philadelphia.[13] The Consensus survey con-

[10]See for data download Michigan.

[11]See Curtin (1996) for further details on the construction of the aggregate expectation series.

[12]See Consensus.

[13]See SPF

tains *fixed event forecasts*, i.e., each month, the respondents are asked to provide their forecast for the current and the next year. We follow Dovern et al. (2012) in adjusting the micro data and computing *fixed horizon forecasts* - one-year-ahead forecasts provided each month - in order to make the date comparable with households' inflation expectations. Again, we use both the mean and the median, together with the variance and the IQR of experts' inflation expectations. Unfortunately, the Consensus data to which we have access only reaches until March 2010, whereas the media variables are available until May 2011. Given that we want to keep as many observations as possible, we use the quarterly time series of the SPF to extend the monthly expectations series of the Consensus Economics Survey, following a procedure suggested by Luoma and Luoto (2009). First, we interpolate linearly the quarterly SPF series, before we regress the monthly inflation expectation of professional forecasters on a constant and on the interpolated series from the SPF.[14] We then use the fitted values of this regression for the missing time span April 2010 - May 2011 in our analysis. Since the data provided by the SPF does not contain a variance of professional forecasters' expectations, in what follows, we run estimations using the mean expectations using only data up to March 2010.

The newspaper and television data stem form the media institute *Media Tenor*[15]. First, all articles published in *The New York Times* between January 1998 and May 2011 have been searched for the keywords "inflation", "deflation", "price increase", "price cut", "price stability" and "oil price". The same has been done for the evening news of the four major U.S. television channels: ABC World News, CBS Evening News, Fox: Special Reports, and NBC Nightly News over the period January 2005 until May 2012. The time span covered and the choice of the newspaper and the TV channels is due to data availability. In a second step, each of the articles is evaluated by means of a human-based content analysis which excludes articles that use one of the keywords in a context different from its economic meaning. Furthermore, we can distinguish articles that deal with inflation in the U.S. from reports about inflation in foreign countries.[16] Summing all articles that contain at least one of these search terms gives us the total number of articles and television reports on inflation in the U.S.: NYT_US and TV_US. Furthermore, we can distinguish the news reports according to the main topic to which the article and the TV report refer to. We use this information to create five additional media variables: reports on rising inflation, INC, reports on decreasing inflation, DEC, reports on decreasing inflation and deflation, $DECFL$, reports on oil and energy prices narrowly defined, OIL_NARROW, and reports on oil and energy prices broadly defined, OIL_BROAD. Table (4.2) provides a detailed overview of all coded price categories

[14] The parameter estimates were 1.14 for the mean series, and 1.18 for the median series, both being highly significant and with an adjusted R^2 of 0.74. The fit of the regression for the IQR is slightly worse, with an estimated parameter of 0.48, and an adjusted R^2 of 0.27. Still, the interpolated IQR taken from the SPF and the IQR from Consensus Economics are correlated with a coefficient of 0.56 in our sample.

[15] See Media Tenor.

[16] See Menz (2012) for a detailed overview of the entire coding of the newspaper articles.

and the corresponding number of news reports, calculated both for all countries and for the US only. The most important difference between the articles published in *The New York Times* and the television broadcasts lies in the relative weight of the coded categories. Whereas the number of articles dealing with increasing inflation, NYT_INC, adds up to 15-20% in *The News York Times*, the same topic comprises only 5% of television broadcasts. While the same holds true for articles about falling inflation, the by far largest fraction of television reports deals with energy prices: narrowly defined, it amounts to to nearly 40%, and broadly defined, to nearly 70%. By contrast, only 25% of articles published in *The New York Times* address this issue.

Table 4.2: The Content of Newspaper Articles, TV Broadcasts, and Google Searches

	TV				**NYT**				**Google**
	Sum		%		Sum		%		%
	all	US	all	US	all	US	all	US	
VOL_ALL	**5565**	**5251**		94.4	**3936**	**2722**		69.2	
Consumer price index	132	127	2.4	2.4	288	213	7.3	7.8	
Price indicators (e.g. inflation rate) in general	294	285	5.3	5.4	686	542	17.4	19.9	
Increasing Inflation	8	8	0.1	0.2	585	342	14.9	12.6	
Increasing inflation or high level	244	193	4.4	3.7	0	0	0.0	0.0	
Inflation: high level	12	6	0.2	0.1	208	62	5.3	2.3	
Decreasing inflation	5	5	0.1	0.1	95	59	2.4	2.2	
Decreasing inflation or low level	20	18	0.4	0.3	0	0	0.0	0.0	
Inflation: low level	14	14	0.3	0.3	115	97	2.9	3.6	
Deflation	16	15	0.3	0.3	77	45	2.0	1.7	
Wages in general	15	12	0.3	0.2	67	47	1.7	1.7	
Wage level	116	107	2.1	2.0	40	33	1.0	1.2	
Rising labor costs	2	1	0.0	0.0	0	0	0.0	0.0	
Increasing labor costs or high level	12	11	0.2	0.2	6	3	0.2	0.1	
Decreasing labor costs or low level	1	0	0.0	0.0	1	1	0.0	0.0	
Commodity price, other	102	82	1.8	1.6	94	49	2.4	1.8	
Energy costs/prices	1464	1421	26.3	27.1	325	273	8.3	10.0	
Energy prices in general	105	95	1.9	1.8	0	0	0.0	0.0	
Food price	132	110	2.4	2.1	190	91	4.8	3.3	
Gold price	33	31	0.6	0.6	23	19	0.6	0.7	
Housing prices	706	693	12.7	13.2	139	80	3.5	2.9	
Natural gas price	7	7	0.1	0.1	13	9	0.3	0.3	
Oil price	569	507	10.2	9.7	300	219	7.6	8.0	
Impact of oil/energy price effect on companies	1	1	0.0	0.0	32	31	0.8	1.1	
Perceived inflation	1	1	0.0	0.0	26	26	0.7	1.0	
Producer Pries, other	31	31	0.6	0.6	81	69	2.1	2.5	

(continued)

	TV				NYT				Google
	Sum		%		Sum		%		%
	all	US	all	US	all	US	all	US	
Gas/diesel price / Petrol price	761	741	13.7	14.1	104	81	2.6	3.0	
Impact of gas or energy prices	762	729	13.7	13.9	332	247	8.4	9.1	
Inflation (as an effect of the Euro)	0	0	0.0	0.0	1	1	0.0	0.0	
Salaries, wages	0	0	0.0	0.0	0	0	0.0	0.0	
(T)Euro: price development after Euro introduction	0	0	0.0	0.0	1	0	0.0	0.0	
Purchasing prices	0	0	0.0	0.0	2	2	0.1	0.1	
Rising wages / high level	0	0	0.0	0.0	46	33	1.2	1.2	
Dropping wages / low level	0	0	0.0	0.0	31	27	0.8	1.0	
Total labor costs in general	0	0	0.0	0.0	0	0	0.0	0.0	
Non-wage Labor costs	0	0	0.0	0.0	2	1	0.1	0.0	
Social effects of food prices	0	0	0.0	0.0	4	2	0.1	0.1	
Drug prices	0	0	0.0	0.0	18	15	0.5	0.6	
Rent in general	0	0	0.0	0.0	4	3	0.1	0.1	
VOL_INC	**278**	**219**	**5.0**	**4.2**	**845**	**440**	**21.5**	**16.2**	**52.2**
Increasing Inflation	8	8	2.9	3.7	585	342	69.2	77.7	
Increasing inflation or high level	244	193	87.8	88.1	0	0	0.0	0.0	
Inflation: high level	12	6	4.3	2.7	208	62	24.6	14.1	
Rising wages / high level	0	0	0.0	0.0	46	33	5.4	7.5	
Rising labor costs	2	1	0.7	0.5	0	0	0.0	0.0	
Increasing labor costs or high level	12	11	4.3	5.0	6	3	0.7	0.7	
VOL_DEC	**40**	**37**	**0.7**	**0.7**	**211**	**157**	**5.4**	**5.8**	**2.6**
Decreasing inflation	5	5	12.5	13.5	95	59	45.0	37.6	
Decreasing inflation or low level	20	18	50.0	48.6	0	0	0.0	0.0	
Inflation: low level	14	14	35.0	37.8	115	97	54.5	61.8	
Decreasing labor costs or low level	1	0	2.5	0.0	1	1	0.5	0.6	
Dropping wages / low level	0	0	0.0	0.0	31	27			
VOL_DECFL	**56**	**52**	**1.0**	**1.0**	**288**	**202**	**7.3**	**7.4**	**5.4**
VOL_DEC	40	37	71.4	71.2	211	157	73.3	77.7	
Deflation	16	15	28.6	28.8	77	45	26.7	22.3	
VOL_ENERGY_NARROW	**2145**	**2030**	**38.5**	**38.7**	**638**	**501**	**16.2**	**18.4**	**7**
Natural gas price	7	7	0.3	0.3	13	9	2.0	1.8	
Oil price	569	507	26.5	25.0	300	219	47.0	43.7	
Energy costs/prices	1464	1421	68.3	70.0	325	273	50.9	54.5	
Energy prices in general	105	95	4.9	4.7	0	0	0.0	0.0	
VOL_ENERGY_BROAD	**3669**	**3501**	**65.9**	**66.7**	**1106**	**860**	**28.1**	**31.6**	**11.6**
VOL_ENERGY_NARROW	2145	2030	58.5	58.0	638	501	57.7	58.3	
Oil/energy price effect on companies	1	1	0.0	0.0	32	31	2.9	3.6	

(continued)

	TV				NYT				Google
	Sum		%		Sum		%		%
	all	US	all	US	all	US	all	US	
Gas/diesel price	761	741	20.7	21.2	104	81	9.4	9.4	
Impact of gas or energy prices	762	729	20.8	20.8	332	247	30.0	28.7	

Note: The percentages for the Google series are given as the average fraction of search terms over time relative to the week with the largest number of search requests, 2008w47.

Finally, given the novelty of the Google data, we now describe the construction of the internet search data in some detail.

In August 2008, Google Inc. (2008) introduced *Google Trends* which allows the comparison and analysis of web searches conducted by Google users. On its web page[17], Google Trends offers a free download of the volume of searches Google users have conducted for any keyword one might think of. The Google series are collected on a weekly frequency and are available from January 2004 until present. In order to avoid ambiguity of different words, a category filter is employed which distinguishes the brand "Apple" from the fruit "apple". Furthermore, Google Insight for Search enables users to narrow down the results of the queries to a geographic region, different time ranges and categories. The geographical region is identified via the IP addresses of Google users, hence our data set consists of search requests for inflation carried out by internet users in the U.S.. Google does not publish the total amount of searches for a specific keyword in a given time period, but provides a so called *query index*. This index is calculated by both normalizing and scaling the number of searches for a particular keyword. The data is normalized by dividing the volume of search queries for each keyword by the total volume of search queries for the requested time period and region:

$$\text{Normalized Value} = \frac{\text{actual search term volume}}{\text{total search volume}} \tag{4.3}$$

Google argues that the normalization corrects for for the growing number of internet and Google users, and allows comparing the search intensity between regions with low and high user densities. In a second step the query share is scaled in a range of 0 to 100 by dividing each data point by the search peak of all requested normalized parameters during the requested time span (Choi and Varian, 2009a).[18]

$$\text{Scaled Value} \equiv \text{Google Index} = \left(\frac{\text{normalized value}}{\text{maximum normalized value}}\right) \cdot 100 \tag{4.4}$$

[17] See Google Trends.

[18] Anvik and Gjelstad (2010) describe the Google data in more detail.

Using the Google search data comes with a problem, though: We do not know, a priori, the search terms Google users have in mind when they turn to the internet in order to collect information about current or future price developments. Da et al. (2011) had the advantage of using stock market tickers to identify Google searches for stocks of a single firm. Moreover, they asked students to list the words they would type into Google if they searched for a stock of a particular firm and checked whether the resulting search items differed across students. As we have discussed in the literature review, there does not yet exist a common method to choose the appropriate keywords for the construction of the Google series. We avoid this problem by taking the keywords that have been used by Media Tenor to search the newspaper articles and television reports which should leave us with a data set with the best possible comparability of the individual series. We defined 6 parameters to measure the peoples' online news-demand regarding inflation following the categories in Table (4.2).[19] Similar to the media variables, for each search category, we calculate the percentage of the total number of search requests. Since the Google variables are already normalized with respect to the total number of searches, the fractions are calculated as averages over time. Comparing the relative search requests with the relative fractions of media reports in Table (4.2) yields some marked differences. About 50% of search requests deal with rising inflation, which is 3 times more than newspaper articles and even 12 times more than television reports. Accordingly, Google users are much less interested in energy prices, which only make up a 7% and 12% of all searches.

Finally, it is important to note that the Google search data can vary with the date of the data download. This is due to the fact that Google computes its series by drawing random subsamples of all Google users asking: "What is the likelihood of a random user to search for a particular term from a certain location at a certain point in time?". Since the drawn subsamples are not representative, the resulting data might be scaled based on different peaks in the total number of searches in the period of interest. We follow Carrière-Swallow and Labbé (2013) and downloaded the Google series for the time span 2005w1-2001w18 several times during a couple of months. We then calculate the cross-section average out of these repeated draws and check if the signal-to-noise ratio exceeds a value of 5. Since the Google series look pretty stable, and given the fact that the signal-to-noise ratio is generally larger than 5, we are confident that our results are not subject to sample error.[20]

Summing up, our data set consists of five main variables: the number of articles about inflation published in *The New York Times*, the total number of news reports shown in four major television news broadcasts, an index of Google searches for inflation, and the inflation expectations of households and professional forecasters. Figure (4.1) plots these variables over time using monthly data, whereas the more volatile weekly series are found in Figure (C.1)

[19] Since the series containing all search requests for inflation, *GOOGLE_ALL*, is always the series with the largest search volume, we could add more than 5 parameters without rescaling our data.

[20] The repeated drawing is implemented by adjusting a R-code kindly provided by Dan Knoepfle, see Knoepfle.

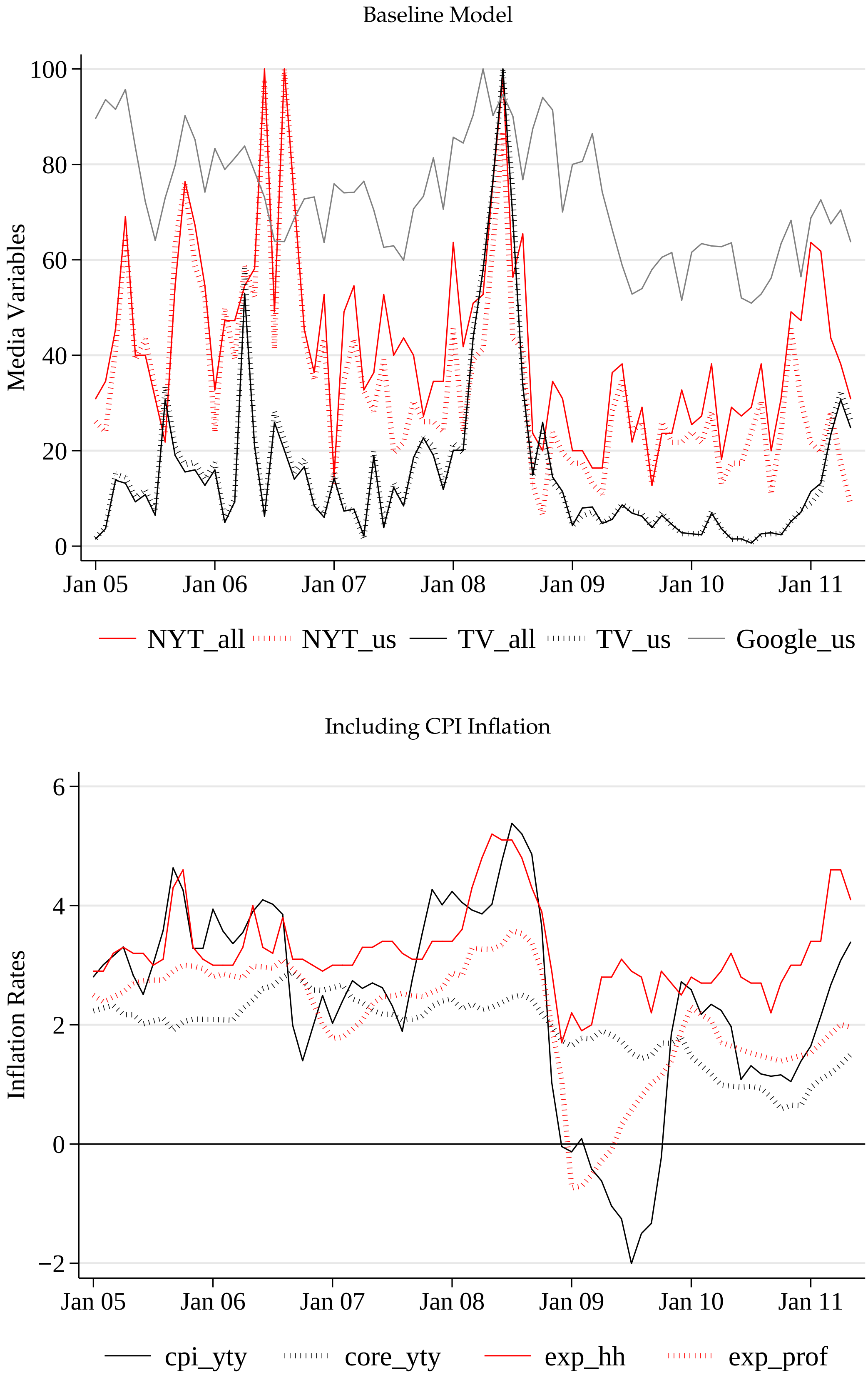
in the appendix. Note that in this plot as well as in the following analysis, we have scaled the news media variables by its maximum to make the data comparable with the Google series. Together, these data cover the time period 2005m1-2011m5. The upper panel of Figure (4.1) plots the newspaper articles, TV reports, and the Google series, where the news media variables are split into the overall number of news reports and the number of reports that only deal with inflation in the U.S.. Comparing print media coverage with TV coverage, the former is found to be much more stable, while the latter spikes only at some points in time. Overall, the correlation between TV and newspaper reports is about 0.45.

Turning to the Google series, the figure shows that the internet searches for inflation first decrease, then start increasing rapidly in the mid of 2007, before falling below their previous level at the beginning of 2009. Interestingly, the drop in Google searches in the first part of the sample has not been interrupted by the simultaneous increase in newspaper and TV reports. From mid 2007 onwards, Google searches move in line with the media series, albeit it takes much longer until internet users loose interest in inflation than it takes the media to reduce the number of reports in 2009. It is worth noting that the correlation of Google searches with the NYT is only 0.3, whereas it is 0.5 for television reports.

In the lower panel of Figure (4.1), we plot the inflation expectations of households and professional forecasters together with annual headline and core inflation. Google searches are only loosely linked to households' inflation expectations, with a correlation of 0.4 compared to 0.6 for articles in *The New York Times* and 0.75 for TV broadcasts. A final note concerns the possibility that our time series are nonstationary. Applying the Dicky-Fuller-GLS tests presented by Elliott et al. (1996), we can reject the null of a unit root for all of the variables in Figure (4.1).[21] Only the monthly Google series is found to be nonstationary for some lag lengths, however, the sample size with 76 observations is rather small. We thus proceed our estimations using the data in levels.

[21] Results are not shown but are available upon request.

Figure 4.1: Media Reports, Google Search Requests, and Inflation Expectations



Note: The upper panel plots the scaled number of articles about inflation published in *The New York Times* and the number of TV reports mentioning "inflation", both for all countries and for the U.S. only. Google searches are shown for the U.S. only. The lower panel plots the annual headline inflation, *cpi_yty*, annual core inflation, *core_yty*, together with households' (*exp_hh*) and professional forecasters' inflation expectations (*exp_prof*).

4.5 Results

The next section starts with the analysis of the information content of the articles in *The New York Times* and the different television channels, together with the Google search requests. We present results both for monthly and for weekly data. We then take a closer look at the interactions between the news media and web queries by evaluating several VAR models.

4.5.1 Information Content

Table (4.3) shows the estimation results of equations (4.1) and (4.2) using monthly data, Table (4.4) displays the results using weekly data. Models (1) and (2) use the median and the interquartile range of households' and professional forecasters' expectations, whereas models (3) and (4) are estimated using the mean and the variance of the expectation series. Also note that model (4) is only estimated over the time span 2005m1-2010m4, since it was not possible to extrapolate the variance of experts' inflation expectations. Hence, this specification allows us to check whether our results are robust with respect to the data extrapolation. In addition, we estimate rolling window regressions to take into account structural breaks in our sample.

Table 4.3: Results: News Content - Monthly Data

	NYT				TV				Google			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
π_t^{ALL}	1.16*	1.91***	1.15**	1.66**	0.27	0.84	0.13	0.61	-0.58	-0.19	-0.71*	-0.56
	(0.61)	(0.69)	(0.58)	(0.76)	(0.46)	(0.52)	(0.48)	(0.63)	(0.39)	(0.43)	(0.38)	(0.47)
$\Delta^+\pi_t^{ALL}$	-0.09	-0.6	0.13	-0.09	0.2	-0.2	0.34	0.05	0.04	-0.36	0.1	0.1
	(0.72)	(0.75)	(0.72)	(0.90)	(0.50)	(0.52)	(0.54)	(0.65)	(0.44)	(0.46)	(0.46)	(0.51)
$\Delta^-\pi_t^{ALL}$	-0.18	-0.6	0.01	-0.35	1.02*	0.7	1.13**	0.38	0.49	0.29	0.44	-0.14
	(0.73)	(0.74)	(0.73)	(0.78)	(0.52)	(0.52)	(0.55)	(0.60)	(0.43)	(0.42)	(0.44)	(0.45)
π_t^{CORE}	0.57	2.36	-0.29	3.09	2.45	3.81*	1.15	3.49	3.86**	4.94***	3.75*	5.90***
	(2.89)	(2.84)	(2.93)	(2.75)	(2.11)	(2.14)	(2.42)	(2.51)	(1.72)	(1.69)	(1.93)	(1.91)
$\Delta^+\pi_t^{CORE}$	2.33	-0.4	3.77	-0.27	1.38	-0.77	3.68	-0.84	4.30**	2.07	5.00***	0.72
	(3.10)	(3.33)	(3.10)	(3.51)	(2.21)	(2.31)	(2.32)	(2.54)	(1.82)	(1.88)	(1.86)	(2.00)
$\Delta^-\pi_t^{CORE}$	-2.28	-2.7	-1.82	-1.5	-3.62	-3.79	-3.73	-2.56	-7.17***	-6.35***	-7.63***	-5.67***
	(3.36)	(3.34)	(3.27)	(3.31)	(2.41)	(2.36)	(2.45)	(2.60)	(2.03)	(1.94)	(2.06)	(1.91)
$price\ variability$	-5.21**	-4.69**	-4.44**	-6.60***	-1.08	-0.82	-0.44	-3.32	-0.34	-1.05	-0.17	-3.36**
	(2.15)	(2.16)	(2.05)	(2.26)	(1.52)	(1.64)	(1.64)	(2.05)	(1.40)	(1.39)	(1.43)	(1.57)
$\pi_t\ above\ average$	-1.11	-0.99	-1.51	-0.1	5.52	5.70*	6.43*	10.46***	5.07*	5.77**	5.83**	7.57**
	(4.69)	(4.63)	(4.61)	(4.99)	(3.59)	(3.39)	(3.68)	(3.99)	(2.85)	(2.79)	(2.90)	(3.01)
$log(oil\ price)$	17.22	-5.66	11.93	-14.62	28.53***	11.64	29.99***	11.39	4.97	-5.08	9.66	5.61
	(17.07)	(17.97)	(16.46)	(21.68)	(10.17)	(12.10)	(11.63)	(15.26)	(10.49)	(11.53)	(11.24)	(12.78)
$log(S\&P500)$	-74.41**	-69.83**	-98.86***	-81.27***	-32.24	-29.69	-48.73**	-23.56	-21.34	-25.52	-23.28	-17.03
	(31.26)	(28.81)	(29.83)	(30.60)	(20.87)	(20.70)	(22.53)	(24.60)	(20.01)	(19.09)	(21.07)	(20.70)
$Fed\ Funds\ Rate$	9.88***	3.4	7.79***	0.21	4.56**	-0.15	3.46*	-3.9	4.59**	2.23	4.76**	0.31
	(2.77)	(3.80)	(2.75)	(4.29)	(1.79)	(2.68)	(2.02)	(3.23)	(1.80)	(2.39)	(1.94)	(2.74)
$FFR\times\ chair$	-1.63	1.1	-0.21	2.39	-0.67	1.33	0.18	2.12	-2.38**	-1.33	-2.47*	-1.71
	(1.76)	(2.00)	(1.74)	(2.21)	(1.09)	(1.40)	(1.27)	(1.63)	(1.18)	(1.31)	(1.28)	(1.42)
$conference\ call$	-6.22	-4.59	-6.34	-10.28	2.00	2.86	0.46	-3.5	7.45**	5.59*	6.68**	1.94
	(5.06)	(5.30)	(5.03)	(6.36)	(3.60)	(3.69)	(3.74)	(4.76)	(3.14)	(3.12)	(3.17)	(3.26)
$statement$	3.09	2.12	3.02	-0.44	2.53	1.77	2.29	-0.16	1.9	1.16	1.92	0.7
	(3.35)	(3.38)	(3.37)	(4.14)	(2.42)	(2.28)	(2.33)	(2.60)	(1.91)	(1.83)	(1.88)	(2.07)
$\pi_{t-1}^{exp}hh$	5.78	1.61	10.69***	10.23**	9.36***	6.18*	10.69***	7.42**	4.17	1.58	2.43	2.62
	(4.74)	(4.84)	(3.85)	(4.79)	(3.22)	(3.46)	(2.88)	(3.61)	(3.01)	(3.10)	(2.71)	(3.12)
$\pi_{t-1}^{exp\ dis}hh$	5.14	6.23*	-0.12	0.12	6.89***	7.67***	0.42	0.56	1.95	2.23	0.21	-0.2
	(3.32)	(3.30)	(0.43)	(0.62)	(2.50)	(2.45)	(0.35)	(0.46)	(2.08)	(1.99)	(0.29)	(0.33)
$\pi_{t-1}^{exp}prof$	-	9.85**		8.37	-	7.49**	-	8.69**	-	5.94**	-	4.2
		(4.43)		(6.27)		(2.97)		(4.40)		(3.01)		(3.83)
$\pi_{t-1}^{exp\ dis}prof$	-	-13.5		-5.95	-	-8.05	-	11.64	-	8.28	-	32.45**
		(13.08)		(24.62)		(9.62)		(18.73)		(8.20)		(14.77)
c	440.99**	506.23***	638.28***	627.04***	48.11	101.78	168.17	83.88	161.6	230.75*	161.89	150.68
	(189.80)	(178.45)	(182.49)	(180.90)	(129.87)	(130.31)	(140.04)	(149.35)	(123.35)	(121.42)	(131.44)	(130.85)
$R^2 adj.$	0.36	0.39	0.37	0.4	0.61	0.63	0.59	0.61	0.41	0.44	0.39	0.41
N	76	76	76	63	76	76	76	63	76	76	76	63

Note: HAC standard errors in parantheses. Models (1) and (2) use median series, (3) and (4) mean series. *<0.1, **<0.05, *** p<0.01. Sample (1), (2), (3): 2005m1-2011m5; (4): 2005m1-2010m4. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 4.4: Results: News Content - Monthly Data

	NYT				TV				Google			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
π_t^{ALL}	-0.02	0.65	-0.19	0.51	-0.16	0.47*	-0.41	0.23	-0.85***	-0.79***	-0.96***	-0.96***
	(0.35)	(0.41)	(0.35)	(0.53)	(0.24)	(0.28)	(0.31)	(0.43)	(0.23)	(0.26)	(0.24)	(0.32)
$\Delta^{+}\pi_t^{ALL}$	0.05	-0.08	0.07	0.53	0.34	0.01	0.69	0.23	0.11	-0.08	0.07	0.26
	(0.50)	(0.55)	(0.51)	(0.67)	(0.35)	(0.36)	(0.44)	(0.54)	(0.32)	(0.35)	(0.34)	(0.41)
$\Delta^{-}\pi_t^{ALL}$	0.1	-0.2	0.39	0.1	0.4	0.14	0.69	0.19	0.75**	0.75**	0.94***	0.61*
	(0.53)	(0.53)	(0.53)	(0.57)	(0.37)	(0.35)	(0.45)	(0.46)	(0.34)	(0.33)	(0.35)	(0.34)
π_t^{CORE}	1.58	3.52*	0.92	5.05**	2.44**	3.81***	1.12	2.51	3.59***	3.32***	2.90**	3.01**
	(1.80)	(1.96)	(1.73)	(2.06)	(1.23)	(1.31)	(1.58)	(1.70)	(1.17)	(1.25)	(1.19)	(1.27)
$\Delta^{+}\pi_t^{CORE}$	0.45	-0.53	0.5	-0.42	-1.23	-2.29	-0.17	-4.31**	-0.84	-1.04	-0.36	-2.72*
	(2.25)	(2.22)	(2.18)	(2.52)	(1.54)	(1.48)	(1.85)	(2.02)	(1.40)	(1.40)	(1.43)	(1.47)
$\Delta^{-}\pi_t^{CORE}$	-0.99	-0.54	-0.99	-2.21	-1.03	-0.03	-1.23	1.29	-3.05**	-2.49*	-3.57**	-0.85
	(2.26)	(2.31)	(2.17)	(2.45)	(1.55)	(1.54)	(1.88)	(1.96)	(1.42)	(1.46)	(1.43)	(1.43)
$price\ variability$	-3.24**	-3.67***	-2.73**	-3.31*	-0.43	-1.13	1.51	-1.39	-1.85**	-2.16**	-1.45	-4.01***
	(1.43)	(1.42)	(1.35)	(1.72)	(0.98)	(0.94)	(1.23)	(1.40)	(0.93)	(0.93)	(0.94)	(1.04)
$\pi_t\ above\ average$	2.88	-2.26	2.00	-5.97	0.42	-4.66	-1.42	-9.57**	6.38**	5.77*	5.21	4.27
	(4.89)	(4.99)	(4.76)	(5.48)	(3.36)	(3.43)	(4.22)	(4.48)	(3.12)	(3.23)	(3.18)	(3.25)
$log(oil\ price)$	9.19	-2.18	6.00	-11.13	15.47*	7.08	24.05**	21.36**	-19.85**	-18.59**	-15.29*	-12.99
	(12.99)	(13.16)	(12.46)	(13.33)	(8.90)	(8.82)	(11.00)	(10.87)	(8.26)	(8.69)	(8.33)	(8.25)
$log(S\&P500)$	-39.27	-53.08**	-49.55**	-49.86**	-6.8	-22.94	-51.19**	-49.28**	7.51	3.59	-0.72	4.15
	(24.92)	(24.36)	(23.48)	(25.05)	(17.06)	(16.57)	(20.80)	(20.63)	(15.88)	(15.93)	(15.83)	(15.57)
$Fed\ Funds\ Rate$	4.28**	2.78	4.31***	1.36	2.34**	1.66	2.64*	0.38	-0.53	0.02	-0.17	-2.25*
	(1.70)	(1.87)	(1.66)	(2.08)	(1.16)	(1.25)	(1.48)	(1.71)	(1.08)	(1.24)	(1.12)	(1.29)
$FFR \times\ chair$	1.28	-0.89	0.79	-2.60*	0.1	-1.35	-0.99	-1.35	1.28*	1.67*	1.31*	1.28
	(1.05)	(1.39)	(1.09)	(1.37)	(0.71)	(0.90)	(1.00)	(1.13)	(0.69)	(0.94)	(0.77)	(0.89)
$conference\ call$	-8.00*	-7.08	-8.19*	-7.93	4.23*	4.86**	2.64	2.98	4.07*	3.92*	3.45	4.42**
	(4.49)	(4.39)	(4.43)	(5.02)	(2.35)	(2.19)	(2.62)	(2.77)	(2.12)	(2.10)	(2.13)	(2.19)
$statement$	4.66**	4.62**	4.66**	4.90*	1.74*	1.77*	1.56*	2.31**	2.13***	2.18***	2.11***	2.64***
	(2.25)	(2.22)	(2.23)	(2.63)	(0.94)	(0.93)	(0.93)	(1.11)	(0.81)	(0.83)	(0.81)	(0.92)
$\pi_{t-1}^{exp}hh$	5.18	1.73	7.84***	2.56	6.16***	3.60	11.35***	13.93***	6.22***	6.59***	6.08***	6.65***
	(3.35)	(3.47)	(2.60)	(3.49)	(2.28)	(2.25)	(2.37)	(2.87)	(2.20)	(2.30)	(1.82)	(2.20)
$\pi_{t-1}^{exp\ dis}hh$	2.37	2.64	0.1	0.67*	9.74***	9.75***	0.14	-0.15	3.87***	3.66**	0.43**	0.18
	(2.25)	(2.24)	(0.30)	(0.38)	(1.54)	(1.46)	(0.27)	(0.31)	(1.46)	(1.42)	(0.21)	(0.24)
$\pi_{t-1}^{exp}prof$	-	9.20***	-	12.96***	-	7.70***	-	3.92	-	-0.19	-	-0.15
		(3.26)		(4.15)		(2.12)		(3.41)		(2.22)		(2.63)
$\pi_{t-1}^{exp\ dis}prof$	-	5.95	-	-15.31	-	12.54**	-	45.46***	-	6.81	-	9.55
		(9.01)		(18.21)		(6.05)		(14.87)		(5.74)		(11.16)
c	228.47	365.94***	309.51**	372.37**	-78.81	61	207.98*	205.75*	58.03	78.01	100.69	76.54
	(141.66)	(141.85)	(136.07)	(146.40)	(96.89)	(95.31)	(122.04)	(121.95)	(91.22)	(94.01)	(93.29)	(93.58)
$R^2 adj.$	0.23	0.25	0.24	0.27	0.69	0.73	0.62	0.69	0.47	0.48	0.46	0.47
N	329	329	329	269	329	329	329	269	329	329	329	269

Note: HAC standard errors in parantheses. Models (1) and (2) use median series, (3) and (4) mean series. *<0.1, **<0.05, *** p<0.01. Sample (1), (2), (3): 2005w1-2011w18; (4): 2005m1-2010m4. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

To start with the effects of price changes on media reports, it is noteworthy that both *The New York Times* and the television channels react only marginally to prices. In case of monthly inflation, a positive increase of headline inflation results in more newspaper articles about inflation, whereas we do not find any evidence for core inflation or for asymmetric reporting. In case of weekly data, by contrast, it is core inflation that exceeds a positive effect on articles in *The New York Times*, whereas headline inflation is not significant anymore. For television broadcasts, we get a positive effect of core inflation for both frequencies, and a positive effect from negative changes of headline inflation. This latter result gives evidence for asymmetric news coverage, albeit in a different way than expected: It is not rising inflation but falling inflation that affects the amount of television reports more than proportional. Next, newspaper and television react differently to changes in relative prices and to inflation rates that exceed a long-run average. Whereas the number of articles in *The New York Times* drops if the price variability increases, changes in relative prices do not affect television reports. Hence, the media do not seem to exaggerate price changes of single goods suggesting possible inflation pressure in the future if the relative price changes are extraordinarily large. In case of above average inflation, we find that television channels do report more on the topic if prices exceed the long-run average, while this does not have an effect on newspaper articles.[22]

It is important to note that we find considerable time variation in the estimated coefficients. The rolling regressions with a window size of 100 of model (2) using weekly data points to a structural break for the effect of prices at the beginning of 2008, both for *The New York Times* and for television reports, see Figures (C.2) and (C.3).[23] The most pronounced break occurs for the effect of negative changes of headline inflation on newspaper articles. Prior to 2008, *The New York Times* increased the amount of coverage if inflation decreased, whereas from 2008 onwards, falling prices resulted in less articles about inflation. The same holds true for core inflation, albeit to a lesser degree. This might be explained by the fact that between 2005 and 2008, falling inflation has been of greater concern for the American public, whereas since 2008, the financial crisis took inflation off the agenda. Interestingly, for television reports, negative changes of headline inflation lead to more news coverage from 2008 onwards which nearly covers the period of negative inflation rates indicated by the gray shaded bars in Figure (C.2) and (C.3). Furthermore, above average inflation leads to more news coverage in both print media and television from 2005 to 2007, subsequently turns negative, before getting positive again in 2009.

With regard to the effects from prices on web queries, we find that the behavior of Google

[22]We interpret the fact that above-average inflation has no effect if we use weekly data in the sense that this variable clearly captures the trend of inflation over the medium run.

[23]We do not show the results of the rolling window estimations for all four models. Hence, in some cases we refer to the general results in the text even if the results are slightly different in the graphs. For example, we find a positive coefficient for oil prices on television reports in all models expect model (2). The additional results are available upon request.

users can be explained fairly well with price changes. For both monthly and weekly data, the number of Google search requests rises with core inflation and with above-average inflation, while it falls with increasing price variability. We also find asymmetries in search requests. Rising core inflation leads to more Google searches, as well as falling core inflation. The latter effect is stronger, however, and also stays valid for weekly data whereas positive changes in core inflation do not remain significant. As in the case of television reports, we find that Google users search less for inflation if headline inflation is falling, while positive changes do not have an effect. Hence, our results suggest that Google users distinguish between headline and core inflation, do not search for additional information if headline inflation is falling but if core inflation is falling. A possible explanation for this might be that from the perspective of a consumer, rising headline inflation is a bad thing hence increasing the users' attention, while falling core inflation might be linked with deflation and and recession thus similarly being considered to be negative for the consumer's economic well-being. The rolling regression estimates in Figure (C.4) provide some support for this interpretation. Generally, i.e. prior to 2008 and after 2010, negative changes in headline inflation reduce the number of Google search requests, while in the time in between, internet users' also payed more attention to prices in general if headline inflation was falling. Finally, the rolling regressions show that the negative sign of headline inflation is driven by the negative inflation rates in 2009: Only at the end of 2009, decreasing headline inflation led to more Google search requests.

Next, we find that *The New York Times* does not link changes in oil prices to news coverage on inflation. However, increasing stock prices decrease the amount of coverage to a very large degree. Hence, it seems that *The New York Times* does not relate stock prices to demand-driven inflation, but that the newspaper simply devotes more space to topics different than inflation in times of a bull market. Television reports as well react negatively to rising stock prices, albeit to a lesser degree than newspaper articles. In addition, the TV stations seem to link rising oil prices to supply-driven inflation: For both monthly and weekly data, we find a strong and significantly positive effect from oil prices on television reports. This pictures changes if we turn to Google search data. Whereas stock prices do not affect internet searches for inflation, rising oil prices are found to reduce users' demand for information on inflation, at least for weekly data.

Again, we find some variation over time, albeit more gradual changes of the estimated coefficients. Starting with *The New York Times*, whereas the oil price effect is generally not significantly different from zero, since 2008, we do find a positive effect similar to the television reports. The Google search requests show a falling trend of the estimated coefficient of stock prices. Wile between 2006 and 2008, rising stock prices have been associated with an increasing demand for information on inflation, the size of the effects became gradually lower over time and turned negative at the end of 2008. With regard to oil prices, the estimated effect moves very much in line with the actual movement of oil prices. While increases in oil prices reduce the search requests for inflation in periods of low oil prices, the coefficient started to increase in line with the rising oil prices from 2007 onwards and eventually turned positive at the beginning over time. With the subsequent fall in oil prices, the estimated coefficient for Google search requests also started to fall.

We now come to the effects of the central bank's policy decisions. With respect to the Federal Funds rate, both newspaper articles, television reports and Google search requests react positively to a rise in the Fed's policy rate, both for monthly and for weekly data. Interacting the Federal Funds rate with the Bernanke dummy generally has no effect with the exception of weekly Google search requests. In this case, however, the Federal Funds rate is not found to be significant. Hence, the results indicate that both journalists and Google users associate rising interest rates with increasing inflation or with the idea that the inflation environment is somewhat problematic. The estimates are relatively stable over time, only for Google search requests, we find a positive trend in the coefficient of the Federal Funds rate. The occurrence of unscheduled policy meetings via conference calls leads to less articles in *The New York Times*. This can be explained by the fact that conference calls are mostly hold to decide on a cut in interest rate following an extraordinary event such as 09/11 that could result in a recession. However, for television reports and Google search requests, we find the contrary effect: months and weeks in which a conference call take place show more news coverage than in normal times. It seems that internet users and television channels react to extraordinary events and link these to possible effects on inflation. Besides these special events, the Fed's regular communication policy also affects the news media and Google users. Overall, we find more news coverage and search requests in months and weeks in which the Fed holds a regular meeting which are followed by a press statement since January 2000. Interestingly, this effect changes over time in case of newspaper and television reports. While the coefficient is significantly positive overall, the media react much stronger between 2007 and 2009 when inflation was low, whereas the reaction of Google search data is fairly stable over time. Hence, the degree to which the central bank's decisions are reported in the media depend on the general economic environment: in times of a positive trend in inflation, the Fed's meetings gain more attention in the news media, however, this does not lead to an increase in consumers' attention.

Finally, the results in Tables (4.3) and (4.4) show that the news media and Google search requests react both to the general public's inflation expectations and to the inflation forecasts of experts. Newspaper articles and television reports react positively to a rise in households' inflation expectations in the previous period, as well as to rising disagreement among households, hence, the news media are clearly linked to the opinion of their readers. Internet users also increase their search intensity if they had previously expected higher inflation. However, this effect only occurs for weekly data which suggests that the demand for additional information tends to be a short-run phenomenon. With respect to professional forecasters' expectations, we indeed find a positive effect on news coverage as claimed by Carroll (2003). Disagreement among professional, by contrast, does not affect news coverage. Interestingly, Google users react positively to a rise in professional forecasters' expectations, but only on a monthly basis. This suggests that consumers search for additional information in response to the best available forecast in the economy, but only, if the forecast prevails over time. Finally, Google users also demand more information if the professional forecasters disagree more on the future path of inflation, due to the resulting uncertainty.

We finally check the overall fit of the estimated equations. For *The New York Times*, we can only account for 30% of the variance, and even less if we use weekly data. However, this is a common result in the literature, given that there are many other variables not included in the estimation that affect news coverage. Interestingly, the fit is twice as large for television reports and also increases if we use weekly data. Yet, this larger fit might stem from the fact that we use four TV stations compared to only one newspaper. Finally, the estimated equations for Google search data have an adjusted R^2 of about 0.5. Plotting the fitted values together with the actual time series shows that the predicted values form our regressions capture the general trend of the news media and Google series fairly well (see Figure (C.5) in the appendix). Only in the middle of 2010, we observe a drop in the Google search requests which is not captured by our estimations.

4.5.2 Interactions

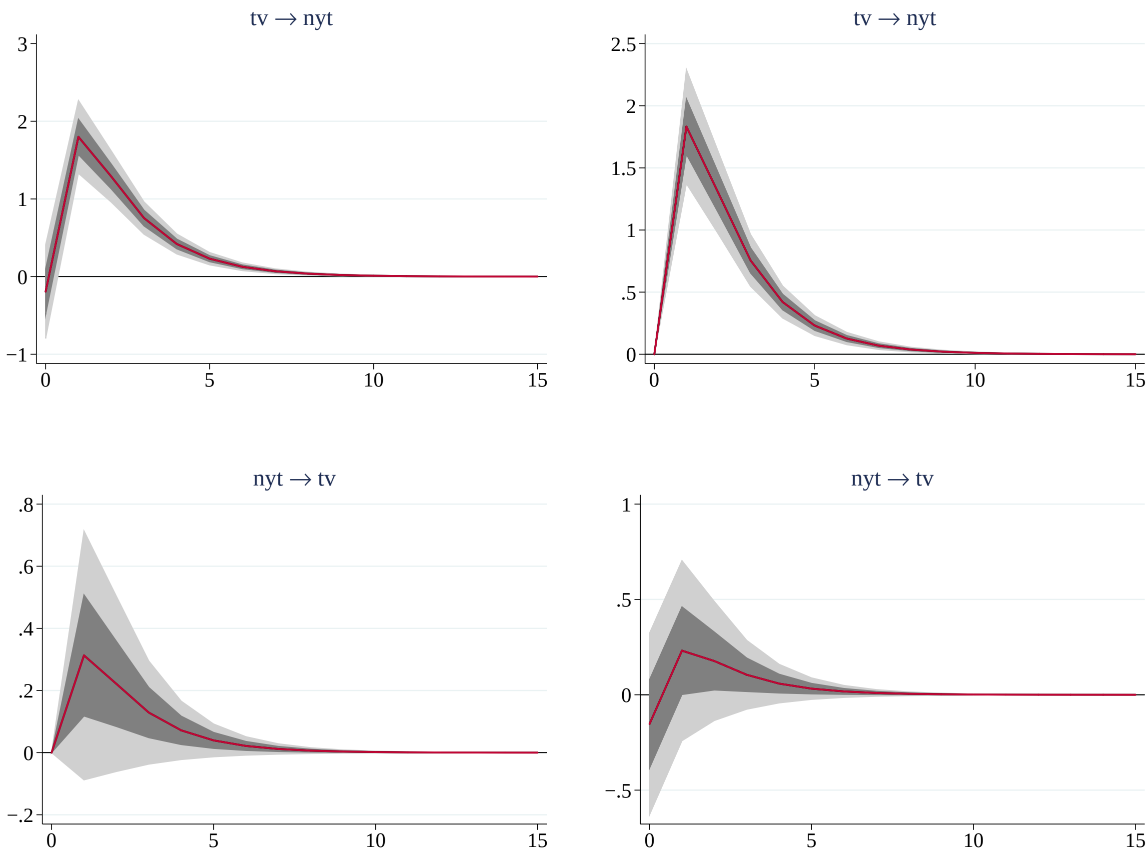
We now discuss the results of three VAR models that take into account the various feedback effects between the news media, Google search requests and inflation expectations.

News Media VAR

We begin with a bivariate VAR using the daily sums of newspaper articles and television reports in order to analyze the causality of these different news media sources. Starting with the choice of the lag order, we find the results of the information criteria rather inconclusive. The final prediction error (FPE) and the Akaike (AIC) suggest a VAR(28), the Hannan-Quinn criterion (HQC) a VAR(14), and the Schwarz Bayesian information criterion (BIC) picks a VAR with 7 lags. Since the estimation of higher-ordered VARs produces quite erratic impulse-response functions, we decide to estimate VARs with only one lag, running the VARs with the lag order suggested by the information criteria as robustness checks. The results (not shown) do not differ qualitatively with regard to the implied response functions within the first days following a shock.

Since we cannot rule out a priori that our results depend on the Cholesky decomposition, we estimate the VAR twice, first assuming that a shock to television reports affect newspaper articles on the same day, but not vice versa, before assuming that newspaper articles affect television reports contemporaneously. The resulting impulse-response functions are shown in Figure (4.2), where the graphs on the left use the Cholesky ordering $TV \rightarrow NYT$. The results confirm our hypothesis that newspapers react to television reports: a positive shock to TV reports raises the number of articles published in *The News York Times*. In addition, television news coverage is also positively affected by newspaper articles, however, the implied impulse-response functions are estimated less precisely and the magnitude of the response is much lower. Noteworthy, the results do not differ much for the different Cholesky orderings, and are also supported by Granger causality tests. TV reports Granger cause newspaper articles, but not vice versa, whereas we find no evidence for instantaneous Granger causality.

Figure 4.2: News Media VAR - Impulse-Response Functions - Daily Data



Note: Response to one s.d. Cholesky innovation. Dark gray area: 66% confidence interval, light gray area: 95% confidence interval, based on 2,000 bootstrap replications. Sample: 01jan2005-26may2011. The figures on the left use the Cholesky ordering $TV \rightarrow NYT$, those on the right use the ordering $NYT \rightarrow TV$.

Summing up, we take this first result as evidence for assuming an immediate effect from television reports on newspaper articles, and only a lagged feedback effect from print media coverage on TV news for the subsequent VAR estimation. This also makes intuitive sense: news broadcast in the evening have the advantage of being more up to date concerning important events that have occurred during the day. On the other hand, newspapers have more space available to present and discuss a topic, hence also being able to pay attention to this topic for a longer time span.

Baseline VAR

Next, we add the Google search data to the news media variables, fitting both a monthly and a weekly VAR. Determining first the optimal lag length of the baseline VAR using information criteria, we find that for monthly data, the HQIC and the BIC choose a VAR(1), while the FPE and the AIC suggest the use of 3 lags, whereas for weekly data, the FPE and the AIC choose even a VAR(2). Using the most parsimonious model, we continue to estimate a VAR with one lag. We motivate the Cholesky ordering with the help of Granger causality tests.

Table 4.5: Granger Causality Tests - Baseline VAR

		monthly		weekly	
Dependent Variable	Excluded	χ^2	p-value	χ^2	p-value
TV	NYT	0.43	0.511	0.13	0.719
	Google	4.96	0.026	4.77	0.029
	all	5.69	0.058	4.78	0.092
NYT	TV	1.09	0.296	5.11	0.024
	Google	2.21	0.137	1.42	0.234
	all	6.43	0.040	10.12	0.006
Google	TV	0.21	0.647	4.15	0.042
	NYT	0.06	0.814	7.43	0.006
	all	0.22	0.897	9.89	0.007

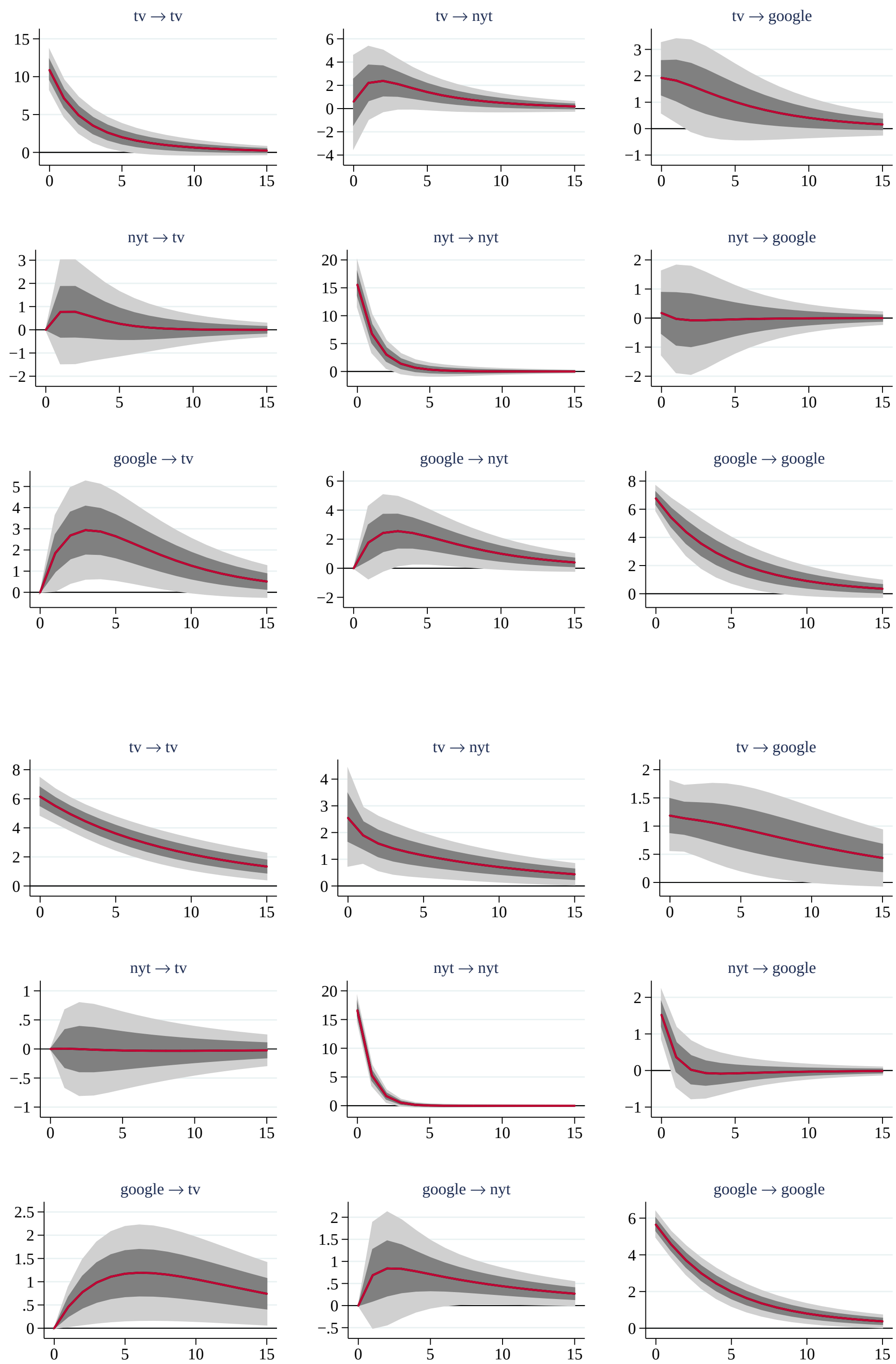
The results in Table (4.5) show that television reports can be predicted by Google searches in the previous period, and that newspaper reports are affected by the common impact of Google searches and television reports, whereas the Google data itself does not seem to be Granger caused by the news variables. The picture looks different, however, if we estimate the VAR using weekly data. Still, we find that television reports are caused by internet searches and that newspaper articles are caused by the combined effect of TV and Google. In addition, the results also show that TV news coverage has predictive power for print media, which confirms our assumed Cholesky ordering already suggested in the previous section. Most importantly, the behavior of internet users is now found to be predicted by both newspaper articles and television reports. The difference between both time frequencies in the Granger causality results supports the idea that media effects are rather short-lived than long-lived. While internet users increase their demand for information in response to a news

shock, this demand is satisfied fairly quickly over time. Based on these results, we allow for a contemporaneous effect from television reports on newspaper articles and Google search requests, and assume that journalists react to an increase in their readers' or viewers' interest only with some time-lag. This ordering is also supported by testing for contemporaneous Granger causality. For both monthly and weekly data, we find that the news media Granger cause Google search requests in the same period.[24]
Figure (4.3) and Figure (4.3) on the next page show the estimated impulse-response functions, where the upper panel uses monthly data and the lower panel weekly data. All variables are find to be persistent, while somewhat surprisingly, a positive shock to TV reports takes up to 15 weeks to die out, compared to only 10 weeks for search requests, and 5 weeks for newspaper articles. With regard to the interaction of print media and television, we find a positive effect from a shock to TV on NYT, but only if we use weekly data, while there is no significant effect from print media to television. Most importantly, the Google search requests react positively to news shocks in the media. Note that the results depend on the media source: The shock on TV is significantly for both monthly and weekly data, and dies out only gradually. By contrast, a shock on NYT to Google searches is only significant for weekly data, and persists for a much shorter period of time. Finally, the estimated impulse response functions document a considerable feedback effect from shocks on web searches to the news media, where again, the results are stronger for TV reports. Moreover, the responses build up gradually, suggesting that it takes some time until journalists pay attention to their readers' need for information.

These findings are also supported by the forecast error variance decompositions shown in Figure (4.4) for monthly data. The FEVD using weekly data are plotted in Figure (C.6) in the appendix, the results are qualitatively similar.
The FEVD of television reports depends largely on itself, while from 6 to 7 months onwards, the feedback effect from Google accounts for about 20%. Articles in *The New York Times* also exhibit this feedback, but the size is only half as large compared to TV reports. In addition, TV broadcasts account for 5% of newspaper articles, while the effect form NYT on the FEVD of TV is virtually zero. Finally, only 5% of Google searches can be accounted for by TV reports, while the NYT does not seem to play any role. If we turn to the results using weekly data in Figure (C.6), however, we do find an impact from NYT which is largest in the first weeks following a shock. Once again, these results show that media effects are very short-lived and that using lower frequency data can significantly downplay the role of the news media in determining the general public's need for information.

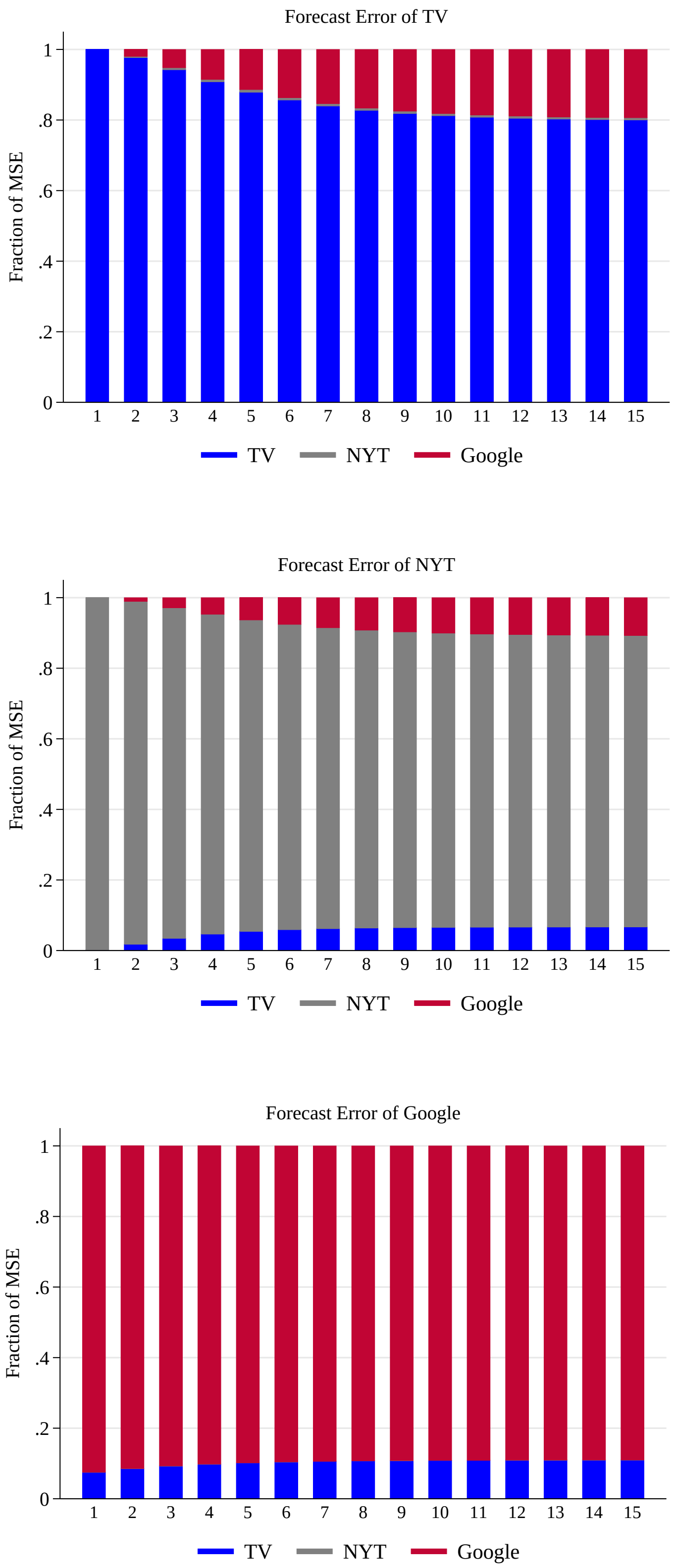
[24] Generally, the results do not change if we put Google searches first in the Cholesky ordering. While the effect from a shock to television reports on Google searches is not find to be significantly different from zero anymore, the sign of the effect does not change. Further results are available upon request.

Figure 4.3: Baseline VAR: Impulse-Response Functions - Monthly and Weekly Data



Note: Response to one s.d. Cholesky innovation. Dark gray area: 66% confidence interval, light gray area: 95% confidence interval, based on 2,000 bootstrap replications. Sample: 2005m1-2011m5 (upper panel) and 2005w1-2011w18 (lower panel). The upper panel uses monthly data, the lower panel weekly data.

Figure 4.4: Baseline VAR - FEVD - Monthly Data



Large VAR

Finally, we augment the baseline VAR with households' and professional forecasters' inflation expectations[25], again adding the annualized monthly inflation rate as exogenous variable. Experts produce the best available forecast of future prices, which is subsequently reported in the news media and transmitted to households that might then search for additional information via the internet and/or adjust their inflation expectations. Various feedback mechanisms are at work in this setting: Not only the news media react to their readers' agenda, but professional forecasters as well take into account inflation expectations of the general public by means of survey data.
Choosing the optimal lag length is not as clear-cut as in the case of the baseline VAR. For monthly data, the BIC suggests the use of a VAR(2) model, whereas the remaining criteria indicate the use of of 11 lags. Ivanaov and Kilian (2005) perform a simulation study evaluating the accuracy of different information criteria for estimating the true impulse response function. They show that for monthly data with a sample size of up to 80 observations, there is virtually no difference in the performance of the AIC, the HQIC, and the BIC criterion. Hence, we select the most parsimonious model, and even decide to estimate a VAR(1) model. For weekly data, the HQIC and the BIC choose a VAR(3), while the FPE and the AIC suggest the use of a VAR(11). Again, we decide to estimate a VAR(1) model, also because the behavior of higher-ordered VARs is rather unstable.[26]

The Granger causality tests for monthly and weekly data are shown in Table (4.6). Beginning with experts' inflation expectations, for both data frequencies, previous values of households' inflation expectations have predictive power, which might stem from the increased use of survey data in forecasting models. In addition, experts' expectations are Granger caused by TV reports. With regard to the news media, television reports are driven by Google search requests and households' inflation expectations, whereas newspaper articles can be predicted by experts' forecasts. In contrast to the previous variables, the results differ for web searches and households' expectations depending on whether we use monthly or weekly data. For the former, Google series are Granger caused by professional forecasters' expectations, whereas it is *The New York Times* that helps predict web searches on a weekly basis. A similar finding emerges for households' expectations: Using monthly data yields Granger causality from experts, television and newspapers, while using weekly data results in Granger causality from newspaper articles and Google search requests. As regards instantaneous Granger causality, using weekly data, only experts' inflation forecast cannot be predicted by contemporaneous changes of the other variables. Households' inflation expectations, are Granger caused by news media coverage and expert opinions of the same month.

[25]The following estimates use the mean of the expectation series. Results do not change if the median series is employed instead.
[26]Results for the VAR(2) are qualitatively similar and are available upon request.

Overall, we take the results from the Granger causality tests as supportive evidence for our theoretically motivated Cholesky ordering $Exp\ Prof \rightarrow TV \rightarrow NYT \rightarrow Google \rightarrow ExpHH$.

Table 4.6: Granger Causality Tests - Large VAR

		monthly		weekly	
Dependent Variable	Excluded	χ^2	p-value	χ^2	p-value
Exp Prof	TV	4.08	0.044	23.18	0.000
	NYT	0.46	0.497	0.11	0.742
	Google	0.13	0.715	0.28	0.594
	Exp HH	12.70	0.000	24.61	0.000
	all	14.70	0.005	32.03	0.000
TV	Exp Prof	0.20	0.652	0.22	0.641
	NYT	0.12	0.732	0.85	0.357
	Google	5.37	0.020	3.43	0.064
	Exp HH	5.01	0.025	8.59	0.003
	all	13.84	0.008	16.44	0.002
NYT	Exp Prof	2.95	0.086	9.09	0.003
	TV	0.23	0.633	0.01	0.905
	Google	1.55	0.214	0.50	0.478
	Exp HH	0.02	0.892	2.15	0.143
	all	10.15	0.038	37.41	0.000
Google	Exp Prof	5.62	0.018	1.93	0.164
	TV	0.11	0.738	0.22	0.640
	NYT	1.01	0.315	11.61	0.001
	Exp HH	0.58	0.445	1.45	0.229
	all	5.95	0.203	18.55	0.001
Exp HH	Exp Prof	4.42	0.036	1.90	0.168
	TV	2.72	0.099	0.48	0.489
	NYT	6.77	0.009	18.92	0.000
	Google	1.48	0.224	17.70	0.000
	all	15.51	0.004	31.17	0.000

Next, Figures (4.5) and (4.6) plot the impulse response functions using monthly and weekly data. The results for the responses of professional expectations in the first column further support the assumption that experts' forecasts are exogenous, with the exception of a positive effect from households' inflation expectations.[27]

With regard to the responses of the news media, the large VAR replicates the results from the news media VAR described above. While a positive shock to television reports increase news coverage in *The News York Times*, there is no such effect in the opposite direction. In addition, television broadcasts react positively to their readers' views: shocks to both Google search requests and to households' inflation expectations increase news coverage. By contrast, the print media reacts mainly to experts' forecasts and only to a lesser degree to peoples' demand for information measured by web searches.

Turning to the responses of Google search requests, for both monthly and weekly data, positive shocks to TV reports increase search intensity, whereas the effect from newspaper articles is significantly positive only in the first couple of weeks. While we find a positive effect from households' expectations on Google search requests, the estimated responses are rather small and only significantly different from zero if we use a 66%-confidence interval. The expectations of professional forecasters do not affect web searches and households' inflation expectations directly. Rather, the latter are driven by TV reports, newspaper articles, and Google searches. Note that the effect from shocks on web searches to expectations is estimated more efficiently for weekly data, suggesting that users' demand for additional information has a rather short-run impact on peoples' expectations.

Finally, Figure (4.7) shows the FEVD for Google search requests and households' expectations for both monthly and weekly data. Results for the remaining variables generally replicate the findings of the previous VARs and can be found in the appendix in Figure (C.7).

Starting with the FEVD of the monthly Google search requests, we find that most of the variation is explained by shocks to the Google series itself. Nevertheless, we find that TV reports, especially in the short run, and the forecasts of experts contribute to the FEVD of web searches, even if their impact only adds up to 10% after 15 months and weeks. Furthermore, only if we use use monthly data, households' inflation expectations account for some variation in Google search requests. Finally, turning to households' expectations, we can explain a much larger fraction of the FEVD. TV reports account for about 15% already in the short run, while the impact of newspaper articles is about half as large. Furthermore, experts' expectations explain 5% in the short run, which increases to 10% after 15 months. Most importantly, we find that Google searches, after having only a marginal impact in the first three months, contribute for about 5% of the variation of households' inflation expectations after 6 months.

Turning to the results for weekly data, note that only about 10% Google searches can be explained by newspaper articles and TV reports over a 30 week horizon. By contrast, Google

[27] In the case of weekly data, a positive shock to TV reports is found to decrease the forecast of experts.

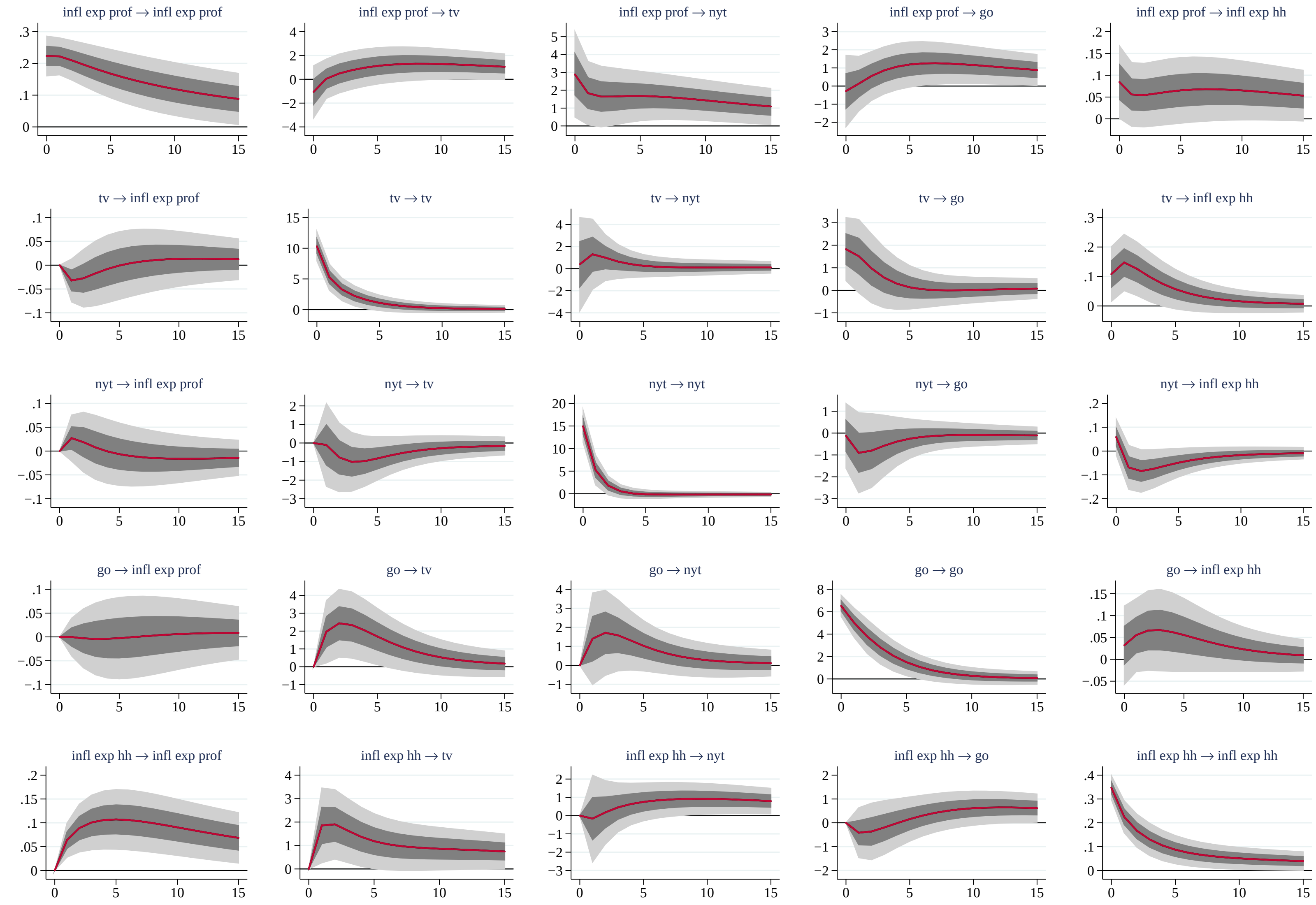
searches explain about 20% of the variation in households' inflation expectations, while newspaper articles as well contribute to 10% of the FEVD of expectations.

Summing up, the estimation of the three VAR models delivers a number of interesting insights. Starting with the interaction between the print media and television, the link is stronger from television news coverage to newspaper articles. Adding Google searches, we find that users indeed demand more information after having heard about inflation in the news while the effect lasts longer for TV news. At the same time, we find that news coverage is also affected by readers' and viewers' interest in inflation. Furthermore, our results suggest that experts' forecasts are indeed transmitted via the news media, there is no direct link from professional forecasters' expectations to Google searches or households' inflation expectations. Most interestingly, we find that households' adjust their expectations after having searched for information, while the opposite link is less strong.

In the previous section, we have found that the estimated parameters change over the sample period. Therefore, we have also applied the Chow test for unknown structural breaks to the different VAR estimates. Across the different model specifications, we find a break approximately in October 2008, i.e., one month after the collapse of Lehman Brothers. Overall, our results remain qualitatively the same if we exclude the financial crisis from the sample.[28] The positive and significant effect from Google search requests on household expectations is even found to be slightly stronger.

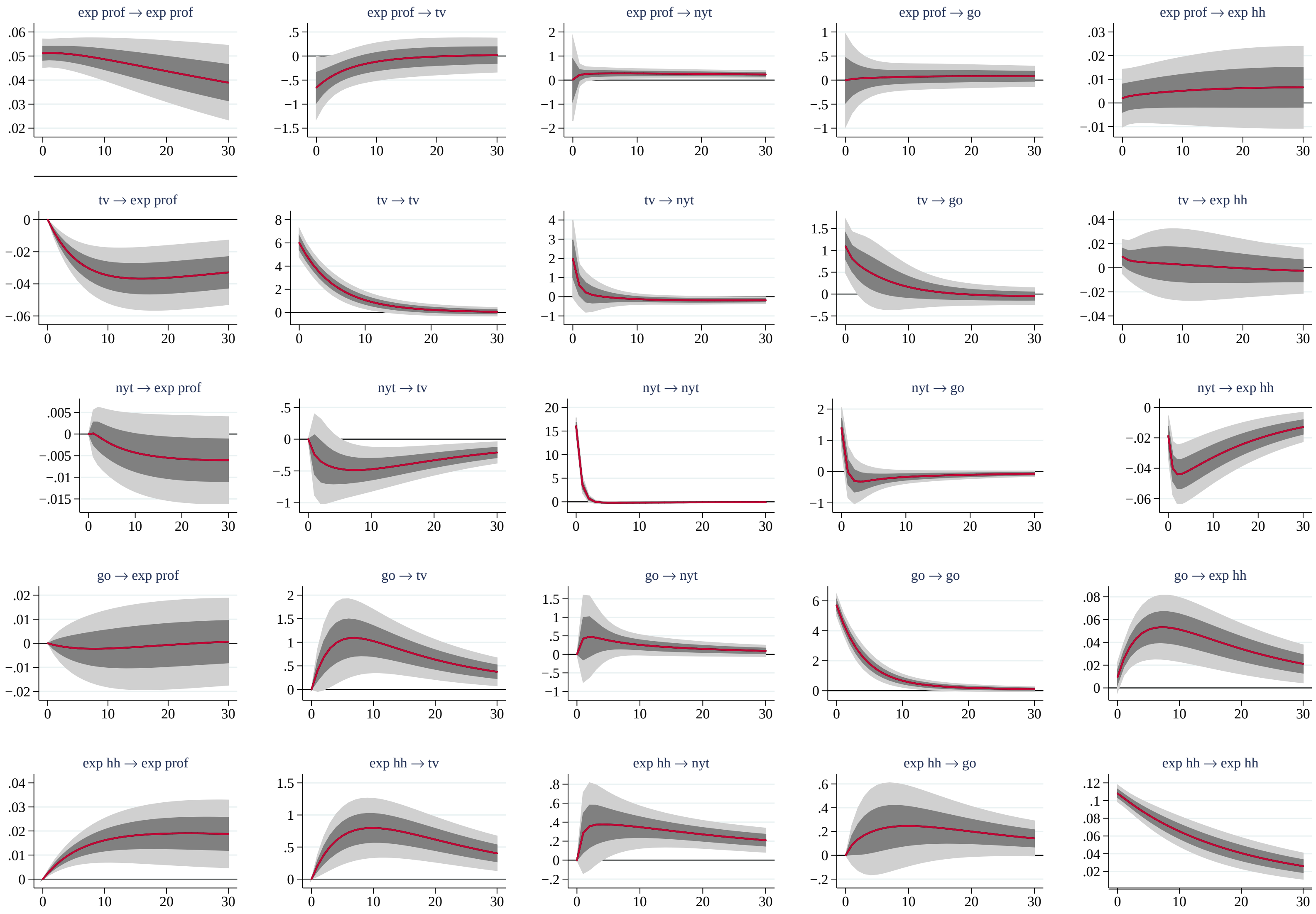
[28]The results of the restricted sample and of the structural break tests are not shown but are available upon request.

Figure 4.5: Large VAR: Impulse Response Functions - Monthly Data



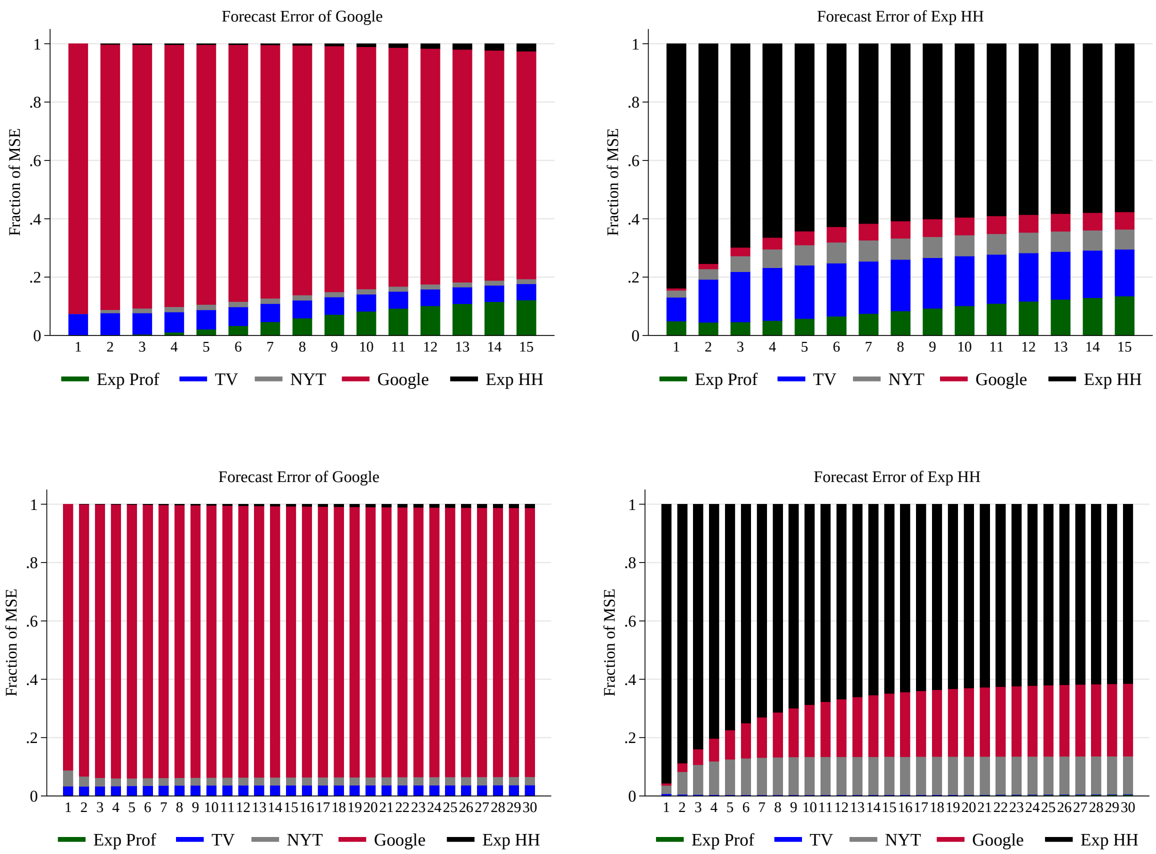
Response to one s.d. Cholesky innovation. Dark gray area: 66% confidence interval, light gray area: 95% confidence interval, based on 2,000 bootstrap replications. Sample 2005m1−2011m5.

Figure 4.6: Large VAR: Impulse Response Functions - Weekly Data



Response to one s.d. Cholesky innovation. Dark gray area: 66% confidence interval, light gray area: 95% confidence interval, based on 2,000 bootstrap replications. Sample 2005w1−2011w18.

Figure 4.7: Large VAR - FEVD - Monthly and Weekly Data



4.6 Conclusion

In this chapter, we have shown that Google data can serve as a valuable supplement to survey data on inflation expectations. Our empirical results indicate that Google search requests react in a robust and reasonable way to economic variables. Furthermore, our VAR estimates show that web queries account for a significant part of the forecast error variance of households' inflation expectations.

These findings provide the following insights for future research. First, one could expect that individuals' forecast error is lower in times of high search intensity following the epidemiology model of Carroll (2003) in which a rising number of media reports results in a closer link between households' and experts' expectations. As we have argued in this paper, Google search data can be used as a proxy for agents' demand for information and thus provide a more direct way to test whether households indeed adjust to the best available forecast in an economy. Second, it seems promising to explore the forecasting performance of Google search requests with regard to inflation and inflation expectations, a route that has so far only been followed by Guzmán (2011). The fact that the Google data is available on a weekly basis can be particular useful in this respect. The recently developed mixed frequency approaches such as the mixed data sampling regression models (MIDAS) by Ghysels et al. (2005, 2006) and the Mixed Frequency Bayesian VAR by Chiu et al. (2011) might be fruitfully applied in this area. Finally, more work should be done on how to best choose the keywords users have in mind when searching for inflation. In this regard, principal component analysis to merge searches for different subcategories as applied by Kholodilin et al. (2010) to nowcast private consumption might be promising.

Chapter 5

A Unifying Discussion

5.1 Summary and Interpretation of the Links between Media Reports and Inflation Expectations

The analysis conducted in this dissertation delivers various insights regard the links between media coverage and inflation expectations.

First, we have found supportive evidence that households form expectations according to the epidemiology model proposed by Carroll (2003). Households partly rely on experts and partly on their own past forecast when forming beliefs about future inflation. This relationship, however, is not stable over time. Households are more prone to adjust to experts in periods of low inflation and during economic crisis. The news effect also varies over time. Whereas in general, a rising number of news reports lowers the difference between households and experts, this effect mainly arises in times of disinflation. By contrast, more news coverage widens the gap during economic crisis.

Using both macro and micro level survey data, we typically find lower updating coefficients on the micro level, whereas the news effect, by contrast, is found to be *larger* on the micro level. This result might be due to the fact that the fraction of individuals which forms expectations similar to the prediction of experts is probably lower than the fraction that follows the news media. Still, this issue deserves a more detailed treatment taking into account interactions between different individuals.

Distinguishing households according to their self-reported information set results in the surprising observation that those who have perceived news on inflation are worse in forecasting inflation compared to other households. Moreover, a larger fraction of these households expect inflation rates that are above 15% or below -5%. Since the largest media effects are found for households who have heard some news or bad news about inflation it seems that the higher forecast errors of these groups stem from an overreaction to news media coverage.

Moreover, we find robust evidence for non-linear news effects. Households in general adjust more to the best available forecast if news coverage moves beyond its normal level where the adjustment is generally slow. Looking at different types of households, we find that those who have heard some news or bad news on inflation are much quicker in adjusting to rising news coverage. Moreover, these types of households seem to have higher attention levels as they already react more to experts if the general level of news coverage is still low.

Second, the results of the dissertation suggest that news coverage can explain why the expectations of households depend on their socioeconomic background. In line with previous findings in the literature, we observe that in Germany as well, inflation expectations are higher for households with low income, for young households and for the unemployed. Moreover, the same types of households show larger deviations from the best available forecast. We have shown that the higher expectation gaps of young and old households as well as the rising deviation with lower income levels can be explained by higher inflation rates of these groups, while no such effect can be observed for occupation groups. Across all household groups, inflation perceptions do not play a role in determining inflation expectations. With regard to the news media, we observe considerable heterogeneity in news consumption of different newspapers and TV news shows for income, age and occupation groups. It thus seems that media coverage offers some explanation on why households with a different socioeconomic background disagree on the future path of prices. Furthermore, we find that constructing an index of news reports by aggregating all available newspaper and TV reports can be misleading. Coverage of inflation in *Tagesschau*, Germany's most influential TV evening news show, is found to increase the gap between households and professional forecasters, while a rising number of articles published in *BILD*, Germany's most prominent tabloid, brings households closer to the best available forecast. Finally, it is important to distinguish between the effects of a rise in the number of news reports and a change in the journalists' judgment of inflation. Whereas households' expectation gaps increase if *BILD* presents inflation in a negative way thereby possibly inducing a media bias, more negative coverage in *Tagesschau* narrows the gap between households and professional forecasters.
These results raise important implications for communication strategies of central banks. If some household groups show systematic biases in inflation expectations and forecast errors, and if these differences are related to specific newspaper consumption, "the ideal communication strategy might then be multi-tiered" (Sims, 2009). Central bankers rarely appear on television, but if it is TV reports that systematically raise the forecasts of some household groups, this might be problematic. Furthermore, if some households rely more on their group-specific inflation rate instead of overall inflation, the credibility of the central bank might be undermined.

Finally, we have shown that the number of Google search requests for inflation can act as a useful complement to household survey data. Internet search data can be understood as a proxy for the demand of information in the sense that households will search more for inflation on the web if they need do know more about the current or future price environment. Internet search data could also serve as a complement to inflation expectations measured by surveys. Whereas surveys suffer from the "cheap talk"-problem arising from the fact that respondents do not have an incentive to provide their best forecast, households will only search for inflation if they really want to use this information. Our analysis shows that the number of Google search requests reacts in a meaningful way to fundamental economic data. Google users distinguish between headline and core inflation and they react asymmetrically: the demand for information increases if core inflation falls whereas in periods of historically high inflation rates, the number of search requests is significantly larger. Estimating various Vector Autoregressive Models, we find that households' inflation forecasts are driven by TV reports, newspaper articles, and Google search requests, while the feedback effect from expectations on web searches is rather small and estimated less precisely. About 20% of the forecast error variance decomposition of households' inflation expectations can be explained by Google search requests.

5.2 Limitations and Further Research

The analysis conducted in this dissertation is facing a number of caveats. Above all, it is important to keep in mind that both inflation expectations of households as well as the information content of news media coverage are unobservable variables. Professional forecasters construct econometric models and use expert judgment resulting in a precise quantitative estimate for future inflation. And in case of central banks, these estimates have a direct impact on policy decisions such as the setting of interest rates. As regards households, the picture is less clear-cut. We do not really know how economic agents arrive at the expected inflation rates which are provided in survey data or which kind of information they use in the process of expectation formation. Moreover, we cannot be sure whether survey responses are indeed the best proxy for households' beliefs on future inflation. As we have discussed in Chapter (4), surveys suffer from the "cheap talk"-problem and can be subject to wording and framing effects. In this respect, the use of Google search requests could serve as a promising complementary variable to measure agents' beliefs, however, internet data also has its limitations. In particular, we do not know whether users actually type "inflation" if they seek for information on future price changes, or whether they use more specific keywords when thinking about investing money or buying a particular product.

The measurement of the news media agenda also comes with some problems. Counting how many times the term "inflation" has appeared in a given newspaper seems to be fine as a first approximation. However, this variable neglects the varying size of newspapers so that less can be said about the relative attention inflation receives. In addition, as we have shown in Chapter (3), it is important not only to capture the amount of news coverage but also its content. Since available software has not yet proven to successfully detect the meaning of written articles, such data still has to be compiled by human resources making it expensive and less readily available. Finally, we would need more precise data on individual news consumption pattern. We have shown that news coverage of inflation by the German television channel *RTL* has a much larger effect on younger households, which corresponds to the fact that the average viewer of *RTL* is typically younger than viewers of *Tagesschau*. However, we know much less about the news preferences of different income, education, or occupation groups.

Based on the results proposed in this dissertation, a number of further research questions seem to be worth investigating.
First, we think that it is important to test whether households actually "act on their beliefs" (Armantier et al., 2012), i.e. whether the expected inflation rates stated in surveys affect consumption or saving decisions. Quantifying the link from beliefs to actions is important since macroeconomic models assign a prominent role to inflation expectations in explaining, for example, the zero lower bound (Eggertsson and Woodford (2003)) or the jobless recovery in the U.S. (Schmitt-Grohe and Uribe, 2013). Only a few studies have so far followed this research direction. Souleles (2004) has shown that consumer sentiment[1] can be useful in forecasting actual consumption spending. Combining a survey with an experiment, Armantier et al. (2012) offer evidence that agents rely on their stated inflation forecasts when deciding on future investments in the experiment. Finally, Bachmann et al. (2012) use inflation expectations from the Michigan survey and test whether these affect respondents' "readiness to spend", which is also captured by the survey.[2]. According to their analysis, higher expected inflation indeed leads to a fall in respondents' readiness to spend. However, Bachmann et al. (2012) do not show whether households' self-reported consumption plans are actually materialized.

Second, we think that more research could be done with respect to the policy implications of our results. For example, the analysis in Chapter (3) has shown that TV news have a much larger impact on households' inflation expectations than newspaper articles, and that different households rely on different news sources when forming expectations. It could be worth exploring whether announcements of central banks actually affect all kinds of households in a similar way, or whether some groups, by following different news media, are affected differently by policy statements than others. More broadly, the way expectations are formed can determine policy conclusions derived from macroeconomic models. In Fritsche et al.

[1]Consumer sentiment is typically measured by survey questions such as the one in the Michigan survey: "Now turning to business conditions in the country as a whole do you think that during the next 12 months we'll have good times financially, or bad times, or what?".

[2]Using the question "About the big things people buy for their homes - such as furniture, a refrigerator, stove, television, and things like that. Generally speaking, do you think now is a good or a bad time for people to buy major household items?"

(2014), we have shown that professional forecasters believe that central banks tolerate small deviations from the point inflation target by following a target zone. As a result, central banks do not have to change interest rates that much if inflation approaches the boundaries of the target zone, since professional forecasters already lower their inflation expectations in advance thereby dampening a possible acceleration of inflation. In this context, it could be worth exploring how an adjustment of the inflation target of central banks affects inflation expectations. In policy debates, it is now often argued that raising the official inflation target to 4% would be beneficial for the economy as a whole (Blanchard et al., 2010, Ball, 2013). However, it is unknown how households would react to such a policy change.

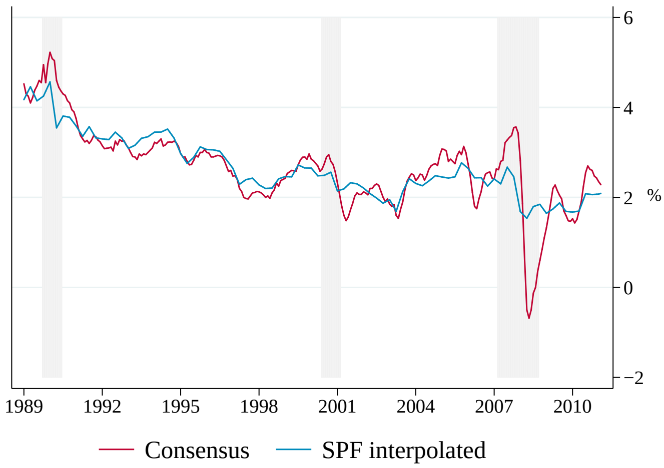
Third, the role of Google search requests might be successfully explored further. Since a number of studies have documented its very good forecasting performance with respect to consumption, it could be worth testing whether this property also extends to inflation. In addition, internet search data can be used as a proxy for households' inflation expectations, especially in countries that do not benefit from the existence of high quality survey data. This applies to Germany, in particular, where the only available survey on inflation expectations does not ask respondents to provide a quantitative estimate of future price changes, and where the underlying micro data are not accessible. Finally, the role of social interaction in determining inflation expectations should be investigated in more detail, not least because this can give rise to "information cascades" (Akerlof and Shiller, 2009). Since only a fraction of the population actively follows the news (Blinder and Krueger, 2004), information could be processed by social interaction, i.e. by the transfer of the news media agenda from "news-followers" to "un-followers". This research direction would also build the bridge to agenda setting theory stressing the need to consider both cognitive information processing that operates on the individual level, and developments in a broader social context such as group identification or social interaction.

Chapter A

Appendix to Chapter 2

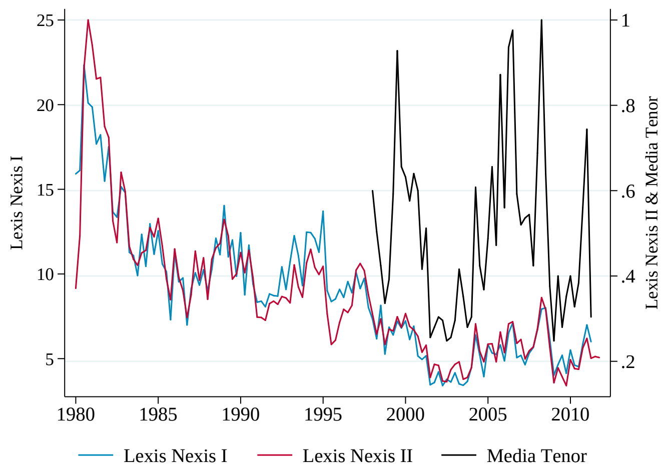
A.1 Data

Figure A.1: Professional Forecasters' Inflation Expectations: SPF and Consensus



Note: *Consensus* denotes the cross-sectional mean inflation forecast of professional forecasters surveyed by *Consensus Economics* each month, and *SPF interpolated* gives the interpolated quarterly forecast by the *Survey of Professional Forecasters*. The gray shaded areas denote NBER recessions.

Figure A.2: News Reports on Inflation



Note: The graph shows the number of articles published about inflation. Using *The New York Times* and *The Washington Post*, *Lexis Nexis I* is scaled by the total number of articles in each quarter, whereas *Lexis Nexis II* is scaled by the maximum number of articles on inflation published in any quarter of the sample. *Media Tenor* uses only articles published in *The New York Times*, and is also scaled by the maximum number of articles.

A.2 Unit Root and Cointegration Tests

Table A.1: Unit Root Tests

	test statistic				critical values		
	$\pi_t^{exp,hh}$	$\pi_t^{exp,prof}$	π_t	$Media_t$	0.1	0.05	0.01
DFG-GLS							
MAIC1	0.214	0.334	-0.010	-3.776	-1.667	-1.985	-2.580
BIC1	-0.067	0.130	-0.010	-3.739	-1.675	-1.994	-2.580
seqt1	0.382	0.334	-0.010	-4.060	-1.659	-1.976	-2.580
DFG-GLS-trend							
MAIC2	-0.728	-1.018	-0.948	-3.919	-2.587	-2.873	-3.480
BIC2	-1.061	-1.267	-0.948	-3.901	-2.601	-2.889	-3.480
seqt2	-0.611	-1.018	-0.948	-4.172	-2.557	-2.841	-3.480
PPerron-t	-4.682	-3.075	-3.694	-3.761	-2.570	-2.875	-3.449
PPerron-rho	-23.338	-6.034	-14.157	-26.722	-11.200	-14.000	-20.406
PPerron-trend-t	-4.763	-2.811	-3.388	-7.399	-3.130	-3.425	-3.985
PPerron-trend-rho	-29.834	-10.500	-16.070	-90.877	-18.053	-21.406	-28.666
KPSS	1.128	1.912	1.046	1.755	0.347	0.463	0.739
KPSS-trend	0.289	0.267	0.216	0.182	0.119	0.146	0.216
ZAndrews-const	-6.183	-4.063	-4.127	-6.317	.	-4.800	-5.430
ZAndrews-trend	-6.012	-3.748	-4.015	-9.632	.	-4.420	-4.930
ZAndrews-both	-6.166	-3.977	-4.369	-9.913	.	-5.080	-5.570
Clemente-Montanes-Reyes							
AO2	-4.028	-3.270	-4.033	-4.545	.	-5.490	.
AO1	-3.438	-2.266	-3.252	-3.655	.	-3.560	.
IO2	-6.491	-5.751	-5.409	-4.756	.	-5.490	.
IO1	-5.606	-5.328	-5.398	-3.785	.	-4.270	.

Note: All tests employ the null hypothesis of non-stationarity, with the exception of the KPSS-test. See Section (A.6.2) for further details. The critical values vary marginally depending on the number of lags included in the test equations of the different variables. However, this does not affect the results. Sample: 1980:01-2011:11.

Table A.2: Cointegration Tests I: Inflation Expectations

	ADF-Test		Gregory-Hansen-Test			
	const	no const	level	trend	regime	regimetrend
$\pi_t^{exp,hh}, \pi_t^{exp,prof}$	-4.172	-2.277	-7.912	-8.341	-7.895	-8.613
Critical Values Phillips and Ouliaris (1990)			Critical Values			
0.10	-3.066	-2.451	-4.340	-4.720	-4.680	-5.240
0.05	-3.365	-2.762	-4.610	-4.990	-4.950	-5.500
0.01	-3.962	-3.387	-5.130	-5.450	-5.470	-6.020
Critical Values: MacKinnon (2010)						
0.10	-3.045	-1.617	.	.	.	.
0.05	-3.336	-1.941	.	.	.	.
0.01	-3.896	-2.566	.	.	.	.

Note: Cointegrating regression: $\pi_t^{exp,hh} = \alpha_1 + \alpha_2 \pi_t^{exp,prof} + \varepsilon_t$. Both tests apply the null hypothesis of no cointegration. The table shows the ADF version of the Gregory-Hansen test, where the optimal lag length is chosen with the BIC criterion. See Section (A.6.2) for further details. Sample: 1980:01-2011:11.

Table A.3: Cointegration Tests II: Expectations and Inflation

	ADF-Test	Gregory-Hansen-Test			
		level	trend	regime	regimetrend
$\pi_t^{exp,hh}, \pi_t^{exp,prof}, \pi_t$	-4.840	-8.196	-7.931	-8.237	-8.148
Critical Values Phillips and Ouliaris (1990)		Critical Values			
0.10	-3.449	-4.690	-5.030	-5.230	-5.720
0.05	-3.768	-4.920	-5.290	-5.500	-5.960
0.01	-4.308	-5.440	-5.800	-5.970	-6.450
Critical Values: MacKinnon (2010)					
0.10	-3.452	.	.	.	.
0.05	-3.741	.	.	.	.
0.01	-4.294	.	.	.	.

Note: Cointegrating regression: $\pi_t^{exp,hh} = \alpha_1 + \alpha_2 \pi_t^{exp,prof} + \alpha_3 \pi_t + \varepsilon_t$. Both tests apply the null hypothesis of no cointegration. The table shows the ADF version of the Gregory-Hansen test, where the optimal lag length is chosen with the BIC criterion. See Section (A.6.2) for further details. Sample: 1980:01-2011:11.

A.3 Additional Results: The Epidemiology Model Without News

Table A.4: Test of Structural Breaks - Model without News - Aggregate Data

	Models					
Estimated Sample	base	base rec	π^{CPI}	π^{CORE}	π^{CPI} rec	π^{CORE} rec
1980m1-2011m11/2009m6	2003m7	1993m1	2007m2	2003m7	2004m9	1993m1
1980m1-1st break	1989m12	1981m12	1993m1	1989m12	1993m1	1981m12
1st break - 2011m11/2009m6	2008m2	2007m2	2003m7	2008m2	none	2003m7

Note: Structural Breaks suggested by QLR test of equations (2.6), (2.6a) and (2.6b). Rejection of the null of no break at least at the 5%-level.

Table A.5: Results: Aggregate Data - Including Core Inflation

	80/1-11/11	80/1-89/11	89/12-03/6	03/7-08/1	08/2-11/11	80/1-11/11	80/1-81/11	81/12-92/12	93/1-01/2	01/3-09/6
$\pi_t^{exp,prof}$	0.23***	0.60***	0.66***	0.76	2.00***	0.26***	2.72**	0.62***	0.67***	1.47***
	(0.06)	(0.14)	(0.12)	(0.52)	(0.29)	(0.08)	(1.33)	(0.12)	(0.17)	(0.33)
$\pi_{t-1}^{exp,hh}$	0.76***	0.53***	0.51***	0.53***	0.57***	0.68***	0.52*	0.26***	0.41***	0.41***
	(0.03)	(0.08)	(0.07)	(0.13)	(0.08)	(0.05)	(0.28)	(0.09)	(0.09)	(0.12)
π_{t-1}^{CORE}	-0.02	-0.02	-0.11*	-0.12	-0.43***	0.01	-0.21	-0.05	0.02	-0.40**
	(0.04)	(0.06)	(0.06)	(0.23)	(0.09)	(0.06)	(0.38)	(0.10)	(0.13)	(0.18)
$\pi_t^{exp,prof} rec$						0.14***	0.22***	0.08	-0.08	0.29*
						(0.03)	(0.08)	(0.06)	(0.12)	(0.17)
$\pi_{t-1}^{exp,hh} rec$						0.04	-0.38	0.20*	0.32**	0.31*
						(0.05)	(0.34)	(0.11)	(0.16)	(0.16)
$\pi_{t-1}^{CORE} rec$						-0.12**	0.30	-0.18	-0.38**	-0.56
						(0.05)	(0.33)	(0.11)	(0.17)	(0.34)
$cons$	0.32***	-0.46	0.18	0.27	-1.45***	0.38***	-16.14**	0.89**	0.21	-0.37
	(0.10)	(0.30)	(0.15)	(0.61)	(0.30)	(0.13)	(7.99)	(0.45)	(0.21)	(0.43)
R^2	0.89	0.91	0.77	0.52	0.88	0.90	0.66	0.69	0.75	0.78
N	383	119	163	55	46	354	23	133	126	72
Wald	0.192	0.062	0.181	0.538	0.000	0.747	0.045	0.493	0.642	0.064

Note: HAC standard error in parentheses. Wald denotes p-value of the test $H_0 : \alpha_1 + \alpha_2 + \alpha_5 = 1$, $H_0 : \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 1$, and $H_0 : \alpha_1 + ... + \alpha_6 = 1$. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A.6: Results: Micro Data I

	80/1-11/11	80/1-89/11	89/12-03/6	03/7-08/1	08/1-11/11	80/1-11/11	80/1-89/11	89/12-03/6	03/7-08/1	08/2-11/11
$\pi_t^{exp,prof}$	-0.01	0.31***	0.24***	0.42***	1.71***	0.08***	0.42***	0.24***	0.35**	1.90***
	(0.02)	(0.04)	(0.05)	(0.13)	(0.17)	(0.02)	(0.04)	(0.05)	(0.16)	(0.20)
$\pi_{t-1}^{exp,hh}$	0.72***	0.49***	0.56***	0.61***	0.62***					
	(0.01)	(0.03)	(0.04)	(0.05)	(0.05)					
$cons$	2.01***	1.18***	1.77***	0.47	-1.66***	1.46***	0.17	1.33***	0.61	-1.58***
	(0.12)	(0.23)	(0.19)	(0.34)	(0.46)	(0.16)	(0.30)	(0.25)	(0.44)	(0.58)
R^2	0.08	0.08	0.04	0.03	0.06	0.09	0.10	0.04	0.03	0.07
N	184886	68408	72276	23979	20223	111494	41106	41485	14375	14528
Wald	0.000	0.000	0.000	0.789	0.000	0.000	0.597	0.015	0.757	0.000

Note: Standard error in parentheses. Wald denotes p-value of the test $H_0 : \alpha_1 + \alpha_2 = 1$. Demographic control variables such as gender, education, race, and age are included in each regression. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A.7: Results: Micro Data II

	80/1-11/11	80/1-89/11	89/12-03/6	03/7-08/1	08/2-11/11	80/1-11/11	80/1-89/11	89/12-03/6	03/7-08/1	08/1-11/11	80/1-11/11	80/1-89/11	89/12-03/6	03/7-08/1	08/2-11/11
$\pi_t^{exp,prof}$	0.27***	0.95***	0.80***	-0.44	2.00***	0.49***	0.67***	0.93***	-1.04	2.57	0.47***	1.06***	0.79***	0.05	1.99***
	(0.04)	(0.12)	(0.17)	(0.46)	(0.47)	(0.07)	(0.11)	(0.32)	(1.01)	(1.97)	(0.05)	(0.13)	(0.19)	(0.51)	(0.43)
$\bar{\pi}_{t-1}^{exp,NINFL}$	0.57***	0.26***	0.22***	0.30***	0.48***										
	(0.03)	(0.08)	(0.07)	(0.09)	(0.13)										
$\bar{\pi}_{t-1}^{exp,NGOOD}$						0.21***	0.18**	0.21**	-0.10	0.12					
						(0.05)	(0.08)	(0.10)	(0.14)	(0.19)					
$\bar{\pi}_{t-1}^{exp,NBAD}$											0.35***	0.13	0.09	0.15	0.37***
											(0.04)	(0.08)	(0.08)	(0.11)	(0.12)
$cons$	1.69***	-2.35**	3.24***	4.31***	-0.97	1.31	-1.02	1.71	7.25**	5.75	2.43***	-2.01	4.76***	3.90**	-1.15
	(0.56)	(1.01)	(1.21)	(1.44)	(1.75)	(0.91)	(1.22)	(2.02)	(3.24)	(6.76)	(0.68)	(1.39)	(1.46)	(1.60)	(1.85)
R^2	0.11	0.15	0.07	0.03	0.07	0.07	0.10	0.07	0.07	0.17	0.10	0.12	0.06	0.03	0.06
N	10820	4363	2303	2625	1529	2670	1622	576	316	156	8165	2780	1709	2304	1372
Wald	0.000	0.000	0.875	0.015	0.000	0.000	0.080	0.646	0.037	0.393	0.000	0.014	0.489	0.126	0.001

Note: Standard error in parentheses. Wald denotes p-value of the test $H_0 : \alpha_1 + \alpha_2 = 1$. Demographic control variables such as gender, education, race, and age are included in each regression. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

A.4 Additional Results: Including News I - Expectation Gaps

Table A.8: QLR-Test of Structural Breaks - GAP Regressions - Aggregate Data

	Models					
Estimated Sample	base	base rec	π^{CPI}	π^{CORE}	π^{CPI} rec	π^{CORE} rec
1980m1-2011m11/2009m6	2007m2	2004m3	2007m1	2005m8	2004m3	1991m10
1980m1-1st break date	1984m1	1983m3	1993m1	1984m8	1990m7	1981m9
1st break - 2nd break	2003m7	1993m1	2003m8	2000m9	1993m1	2000m10

Note: Structural breaks suggested by QLR test, rejection of the null of no break at least at the 5%-level.

Table A.9: Results: Expectation Gaps: Aggregate Data II

	80/1-11/11	80/1-83/12	84/1-03/6	03/7-07/1	07/2-11/11	80/1-09/6	80/1-83/2	83/3-92/12	93/1-04/2	04/3-09/6
$MEDIA_t$	-2.58	-7.14	-0.99*	12.92*	43.33***	-3.65*	-11.67**	0.34	0.12	30.41***
	(2.08)	(5.14)	(0.53)	(6.60)	(8.88)	(1.86)	(5.09)	(0.40)	(0.93)	(8.86)
$MEDIA_t rec$						5.02***	-1.39	3.60***	1.80	13.53***
						(1.56)	(3.11)	(0.29)	(1.75)	(2.97)
$cons$	2.39***	6.47*	0.91***	-0.72	-4.82**	2.24***	10.89***	0.10	0.74***	-4.71**
	(0.81)	(3.43)	(0.19)	(1.45)	(1.95)	(0.71)	(3.62)	(0.16)	(0.28)	(2.03)
R^2	0.02	0.09	0.02	0.11	0.41	0.10	0.20	0.52	0.02	0.50
N	383	48	234	43	58	354	38	118	134	64

Note: HAC standard error in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table A.10: Results: Expectation Gaps - Micro Data II

	all households					news heard				
	80/1-11/11	80/1-83/12	84/1-03/6	03/7-07/1	07/2-11/11	80/1-11/11	80/1-83/12	84/1-03/6	03/7-07/1	07/2-11/11
$MEDIA_t$	51.06***	12.59***	34.27***	36.46***	22.90***	45.78***	9.88***	23.99***	44.01***	26.93***
	(1.17)	(3.08)	(2.61)	(6.07)	(6.85)	(1.39)	(3.65)	(3.10)	(7.45)	(8.12)
$cons$	33.50***	64.91***	41.45***	21.39***	33.07***	39.04***	69.13***	50.74***	20.44***	38.09***
	(1.92)	(6.22)	(2.58)	(3.24)	(4.58)	(2.46)	(7.64)	(3.29)	(4.30)	(5.89)
R^2	0.04	0.03	0.03	0.02	0.01	0.04	0.04	0.03	0.03	0.01
N	184886	29360	111324	18742	25460	111494	19438	63153	11162	17741

	news inflation					good news					bad news				
	80/1-11/11	80/1-83/12	84/1-03/6	03/7-07/1	07/2-11/11	80/1-11/11	80/1-83/12	84/1-03/6	03/7-07/1	07/2-11/11	80/1-11/11	80/1-83/12	84/1-03/6	03/7-07/1	07/2-11/11
$MEDIA_t$	49.57***	6.35	42.35***	26.30	89.04***	44.04***	47.48***	4.82	5.55	25.06	53.70***	-6.54	65.58***	22.93	94.70***
	(4.39)	(9.44)	(14.46)	(22.08)	(22.29)	(7.53)	(15.70)	(18.54)	(41.62)	(78.98)	(5.27)	(11.73)	(19.40)	(24.30)	(23.98)
$cons$	59.91***	114.54***	70.86***	33.68***	21.83	43.75***	52.37*	75.02***	15.36	12.60	67.21***	135.84***	65.70***	38.16***	19.84
	(8.59)	(21.81)	(14.97)	(12.57)	(17.85)	(13.10)	(29.56)	(17.45)	(21.39)	(54.85)	(10.68)	(28.77)	(20.87)	(14.01)	(18.98)
R^2	0.06	0.06	0.06	0.04	0.03	0.06	0.09	0.05	0.10	0.05	0.06	0.05	0.06	0.04	0.03
N	10820	2496	4170	2013	2141	2711	767	1457	276	211	8171	1749	2746	1742	1934

Note: Standard error in parentheses. Demographic control variables such as gender, income, education, race, and age are included in each regression. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

A.5 Additional Results: STAR Results for Different Subsamples

Table A.11: STAR Results - All Households

	80/1-11/11	80/1-92/12	93/1-03/7	03/8-06/12	07/1-11/11
			linear part		
$cons$	0.62***	0.10	0.23	0.33	-0.05
	(0.11)	(0.27)	(0.19)	(0.73)	(0.41)
$\pi_t^{exp,prof}$	0.14***	0.50***	0.41*	0.47	0.57**
	(0.04)	(0.10)	(0.22)	(0.39)	(0.24)
$\pi_{t-1}^{exp,hh}$	0.71***	0.40***	0.79***	0.60***	0.72***
	(0.05)	(0.08)	(0.18)	*'(0.15)	(0.08)
π_{t-1}	0.04	0.13**	-0.44***	-	-
	(0.03)	(0.05)	(0.15)		
			nonlinear part		
$consG$	-	-			
		-	-	-	-
$\pi_t^{exp,prof}G$	0.52**	0.36*	0.15	2.60**	1.29***
	(0.22)	(0.19)	(0.23)	(1.09)	(0.47)
$\pi_{t-1}^{exp,hh}G$	-0.89***	-0.62***	-0.42**	-1.54**	-0.45**
	(0.26)	(0.19)	(0.23)	(0.68)	(0.03)
$\pi_{t-1}G$	0.35***	0.20*	0.64***	-	-
	(0.11)	(0.11)	(0.16)		
γ	2.13*	113.53	4363.66	11.61	10.60
	(1.24)	(1.30E+06)	(2.81E+14)	(24.75)	(20.03)
c_1	0.60***	0.63	0.17	0.18***	0.30***
	(0.09)	(428.78)	(7.74E+07)	(0.01)	(0.02)
c_2	-	-	-	0.31***	-
				(0.01)	
AIC	-1.42	-1.22	-2.21	-1.47	-1.61
SBIC	-1.33	-1.05	-2.01	-1.14	-1.37
HQ	-1.39	-1.15	-2.13	-1.35	-1.52
$R^2 adj.$	0.91	0.93	0.80	0.58	0.84
Linearity Test					
w.r.t. $NEWS_t$	0.000	0.021	0.018	0.382	0.138
w.r.t. $TIME$	0.000	0.001	0.125	0.332	0.000

Note: Standard errors in brackets. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Values of linearity tests are p-values derived from the F-statistic.

Table A.12: STAR Results - News about Economic Issues in General

	80/1-11/11	80/1-92/12	93/1-03/7	03/8-06/12	07/1-11/11
			linear part		
$cons$	0.57***	-0.21	-0.01	0.83	-0.31
	(0.10)	(0.25)	(0.25)	(0.67)	(0.52)
$\pi_t^{exp,prof}$	0.08**	0.55***	0.48***	0.65	0.69**
	(0.03)	(0.12)	(0.13)	(0.40)	(0.32)
$\pi_{t-1}^{exp,NH}$	0.74***	0.53***	0.62***	0.33*	0.70***
	(0.04)	(0.07)	(0.08)	(0.17)	(0.10)
π_{t-1}	0.04	-	-	-	-
	(0.03)				
			nonlinear part		
$consG$	-	-1.47	-	-	-
		(1.04)			
$\pi_t^{exp,prof}G$	2.61****	0.74***	0.67***	-1.13*	1.31**
	(0.91)	(0.27)	(0.24)	(0.66)	(0.63)
$\pi_{t-1}^{exp,NH}G$	-2.60***	-0.52***	-0.64***	0.77*	-0.47*
	(0.91)	(0.18)	(0.21)	(0.40)	(0.27)
$\pi_{t-1}G$	-	-	-	-	-
γ	13.56*	17.96	710.78	224.32	5.18
	(8.18)	(51.91)	(4.20E+09)	(2.55E+06)	(7.71)
c_1	0.90***	0.62***	0.35	0.26	0.29***
	(0.01)	(0.08)	(2.86E+04)	(0.61)	(0.02)
AIC	-1.39	-1.14	-1.90	-1.47	-1.15
SBIC	-1.31	-0.98	-1.74	-1.17	-0.90
HQ	-1.36	-1.07	-1.83	-1.36	-1.05
$R^2 adj.$	0.88	0.92	0.68	0.60	0.81
Linearity Test					
w.r.t. $NEWS_t$	0.000	0.000	0.019	0.576	0.222
w.r.t. $TIME$	0.000	0.000	0.059	0.210	0.000

Note: Standard errors in brackets. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Values of linearity tests are p-values derived from the F-statistic.

Table A.13: STAR Results - News about Inflation

	80/1- 11/11	80/1- 92/12	93/1- 03/7	03/8- 06/12	07/1- 11/11
			linear part		
$cons$	1.71***	-0.42	0.06	0.16	-0.25
	(0.31)	(0.38)	(1.20)	(2.68)	(1.19)
$\pi_t^{exp,prof}$	-0.21	1.19***	0.59	3.36**	2.06**
	(0.22)	(0.21)	(0.55)	(1.43)	(0.86)
$\pi_{t-1}^{exp,NINFL}$	0.47***	0.10	0.45***	-0.80***	0.15
	(0.09)	(0.14)	(0.14)	(0.25)	(0.20)
π_{t-1}	0.29***	-	-	-0.02	-
	(0.07)			(0.22)	
			nonlinear part		
$consG$	-	-0.04	-	1.77	-
		(0.18)		(3.19)	
$\pi_t^{exp,prof}G$	0.40**	-0.17	0.55*	-3.28**	-1.10
	(0.17)	(0.23)	(0.28)	(1.57)	(0.66)
$\pi_{t-1}^{exp,NINFL}G$	-0.27***	-0.04	-0.31*	1.27***	0.46*
	(0.10)	(0.18)	(0.18)	(0.32)	(0.27)
$\pi_{t-1}G$	-	-	-	-	-
γ	35.60	69.44	2114.35	127.41	102.52
	(42.91)	(131.07)	(3.16E+08)	(2.83E+04)	(418.94)
c_1	0.28***	0.32***	0.27	0.18**	0.18***
	(0.01)	(0.01)	(180.18)	(0.07)	(0.003)
AIC	0.76	0.58	1.17	-0.22	0.26
SBIC	0.85	0.71	1.33	0.16	0.51
HQ	0.80	0.63	1.23	-0.08	0.36
$R^2adj.$	0.46	0.64	0.22	0.44	0.38
Linearity Test					
w.r.t. $NEWS_t$	0.177	0.073	0.647	0.084	0.568
w.r.t. $TIME$	0.000	0.012	0.039	0.026	0.240

Note: Standard errors in brackets. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Values of linearity tests are p-values derived from the F-statistic.

Table A.14: STAR Results - Good News about Inflation

	80/1- 11/11	80/1- 92/12	93/1- 03/7	03/8- 06/12	07/1- 11/11
			linear part		
$cons$	0.57**	-0.92	-0.29	-1.50	5.22
	(0.25)	(1.20)	(1.20)	(2.31)	(3.81)
$\pi_t^{exp,prof}$	0.52***	0.20	0.20	1.75	-2.09
	(0.08)	(0.57)	(0.57)	(1.04)	(2.07)
$\pi_{t-1}^{exp,NGOOD}$	0.20***	0.55***	0.55***	-0.02	0.03
	(0.05)	(0.12)	(0.12)	(0.17)	(0.17)
π_{t-1}		0.31	0.31	-	-
		(0.29)	(0.29)		
			nonlinear part		
$consG$	75.29***	-	-	-	-4.66
	(24.72)				(5.07)
$\pi_t^{exp,prof}G$	-8.86***	0.55**	0.55**	-1.37	3.03
	(2.91)	(0.22)	(0.22)	(4.55)	(2.61)
$\pi_{t-1}^{exp,NGOOD}G$	0.24	-0.58***	-0.58***	2.17	0.32
	(0.38)	(0.17)	(0.17)	(5.01)	(0.26)
$\pi_{t-1}G$	-	-	-	-	-
γ	590.96	911.21	911.21	29.14	3329.28
	(3.10E+08)	(3.68E+09)	(3.68E+09)	(104.80)	(4.92E+17)
c_1	0.89	0.27	0.27	0.31***	0.203
	(4.46E+03)	(1.49E+04)	(1.49E+04)	(0.01)	(1.78E+11)
AIC	1.31	1.21	1.21	0.97	2.04
SBIC	1.39	1.39	1.39	1.26	2.32
HQ	1.35	1.28	1.28	1.08	2.15
$R^2 adj.$	0.32	0.25	0.25	0.25	0.24
Linearity Test					
w.r.t. $NEWS_t$	0.924	0.500	0.427	0.445	0.544
w.r.t. $TIME$	0.329	0.052	0.001	0.254	0.276

Note: Standard errors in brackets. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Values of linearity tests are p-values derived from the F-statistic.

Table A.15: STAR Results - Bad News about Inflation

	80/1-11/11	80/1-92/12	93/1-03/7	03/8-06/12	07/1-11/11
			linear part		
$cons$	2.09***	1.47	1.69	1.48	-36.78***
	(0.36)	(1.18)	(1.66)	(3.00)	(13.37)
$\pi_t^{exp,prof}$	0.42***	0.49	0.90	2.44*	25.63***
	(0.15)	(0.411)	(0.68)	(1.42)	(8.79)
$\pi_{t-1}^{exp,NBAD}$	0.21***	0.41**	-0.01	0.57**	-0.58
	(0.05)	(0.16)	(0.11)	(0.28)	(0.48)
π_{t-1}	0.14*	-0.02	-	-	-0.14
	(0.08)	(0.14)			(0.14)
			nonlinear part		
$consG$	-	-1.55)	-	1.91	40.07***
		(1.16)		(3.59)	(13.49)
$\pi_t^{exp,prof}G$	0.35*	0.78**	0.00	-2.53	-24.88***
	(0.20)	(0.35)	(0.40)	(1.69)	(8.86)
$\pi_{t-1}^{exp,NBAD}G$	-0.34**	-0.48***	0.24	0.79**	0.71
	(0.16)	(0.18)	(0.19)	(0.35)	(0.50)
$\pi_{t-1}G$	-	-	-	-	-
γ	1004.50	1374.64	103.60	129.58	345.82
	(2.63E+05)	(6.89E+05)	(284.98)	(4.25E+04)	(3.31E+04)
c_1	0.46***	0.32*	0.34***	0.18**	0.15***
	(0.03)	(0.19)	(0.001)	(0.08)	(0.02)
AIC	1.38	1.21	1.79	-0.03	0.90
SBIC	1.46	1.39	1.95	0.31	1.22
HQ	1.41	1.28	1.85	0.09	1.03
$R^2 adj.$	0.29	0.49	0.15	0.22	0.27
Linearity Test					
w.r.t. $NEWS_t$	0.618	0.140	0.828	0.512	0.170
w.r.t. $TIME$	0.000	0.094	0.444	0.161	0.069

Note: Standard errors in brackets. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Values of linearity tests are p-values derived from the F-statistic.

A.6 Theoretical Background

A.6.1 Derivation of Carroll's Equation

Carroll (2001, 2003, 2005) derives a link between media reports about inflation and individuals' inflation expectations using an epidemiology model that describes the spread of a disease by infection. The idea behind this analogy is that information about inflation disseminates gradually through the economy, since only some agents receive new information immediately whereas it takes some time until all agents are infected, i.e., well-informed. Individuals receive information, or get infected, by reading newspaper articles.
Denoting the newly infected individuals in a given period t by I_t, the individuals that are susceptible to be infected as S_t, and the fixed probability that those latter individuals actually get the information as λ, we have:

$$I_t = \lambda S_t \tag{A.1}$$

Next, one has to determine how individuals get susceptible. In epidemiology models used in public health, this depends on the fraction of people who have already caught the disease: the more individuals are infected, the larger the probability that others catch the disease as well. However, referring to the "Legionnaire's disease" which is transmitted by a common source of infection, the air conditioning system of hotels, Carroll (2005) assumes a simpler framework. Every individual has a fixed probability of reading a newspaper report, no matter of how many individuals have already read the newspaper.
This being said, the dynamics of the spread of information through the economy are as follows. In the first period, a fraction λ receives new information and updates their expectations, leaving $1-\lambda$ individuals who stick to their expectations build in the previous period. In the next period, again a fraction λ of the remaining uninformed individuals $1-\lambda$ gets the information, leading to $\lambda(1-\lambda)$ of newly infected individuals. In total, after two periods, $\lambda+\lambda(1-\lambda)$ individuals will be infected, leaving $1-[\lambda+\lambda(1-\lambda)] = 1-2\lambda+\lambda^2 = (1-\lambda)^2$ of uninfected individuals. This leads to $\lambda(1-\lambda)^2$ newly infected individuals in period 3, and so on. Hence, one can write the total fraction of informed individuals over time as:

$$\begin{aligned} \text{fraction infected} &= \lambda + \lambda(1-\lambda) + \lambda(1-\lambda)^2 + ... + \lambda(1-\lambda)^t \\ &= \lambda \sum_{s=0}^{t} \frac{1-(1-\lambda)^{s+1}}{\lambda} \end{aligned} \tag{A.2}$$

If $t \to \infty$, fraction infected $= \lambda/\lambda = 1$, hence, all agents will be infected as time goes by.[1] Next, Carroll (2005) adds the news coverage on inflation to his model. Assuming that newspapers print identical and complete forecasts, he interprets the parameter λ as the fixed probability of an individual to read an newspaper article about inflation and subsequently adopt the recent forecast on future inflation contained in this article. Then, the remaining fraction $1-\lambda$ of individuals sticks to the latest information they have received in the past. Hence, denoting the operator that yields the cross-sectional mean of inflation expectations with M_t, and the newspaper forecast with N_t, one can rewrite equation (A.2) for the first period as

$$M_t[\pi_{t+1}] = \lambda N_t[\pi_{t+1}] + (1-\lambda)N_{t-1}[\pi_{t+1}] \tag{A.3}$$

can be derived as follows To start, note that equation (A.3) can be decomposed further. Of the fraction $(1-\lambda)$ who did not receive new information in period t, part of this fraction had received new information in the previous period $t-1$, whereas the other part did not. Hence,

$$N_{t-1}[\pi_{t+1}] = \lambda N_{t-1}[\pi_{t+1}] + (1-\lambda)N_{t-2}[\pi_{t+1}] \tag{A.4}$$

Using this in equation (A.3) yields[2,3]

$$\begin{aligned} M_t[\pi_{t+1}] &= \lambda N_t[\pi_{t+1}] + (1-\lambda)\left[\lambda N_{t-1}[\pi_{t+1}] + (1-\lambda)N_{t-2}[\pi_{t+1}]\right] \\ &= \lambda N_t[\pi_{t+1}] + (1-\lambda)\left[\lambda N_{t-1}[\pi_{t+1}] + (1-\lambda)\{\lambda N_{t-2}[\pi_{t+1} + (1-\lambda)N_{t-3}[\pi_{t+1}] + ...\}\right] \end{aligned} \tag{A.5}$$

Next, Carroll has to make some adjustments, given that so far, $M_t[\pi_{t+1}]$ describes today's forecast of inflation in the next period. While this might be justified by the fact that in the U.S., newspapers mostly mention the period by period change in the price level (Branch, 2004), households participating in surveys on inflation expectations are typically asked to state their estimate for inflation in the next year. Carroll thus moves on by assuming that

[1] To get the second line of equation (A.2), write the first line as frac inf $= \lambda\left[1 + (1-\lambda) + (1-\lambda)^2 + ... + (1-\lambda)^t\right]$, multiply by $(1-\lambda)$ to get $(1-\lambda)$frac inf $= \lambda\left[(1-\lambda) + (1-\lambda)^2 + (1-\lambda)^3 + ... + (1-\lambda)^{t+1}\right]$, and subtract the latter expression from the former one. This yields frac inf $= \lambda\frac{1-(1-\lambda)^{t+1}}{1-(1-\lambda)}$.

[2] Carroll (2005) mentions a similar formulation by Roberts (1997, 1998), who has used past realizations of the inflation rate instead of forecasts made in the past. Note that the formulation by Roberts either requires adaptive expectations or the assumption that newspapers only report the recent inflation rate without mentioning forecasts of future price developments.

[3] Carroll implicitly uses forecasts made in the past of tomorrow's inflation rate, rather than forecasts made today about tomorrow's inflation rate using past information. The latter formulation leads to the one by Mankiw and Reis (2007).

1. People believe that at any point in time, there exists a "fundamental inflation rate" π_t^f.

2. People think that the fundamental inflation rate cannot be forecasted by more than one period, but follows a random walk. Permanent shocks to the fundamental inflation rate are denoted with ν_t

3. People think that the actual inflation rate is the sum of fundamental inflation and a transitory shock ε_t.

4. At period t, ν_{t+2} and ε_{t+1} cannot be forecasted.

Then, the actual inflation rate π_t becomes

$$\pi_t = \pi_t^f + \varepsilon_t, \tag{A.6}$$

where fundamental inflation is given by

$$\pi_{t+1}^f = \pi_t^f + \nu_{t+1} \tag{A.7}$$

In other words, Carroll (2005) assumes that inflation follows a random walk plus transitory shocks.
Next, denoting the annualized quarterly inflation rate between t and t+1 as

$$\pi_{t,t+1} = 4(\log p_{t+1} - \log p_t) = 4\pi_{t+1}, \tag{A.8}$$

and using the assumed process (A.6) and (A.7) for inflation, the annual inflation rate between t and $t+4$ can be written as the sum of the intermediate quarterly inflation rates:

$$\begin{aligned}
\pi_{t,t+4} &= \pi_{t+1} + \pi_{t+2} + \pi_{t+3} + \pi_{t+4} \\
&= \pi_{t+1}^f + \varepsilon_{t+1} + \underbrace{\pi_{t+2}^f}_{=\pi_{t+1}+\nu_{t+2}} + \varepsilon_{t+2} + \underbrace{\pi_{t+3}^f}_{=\pi_{t+2}^f+\nu_{t+3}=\pi_{t+1}^f+\nu_{t+2}+\nu_{t+3}} + \varepsilon_{t+3} + \pi_{t+4}^f + \varepsilon_{t+4} \\
&= \pi_{t+1}^f + \varepsilon_{t+1} + \pi_{t+1}^f + \nu_{t+2} + \varepsilon_{t+2} + \pi_{t+1}^f + \nu_{t+2} + \nu_{t+3} + \varepsilon_{t+3} \\
&+ \pi_{t+1}^f + \nu_{t+2} + \nu_{t+3} + \nu_{t+4} + \varepsilon_{t+4}
\end{aligned} \tag{A.9}$$

Next, define $F_t[s]$ as the forecast made in period t for inflation in period s. Assuming that agents in t cannot forecast the transitory shock ε_{t+1} and permanent shock ν_{t+2}, one has, for all $n > 0$:

$$F_t[\varepsilon_{t+n}] = F_t[\nu_{t+n+1}] = 0 \tag{A.10}$$

Using this in equation (A.9), the shocks drop out and one gets the results that that expected annual inflation rate $F_t[\pi_{t,t+4}]$ is simply four times the expected annualized quarterly inflation rate:[4]

$$F_t[\pi_{t,t+4}] = 4F_t[\pi^f_{t+1}] = F_t[\pi^f_{t,t+1}] \tag{A.11}$$

In a similar vein, the annual expected inflation forecast of newspapers is given as

$$N_t[\pi_{t,t+4}] = 4N_t[\pi^f_{t+1}] = N_t[\pi^f_{t,t+1}] \tag{A.12}$$

Next, note that with using equation (A.10) and (A.11), one has

$$F_t[\pi_{t,t+1}] = F_t[\pi_{t,t+4}] = F_t[\pi^f_{t,t+4}] \tag{A.13}$$

Assuming that an individual will update his expectation adopting the newspaper forecast:

$$F_t[\pi_{t,t+4}] = N_t[\pi^f_{t,t+4}] \tag{A.14}$$

and, under the assumption that also newspapers have no information about inflation in periods $t+n, n>0$,

$$F_t[\pi_{t,t+4}] = N_t[\pi^f_{t,t+4}] = N_t[\pi_{t,t+4}] \tag{A.15}$$

Next, note that the assumption that agents cannot forecast changes in inflation beyond $t+1$ leads to:

$$F_t[\pi_{t,t+4}] = F_t[\pi_{t+1,t+5}] \tag{A.16}$$

However, up to period t, individuals can forecast inflation. Hence, one has:

$$F_{t-1}[\pi_{t-1,t+3}] = F_{t-1}[\pi_{t,t+4}] \tag{A.17}$$

and

$$F_{t-2}[\pi_{t-2,t+2}] = F_{t-2}[\pi_{t,t+4}] \tag{A.18}$$

Using the epidemiology process (A.5), one can write for annual inflation expectations:

[4]Where the final expression stems from the notation in equation (A.8) of the annualized quarterly inflation rate, $\pi_{t,t+1} = 4\pi_{t+1}$.

$$\begin{aligned} M_t[\pi_{t,t+4}] &= \lambda F_t[\pi_{t,t+4}] + (1-\lambda)F_{t-1}[\pi_{t,t+4}] \\ &= \lambda F_t[\pi_{t,t+4}] + (1-\lambda)\left[\lambda F_{t-1}[\pi_{t,t+4}] + (1-\lambda)F_{t-2}[\pi_{t,t+4}]\right] \\ &= \lambda F_t[\pi_{t,t+4}] + (1-\lambda)\left[\lambda F_{t-1}[\pi_{t,t+4}] + (1-\lambda)\{\lambda F_{t-2}[\pi_{t,t+4}] + (1-\lambda)F_{t-3}[\pi_{t,t+4}] + ...\}\right] \end{aligned} \tag{A.19}$$

With the help of (A.17) and (A.18), this expression can be further simplified. First, write

$$\begin{aligned} M_t[\pi_{t,t+4}] &= \lambda F_t[\pi_{t,t+4}] + (1-\lambda)F_{t-1}[\pi_{t,t+4}] \\ &= \lambda F_t[\pi_{t,t+4}] + (1-\lambda)\left[\lambda F_{t-1}[\pi_{t-1,t+3}] + (1-\lambda)F_{t-2}[\pi_{t,t+4}]\right] \end{aligned} \tag{A.20}$$

Second, move the first line of (A.20) backwards by one period and rearrange:

$$\begin{aligned} M_{t-1}[\pi_{t-1,t+3}] &= \lambda F_{t-1}[\pi_{t-1,t+3}] + (1-\lambda)F_{t-2}[\pi_{t,t+4}] \\ \Leftrightarrow \lambda F_{t-1}[\pi_{t-1,t+3}] &= M_{t-1}[\pi_{t-1,t+3}] - (1-\lambda)F_{t-2}[\pi_{t,t+4}] \end{aligned} \tag{A.21}$$

Third, use (A.21) in (A.20), and replace F_t by the newspaper forecast N_t to receive the final equation

$$\begin{aligned} M_t[\pi_{t,t+4}] &= \lambda F_t[\pi_{t,t+4}] + (1-\lambda)\left[M_{t-1}[\pi_{t-1,t+3}] - (1-\lambda)F_{t-2}[\pi_{t,t+4}] + (1-\lambda)F_{t-2}[\pi_{t,t+4}]\right] \\ &= \lambda N_t[\pi_{t,t+4}] + (1-\lambda)M_{t-1}[\pi_{t-1,t+3}], \end{aligned}$$

or, applied to monthly data and using my own notation:

$$\pi_t^{exp,hh} = \lambda\pi_t^{exp,prof} + (1-\lambda)\pi_{t-1}^{exp,hh} \tag{2.1}$$

Carroll (2003) notes that without the assumption about the underlying process driving the inflation rate, one would get

$$M_t[\pi_{t,t+4}] = \lambda N_t[\pi_{t,t+4}] + (1-\lambda)M_{t-1}[\pi_{t,t+4}], \tag{A.22}$$

The only difference is the term inside of M_{t-1} which leads to the result that today's inflation expectations cannot be written anymore as a weighted average of the newspaper forecast and households' one-year-ahead forecast made in the previous period. Instead, one has a forecast made yesterday for inflation between today and one year ahead. However, Carroll considers this to be of minor importance, claiming that in reality the two differing forecasts are highly correlated.

A.6.2 Unit Root and Cointegration Tests

This section briefly reviews the tests for unit roots and cointegration applied to the epidemiology model. The most widely used Dickey-Fuller-test (Dickey and Fuller, 1979) starts from the assumption that the true process of a non-stationary time series y_t is given by

$$y_t = \alpha_0 + \rho_0 y_{t-1} \, (+\delta_0 t) + \varepsilon_t, \tag{A.23}$$

where ε_t is white noise, and $\delta_0 t$ captures a deterministic time trend.

Then, one uses the equation

$$\Delta y_t = \alpha_0 + \beta_0 y_{t-1} + \sum_{j=1}^{k} \gamma_j \Delta y_{t-k} \, (+\delta_0 t) + \varepsilon_t \tag{A.24}$$

to test $H_0 : \beta_0 \equiv \rho_0 - 1 = 0$, i.e. whether the time series follows a unit root.

The DF-GLS test proposed by Elliott et al. (1996) first runs a GLS estimation of the equation

$$\tilde{y}_t = \tau_0 x_t + \tau_1 z_t + \varepsilon_t, \tag{A.25}$$

where $\tilde{y}_1 = y_1$, $\tilde{y}_t = y_t - \alpha^* y_{t-1}, t = 2, \cdots, T$, $x_1 = 1$, $x_t = 1 - \alpha^*, t = 2, \cdots, T$, $z_1 = 1$, $z_t = t - \alpha^*(t-1)$, and $\alpha^* = 1 - (13.5/T)$.[5]
Next, the estimates $\hat{\tau}_0$ and $\hat{\tau}_1$ are used to detrend the original variable y_t by

$$y^* = y_t - (\hat{\tau}_0 + \hat{\tau}_1 t) \tag{A.26}$$

Finally, the detrended variable y_t^* is used in equation (A.24), testing the same null hypothesis as in the original Dickey-Fuller test. Elliott et al. (1996) show that this modification of the Dickey-Fuller-test has more power than its original version. The DF-GLS code provided by Stata uses a criterion developed by Schwert (1989) to set the maximum lag length to be included in the test equation. Also, three information criteria are available to chose the optimal lag length: the Schwarz information criterion (SIC), the modified Akaike (MAIC) suggested by Ng and Perron (2000) and the sequential t (seqt) proposed by Ng and Perron (1995).

Similar to the ADF-test, the Phillips-Perron test (Phillips and Perron, 1988) uses the null-hypothesis of non-stationarity, but accounts for potential serial correlation in the underlying time series by applying Newey-West-standard errors instead of augmenting the test equation with additional lags. Two test statistics, Z_t and Z_ρ, are available, whereas the critical values are the same as for the Dickey-Fuller test.

[5]Where the value 13.5 is the optimal value gained via simulation.

Instead of using the null hypothesis of non-stationarity, the KPSS-test developed by Kwiatkowski et al. (1992) uses the null of stationarity. In this test, equation (A.23) is replaced with

$$\begin{aligned} y_t &= \alpha_{0,t}\,(+\delta_0 t) + \varepsilon_t \\ \alpha_{0,t} &= \alpha_{0,t-1} + u_t,\; u_t \sim N(0, \sigma_u^2) \end{aligned} \tag{A.27}$$

The null hypothesis is then $H_0 : \sigma_u^2 = 0$, under which the time series y_t is stationary. The KPSS test statistic is computed as the LM test for the residuals e_t from the estimated equation

$$y_t = \alpha_0\,(+\delta_0 t) + \varepsilon_t \tag{A.28}$$

$$KPSS = \sum_{t=1}^{T} S_t^2/\hat{\sigma}_u^2, \tag{A.29}$$

where $S_t = \sum_{i=1}^{t} e_i$, and $\hat{\sigma}_u^2$ is the sum of squared residuals.

The Zivot-Andrews-test (Zivot and Andrews, 1992) allows for a structural break in the time series, both in the intercept and in the deterministic trend. As it has been shown by Perron (1990), if the series y_t is $I(0)$, but exhibits a structural break in the mean at time λT such that

$$y_t = \begin{cases} y_{1,t},\; t < \lambda T \\ y_{2,t} + \mu,\; t \geq \lambda T \end{cases}, \lambda \in (0,1), \tag{A.30}$$

the estimated $\hat{\rho}$ in the ADF-test (A.24) converges to one the larger the structural break μ. Hence, the wrong null hypothesis of a unit root is rejected less frequently in case of a considerable structural break. Perron (1989) and Perron and Vogelsang (1992) have thus proposed to take into account a structural break when testing for a unit root in a time series. Using the break dummy

$$s_t(\lambda) = \begin{cases} 0,\; t < \lambda T \\ 1,\; t \geq \lambda T \end{cases}, \tag{A.31}$$

they allow for three variants of a structural break in the test equation, a break in the the intercept, a break in the slope of the linear trend, and a break in both the intercept and the linear trend. The test consists of a two-step-approach: first, the deterministic components including the break dummies are removed with a first regression, before the residuals are tested for a unit root. More precisely, the first equation in case of a break in the intercept only reads as

$$y_t = \tilde{\alpha}_0 + \tilde{\alpha}_1 s_t(\lambda) + \varepsilon_t \tag{A.32}$$

In case of a break in the linear trend, we have

$$y_t = \tilde{\alpha}_0 + \tilde{\delta}_0 t + \tilde{\delta}_1 s_t(\lambda) t + \varepsilon_t, \tag{A.33}$$

whereas for a break in both components, we get

$$y_t = \tilde{\alpha}_0 + \tilde{\alpha}_1 + s_t(\lambda)\tilde{\delta}_0 t + \tilde{\delta}_1 s_t(\lambda) t + \varepsilon_t \tag{A.34}$$

Then, the residuals e_t of these first equations are used in the test equation, which is given by

$$\Delta e_t = (\rho - 1)\, e_{t-1} + \sum_{j=1}^{k} \gamma_j \Delta e_{t-j} + \sum_{j=0}^{k} g_j \Delta s_{t-j} + u_t \tag{A.35}$$

in case of a break in the intercept. The test proposed by Perron and Vogelsang (1992) assumes that the break date λ is known. Zivot and Andrews (1992) have extended this test to the case where the break date is unknown a priori. They propose to run the testing procedure for each possible break date λ and to test for a unit root using the lowest t-statistic (the highest negative value) from the previous regressions, i.e. the largest rejection of the null hypothesis of a unit root.

Finally, Clemente et al. (1998) proposed a test that allows for two breaks in the time series. In one version of their test, the *Additive-Outlier (AO)*-test, they model a sudden shift in the mean, whereas a second versions assumes that the mean changes gradually over time (*Innovational Outlier test*). As in the test suggested by Perron and Vogelsang (1992) and Zivot and Andrews (1992), the Additive-Outlier (AI)-test is carried out in a two step procedure, now allowing for two breaks λ_1 and λ_2. The first regression becomes

$$y_t = \tilde{\alpha}_0 + \tilde{\alpha}_1 s_{1,t}(\lambda_1) + \tilde{\alpha}_2 s_{2,t}(\lambda_2) + \varepsilon_t, \tag{A.36}$$

followed by the test equation that is searched for the lowest t statistic given the two break dates λ_1 and λ_2

$$e_t = \sum_{j=0}^{k} \omega_{1,j} \tilde{s}_{1,t-j} + \sum_{j=0}^{k} \omega_{2,j} \tilde{s}_{2,t-j} + \rho_0 e_{t-1} + \sum_{j=1}^{k} \gamma_j \Delta e_{t-1} + u_t, \tag{A.37}$$

where $\tilde{s}_{1/2,t} = 1$ for $s_{1/2,t} + 1$.
In case of the Innovational Outlier (IO)-test that allows for a gradual shift in the linear trend, the first regression is given by:

$$y_t + \alpha_0 + \alpha_1 s_{1,t}(\lambda_1) + \alpha_2 s_{2,t}(\lambda_2) + \alpha_3 \tilde{s}_{1,t}(\lambda_1) + \alpha_4 \tilde{s}_{2,t}(\lambda_2) + \rho_0 e_{t-1} + \sum_{j=1}^{k} \gamma_j \Delta e_{t-j} + u_t \quad \text{(A.38)}$$

With regard to cointegration, Gregory and Hansen (1996a,b) propose a test that allows for a structural break in the relationship between the variables that are possibly cointegrated. If two variables y_t and x_t are cointegrated, a long-run relationship exists such that e_t is I(0):

$$y_t = \alpha_0 + \beta_0 x_t + \varepsilon_t \quad \text{(A.39)}$$

Using the break dummy defined in (A.31), the Gregory-Hansen test models four possible versions of cointegration including structural breaks. First, the level of the long-run relationship shifts over time, while the slope coefficient β_0 is not affected:

$$y_t = \alpha_0 + \alpha_1 s_t(\lambda) + \beta_0 x_t + \varepsilon_t \quad \text{(A.40)}$$

Second, the cointegration relationship can also include a a linear time trend:

$$y_t = \alpha_0 + \alpha_1 s_t(\lambda) + \beta_0 x_t + \delta_0 t + \varepsilon_t \quad \text{(A.41)}$$

Next, one can also imagine that the slope coefficient of the long-run relationship also shifts over time. In such as *regime shift* model, we have

$$y_t = \alpha_0 + \alpha_1 s_t(\lambda) + \beta_0 x_t + \beta_1 s_t(\lambda) x_t + \varepsilon_t \quad \text{(A.42)}$$

Finally, Gregory and Hansen (1996b) extend the regime shift model allowing for shifts in the linear time trend as well:

$$y_t = \alpha_0 + \alpha_1 s_t(\lambda) + \beta_0 x_t + \beta_1 s_t(\lambda) x_t + \delta_0 t + \delta_1 s_t(\lambda) t + \varepsilon_t \quad \text{(A.43)}$$

The Gregory-Hansen test allows for a structural beak of unknown date following the procedure of Zivot and Andrews (1992). The test statistic is computed for each possible break date λ, and the smallest, i.e. the largest negative negative value is used for the cointegration test. As in the standard ADF test of cointegration, the first differences of the residuals e_t from the equations (A.40) - (A.43) are regressed on e_{t-1} and lagged differences $\Delta e_{t-1}, \Delta e_{t-2}, ...$, and the t-statistic for e_{t-1} is used to test the null of no cointegration.

Chapter B

Appendix to Chapter 3

B.1 Literature Overview: Demographics and Inflation Expectations

The Literature Reporting Demographic Differences in Inflation Expectations

A number of studies, often conducted by central banks, have documented a direct impact of demographic characteristics on households' inflation expectations. We briefly summarize the results and refer to Table (B.1) on the next page for a more detailed overview.
Bryan and Venkatu (2001b) conduct telephone interviews in the U.S.-state of Ohio asking respondents for their perceived and expected inflation. They report higher inflation expectations for less educated, low-income, young and old people compared to middle-age survey participants, in addition to women, singles and nonwhites. Across all groups, differences in perceived inflation are larger compared to expected inflation. In a representative survey conducted in New Zealand, Leung (2009) reports higher forecast errors for the young, individuals with a non-European background, lower income levels, females, low-skilled workers and respondents from rural areas. As it turns out, those groups which overpredict inflation correspond to those that have a higher probability of not answering the survey, hence, aggregate survey measures might be biased. Brischetto and de Brouwer (1999) offer results for Australia and report higher expectations of low-income groups and younger individuals as well. In addition, predictions were higher for the unemployed and for people with a lower education level. Respondents' political views seem to matter as well: expectations are higher for participants who claimed to support the Labor Party and the Greens. Blanchflower and MacCoille (2009) use two different surveys for the UK, one with quantitative answers and another one with qualitative responses. In both surveys, the better educated have lower expectations, whereas expectations rise with age. However, computing forecast errors over a shorter time span, people tend to better forecast inflation if they

grow older. Moreover, females, unemployed and home owners are worse in forecasting inflation. Palmqvist and Strömberg (2004) analyze survey data for Sweden, observing higher expectations for the young and the old compared to middle-age households, females, unemployed, tenants, singles and households with children. By contrast, inflation rates fall with rising education and income, and if households live in urban areas. The most comprehensive study is offered by Souleles (2004). Using micro-level data for the U.S. from December 1978 to June 1996, he computes three different forecast errors. Two measures compare expectations with inflation perceptions of the same household six months later (using qualitative and quantitative survey responses), and one measure compares expectations with realized inflation. For all three measures, Souleles (2004) reports larger forecast errors for the elderly, females, less educated and poor households, blacks and households with a growing number of children. Finally, Bruine de Bruin et al. (2010) conduct a representative survey in the U.S in 2007 and find higher expectations for females, older people, and singles, while better educated, poorer households, as well as whites report lower forecasts. Pfajfar and Santoro (2009) provide the only study using group-level data for households in the U.S.. In line with the evidence quoted previously, they find that inflation expectations and forecast errors are higher for females, younger households, less educated, and individuals with lower levels of education.

Table B.1: Studies Documenting Demographic Effects on Inflation Expectations

Paper		Bryan and Venkatu (2001b)	Leung (2009)	Brischetto and de Brouwer (1999)	Palmqvist and Strömberg (2004)	Souleles (2004)
Country		US (Ohio)	NZ	AU	SE	US
Survey		Cleveland Fed	Reserve Bank of NZ	Melbourne Institute	Konjunkturinstitutet	Michigan Survey
Survey Level		micro	micro	micro	micro	micro
Time Span		1998m8-2001m11	1998q3-2008q3	1995m1-1998m12	2001m11-2004m5	1978m12-1996m6
Expectations		quantitative	quantitative	quantitative	quantitative	qualitative and qualitative
Dependent Variable		expectations	forecast error	expectations	expectations	forecast errors: perceptions - expectations inflation - expectations
Groups	Age	young +, old +	-	-	young +, old +	+
	Gender	female +	female +	female +	female +	female +
	Education	-	na	-	-	-
	Income	-	-	-	-	-
	Employment	na	low skilled +	unemp +	unemployed +	na
	housing	na	na	na	rent +	na
	Region	na	city -	city -	city -	0
	Race	nonwhite +	white -	na	na	white -
	Relationship Status	single +	na	na	single +	0
	Political Tendency	na	na	Labor, Greens +	na	na
	Children in Household	na	na	na	children +	children +
Explanation		none	none	none	none	none

	Blanchflower and MacCoille (2009)			Pfajfar and Santoro (2009)	Burke and Manz (2011)	Bruine de Bruin et al. (2010)
	UK	UK	UK		US	US
	Bank of England	GfK	Eurobarometer		Harvard University	own survey
	micro	micro	micro	group-level	micro	micro
	2001q1-2009q2	1996m1-2008m10	2005-2007	expectations	2009m12	2007
	quantitative, ranges	qualitative	quant, ranges	1978m1-2005m2	quantitative	quantitative
	expectations	expectations	forecast error	expectations forecast error	forecast error expectations	expectations
Age	+	+	-	-	+ (> 32)	+
Gender	female -	female +	female +	female +	0	female+
Education	-	-	-	-	0	-
Income	na	na	na	-	0	-
Employment	0	self-employed -	unemp +	na	0	na
Housing	rent +	na	rent +	na	0	na
Region	na	city +	na	0	0	na
Race	na	na	na	na	white -	white -
Relationship Status	na	na	na	na	na	single +
Political Tendency	na	na	na	na	na	na
Children in Household	na	na	na	na	na	na
	infl perceptions: more education, less effect from perceptions satisfaction with BoE: more satisfied, lower expectations (not for age)	perceptions	none	news consumption hh-specific inflation	financial literacy	hh-specific inflation financial literacy

Note: + (-) means above (below) average inflation expectations or forecast errors. 0 denotes no significant effect, and na means that the category is not included in the survey.

The Impact of Demographics on Inflation Expectations: Explanations in the Literature

This section classifies the various determinants of inflation expectations disagreement[1] of households proposed in the literature.[2] We illustrate our proposed summary in Figure (B.1). In general, households' socioeconomic background can affect expectations via four channels. First, personal attributes such as individual processing capacities vary between households, resulting in different expectations. Second, households might hold different beliefs on future prices because they find themselves in different microeconomic situations. Third, individuals might react differently to the macroeconomic environment. Fourth, different news media report differently on inflation, and since households consume different newspapers and TV shows, this results in heterogeneous inflation expectations. Note that the media effect works both directly (e.g., because old people spend more time readings newspapers than the young) and indirectly (if households with large asset holdings read newspapers specialized on economic issues, for example). We will briefly explain each of these channels, and present the results of studies that have made use of these channels in order to explain demographic differences in inflation expectations.

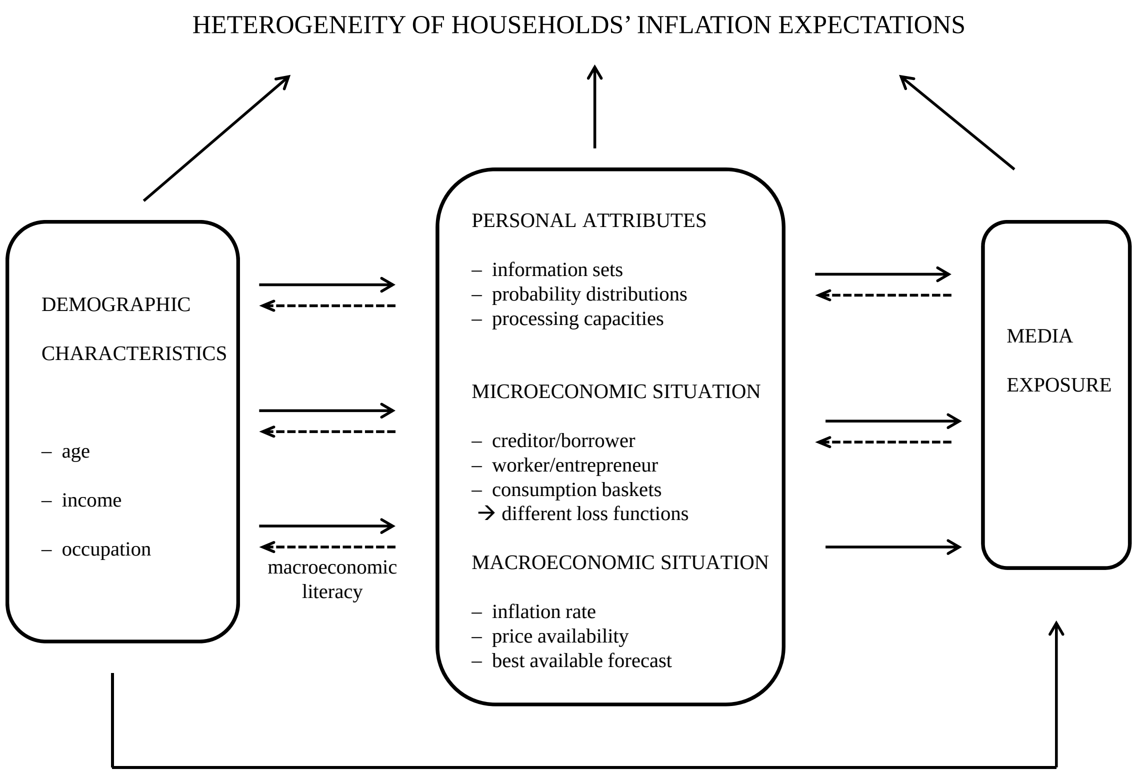


Figure B.1: Driving Forces of Households' Disagreement on Inflation Expectations

[1] In what follows, we use the terms " disagreement" and "heterogeneity" interchangeably.

[2] The disagreement of professional forecasters raises additional questions, since factors such as herding behavior are found to play an important role (Gallo et al., 2002).

The Influence of Personal Attributes To put it simple: inflation expectations are different because individuals are different. They use different information sets, spend a different amount of time to interpret incoming news, have different capacities of processing information, and use more or less sophisticated models of expectation formation. As it is shown in a number of recent papers, each of these personal attributes result in disagreement in individuals' inflation expectations. The sticky information model of Mankiw and Reis (2002, 2007) assumes that acquiring information is costly, leading to the result that only a fraction of individuals makes use of all the information available while the remaining fraction sticks to information sets collected in the past. Relying on the assumption that information processing capacities are limited, Sims (2003) shows that some individuals will rationally choose not to updated to the latest available information sets, while Branch (2004) argues that individuals might even switch between different expectation formation models. Likewise, in the context of learning models (Evans and Honkapohja, 2001), people will more or less quickly converge to the rational expectations benchmark, if their learning curves are different. And Capistran and Timmermann (2009) argue that households have heterogeneous and asymmetric loss functions, thereby weighting the costs of over- and underpredicting inflation differently.

Each of these models makes a microeconomic assumption on individuals' personal attributes and analyze the implied impact on the heterogeneity of inflation expectations on the macroeconomic level. The assumptions on information acquisition and processing can be related to specific household characteristics thus explaining the effect from demographics on inflation expectations. For example, older households might have more experience in understanding the concept of inflation resulting in faster updating and learning pattern. However, it might also be the case that younger households are better in adjusting to new information technologies and policy regimes resulting in more rational expectations of households in younger age. Similarly, unemployed individuals might be less familiar with every-day economic decision making compared to employees or self-employed individuals who are used to do their own book-keeping. Finally, with regard to education, individuals with a high-school degree are expected to better understand the determinants of inflation thus leading to better inflation forecasts if households reach higher education levels.

These possible links between models of information formation and heterogeneous inflation expectations arising from households' socioeconomic backgrounds are rarely tested, though. In two cross-section studies, Burke and Manz (2011) and Bruine de Bruin et al. (2010) argue that the demographic differences of inflation expectations can be explained by households' degree of financial literacy (Lusardi and Mitchell, 2008). They show that individuals' demographic characteristics determine the financial literacy score of individuals which turns out to significantly improve households' inflation forecasts. However, both papers suffer from the fact that they do not find large effects from demographics in the first place, which might be due to the small cross-section dimension.[3] Hence, only some demographic effects can be explained by financial literacy: Burke and Manz (2011) can account for the impact of race (the higher expectations of black survey respondents), while Bruine de Bruin et al. (2010) find lower point estimates for all demographic variables if financial literacy is included, however, the demographic effects are already found to be insignificant without financial literacy. A third paper shows that demographic differences between individuals' expectations are reduced by learning (Anderson et al., 2010). Exploiting the short panel dimension of the Michigan survey[4], those groups that show the largest forecast error in the first interview (low income, female, non-white, young, households with children) show larger reductions of their expectation errors than other groups. Hence, even if Anderson et al. (2010) cannot explain why households' expectations differ in the first place, their results suggest that heterogeneity can be reduced by appropriate communication policies of the central bank or increased news coverage.

The Role of Households' Microeconomic Situation Apart from psychological reasons or different personal attributes, the expectation formation models quoted above can also be linked to the microeconomic situation of households. For example, indebted households might consider inflation as a gain whereas individuals with large asset holdings are expected to spend more time and effort to forecast expectations in order to protect the real value of their wealth. Here, the argument is that households will rationally weight costs and benefits of making a good forecast, and that the cost-benefit analysis depends on their socioeconomic background. Following this reasoning, conflicting conclusions might arise. Whereas old agents are expected to make better forecasts due to higher asset holdings, they could also provide less accurate forecasts since they face higher opportunity costs due to a shorter remaining lifetime (Fishe and Idson, 1990). Empirically, the hypothesis that the dependence of inflation expectations on demographic characteristics stems from households' microeconomic situation is tested by using household-specific inflation rates and inflation perceptions.

The overall Consumer Price Index (CPI) is calculated for consumption goods of a representative individual. Hence, if some households consistently consume more or less of the goods that are included in the CPI, their group-specific inflation rate will differ from overall inflation.[5] A number of papers has documented households' inflation differentials arguing that these can be related to individuals' socioeconomic background. Overall, households with low income, low education levels and older households face higher inflation rates. Results for the U.S. are provided by Michael (1979), Hagemann (1982), Hobijn and Lagakos (2005), and McGranahan and Paulson (2006), while Colavecchio et al. (2011) offer results for a panel

[3] For example, the highest age category used by Burke and Manz (2011) is "older than 32".
[4] 40% of respondents are interviewed a second time six months after the first interview.
[5] Indeed, Inoue et al. (2009) show that inflation expectations derived form households' spending pattern outperform survey measures in forecasting CPI inflation.

of 15 European countries. We refer to the latter study for a comprehensive literature review. For Germany, there exists only one unpublished study quoted by Colavecchio et al. (2011), suggesting higher inflation rates for the elderly and for households with high income levels. Jonung (1981) was among the first to suggest that the differences in group-specific inflation rats can account for the differences in inflation expectations, especially the higher inflation expectations of women compared to men. Since women were thought to be mainly responsible for food purchases, and since food prices were rising faster than CPI at the time of his survey, females reported higher inflation expectations. However, Bryan and Venkatu (2001a) could not support this hypothesis, leaving the gender inflation differential an open research question. More generally, Pfajfar and Santoro (2009) provide some support for the view that households are better in forecasting their group-specific inflation rate instead of CPI inflation. They find that for low and middle income households, the forecast error is smaller if household-specific inflation is used, while richer households are better in forecasting overall inflation. However, separating households with respect to education always yields lower forecast errors for aggregate inflation, while the results are mixed for the elderly. Bruine de Bruin et al. (2010) ask participants in a survey conducted at the end of 2007 about their thoughts when forming their inflation expectations. Including the responses "thoughts about prices you pay" and "thoughts about how to cover expenses" makes the initial effect from education insignificant. This suggests that individuals with lower education levels think more of their group-specific inflation rate instead of overall CPI inflation. Anderson et al. (2012) proxy household-specific inflation rates with inflation rates at the top-level item categories in the U.S.-CPI. They argue that poor households spend a larger fraction of their overall expenditure on housing, thus above average price changes in this category should impact more on households with lower income levels. However, splitting the CPI into its components does not help explain that some households report higher expectations than others.[6]

Apart from different cost-benefit-analysis arising from the household's microeconomic situation, households' dependence on individual inflation rates can also be explained by psychological effects. According to the availability hypothesis (Tversky and Kahneman, 1973), people have a better memory for prices of goods they buy more frequently. Hence, if survey participants are asked for their price expectations, they might implicitly use a goods basket as reference point that relates more to their individual consumption. It is by no means clear, however, that consumers indeed rely on household-specific inflation rates. Research in psychology summarized by Ranyard et al. (2008) shows that households have difficulties in recalling prices they have paid, even of goods they bought recently. If this is true, households would not base their expectations on actual group-specific inflation rates, but instead use an estimate of past prices, the so-called perceived inflation rate. Since the ability

[6] This might stem from the fact that the CPI categories are not precise enough in measuring household-specific consumption spending.

of retrospection might be systematically related to households' demographic characteristics, households with lower income levels might perceive their own inflation rate much stronger than other households, which subsequently feeds into larger expectation differentials. Blanchflower and MacCoille (2009) provide the only study that tests the impact of inflation perceptions on households' expectations. However, demographic differences in inflation expectations still prevail if perceived inflation is included as explanatory variable. Only with respect to education, their results suggest that more educated individuals tend to rely less on perceptions when forecasting inflation.

The Macroeconomic Environment In the near-rationality model of Akerlof et al. (1996, 2000), the heterogeneity of inflation expectations depends on the level of the overall inflation rate. In a low-inflation environment, most agents tend to ignore latest news on inflation, while as soon as inflation picks up, a growing number of individuals starts forming expectations rationally until inflation reaches a level where again, all households share the same beliefs on future prices. Mankiw et al. (2003) test the impact of the macroeconomic environment on expectation disagreement, using the level and the change of overall inflation, relative price variability and the output gap as explanatory variables. Gnan et al. (2011), using group level data for a panel of 12 Euro Area countries, repeat their analysis and test whether the within-group forecast disagreement is different between demographic groups. Across all groups, a positive output gap and rising inflation lowers the disagreement of households in the same group, while an increase in relative price variability leads to more disagreement. With regard to differences between household groups, their results suggest that the richer the households the more they tend to agree on expectations if inflation rises. The same holds true for young and old households, households with higher education and males, while no clear pattern emerges for the price variability and the output-gap. However, since the authors do not report how the within-group disagreement varies between groups, it remains unanswered whether the demographic differences in households' inflation expectations can be explained with different reactions to macroeconomic conditions. Instead of referring to real economic data, Blanchflower and MacCoille (2009) claim that it is households' trust in the policy of the central bank that leads to different expectations between household groups. Generally, they find that individuals who are more satisfied with the conduct of monetary policy report lower inflation expectations compared to dissatisfied households. Only for age groups, they observe higher expectations for the elderly even if these have greater confidence in the central bank. Instead of trusting in the central bank, households might rely on the expectations of professional forecasters serving as a proxy for the best available forecast in an economy. Carroll (2003) has proposed that on aggregate, households only sluggishly update their expectations in line with those of professional forecasters. Pfajfar and Santoro (2009) apply this framework to households' inflation expectations differentiated by demographic characteristics. They find that males as well as younger and older households rely more on expert forecasts than others. Also, households in the lowest income and lowest education group react least to the best available forecast. However, the results that rising income and education leads to lower inflation expectations and forecast errors cannot be explained by increased attention to expert forecasts. Finally, Malmendier and Nagel (2013) test whether households rely on inflation experiences in their lifetimes when forming their expectations. Younger households should be affected more by recent price developments than older households whose information sets reach back further in the past. Hence, individuals who have experienced the high-inflation period in the 1970s should be slower in adjusting their expectations to the following low-inflation period. Their empirical analysis indeed supports this view of "learning by experience".

Household-Specific Media Exposure Pfajfar and Santoro (2009) investigate the role of the news media for explaining the dependence of inflation expectations on demographic characteristics. They do not use a media measure for news coverage such as the number of articles in a given newspaper, but employ the answers to a question included in the Michigan Survey. Households are asked whether they have heard (favorable and unfavorable) news about prices within the past months. It turns out that the better educated and the richer the households, the higher the fraction of respondents who have heard news about prices. The same holds true for men, while with regard to age, middle-age households report to be better informed than others. Hence, with the exception of age, it seems that the higher forecast errors of some household groups stem from the fact that they do not pay enough attention to news. In a second step, Pfajfar and Santoro (2009) test whether the fact that households have heard news about inflation affects the distance of their expectations from professional forecasters' expectations, as suggested by Carroll (2003). For example, if a piece of news has a larger impact on this expectation gap for low income households compared to high income households, one could attribute the demographic differences in expectations to different news reception. Generally, however, their results do not support this hypothesis. With regard to the overall number of news heard, they find larger news effects for the young, the better educated, males, and the rich, but since the media effect is always found to be positive, this means that these households deviate more from the expert forecast if they receive news on inflation.[7] Distinguishing favorable news from unfavorable news, the same picture emerges. While more positive news make households to be more in line with experts, the effect is stronger for the less educated and poorer households. Conversely, more negative news increase the expectation gap more strongly for better educated and richer households. The same pattern holds true for gender. Anderson et al. (2012) also exploit the "news heard"-question from the Michigan survey, but add news heard about government

[7] Pfajfar and Santoro (2009) do not say whether those groups with higher forecast errors correspond to those with the largest deviation from professional forecasters' expectations. Implicitly, they seem to assume that this is the case.

spending, employment, and money and profits to news about inflation. Part of their results support the hypothesis that news drive expectation differentials. Females more than proportionally increase their inflation expectations if they hear positive news on government spending, while the effect from news about inflation does not differ between sexes. Similarly, the least educated households raise their expectations in response to positive news on fiscal spending, and in response to negative news on inflation. A slightly stronger news effect is observed for young and old households compared to middle-aged individuals, while the results are less supportive for income groups: news on inflation do not have a heterogeneous effect, only positive news about employment increase the expectations of low income households relative to households with higher income. Finally, Lamla and Maag (2012) find that more negative news reports on inflation reduces the within-group disagreement of German households. Differentiating households only with respect to education, the media effect rises with the education level of households.

B.2 Quantification Technique

This section describes the probability method used to to quantify the qualitative survey responses, where we follow Nielsen (2003) who applies the method to the Consumer Survey of the European Commission. Remember that survey participants have six possible answer categories to the question on how they think consumer prices will develop in the future:

pp : “rise a lot”

p : “rise moderately”

e : “rise slightly”

m : “stay about the same”

mm : “fall”

dn : “don't know”

Thus, for each month, the survey provides the fractions of respondents choosing one of the above answer categories. In a first step, we proportionally add the fraction of “don't know”-answers to the remaining five categories, such that

$$ffrac_i = frac_i + frac_i/5, \text{ where } frac_i \in \{mm, m, e, p, pp\} \tag{.1}$$

Next, using the notation of Nielsen (2003), we assume an interval $(-\delta_t^L, \delta_t^U)$ around 0, which defines those inflation rates that individuals associate with stable prices. Similarly, we assume an interval $(\tilde{\mu}_t - \varepsilon_t^L, \tilde{\mu}_t + \varepsilon_t^U)$ which captures inflation rates that are associated with prices thought to “increase at the same rate”. Applying these assumptions to the reaming answer categories, we get

$$
\begin{array}{lll}
\textit{prices will...} & & \\
\quad \text{fall slightly} & \text{if} & \pi^e_{t+1} \leq -\delta^L_t \\
\quad \text{be stable} & \text{if} & -\delta^L_t < \pi^e_{t+1} \leq \delta^U_t \\
\quad \text{increase at slower rate} & \text{if} & \delta^U_t < \pi^e_{t+1} \leq \tilde{\mu}_t - \varepsilon^L_t \\
\quad \text{increase at same rate} & \text{if} & \tilde{\mu}_t - \varepsilon^L_t < \pi^e_{t+1} < \tilde{\mu}_t + \varepsilon^U_t \\
\quad \text{increase more rapidly} & \text{if} & \tilde{\mu}_t + \varepsilon^U_t \leq \pi^e_{t+1}
\end{array}
\tag{.2}
$$

Next, we use the fractions of the answer categories fmm, fm, fe, fp, fpp, and express the intervals in terms of the cumulative standard normal distribution function ϕ:

$$qmm_{t,t+1} = \phi^{-t}(fmm_{t,t+1}) \tag{.3}$$
$$qm_{t,t+1} = \phi^{-t}(fmm_{t,t+1} + fm_{t,t+1}) \tag{.4}$$
$$qe_{t,t+1} = \phi^{-t}(fmm_{t,t+1} + fm_{t,t+1} + fe_{t,t+1}) \tag{.5}$$
$$ep_{t,t+1} = \phi^{-t}(fmm_{t,t+1} + fm_{t,t+1} + fe_{t,t+1} + fp_{t,t+1}) \tag{.6}$$

Finally, Nielsen (2003) shows that the quantified mean inflation expectation is given by

$$\mu_{t,t+1} = \frac{\tilde{\mu}_t(qmm_{t,t+1} + qm_{t,t+1})}{q_{t,t+1}}, \tag{.7}$$

where $q_{t,t+1}$ is defined as $q_{t,t+1} = qmm_{t,t+1} + qm_{t,t+1} - qe_{t,t+1} - qp_{t,t+1}$. Hence, the only unknown parameter in the equation of households' quantitative inflation expectations is the perceived inflation rate $\tilde{\mu}_t$. We replace $\tilde{\mu}_t$ with the HP-filter of households' group-specific inflation rate, whereas the filter is calculated recursively over 20 months. Using different lag lengths does not qualitatively change the results for the quantified rate of inflation expectations.

B.2.1 Additional Tables and Figures

Table B.2: Data Sources

Data	Start Date	End Date	Source	Link
Households' Expectations and Perceptions	1998M09	2010M05	European Commission (EC)	EC
Household-specific Inflation	1997M01	2010M06	EC Household Budged Surveys (HBS)	HBS
Professional Forecasters' Expectations	1989M10	2010M03	Consensus Economics	Consensus
Inflation Rates (HICP)	1997M01	2012M03	Eurostat	Eurostat
Media Coverage	1998M01	2011M02	Media Tenor	Media Tenor
Media Circulation (TV)	1998Q1	2011Q4	Media Perspektiven (MP)	MP
Media Circulation (Print)	1998Q1	2011Q4	Informationsgemeinschaft zur Feststellung der Verbreitung von Werbeträgern e.V. (IVW)	IVW

Figure B.2: Inflation Expectations of Different Household Groups

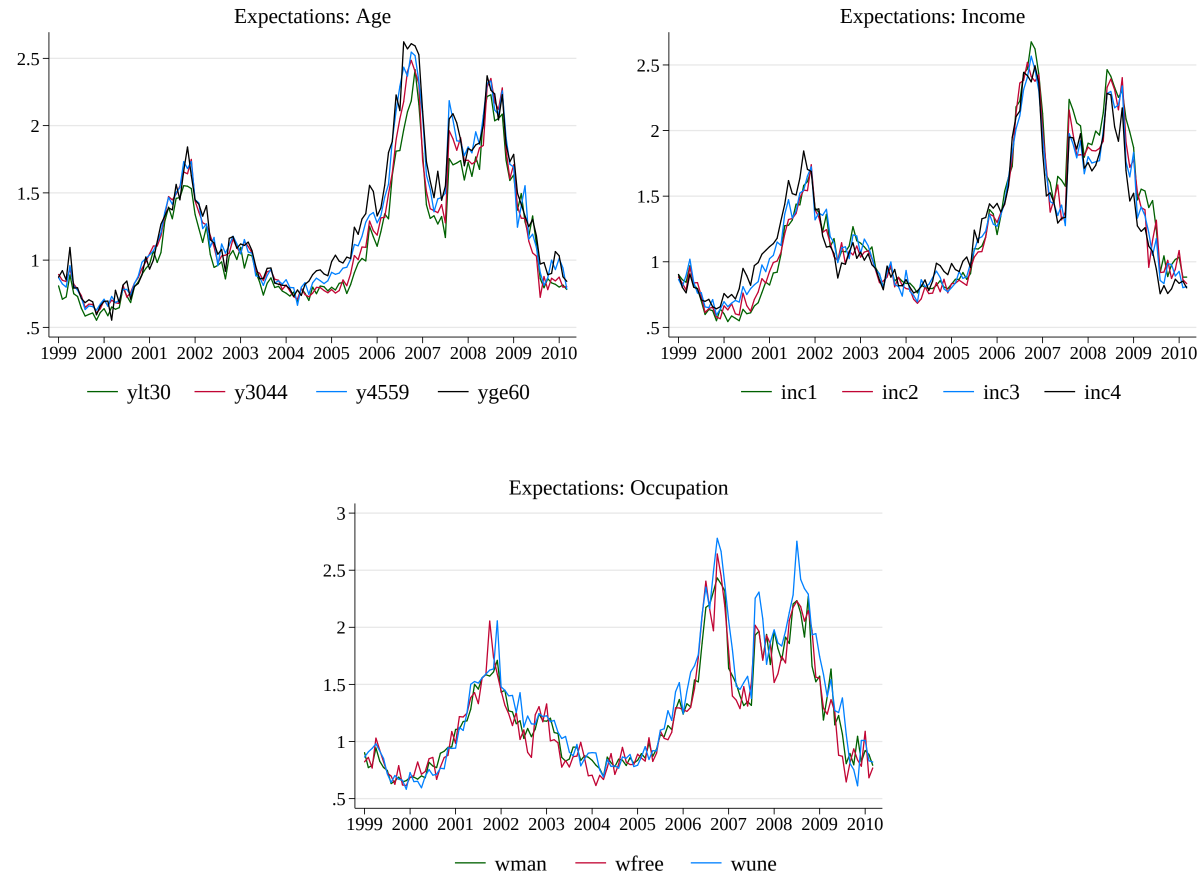


Figure B.3: Print Run and TV Audience

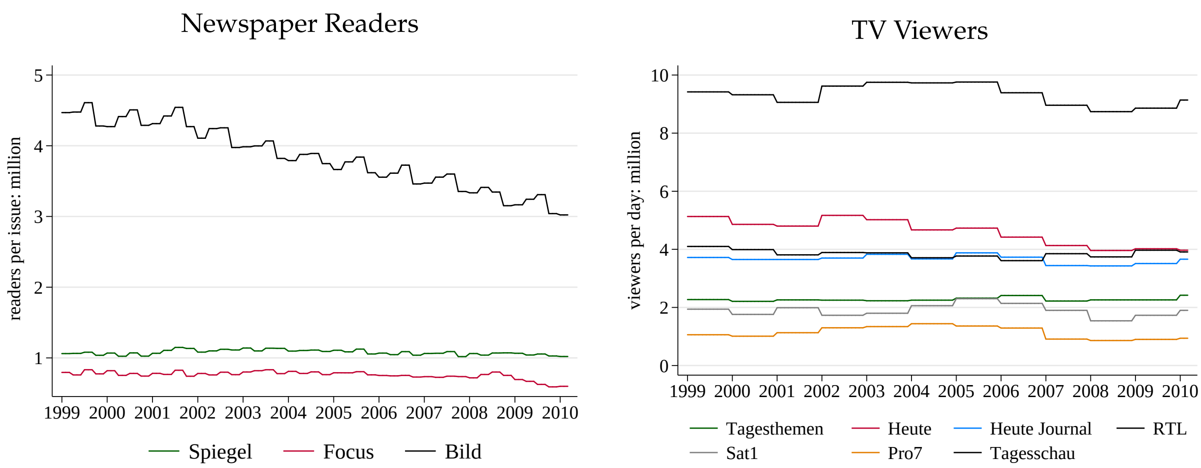


Figure B.4: Media Coverage Ia: Number of News Reports About Inflation per Month

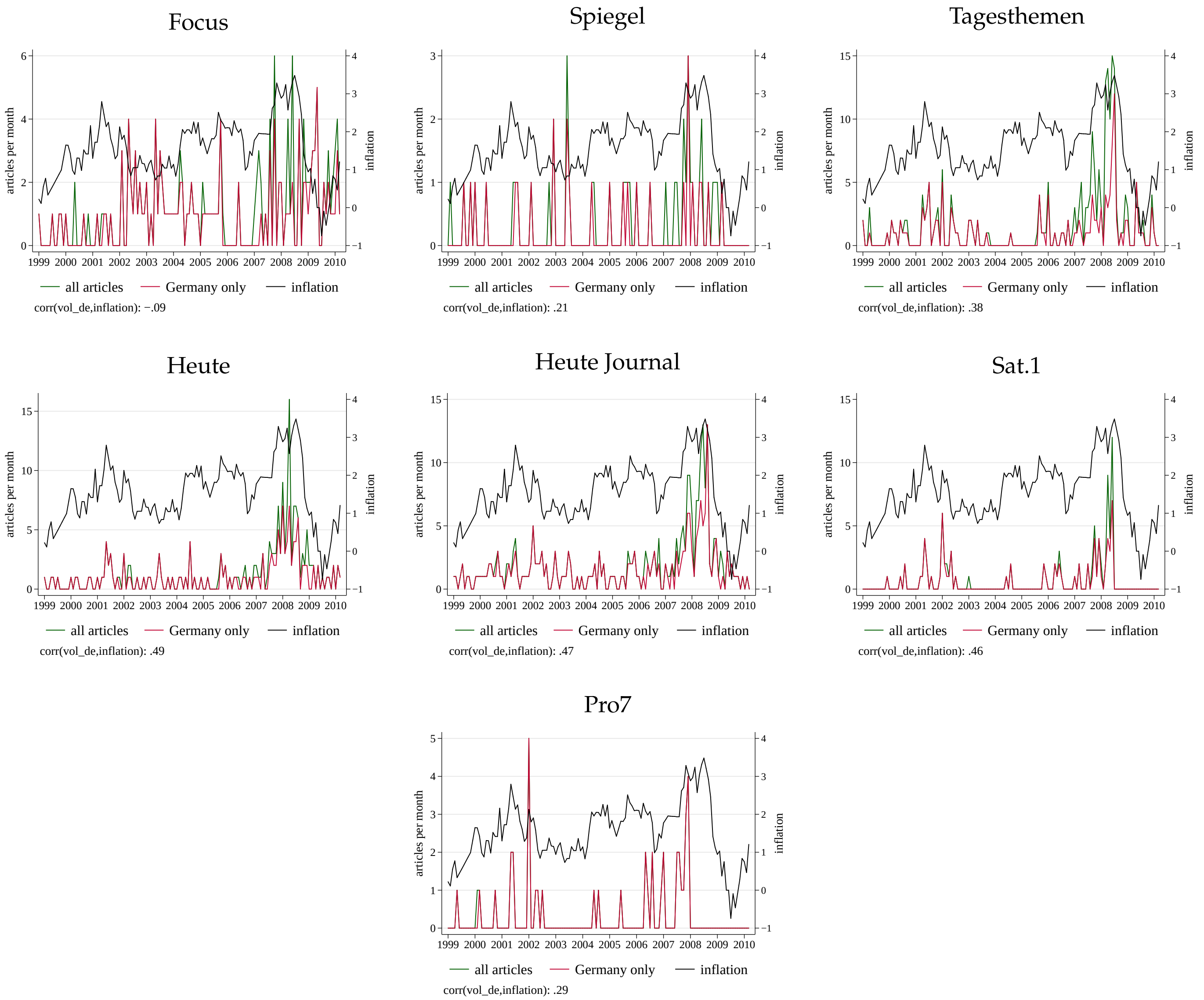


Figure B.5: Media Coverage IIa: Number of Positive and Negative News About Inflation per Month - Context

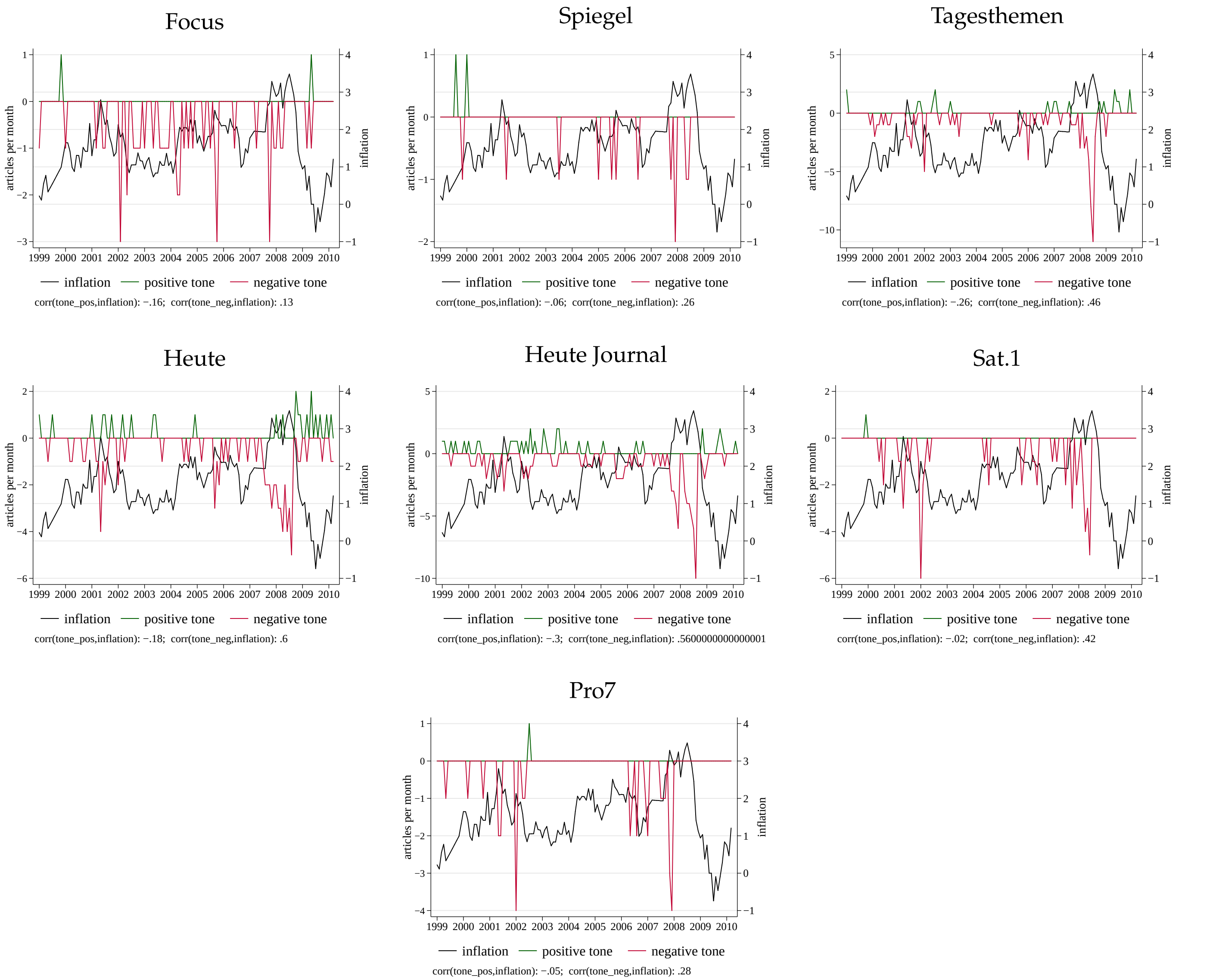
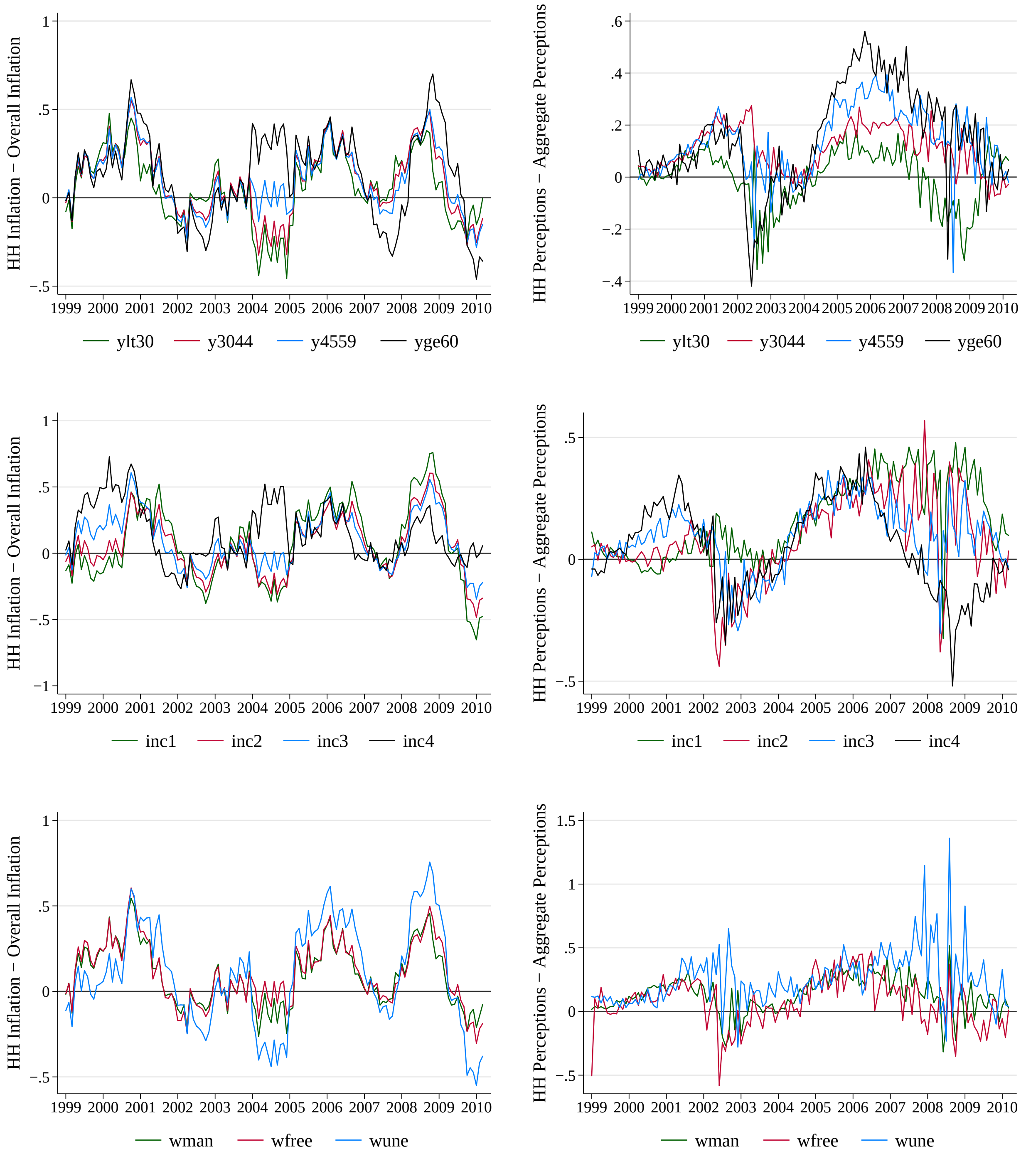


Figure B.6: Differentials of HH-Inflation and HH-Perceptions



B.2.2 Results Assuming Exogeneity of Media Variables

Table B.3: Results: Aggregate Volume of Media Reports - SUR Regression

	ylt30	y3044	y4559	yge60	inc1	inc2	inc3	inc4	wman	wfree	wune
π_{t-1}	0.05	0.04	-0.00	-0.02	-0.06	0.00	0.01	0.01	0.01	0.02	-0.06*
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)
$News_t^{pr\ index}$	-0.23*	-0.25**	-0.21*	-0.20*	-0.33**	-0.25*	-0.21*	-0.10	-0.18	-0.12	-0.33**
	(0.13)	(0.12)	(0.12)	(0.12)	(0.17)	(0.15)	(0.12)	(0.09)	(0.12)	(0.12)	(0.14)
$News_t^{tv\ index}$	-0.13	-0.11	0.01	0.07	0.06	-0.07	-0.04	-0.01	-0.07	-0.03	0.09
	(0.14)	(0.12)	(0.12)	(0.12)	(0.17)	(0.15)	(0.12)	(0.09)	(0.12)	(0.12)	(0.14)
$\pi_{j,t} - \pi_t$	0.13*	0.07	0.12*	0.20***	0.24***	0.18***	0.16***	0.16***	0.27***	0.28***	0.23***
	(0.07)	(0.06)	(0.06)	(0.05)	(0.06)	(0.07)	(0.05)	(0.04)	(0.08)	(0.08)	(0.06)
$perc_{j,t} - perc_t$	0.06	0.04	-0.02	-0.10*	0.03	-0.08	0.01	-0.05	-0.03	-0.04	0.01
	(0.08)	(0.09)	(0.06)	(0.06)	(0.07)	(0.05)	(0.06)	(0.04)	(0.06)	(0.05)	(0.05)
cons	0.27***	0.23***	0.25***	0.25***	0.40***	0.31***	0.24***	0.16***	0.24***	0.20***	0.34***
	(0.05)	(0.04)	(0.04)	(0.04)	(0.06)	(0.05)	(0.04)	(0.03)	(0.04)	(0.04)	(0.05)
R^2	0.082	0.080	0.090	0.188	0.080	0.083	0.122	0.159	0.118	0.144	0.110
N	134				134				134		

Note: Unconstrained SUR regressions. *<0.1, **<0.05, *** p<0.01. Sample 1999M1-2010M3. S.e.'s in brackets.

Table B.4: Results: Disaggregate Volume of Media Reports - SUR Regressions

	ylt30	y3044	y4559	yge60	inc1	inc2	inc3	inc4	wman	wfree	wune
π_{t-1}	0.05*	0.04	0.00	-0.01	-0.05	0.00	0.01	0.02	0.01	0.03	-0.06*
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)
$News_t^{Bild}$	-0.24*	-0.24**	-0.23**	-0.21*	-0.36**	-0.25*	-0.22*	-0.09	-0.21*	-0.12	-0.35***
	(0.13)	(0.12)	(0.11)	(0.11)	(0.16)	(0.14)	(0.11)	(0.08)	(0.12)	(0.11)	(0.13)
$News_t^{Tag}$	0.21	0.21*	0.29***	0.30***	0.44***	0.31**	0.23**	0.21**	0.17	0.29***	0.37***
	(0.13)	(0.12)	(0.11)	(0.11)	(0.17)	(0.15)	(0.12)	(0.08)	(0.12)	(0.11)	(0.14)
$News_t^{RTL}$	-0.25**	-0.23**	-0.18*	-0.13	-0.20	-0.23*	-0.18*	-0.16**	-0.16	-0.26***	-0.09
	(0.11)	(0.10)	(0.10)	(0.10)	(0.14)	(0.12)	(0.10)	(0.07)	(0.10)	(0.09)	(0.12)
$\pi_{j,t} - \pi_t$	0.11*	0.06	0.10	0.19***	0.23***	0.17***	0.15***	0.16***	0.24***	0.25***	0.22***
	(0.06)	(0.06)	(0.06)	(0.05)	(0.06)	(0.06)	(0.05)	(0.04)	(0.08)	(0.08)	(0.06)
$perc_{j,t} - perc_t$	0.06	0.05	-0.02	-0.10	0.04	-0.07	0.01	-0.06	-0.04	-0.03	0.03
	(0.08)	(0.09)	(0.06)	(0.06)	(0.07)	(0.05)	(0.05)	(0.04)	(0.06)	(0.05)	(0.05)
cons	0.23***	0.20***	0.22***	0.22***	0.34***	0.26***	0.21***	0.14***	0.21***	0.16***	0.29***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.06)	(0.05)	(0.04)	(0.03)	(0.04)	(0.04)	(0.05)
R^2	0.128	0.125	0.141	0.222	0.137	0.130	0.163	0.205	0.147	0.213	0.158
N	134				134				134		

Note: Unconstrained SUR regressions. *<0.1, **<0.05, *** p<0.01. Sample 1999M1-2010M3. S.e.'s in brackets.

Table B.5: Results: Aggregate Tone of Media Reports - SUR Regressions

	ylt30	y3044	y4559	yge60	inc1	inc2	inc3	inc4	wman	wfree	wune
π_{t-1}	0.02	0.01	-0.03	-0.04	-0.10**	-0.03	-0.02	0.01	-0.02	0.01	-0.08**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)
$News_t^{pos\ con}$	-0.12	-0.15	-0.08	-0.06	-0.10	-0.19	-0.09	-0.00	-0.14	-0.07	-0.08
	(0.13)	(0.11)	(0.11)	(0.11)	(0.16)	(0.14)	(0.11)	(0.08)	(0.12)	(0.11)	(0.14)
$News_t^{neg\ con}$	-0.31	-0.34	-0.48**	-0.51**	-0.76**	-0.49*	-0.37*	-0.30*	-0.27	-0.26	-0.57**
	(0.25)	(0.22)	(0.21)	(0.21)	(0.31)	(0.27)	(0.22)	(0.16)	(0.22)	(0.21)	(0.26)
$News_t^{pos\ val}$	0.15	0.15	0.15	0.07	0.17	0.21	0.11	0.12	0.11	0.18*	0.14
	(0.13)	(0.11)	(0.11)	(0.11)	(0.16)	(0.14)	(0.11)	(0.08)	(0.11)	(0.10)	(0.13)
$News_t^{neg\ val}$	0.51**	0.52**	0.57***	0.51**	0.84***	0.62**	0.47**	0.38***	0.39*	0.36*	0.68***
	(0.24)	(0.21)	(0.20)	(0.20)	(0.30)	(0.26)	(0.21)	(0.15)	(0.21)	(0.20)	(0.25)
$\pi_{j,t} - \pi_t$	0.15**	0.10	0.13**	0.21***	0.25***	0.20***	0.17***	0.15***	0.27***	0.28***	0.24***
	(0.06)	(0.06)	(0.06)	(0.05)	(0.06)	(0.07)	(0.05)	(0.04)	(0.08)	(0.08)	(0.06)
$perc_{j,t} - perc_t$	0.07	0.05	-0.01	-0.10	0.02	-0.08	0.02	-0.05	-0.03	-0.03	0.01
	(0.08)	(0.09)	(0.06)	(0.06)	(0.07)	(0.05)	(0.05)	(0.04)	(0.06)	(0.05)	(0.05)
cons	0.25***	0.22***	0.24***	0.25***	0.39***	0.31***	0.23***	0.14***	0.24***	0.19***	0.33***
	(0.05)	(0.04)	(0.04)	(0.04)	(0.06)	(0.06)	(0.04)	(0.03)	(0.04)	(0.04)	(0.05)
R^2	0.102	0.110	0.135	0.215	0.117	0.124	0.145	0.189	0.134	0.175	0.131
N	134				134				134		

Note: Unconstrained SUR regressions. *<0.1, **<0.05, *** p<0.01. Sample 1999M1-2010M3. S.e.'s in brackets.

Table B.6: Results: Disaggregate Positive Tone of Media Reports - SUR Regressions

	ylt30	y3044	y4559	yge60	inc1	inc2	inc3	inc4	wman	wfree	wune
π_{t-1}	0.02	0.01	-0.01	-0.02	-0.08**	-0.02	-0.01	0.01	-0.01	0.01	-0.07**
	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
$News_t^{Bild\ con\ pos}$	-0.21*	-0.17*	-0.17*	-0.16*	-0.28**	-0.22*	-0.17*	-0.04	-0.18*	-0.10	-0.18
	(0.11)	(0.10)	(0.09)	(0.09)	(0.14)	(0.12)	(0.10)	(0.07)	(0.10)	(0.09)	(0.12)
$News_t^{Tag\ con\ pos}$	0.23**	0.19*	0.21**	0.16	0.29**	0.22*	0.17*	0.19***	0.17*	0.29***	0.25**
	(0.11)	(0.10)	(0.10)	(0.10)	(0.14)	(0.13)	(0.10)	(0.07)	(0.10)	(0.09)	(0.12)
$News_t^{RTL\ con\ pos}$	-0.04	-0.11	-0.07	-0.09	-0.10	-0.06	-0.05	-0.02	-0.06	-0.13	-0.08
	(0.14)	(0.13)	(0.12)	(0.12)	(0.18)	(0.16)	(0.12)	(0.09)	(0.12)	(0.11)	(0.15)
$\pi_{j,t} - \pi_t$	0.11*	0.06	0.09	0.18***	0.24***	0.18***	0.17***	0.17***	0.23***	0.22***	0.22***
	(0.07)	(0.06)	(0.06)	(0.05)	(0.06)	(0.07)	(0.05)	(0.03)	(0.08)	(0.08)	(0.06)
$perc_{j,t} - perc_t$	0.09	0.07	-0.03	-0.13**	0.04	-0.07	0.02	-0.04	-0.02	0.01	0.02
	(0.08)	(0.10)	(0.06)	(0.06)	(0.07)	(0.05)	(0.05)	(0.04)	(0.06)	(0.05)	(0.06)
cons	0.24***	0.20***	0.23***	0.24***	0.36***	0.28***	0.22***	0.13***	0.22***	0.17***	0.30***
	(0.05)	(0.04)	(0.04)	(0.04)	(0.06)	(0.05)	(0.04)	(0.03)	(0.04)	(0.04)	(0.05)
R^2	0.095	0.083	0.100	0.190	0.104	0.098	0.140	0.201	0.129	0.194	0.115
N	134				134				134		

Note: Unconstrained SUR regressions. *<0.1, **<0.05, *** p<0.01. Sample 1999M1-2010M3. S.e.'s in brackets.

Table B.7: Results: Disaggregate Negative Tone of Media Reports

	ylt30	y3044	y4559	yge60	inc1	inc2	inc3	inc4	wman	wfree	wune
π_{t-1}	0.03	0.02	-0.01	-0.02	-0.08**	-0.02	-0.00	0.02	-0.00	0.01	-0.07**
	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)
$News_t^{Bild\ con\ neg}$	0.12	0.10	0.05	-0.01	0.05	-0.01	0.09	0.01	0.09	-0.02	0.14
	(0.17)	(0.15)	(0.14)	(0.14)	(0.21)	(0.18)	(0.15)	(0.11)	(0.15)	(0.14)	(0.17)
$News_t^{Tag\ con\ neg}$	-0.20	-0.18	-0.25*	-0.32**	-0.44**	-0.29	-0.21	-0.11	-0.10	-0.19	-0.30*
	(0.17)	(0.15)	(0.14)	(0.14)	(0.21)	(0.19)	(0.15)	(0.11)	(0.15)	(0.14)	(0.18)
$News_t^{RTL\ con\ neg}$	0.25	0.20	0.23*	0.22*	0.34*	0.30*	0.19	0.17*	0.14	0.22*	0.18
	(0.16)	(0.14)	(0.13)	(0.13)	(0.20)	(0.17)	(0.14)	(0.10)	(0.14)	(0.13)	(0.16)
$\pi_{j,t} - \pi_t$	0.13**	0.08	0.12**	0.21***	0.24***	0.18***	0.16***	0.15***	0.27***	0.28***	0.24***
	(0.06)	(0.06)	(0.06)	(0.05)	(0.06)	(0.07)	(0.05)	(0.04)	(0.08)	(0.08)	(0.06)
$perc_{j,t} - perc_t$	0.06	0.04	-0.04	-0.12**	0.06	-0.06	0.03	-0.06	-0.01	-0.04	0.02
	(0.07)	(0.09)	(0.06)	(0.06)	(0.07)	(0.05)	(0.05)	(0.04)	(0.06)	(0.05)	(0.06)
cons	0.23***	0.20***	0.23***	0.23***	0.36***	0.27***	0.21***	0.14***	0.22***	0.18***	0.31***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.06)	(0.05)	(0.04)	(0.03)	(0.04)	(0.04)	(0.05)
R^2	0.090	0.072	0.104	0.210	0.093	0.090	0.129	0.168	0.117	0.159	0.105
N	134				134				134		

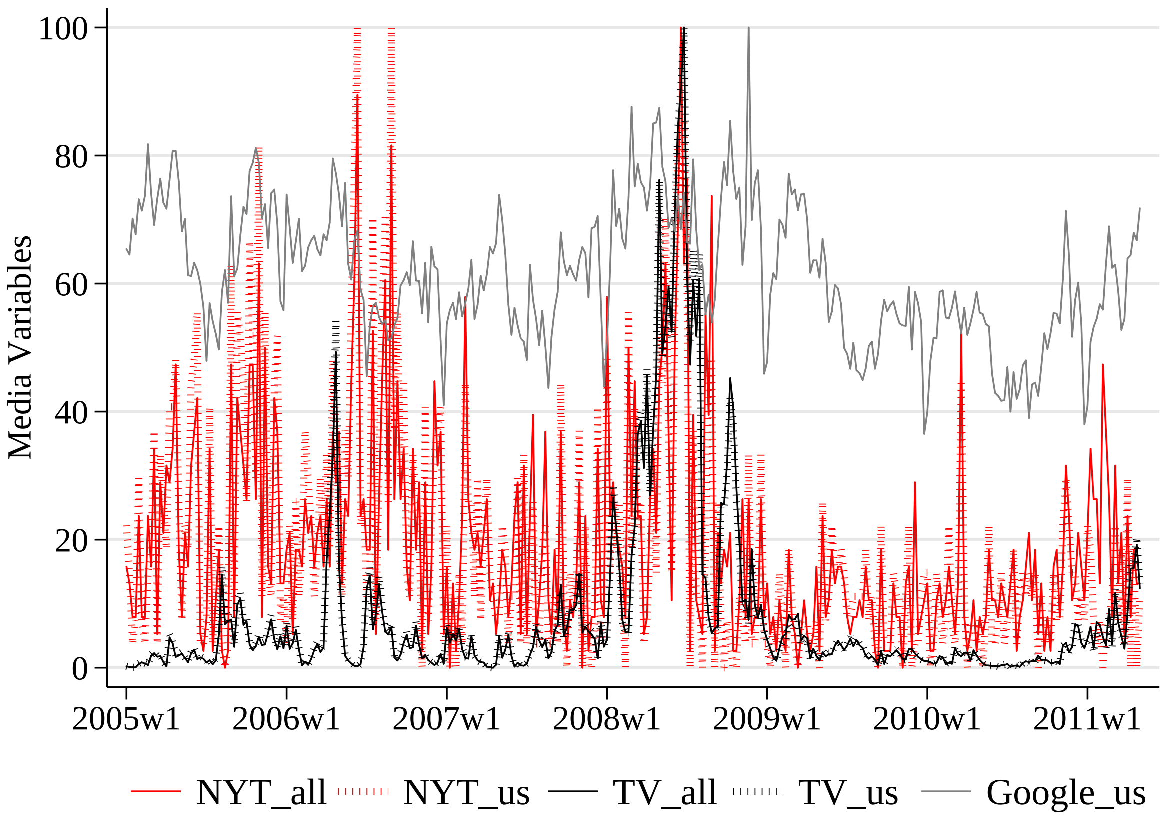
Note: Unconstrained SUR regressions. *<0.1, **<0.05, *** p<0.01. Sample 1999M1-2010M3. S.e.'s in brackets.

Chapter C

Appendix to Chapter 4

C.1 Additional Figures

Figure C.1: Media Variables - Weekly Data



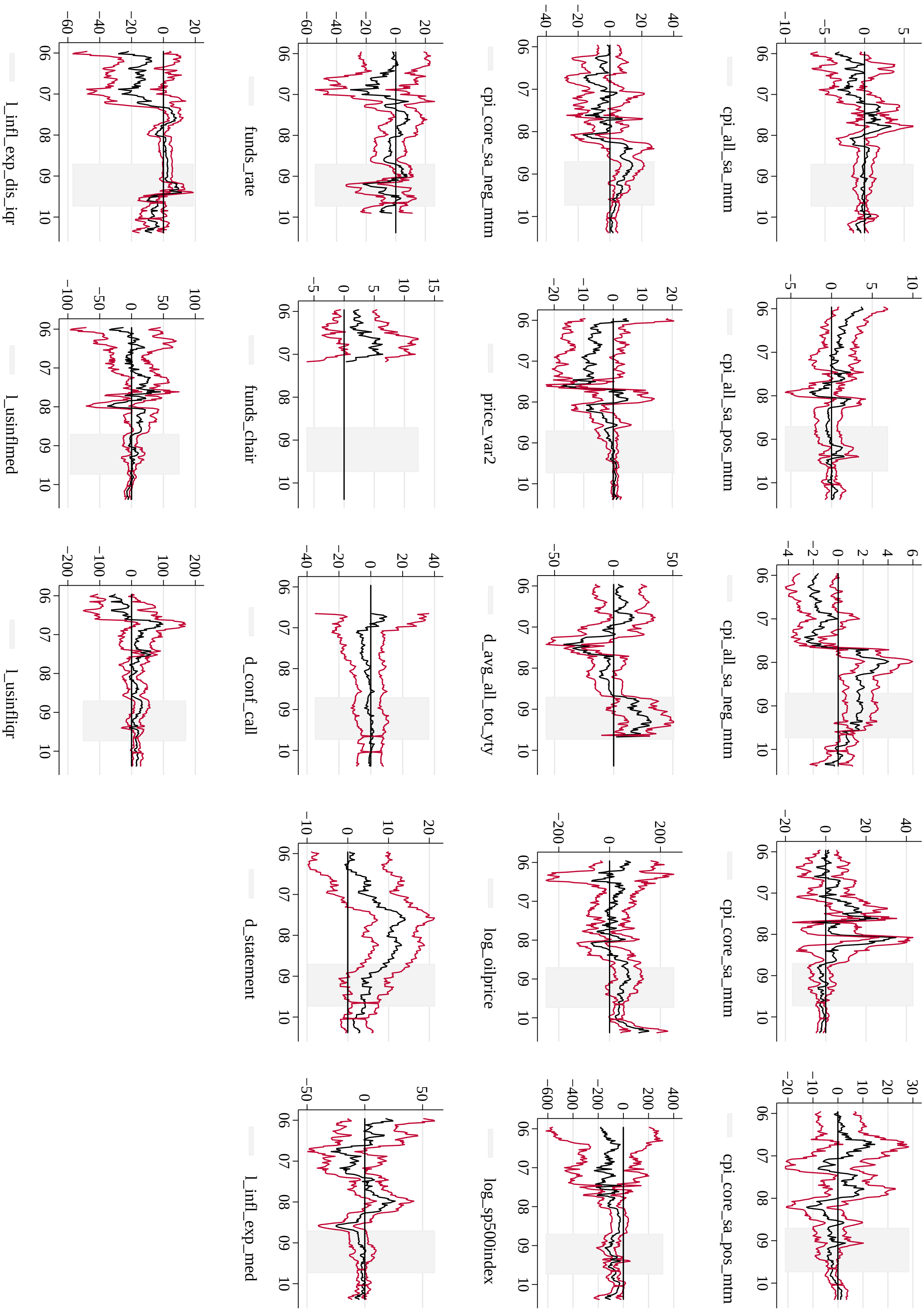


Figure C.2: Rolling Regression - NYT

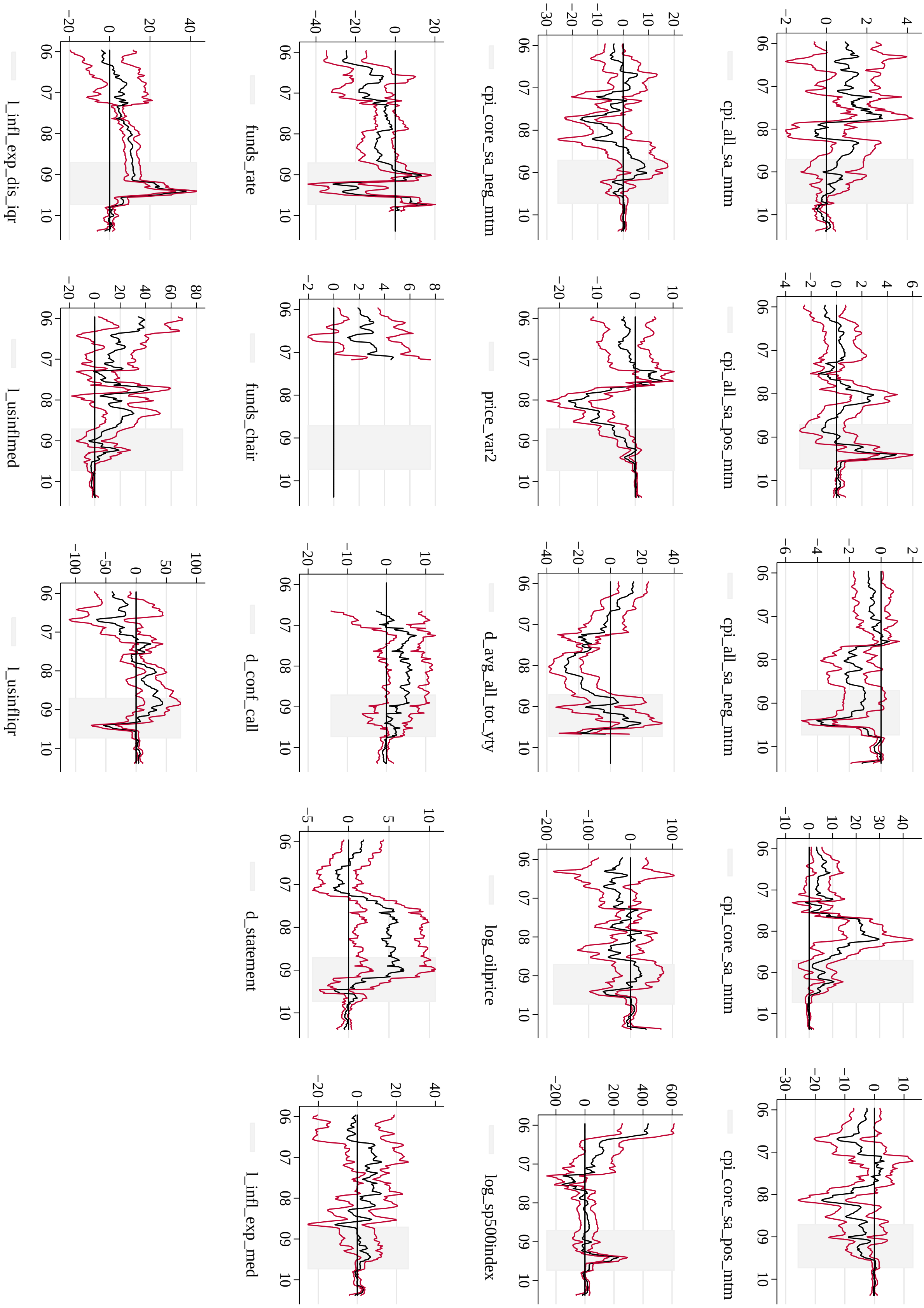


Figure C.3: Rolling Regression - TV

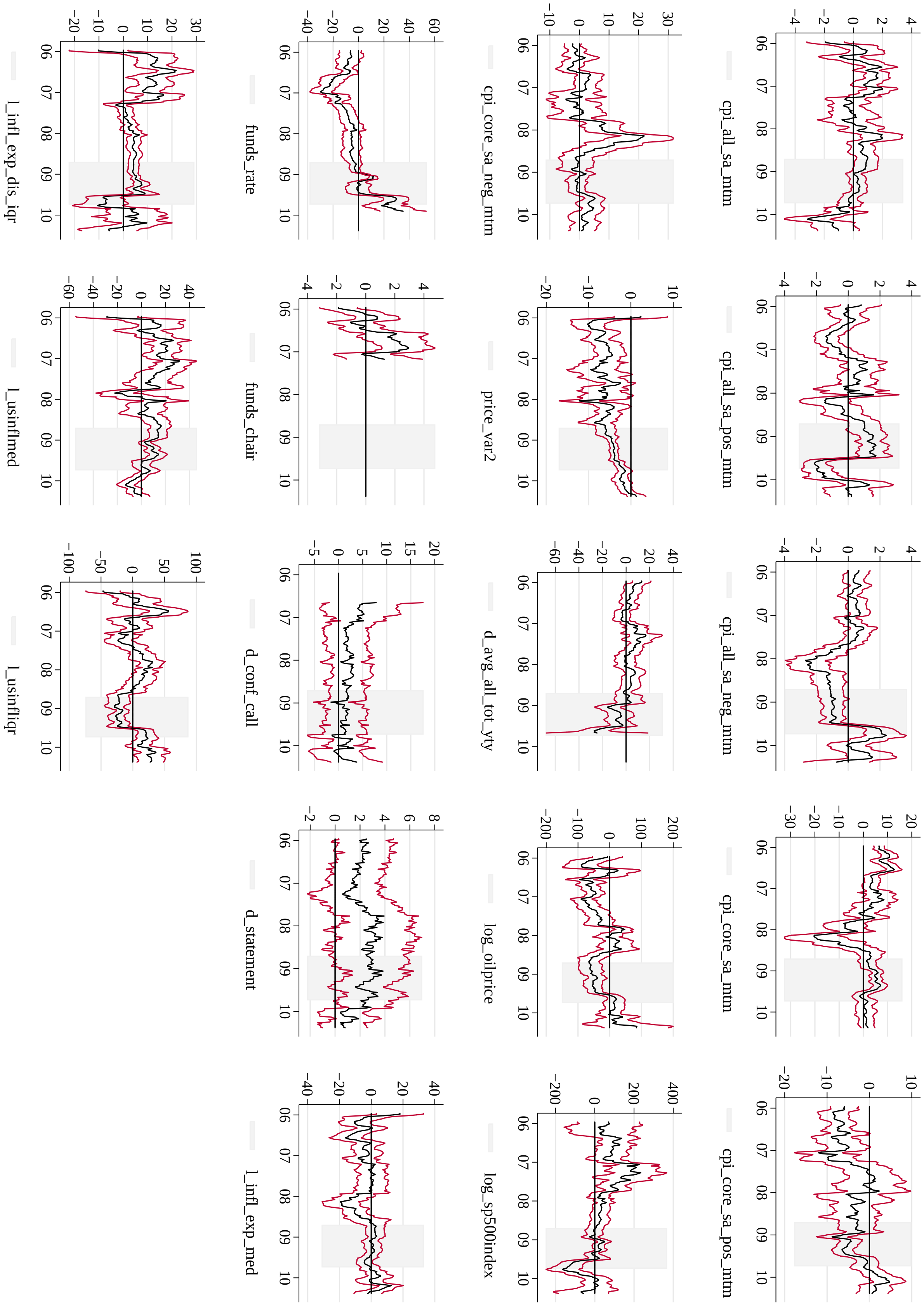


Figure C.4: Rolling Regression - Google

Figure C.5: News Content - Fitted Values

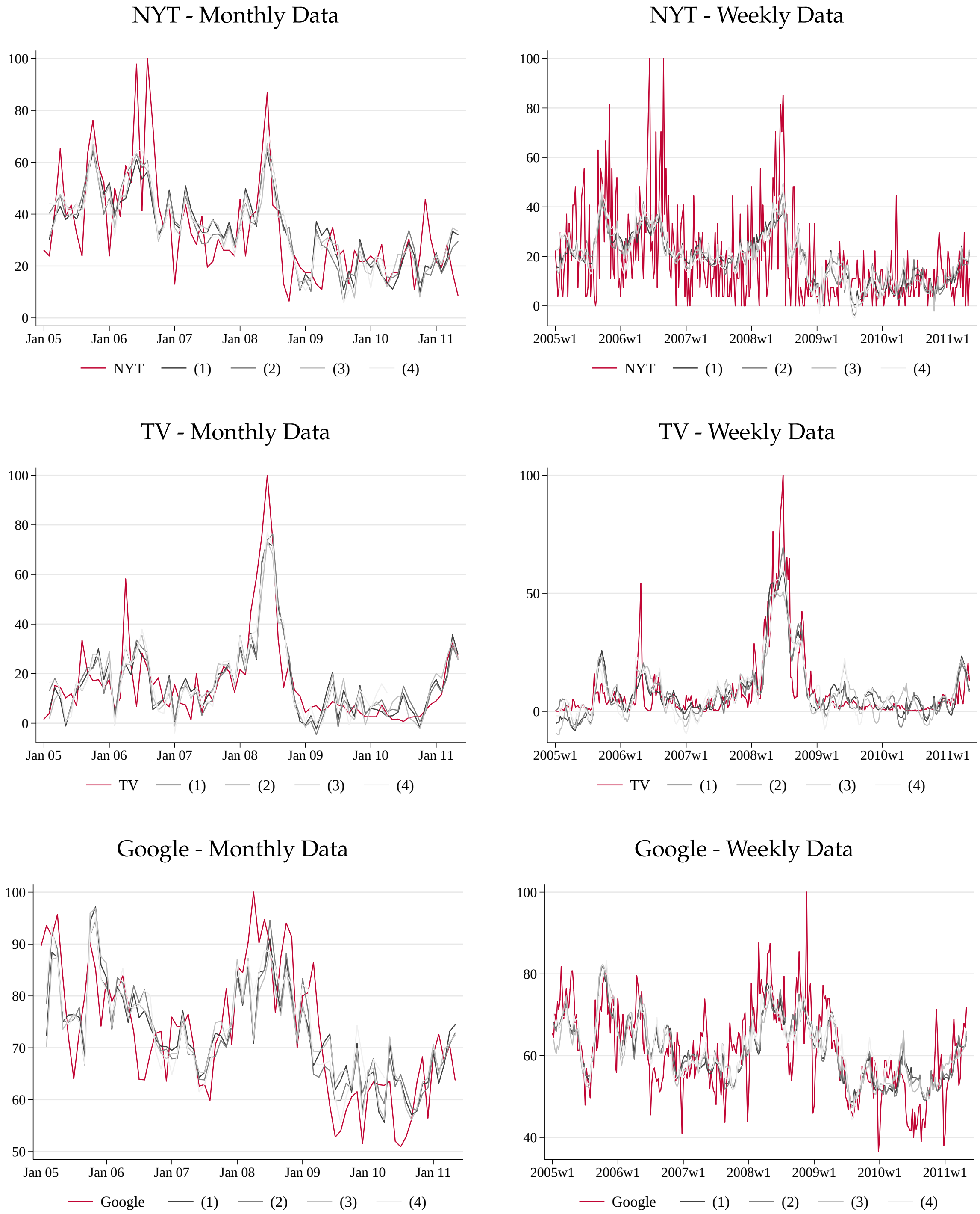
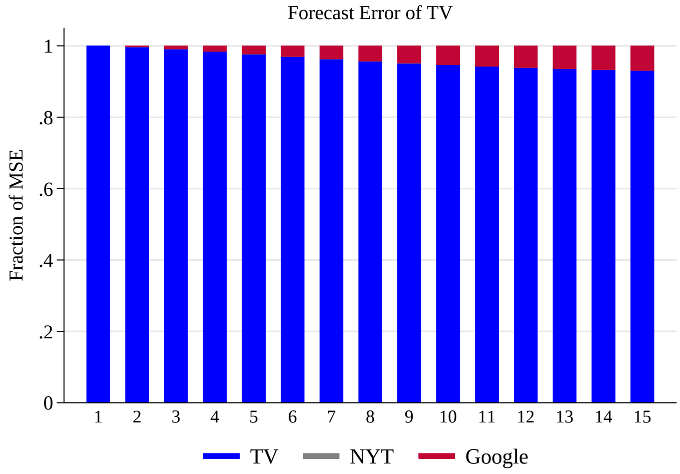
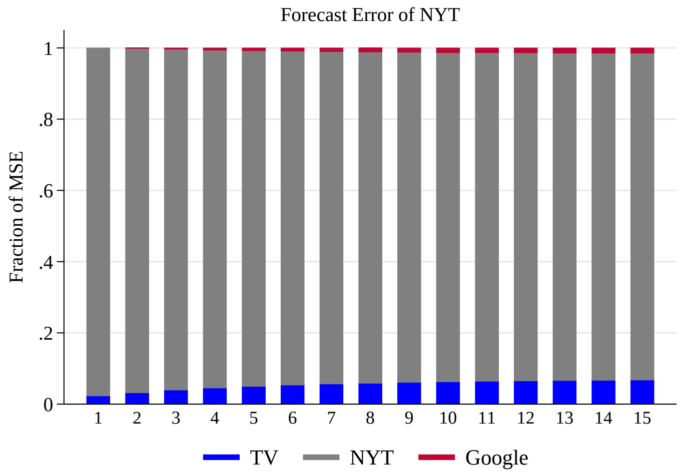


Figure C.6: Baseline VAR - FEVD - Weekly Data

Baseline Model



Including CPI Inflation



Including CPI Inflation

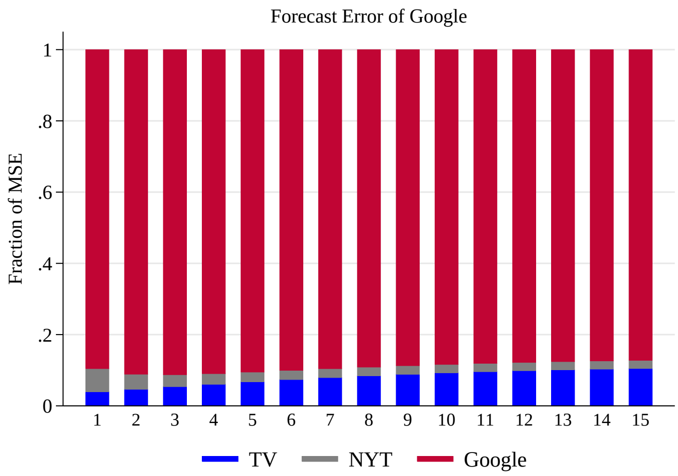
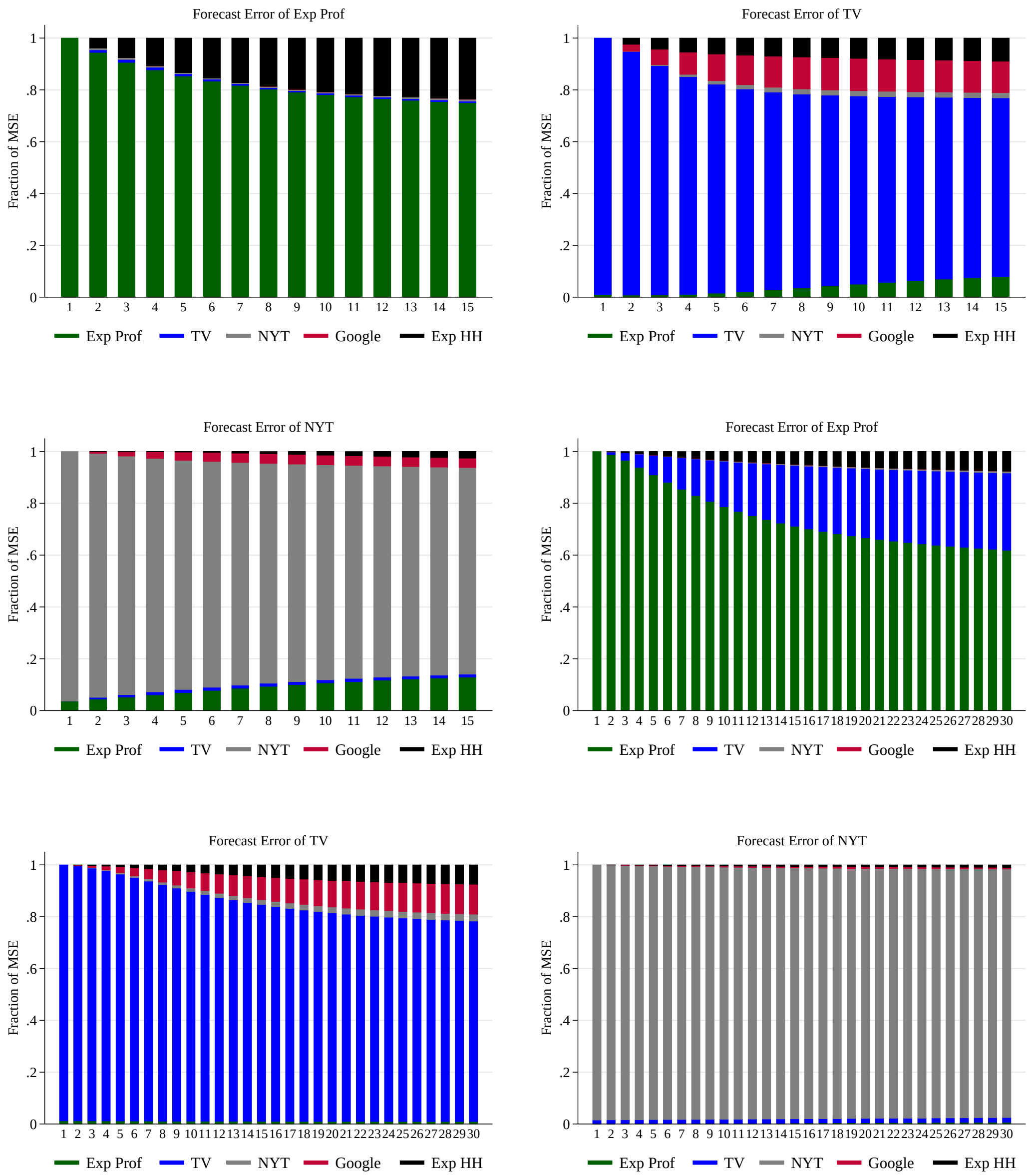


Figure C.7: Large VAR - FEVD - Monthly and Weekly Data



C.2 Literature Overview: Google Search Data in Economic Analysis

Ginsberg et al. (2009) have been the first to introduce the use of Google search requests to scientific work. Using data for the U.S., they show that the spread of influenza infection rates can be predicted by Google search requests such as "indications of flu" two weeks prior to the publication of health surveillance systems.[1] In economics, Choi and Varian (2009b) were the first to employ Google search queries for the purpose of nowcasting retail sales, automotive sales and home sales. Moreover, they successfully used Google data to predict visitor arrivals at Hong Kong. Similarly, Chamberlain (2010) documented the usefulness of Google data to nowcast retail sales in the UK.

In a subsequent paper, Choi and Varian (2009a) employ Google search data to forecast the number of U.S. citizens that subscribe for unemployment benefits each week (initial jobless claims), a figure which is considered as an important leading indicator for the U.S. unemployment rate. Their analysis shows that adding the number of search terms such as "jobs" to an autoregressive model significantly decreases the out-of-sample mean forecast error. Since then, a number of researchers have used internet search data to predict unemployment in different countries. D'Amuri and Marucci (2010) use the initial jobless claims together with the same Google search series of Choi and Varian (2009b) to forecast the U.S. unemployment rate. Their results suggest that the forecasting models using the Google search data show much better forecast performance in terms of mean squared error than models that use only the initial jobless claims. Askitas and Zimmermann (2009) run error-correction models to explain the unemployment rate in Germany. Using the search keywords "unemployment office", "unemployment rate", as well as the names of popular job search engines in Germany, their results suggest a significant and positive impact from Google search requests on future unemployment. Using data for Israel, Suhoy (2009) shows that adding Google search requests helps predict the unemployment rate. Anvik and Gjelstad (2010) employ Google search data for Norway and again find that adding Google search data to an ARMA forecasting model significantly improves forecast performance. The latest study dealing with unemployment is provided by Fondeur and Karamé (2013). Using keyword searches for "emploi" (the French word for "job" and "employment"), the authors also find that Google search data improves the forecast for the youth unemployment rate in France. Most interestingly, this is the only paper at present that does not aggregate the weekly Google series into lower frequencies but keeps the original data by estimating an unobserved component model.

Besides forecasting the unemployment rate, Google search data is also found to have predictive power for private consumption. Kholodilin et al. (2010) show that web query data

[1] In 2009, Google has launched an own web page with a graphical illustration of worldwide search requests for "flu" , see Google Flutrends.

significantly help nowcast U.S. real-time private consumption. The authors also suggest a careful way of constructing the Google series. Selecting first the top ten queries of the 27 main Google search categories, they select all series related to consumption and apply a principal component analysis to merge the large number of Google series into a common factor without loosing too much information or employing too many a priori assumptions for choosing a particular keyword. Schmidt and Vosen (2011) also present empirical evidence that adding Google search data to more traditional models improves the forecast of private consumption in the U.S.. Even if these two papers are fairly similar, they differ with respect to the choice of the categories and the aggregation approach of the weekly Google data, which suggests that the result of the good forecasting performance of Google data does not depend on the specific assumptions with regard to the data collection. Furthermore, Della Penna and Huang (2009) find superior forecasting performance of Google searches for retail sales in the U.S. compared to survey-based consumer sentiment indices by simply adding up single Google series instead of running a principal component analysis. Finally, Schmidt and Vosen (2012) produce supportive evidence for the use of internet search request in nowcasting private consumption in Germany, while Carrière-Swallow and Labbé (2013) get the same finding for car sales in Chile, a remarkable results given the lower internet usage in emerging countries.

Recently, the potential of using Google data for nowcasting economic variables started to be explored in additional areas. Da et al. (2011) analyze the usefulness of search queries to predict share prices. Employing stocks from the Russell 3000 index, they show that Google searches help predict stock returns. While Google searches for different stocks are found to be related to traditional measures of investor attentions such as news reports on individual companies and advertising expenditures, the web-based queries turn out to be the leading variable. McLaren and Shanbhogue (2011), besides documenting good nowcasting performance of internet searches for the UK labor market, show that web queries also help predict house prices. Furthermore, they suggest that the analysis of Google search data might serve as a valuable indicator to assess the general public's judgment of policy measures or consumer sentiment in general. This latter route has been followed by Kahn and Kotchen (2010). Using monthly panel data for U.S. federal states, they show that rising state-level unemployment increases the search requests for "unemployment" and decreases the queries for "global warming".

Bibliography

Akerlof, G. A., W. T. Dickens, and G. L. Perry (1996). The Macroeconomics of Low Inflation. *Brookings Paper on Economic Activity 27*(1996-1), 1–76.

Akerlof, G. A., W. T. Dickens, and G. L. Perry (2000). Near-Rational Wage and Price Setting and the Long-Run Phillips Curve. *Brookings Paper on Economic Activity 31*(2000-1), 1–60.

Akerlof, G. A. and R. J. Shiller (2009). *Animal Spirits: How Human Psychology Drives the Economy, and Why It Matters for Global Capitalism*. Princeton University Press.

Anderson, R., R. Becker, and D. Osborn (2010). Heterogeneity in Consumers' Learning about Inflation. *Working Paper*, 1–27.

Anderson, R., R. Becker, and D. Osborn (2012). Evidence on US Consumer Inflation Forecasting Processes. *Working Paper*, 1–29.

Andrews, D. (1993). Tests for Parameter Instability and Structural Change with Unknown Change Point. *Econometrica 61*, 821–856.

Ang, A., G. Bekaert, and M. Wei (2007). Do Macro Variables, Asset Markets, or Surveys Forecast Inflation Better? *Journal of Monetary Economics 54*(4), 1163–1212.

Anvik, C. and K. Gjelstad (2010). "Just Google it". Forecasting Norwegian Unemployment Figures with Web Queries. *CREAM Publication 11*.

Armantier, O., W. Bruine de Bruin, G. Topa, W. van der Klaauw, and B. Zafar (2012). Inflation Expectations and Behavior: Do Survey Respondents Act on Their Beliefs? *Federal Reserve Bank of New York Staff Reports 509*, 1–51.

Arnold, E. (2013). The Role of Revisions and Uncertainty in Professional Forecasts. *mimeo*.

Askitas, N. and K. Zimmermann (2009). Google Econometrics and Unemployment Forecasting. *Applied Economics Quarterly (formerly: Konjunkturpolitik) 55*(2), 107–120.

Bachl, M. (2008). Wirkung von Wirtschaftsnachrichten - Eine Untersuchung von Medieneffekten auf die Wahrnehmung der Wirtschaftslage in der Bevölkerung. *Hochschule für*

Musik und Theater Hannover (HMTH) Institut für Journalistik und Kommunikationsforschung (IJK) Master-Arbeit.

Bachmann, R., T. Berg, and E. Sims (2012). Inflation Expectations and Readiness to Spend: Cross-Section Evidence. *NBER Working Paper 17958*, 1–41.

Bai, J. (1997). Estimating Multiple Breaks One at a Time. *Econometric Theory 13*, 315–352.

Bai, J. and P. Perron (1998). Estimating and Testing Linear Models with Multiple Structural Changes. *Econometrica 66*, 47–78.

Ball, L. (2013). The Case for Four Percent Inflation. *Central Bank Review 13*, 17–31.

Batchelor, R. A. (1986). The Psychophysics of Inflation. *Journal of Economic Psychology 7*(3), 269–290.

Behr, R. L. and S. Iyengar (1985). Television News, Real-World Cues, and Changes in the Public Agenda. *Public Opinion Quarterly 49*(1), 38–57.

Blanchard, O., G. Dell'Ariccia, and P. Mauro (2010). Rethinking Macroeconomic Policy. *IMF Staff Position Note SPN/10/3*, 1–19.

Blanchflower, D. and C. MacCoille (2009). The Formation of Inflation Expectations: An Empirical Analysis for the UK. *NBER Working Paper 15388*, 1–45.

Blinder, A. and A. Krueger (2004). What Does the Public Know about Economic Policy, and How Does It Know It? *Brookings Paper on Economic Activity* (1), 327–387.

Blinder, A. and R. Reis (2005). Understanding the Greenspan Standard. In *The Greenspan Era: Lessons for the Future*. Jackson Hole: Federal Reserve Bankd of Kansas City.

Bomberger, W. A. (1996). Disagreement as a Measure of Uncertainty. *Journal of Money, Credit and Banking 28*(3), 381–392.

Branch, W. A. (2004). The Theory of Rationally Heterogeneous Expectations: Evidence From Survey Data on Inflation Expectations. *Economic Journal 114*(497), 592–621.

Branch, W. A. (2007). Sticky Information and Model Uncertainty in Survey Data on Inflation Expectations. *Journal of Economic Dynamics and Control 31*(1), 245–276.

Breitung, J. and M. Schmeling (2013). Quantifying Survey Expectations: What's wrong with the Probability Method? *International Journal of Forecasting 29*, 142–154.

Brischetto, A. and G. de Brouwer (1999). Householders' Inflation Expectations. *Reserve Bank of Australia Research Discussion Paper 03*, 1–41.

Bruine de Bruin, W., W. van der Klaauw, J. Downs, B. Fischhoff, G. Topa, and O. Armantier (2010). Expectations of Inflation: The Role of Demographic Variables, Expectation Formation, and Financial Literacy. *The Journal of Consumer Affairs 44*(2), 381–402.

Bruine de Bruin, W., W. van der Klaauw, G. Topa, A. Downs, B. Fischhoff, and O. Armantier (2012). The Effect of Question Wording on Consumers' Reported Inflation Expectations. *Journal of Economic Psychology 33*(4), 749–757.

Bryan, M. and S. Cecchetti (1994). Measuring Core Inflation. In G. Mankiw (Ed.), *Monetary Policy*. University of Chicago Press for NBER.

Bryan, M. and G. Venkatu (2001a). The Curiously Different Inflation Perspectives of Men and Women. *Federal Reserve Bank of Cleveland Commentary 1101*, 1–4.

Bryan, M. and G. Venkatu (2001b). The Demographics of Inflation Opinion Surveys. *Federal Reserve Bank of Cleveland Commentary 1015*, 1–4.

Burke, M. and M. Manz (2011). Economic Literacy and Inflation Expectations: Evidence from a Laboratory Experiment. *Federal Reserve Bank of Boston Public Policy Discussion Papers 11*(8), 1–49.

Calvo, G. (1983). Staggered Prices in a Utility-Maximizing Framework. *Journal of Monetary Economics 12*(3), 383–398.

Capistran, C. and A. Timmermann (2009). Disagreement and Biases in Inflation Expectations. *Journal of Money, Credit and Banking 41*, 365–396.

Carlson, J. and M. Parkin (1975). Inflation Expectations. *Economica 42*(166), 123–138.

Carrière-Swallow, Y. and F. Labbé (2013). Nowcasting with Google Trends in an Emerging Market. *Journal of Forcasting 32*, 289–298.

Carroll, C. D. (2001). The Epidemiology of Macroeconomic Expectations. *NBER Working Paper 8695*, 1–49.

Carroll, C. D. (2003). Macroeconomic Expectations of Households and Professional Forecasters. *Quarterly Journal of Economics 118*(1), 269–298.

Carroll, C. D. (2005). The Epidemiology of Macroeconomic Expectations. In L. E. Blume and S. N. Durlauf (Eds.), *The Economy as an Evolving Complex System, III: Current Perspectives and Future Directions*, pp. 5–30. Oxford: Oxford University Press.

Cavallo, A. (2013). Online and Official Price Indexes: Measuring Argentina's Inflation. *Journal of Monetary Economics 60*, 152–165.

Chamberlain, G. (2010). Googling the Present. *Economic and Labour Market Review 4*(12), 59–95.

Chiu, C. W. J., B. Eraker, A. T. Foerster, K. T. Bong, and H. Seoane (2011). Estimating VAR's Sampled at Mixed or Irregular Spaced Frequencies: A Bayesian Approach. *Federal Reserve Bank of Kansas City Research Working Paper 11-11*, 1–38.

Choi, H. and H. Varian (2009a). Predicting Initial Claims for Unemployment Benefits. *Google Technical Report*, 1–5.

Choi, H. and H. Varian (2009b). Predicting the Present with Google Trends. *Google Technical Report*, 1–23.

Clarida, R., J. Galí, and M. Gertler (1999). The Science of Monetary Policy: A New Keynesian Perspective. *Journal of Economic Literature 37*(4), 1661–1707.

Clemente, J., A. Montanés, and M. Reyes (1998). Testing for a Unit Root in Variables with a Double Chane in the Mean. *Economics Letters 59*, 175–182.

Cohen, B. C. (1963). *The Press and Foreign Policy* (1 ed.). Princeton, New Jersey: Princeton University Press.

Coibion, O. and Y. Gorodnichenko (2012). What Can Survey Forecasts Tell Us About Information Rigidities? *Journal of Political Economy 120*, 116–159.

Colavecchio, R., U. Fritsche, and M. Graff (2011). Inflation Inequality in Europe. *DEP Discussion Papers Macroeconomics and Finance Series 2*, 1–38.

Collado, M. D. (1997). Estimating Dynamic Models from Time Series of Independent Cross-Sections. *Journal of Econometrics 82*, 37–62.

comScore (2012). U.S. Search Enginge Rankings. http://www.comscore.com/Press_Events/Press_Releases/2012/5/comScore_Releases_April_2012_U.S._Search_Engine_Rankings.

Croushore, D. (1993). Introducing: The Survey of Professional Forecasters. *Business Review, Federal Reserve Bank of Philadelphia Nov*, 3–15.

Curtin, R. (1996). Procedure to Estimate Price Expectations. *University of Michigan Working Paper*, 1–56.

Da, Z., J. Engelberg, and P. Gao (2011). In Search of Attention. *Journal of Finance 66*, 1461–1499.

D'Amuri, F. and J. Marucci (2010). "Google it!" Forecasting the U.S. Unemployment Rate with a Google Job Search Index. *Fondazione Eni Enrico Mattei*.

Dearing, J. W. (1989). Setting the Polling Agenda for the Issue of AIDS. *Public Opinion Quarterly 53*(3), 309–329.

Deaton, A. (1985). Panel Data from Time Series of Cross-Sections. *Journal of Econometrics 30*, 109–126.

Della Penna, N. and H. Huang (2009). Constructing Consumer Sentiment Index for U.S. Using Internet Search Patterns. *University of Alberta Working Paper 26*, 1–22.

Demers, D. P., D. Craff, Y.-H. Choi, and B. M. Pessin (1989). Issue Obtrusiveness and the Agenda-Setting Effects of National Network News. *Communication Research 16*(6), 793–812.

Dickey, D. and W. Fuller (1979). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association 74*(366), 427–431.

Doepke, M. and M. Schneider (2006). Inflation and the Redistribution of Nominal Wealth. *Journal of Political Economy 114*(6), 1069–1097.

Döpke, J., J. Dovern, U. Fritsche, and J. Slacalek (2008). The Dynamics of European Inflation Expectations. *B. E. Journal of Macroeconomics 8*(1), 1–21.

Dovern, J., U. Fritsche, P. Loungani, and N. Tamirisa (2013). Information Rigidities in Economic Growth Forecasts: Evidence from a Large International Panel. *IMF Working Paper 13/56*, 1–24.

Dovern, J., U. Fritsche, and J. Slacalek (2012). Disagreement Among Forecasters in G7 Countries. *Review of Economics and Statistics 94*(4), 1081–1096.

Dräger, L. (2011). Inflation Perceptions and Expectations in Sweden - Are Media Reports the 'Missing Link'? *KOF Working Papers 273*, 1–51.

Dräger, L. and U. Fritsche (2013). Don't Worry, Be Right! Survey Wording Effects of Inflation Perceptions and Expectations. *DEP Discussion Paper 8/2013*, 1–32.

Dräger, L. and M. Lamla (2013a). Anchoring of Consumers' Inflation Expectations: Evidence from Microdata. *KOF Working Paper 339*, 1–24.

Dräger, L. and M. Lamla (2013b). Imperfect Information and Inflation Expectations: Evidence from Microdata. *KOF Working Paper 329*, 1–38.

Dräger, L., M. Lamla, and D. Pfajfar (2013). Are Consumer Expectations Theory-Consisten? The Role of Macroeconomic Determinants and Central Bank Communication. *KOF Working Paper 345*, 1–36.

Dräger, L., J.-O. Menz, and U. Fritsche (2009). Prospect Theory and Inflation Perceptions - An Empirical Assessment. *DEP Discussion Paper 3/2009*, 1–80.

Dräger, L., J.-O. Menz, and U. Fritsche (2014). Perceived Inflation under Loss Aversion. *Applied Economics 46*, 282–293.

Edelman, B. (2012). Using Internet Data for Economic Research. *Journal of Economic Perspectives 26*(2), 189–206.

Eggertsson, G. and M. Woodford (2003). The Zero Bound on Interest Rates and Optimal Monetary Policy. *Brookings Papers on Economic Activity 34*, 139–235.

Ejsing, J., J. A. García, and T. Werner (2007). The Term Structure of Euro Area Break-Even Inflation Rates: The Impact of Seasonality. *ECB Working Paper 830*, 1–39.

Elliott, G., T. Rothenberg, and J. Stock (1996). Efficient Tests for an Autoregressive Unit Root. *Econometrica 64*(4), 813–836.

Engle, R. and C. Granger (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica 55*(2), 251–276.

Evans, G. and Honkapohja (2001). *Learning and Expectations in Macroeconomics*. Princeton University Press.

Experian Hitwise (2012). Search Engine Trends. http://www.experian.com/hitwise/online-trends-search-engine.html.

Fishe, R. and T. Idson (1990). Information-Induced Heteroscedasticity in Price Expectations Data. *The Review of Economic and Statistics 72*(2), 304–312.

Fondeur, Y. and F. Karamé (2013). Can Google Data Help Predict French Unemployment? *Economic Modeling 30*, 117–125.

Friedman, M. (1957). *A Theory of the Consumption Function*. Princeton University Press.

Fritsche, U., J.-O. Menz, and C. Pierdzioch (2014). Inflation Zone Targeting and Skewed Inflation Expectations. *Deutsche Bundesbank Discussion Paper*.

Galí, J. (2008). *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*. Princeton, New Jersey: Princeton University Press.

Gallo, G., C. Granger, and Y. Jeon (2002). Copycats and Common Swings: The Impact of the Use of Forecasts in Information Sets. *IMF Staff Papers 49*(1), 4–21.

Garz, M. (2013). *Economic Aspects of Information Processing in the Case of Labor Market News*. Ph. D. thesis, University of Hamburg.

Gentzkow, M. and J. Shapiro (2010). What Drives Media Slant? Evidence from U.S. Daily Newspapers. *Econometrica 78*(1), 35–71.

Gertchev, N. (2007). A Critique of Adaptive and Rational Expectations. *Quarterly Journal of Austrian Economics 10*, 313–329.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2005). There is a Risk-Return Trade-Off After All. *Journal of Financial Economics 76*, 509–548.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2006). Predicting Volatility: Getting the Most Out of Return Data Sampled ad Different Frequencies. *Journal of Econometrics 131*(1-2), 59–95.

Giannetti, M. and Y. Koskinen (2010). Investor Protection, Equity Returns, and Financial Globalization. *Journal of Financial and Quantitative Analysis 45*, 135–168.

Ginsberg, J., M. Mohebbi, R. Patel, L. Brammer, M. Smolinski, and L. Brilliant (2009). Detecting Influenza Epidemics Using Search Engine Query Data. *Nature 457*, 1012–1015.

Glocker, D. and V. Steiner (2007). Self-Employment - A Way to End Unemployment?: Empirical Evidence from German Pseudo-Panel Data. *DIW Discussion Paper 661*, 1–28.

Gnan, E., J. Langthaler, and M. T. Valderrama (2011). Heterogeneity in Euro Area Consumers' Inflation Expectations: Some Stylized Facts and Implications. *Monetary Policy and the Economy Q2/11*, 43–66.

Goidel, R. K. and R. E. Langley (1995). Media Coverage of the Economy and Aggregate Economic Evaluations: Uncovering Evidence of Indirect Media Effects. *Political Research Quarterly 48*(2), 313–328.

Gonzalez, A., T. Teräsvirta, and D. van Dijk (2005). Panel Smooth Transition Regression Models. *SSE/EFI Working Paper Series in Economics and Finance 604*, 1–34.

Google Inc. (2008). Announcing Google Insights for Search. *Google's official blog for news information and tips on AdWords http://adwords.blogspot.com/2008/08/announcing-google-insights-for-search.html*.

Granka, L. (2010). Measuring Agenda Setting with Online Search Traffic: Influences of Online and Traditional Media. *Paper prepared for the Annual Meeting of the American Political Science Association, September 2-5, 2010*.

Greene, W. (2003). *Econometric Analysis*. Uppersaddle River, NJ: Prenctice-Hall.

Gregory, A. and B. Hansen (1996a). Residual-Based Tests for Cointegration in Models with Regime Shifts. *Journal of Econometrics 70*, 99–126.

Gregory, A. and B. Hansen (1996b). Tests for Cointegration in Models with Regime and Trend Shift. *Oxford Bulletin of Economics and Statistics 58*(3), 555–560.

Gregory, A., J. Nason, and D. Watt (1996). Testing for Structural Beaks in Cointegrated Relationships. *Journal of Econometrics 71*, 321–341.

Groen, J. and M. Middeldorp (2013). Creating a History of Inflation Expectations. *Federal Reserve Bank of New York Liberty Street Economics*.

Guzmán, G. (2011). Internet Search Behavior as an Economic Forecasting Tool: The Case of Inflation Expectations. *Journal of Economic and Social Measurement 36*, 119–167.

Hagemann, R. P. (1982). The Variability of Inflation Rates across Household Types. *Journal of Money, Credit and Banking 14*(4 (1)), 494–510.

Hagen, L. M. (2005). *Konjunkturnachrichten, Konjunkturklima und Konjunktur: Wie sich die Wirtschaftsberichterstattung der Massenmedien, Stimmungen der Bevölkerung und die aktuelle Wirtschaftslage wechselseitig beeinflussen - eine transaktionale Analyse* (1 ed.). Cologne: Herbert von Halem Verlag.

Hansen, B. (2001). The New Econometrics of Structural Change: Dating Breaks in U.S. Labor Productivity. *Journal of Economic Perspectives 15*(4), 117–128.

Harrington, D. E. (1989). Economic News on Television: The Determinants of Coverage. *Public Opinion Quarterly 53*(1), 17–40.

Hartley, J. (1997). *The Representative Agent in Macroeconomics*. Routledge Chapman & Hall.

Hobijn, B. and D. Lagakos (2005). Inflation Inequality in the United States. *Review of Income and Wealth 51*(4), 581–606.

Hommes, C. (2006). Heterogeneous Agent Models in Economics and Finance. In L. Tesfatsion and K. Judd (Eds.), *Handbook of Computational Economics*, Volume 2, Chapter 23, pp. 1109–1186. North-Holland.

Inoue, A., L. Kilian, and F. Burcu Kiraz (2009). Do Actions Speak Louder Than Words? Household Expectations of Inflation Based on Micro Consumption Data. *Journal of Money, Credit, and Banking 41*(7), 1331–1363.

Ivanaov, V. and L. Kilian (2005). A Practitioner's Guide to Lag Order Selection for VAR Impulse Response Analysis. *Studies in Nonlinear Dynamics & Econometrics 9*, 1–34.

Jonung, L. (1981). Perceived and Expected Rates of Inflation in Sweden. *American Economic Review 71*(5), 961–968.

Ju, Y. (2008). The Asymmetry in Economic News Coverage and its Impact on Public Perception in South Korea. *International Journal of Public Opinion Research 20*(2), 237–249.

Kahn, M. and M. Kotchen (2010). Environmental Concern and the Business Cycle: The Chilling Effect of Recession. *NBER Working Paper 16241*, 1–28.

Kahneman, D. (2011). *Thinking, Fast and Slow*. Allen Lane: Penguin Books.

Kahneman, D. and A. Tversky (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica 47*(2), 263–291.

Kaldor, N. (1934). A Classificatory Note on the Determination of Equilibrium. *Review of Economic Studies 1*, 122–136.

Kholodilin, K., M. Podstawski, and B. Siliverstovs (2010). Do Google Searches Help in Nowcasting Private Consumption? A Real-Time Evidence fo the U.S. *DIW Discussion Papers 997*.

Kirchgässner, G. and J. Wolters (2007, October). *Introduction to Modern Time Series Analysis*. Berlin: Springer-Verlag.

Krause, B. and V. Gehrau (2007). Das Paradox der Medienwirkung auf Nichtnutzer - Eine Zeitreihenanalyse auf Tagesbasis zu den kurzfristigen Agenda-Setting-Effekten von Fernsehnachrichten. *Publizistik 52*(2), 191–209.

Kwiatkowski, D., P. C. B. Phillips, P. Schmidt, and Y. Shin (1992). Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure are we that Economic Time Series have a Unit Root? *Journal of Econometrics 54*(1-3), 159–178.

Lamla, M. and S. Lein (2010). The Role of Media for Consumers' Inflation Expectation. *Working Paper*, 1–39.

Lamla, M. and T. Maag (2012). The Role of Media for Inflation Forecast Disagreement of Households and Professional Forecasters. *Journal of Money, Credit and Banking 44*(7), 1325–1350.

Lamla, M. and S. Sarferaz (2012). Updating Inflation Expectations. *KOF Swiss Economic Institute Working Paper 301*, 1–39.

Lanne, M., A. Luoma, and J. Luoto (2009). A Naïve Sticky Information Model of Households' Inflation Expectations. *Journal of Economic Dynamics & Control 33*, 1332–1344.

Larcinese, V., R. Puglisi, and J. Snyder (2011). Partisan Bias in Economic News: Evidence on the Agenda-Setting Behavior of U.S. Newspapers. *Journal of Public Economics 95*, 1178–1189.

Leung, C. (2009). The Demographics of Household Inflation Perceptions and Expectations. *Reserve Bank of New Zealand Bulletin 72*(2), 34–42.

Lucas, R. (1976). Econometric Policy Evaluation: A Critique. *Carnegie-Rochester Conference Series on Public Policy 1*(1), 19–46.

Luoma, A. and J. Luoto (2009). Modelling the General Public's Inflation Expectations Using the Michigan Survey Data. *Applied Economics 41*, 1311–1320.

Lusardi, A. and O. Mitchell (2008). Planning and Financial Literacy: How Do Women Fare? *American Economic Review 98*(2), 413–417.

Lütkepohl, H., T. Teräsvirta, and J. Wolters (1999). Investigating Stability and Linearity of a German M1 Money Demand Function. *Journal of Applied Econometrics 14*, 511–525.

MacKinnon, J. (2010). Critical Values of Cointegration Tests. *Queen's Economics Department Working Paper No. 1227*, 1–17.

Malmendier, U. and S. Nagel (2013). Learning from Inflation Experiences. *Working Paper*, 1–54.

Mankiw, N. G. and R. Reis (2002). Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve. *Quarterly Journal of Economics 117*(4), 1295–1328.

Mankiw, N. G. and R. Reis (2003). What Measure of Inflation Should A Central Bank Target? *Journal of the European Economic Association 1*(5), 1058–1086.

Mankiw, N. G. and R. Reis (2006). Pervasive Stickiness. *American Economic Review 96*(2), 164–169.

Mankiw, N. G. and R. Reis (2007). Sticky Information in General Equilibrium. *Journal of the European Economic Association 5*(2-3), 603–613.

Mankiw, N. G., R. Reis, and J. Wolfers (2003). Disagreement about Inflation Expectations. *NBER Macroeconomics Annual 18*, 209–248.

Manski, C. F. (2004). Measuring Expectations. *Econometrica 72*(5), 1329–1376.

Matthes, J. (2006). The Need for Orientation Towards News Media: Revising and Validating a Classic Concept. *International Journal of Public Opinion Research 18*(4), 422–444.

Maurer, M. (2004). Das Paradox der Medienwirkungsforschung. *Publizistik 49*(4), 405–422.

McCombs, M. (2004). *Setting the Agenda: The Mass Media and Public Opinion*. Polity Press.

McCombs, M. E. and D. L. Shaw (1972). The Agenda-Setting Function of Mass Media. *Public Opinion Quarterly 36*(2), 176–187.

McGranahan, L. and A. Paulson (2006). Constructing the Chicago Fed Income Based Economic Index - Consumer Price Index: Inflation Experiences by Demographic Group: 1983-2005. *Federal Reserve Bank of Chicago Working Paper 2005-20*, 1–104.

McKenzie, D. (2004). Asymptotic Theory for Heterogeneous Dynamic Pseudo-Panels. *Journal of Econometrics 120*, 235–262.

McLaren, N. and R. Shanbhogue (2011). Using Internet Search Data as Economic Indicators. *Bank of England Quarterly Bulletin 51*(2), 134–140.

Menz, J.-O. (2012). Analyzing the News Content of Media Reports on Inflation. *mimeo*.

Menz, J.-O. (2013). Unfinished Business in the Epidemiology of Inflation Expectations. *mimeo*.

Menz, J.-O. and P. Poppitz (2013). Households' Disagreement on Inflation Expectations and Socioeconomic Media Exposure in Germany. *Deutsche Bundesbank Discussion Paper 27/2013*, 1–51.

Michael, R. T. (1979). Variation Across Households in the Rate of Inflation. *Journal of Money, Credit and Banking 11*(1), 32–46.

Modigliani, F. and R. Brumberg (1954). Utility Analysis and the Consumption Function: An Interpretation of Cross-Section Data. In K. Kurihara (Ed.), *Postkeynesian Economics*, pp. 388–436. Rutgers University Press.

Moffitt, R. (1993). Identification and Estimation of Dynamic Models with a Time Series of Repeated Cross-Sections. *Journal of Econometrics 59*, 99–123.

Mortensen, D. and C. Pissarides (1994). Job Creation and Job Destruction in the Theory of Unemployment. *Review of Economic Studies 61*(3), 397–415.

Mullainathan, S. and A. Shleifer (2005). The Market for News. *American Economic Review 95*(4), 1031–1053.

Muth, J. (1961). Rational Expectations and the Theory of Price Movements. *Econometrica 29*, 315–335.

Nardo, M. (2003). The Quantification of Qualitative Survey Data: A Critical Assessment. *Journal of Economic Surveys 17*(5), 645–668.

Nerlove, M. (1958). Adaptive Expectations and Cobweb Thenomena. *Quarterly Journal of Economics 72*, 227–240.

Ng, S. and P. Perron (1995). Unit Root Tests in ARMA Models with Data-Dependent Methods for the Selection of the Truncation Lag. *Journal of the American Statistical Association 90*, 268–281.

Ng, S. and P. Perron (2000). Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power. *Econometrica 69*, 1519–1554.

Nielsen, H. (2003). Inflation Expectations in the EU - Results from Survey Data. *Humboldt-Universität zu Berlin, Sonderforschungsbereich 373: Quantification and Simulation of Economic Processes 13*, 1–26.

Ottaviani, M. and P. N. Sorensen (2006). The Strategy of Professional Forecasting. *Journal of Financial Economics 81*, 441–466.

Paciello, L. and M. Wiederholt (forthcoming). Exogenous Information, Endogenous Information and Optimal Monetary Policy. *Review of Economic Studies*.

Palmqvist, S. and L. Strömberg (2004). Households' Inflation Opinions - A Tale of Two Surveys. *Sveriges Riksbank Economic Review 2004:4*, 23–42.

Perron, P. (1989). The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis. *Econometrica 57*, 1362–1401.

Perron, P. (1990). Testing for a Unit Root in a Time Series with a Changing Mean. *Journal of Business and Economic Statistics 8*, 153–162.

Perron, P. and T. Vogelsang (1992). Nonstationarity and Level Shifts with an Application to Purchasing Power Parity. *Journal of Business and Economic Statistics 10*, 301–320.

Pesaran, H. (1987). *The Limits to Rational Expectations*. Blackwell.

Pfajfar, D. and E. Santoro (2009). Asymmetries in Inflation Expectations across Sociodemographic Groups. *Working Paper*, 1–29.

Pfajfar, D. and E. Santoro (2013). News on Inflation and the Epidemiology of Inflation Expectations. *Journal of Money, Credit and Banking*.

Phillips, P. and S. Ouliaris (1990). Asymptotic Properties of Residual Based Tests for Cointegration. *Econometrica 58*(1), 165–193.

Phillips, P. C. B. and P. Perron (1988). Testing for a Unit Root in Time Series Regression. *Biometrika 75*(2), 335–346.

Price, V. (1988). On the Public Aspects of Opinion. Linking Levels of Analysis in Public Opinion Research. *Communication Research 15*, 659–679.

Prior, M. (2009). The Immensely Inflated News Audience: Assessing Bias in Self-Reported News Exposure. *Public Opinion Quarterly 73*, 130–143.

Quandt, R. (1960). Tests of the Hypothesis That a Linear Regression System Obeys Two Separate Regimes. *Journal of the American Statistical Association 55*, 324–330.

Quiring, O. (2004). *Wirtschaftsberichterstattung und Wahlen*. Konstanz: UVK Verlagsgesellschaft mbH.

Ranyard, R., F. Del Missier, N. Bonini, D. Duxbury, and B. Summers (2008). Perceptions and Expectations of Price Changes and Inflation: A Review and Conceptual Framework. *Journal of Economic Psychology 29*, 378–400.

Reis, R. (2006). Inattentive Consumers. *Journal of Monetary Economics 53*(8), 1761–1800.

Roberts, J. M. (1997). Is Inflation Sticky? *Journal of Monetary Economics 39*(2), 173–196.

Roberts, J. M. (1998). Inflation Expectations and the Transmission of Monetary Policy. *Finance and Economics Discussion Series Board of Governors of the Federal Reserve System 43*, 1–38.

Rössler, P. (1999). The Individual Agenda-Designing Process: How Interpersonal Communication, Egocentric Networks, and Mass Media Shape the Perception of Political Issues by Individuals. *Communication Research 26*(6), 666–700.

Schmidt, T. and S. Vosen (2011). Forecasting Private Consumption: Survey-Based Indicators vs. Google Trends. *Journal of Forecasting 30*(6), 565–578.

Schmidt, T. and S. Vosen (2012). A Monthly Consumption Indicator for Germany Based on Internet Serach Query Data. *Applied Economics Letters 19*(7), 683–687.

Schmitt-Grohe, S. and M. Uribe (2013). The Making Of A Great Contraction With A Liquidity Trap and A Jobless Recovery. *Working Paper*.

Schoenbach, K., E. Lauf, J. McLeod, and D. Scheufele (1999). Distinction and Integration: Sociodemographic Determinants of Newspaper Reading in the USA and Germany, 1974-96. *European Journal of Communication 14*(2), 225–239.

Schulz, A. and J. Stapf (2009). Price Discovery on Traded Inflation Expectations: Does the Financial Crisis Matter? *Deutsche Bundesbank Discussion Paper 25/2009*, 1–33.

Schwert, G. (1989). Tests for Unit Roots: A Monte Carlo Investigation. *Journal of Business and Economic Statistics 2*, 147–159.

Sims, C. (2009). Inflation Expectations, Uncertainty, and Monetary Policy. *BIS Working Papers 275*, 1–29.

Sims, C. A. (2003). Implications of Rational Inattention. *Journal of Monetary Economics 50*(3), 665–690.

Soroka, S. N. (2002). Issue Attributes and Agenda-Setting by Media, the Public, and Policymakers in Canada. *International Journal of Public Opinion Research 14*(3), 264–285.

Soroka, S. N. (2006). Good News and Bad News: Asymmetric Responses to Economic Information. *Journal of Politics 68*(2), 372–385.

Souleles, N. S. (2004). Expectations, Heterogenous Forecast Errors, and Consumption: Micro Evidence from the Michigan Consumer Sentiment Surveys. *Journal of Money, Credit, and Banking 36*(1), 40–72.

Stock, J. and M. Watson (2007). *Introduction to Econometrics*. Addison-Wesley.

Suhoy, T. (2009). Query Indices and a 2008 Downturn: Israeli Data. *Bank of Israel Discussion Paper 6*, 1–33.

Teräsvirta, T. (2004). Smooth Transition Regression Modeling. In H. Lütkepohl and M. Krätzig (Eds.), *Applied Time Series Econometrics*, Themes in Modern Economics, Chapter 6, pp. 222–242. Cambridge: Cambridge University Press.

Thaler, R. (1980). Toward a Positive Theory of Consumer Choice. *Journal of Economic Behavior and Organization 1*(1), 39–60.

Thomas, L. (1999). Survey Measures of Expected U. S. Inflation. *Journal of Economic Perspectives 13*(4), 125–144.

Traut-Mattausch, E., T. Greitemeyer, D. Frey, and S. Schulz-Hardt (2007). Illusory Price Increases after the Euro Changeover in Germany: An Expectancy-Consistent Bias. *Journal of Consumer Policy 30*(4), 421–434.

Traut-Mattausch, E., S. Schulz-Hardt, T. Greitemeyer, and D. Frey (2004). Expectancy Confirmation in Spite of Disconfirming Evidence: The Case of Price Increases Due to the Introduction of the Euro. *European Journal of Social Psychology 34*(6), 739–760.

Turnovsky, S. (1970). Empirical Evidence on the Formation of Price Expectations. *Journal of the American Statistical Association 65*(332), 1441–1454.

Tversky, A. and D. Kahneman (1973). Availability: A Heuristic for Judging Frequency and Probability. *Cognitive Psychology 5*(2), 207–232.

Varian, H. (2010). Computer Mediated Transactions. *American Economic Review 100*(2), 1–10.

Verbeek, M. (2012). *A Guide to Modern Econometrics*. John Wiley & Sons, Ltd.

Vliegenthart, R. and S. Walgrave (2008). The Contingency of Intermediate Agenda Setting: A Longitudinal Study in Belgium. *Journalism & Mass Communication Quarterly 85*(4), 860–877.

Walsh, C. E. (2003). *Monetary Theory and Policy* (2 ed.). Cambridge and London: MIT Press.

Wiederholt, M. (2013). Dispersed Inflation Expectations and the Zero Lower Bound. *Working Paper*, 1–17.

Winter, J. P., C. H. Eyal, and A. H. Rogers (1982). Issue-Specific Agenda-Setting: The Whole as Less Than the Sum of the Parts. *Canadian Journal of Communication 8*(2), 1–10.

Woodford, M. (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy* (1 ed.). Princeton, New Jersey: Princeton University Press.

Zhu, J.-H., J. H. Watt, L. B. Snyder, J. Yan, and Y. Jiang (1993). Public Issue Priority Formation: Media Agenda-Setting and Social Interaction. *Journal of Communication 43*(1), 8–29.

Zivot, E. and D. Andrews (1992). Further Evidence on the Great Crash, the Oil-Price shock, and the Unit-Root Hypothesis. *Journal of Business & Economic Statistics 10*(3), 251–270.

Zucker, H. G. (1978). The Variable Nature of News Media Influence. In B. D. Rubin (Ed.), *Communication Yearbook*, Volume 2, pp. 225–245. New Brunswick, New Jersey: Transaction.

Declarations

List of Individual Papers

Chapter (3) is based on Menz and Poppitz (2013) and was largely written while Phillip Poppitz has been working at the chair of Ulrich Fritsche. Phillip Poppitz has collected the various data sources that have been necessary for conducting the analysis of this paper. Furthermore, he has compiled Stata codes performing the estimation. The idea of this paper, the choice of the econometric techniques, the compilation of the literature survey and the revision of the earlier draft during the process of publishing our work as a Deutsche Bundesbank Discussion Paper were my sole responsibility.

Eidesstattliche Versicherung

Hiermit erkläre ich, Jan-Oliver Menz, an Eides statt, dass ich die Dissertation mit dem Titel "Media Reports and Inflation Expectations" selbständig und ohne fremde Hilfe verfasst habe. Andere als die von mir angegebenen Quellen und Hilfsmittel habe ich nicht benutzt. Die den herangezogenen Werken wörtlich oder sinngemäß entnommenen Stellen sind als solche gekennzeichnet. Es wurde keine kommerzielle Promotionsberatung in Anspruch genommen.

Ich habe mich anderweitig noch keiner Doktorprüfung unterzogen oder um Zulassung zu einer solchen beworben. Diese Dissertation hat noch keiner Fachvertreterin, keinem Fachvertreter und keinem Prüfungsausschuss einer anderen Hochschule vorgelegen; sie wurde nicht schon einmal in einem früheren Promotionsverfahren angenommen oder als ungenügend beurteilt.

Ort/Datum: Hamburg, 17.Juli 2014

Unterschrift: Jan-Oliver Menz

* Gemäß § 6 Abs. 4 der Promotionsordnung der Fakultät Wirtschafts- und Sozialwissenschaften der Universität Hamburg vom 24. August 2010.