PhD THESIS DECLARATION

I, the undersigned

FAMILY NAME	Angelico
NAME	Cristina
Student ID no.	1824325

Thesis title:

Essays on Household Beliefs and Behavior

PhD in	Economics and Finance
Cycle	30th
Student's Advisor	Prof. Nicola Gennaioli
Calendar year of thesis defence	2019

DECLARE

under my responsibility:

1) that, according to Italian Republic Presidential Decree no. 445, 28th December 2000, mendacious declarations, falsifying records and the use of false records are punishable under the Italian penal code and related special laws. Should any of the above prove true, all benefits included in this declaration and those of the temporary "embargo" are automatically forfeited from the beginning;

2) that the University has the obligation, according to art. 6, par. 11, Ministerial Decree no. 224, 30th April 1999, to keep a copy of the thesis on deposit at the "Biblioteche Nazionali Centrali" (Italian National Libraries) in Rome and Florence, where consultation will be permitted, unless there is a temporary "embargo" protecting the rights of external bodies and the industrial/commercial exploitation of the thesis;

3) that the Bocconi Library will file the thesis in its "Archivio Istituzionale ad Accesso Aperto" (Institutional Registry) which permits online consultation of the complete text (except in cases of temporary "embargo");

4) that, in order to file the thesis at the Bocconi Library, the University requires that the thesis be submitted online by the student in unalterable format to Società NORMADEC (acting on behalf of the University), and that NORMADEC will indicate in each footnote the following information:

- PhD thesis *Essays on Household Beliefs and Behavior*;
- by Cristina Angelico;
- defended at Università Commerciale "Luigi Bocconi" – Milano in the year 2019;
- the thesis is protected by the regulations governing copyright (Italian law no. 633, 22nd April 1941 and subsequent modifications). The exception is the right of Università Commerciale "Luigi Bocconi" to reproduce the same, quoting the source, for research and teaching purposes;

- **only when a separate "Embargo" Request has been undersigned**: the thesis is subject to "embargo" for (indicate duration of the "embargo") .... months;

5) that the copy of the thesis submitted online to Normadec is identical to the copies handed in/sent to the members of the Thesis Board and to any other paper or digital copy deposited at the University offices, and, as a consequence, the University is absolved from any responsibility regarding errors, inaccuracy or omissions in the contents of the thesis;

6) that the contents and organization of the thesis is an original work carried out by the undersigned and does not in any way compromise the rights of third parties (Italian law, no. 633, 22nd April 1941 and subsequent integrations and modifications), including those regarding security of personal details; therefore the University is in any case absolved from any responsibility whatsoever, civil, administrative or penal, and shall be exempt from any requests or claims from third parties;

7) **choose 7a or 7b**:

7a) that the thesis is not subject to "embargo", i.e. that it is not the result of work included in the regulations governing industrial property; it was not written as part of a project financed by public or private bodies with restrictions on the diffusion of the results; is not subject to patent or protection registrations.

or

7b) that the PhD thesis is subject to "embargo" as per the separate undersigned "PhD Thesis Temporary "Embargo" Request".

Date 16/04/2019,

Cristina Angelico

Abstract

This thesis is composed of three Chapters, each one containing an essay on household beliefs and behaviours. In Chapter 1, I investigate the role of expectations in the household credit cycle. First, I provide empirical evidence that survey data on expectations have strong predictive power for the dynamics of household debt. Optimism on future income predicts an increase in credit, in line with the permanent income hypothesis. Second, I show that beliefs depart from rationality at the aggregate level in a way coherent with the hypothesis of natural expectations (Fuster et al. (2010)). Then, I provide a tractable model that accounts for these two pieces of evidence and investigate the implications of non-rational expectations on credit. I study a consumption-savings model in which a representative agent has natural expectations. The Chapter shows that in this framework a positive shock to income generates a boom-bust cycle in debt as observed in the data. Overall, the model's predictions match well the positive correlation between debt and income which cannot be captured with rational expectations nor with alternative beliefs′ hypotheses. Chapter 2, co-authored with Federica Di Giacomo, (PhD Candidate at Tor Vergata University), suggests a new channel to explain the dispersion in inflation expectations across income groups, using personal shopping experiences. We propose a model of intuitive reasoning in which the agents use private and public signals to predict future prices. Exploiting data on personal spending from the Kilts-Nielsen Consumer Panel and data on households' beliefs from the University of Michigan Survey of Consumers, we empirically test the model using the inflation rates computed at the household level as a proxy for the private signals and the US CPI as a proxy for the public one. We find that agents look at both personal experience and US aggregate inflation when forming their expectations. Further, agents are particularly affected by salient goods. We conclude that the heterogeneity in inflation expectations derives from the heterogeneity in the private signals and in the way agents process the multiple pieces of information. Finally, Chapter 3 analyses how economic background, rights, and cultural norms affect parenting attitudes towards girls and in turn women access to education and gender equality. I present an overlapping generation model in which the socioeconomic environment determines the equilibrium parenting style, which in turn feeds back into daughters′ ambition, investment in schooling, and participation in the labour market. In equilibrium, parents exclude their daughters from the formal labour market when the cost of emancipation is high (i.e. religiosity, stereotypes). Parents instead boost their daughters' ambition if the discrimination is low and benefits from schooling are high. The Chapter concludes showing that the theoretical predictions are in line with the cross-country empirical evidence.

Chapter 1

Household Expectations and the Credit Cycle

1.1 Introduction

The financial crisis has revealed the need to rethink the role of household debt in macroeconomic models, which is especially important given the dramatic expansion in the household credit to GDP ratio over the last 50 years in advanced economies. Recent studies provide empirical evidence on the relationship between credit and economic activity. Mian et al. (2017) show that a rise in the household debt to GDP ratio predicts lower output growth over the medium term, and Jorda et al. (2013) document that intensive credit expansions tend to be followed by deep recessions and slow recoveries. These results inform that policy-makers should look at household credit and incorporate it into a broader policy framework; but what drives credit boom-bust episodes?

The existing literature focuses intensely on the role of credit supply shocks. In particular, researchers in macro-finance rely on rational expectations and exogenous credit supply shocks so that expansions in credit supply lead to an increase in debt and consumption, whereas credit tightening to a reduction in credit growth. This literature shows that variations in the supply of funds play an important part in the dynamics of the credit cycle, but is there something besides than these models are missing? In this paper, I study the role of households' expectations in the demand side and propose a complementary channel to the supply side view. Building on the principle that expectations influence decisions, optimism may generate an expansion in credit and consumption and lead to a crisis when such confidence declines. Looking at the data for the period

1978-2016, indeed I find that the personal saving rate tends to decrease when the Consumer Confidence Index is high and increases when the index drops. On the contrary, the Consumer Confidence Index and the Total Consumer Credit, flow, appear to follow the same pattern (Figure 1.1).[1] Most importantly these correlations survive when controlling for proxies of the supply of credit, suggesting that consumers' confidence can explain some residual variation in household debt that cannot be accounted by the supply side (Figure 1.1, lower Panel).[2] In this paper, I aim at explaining this residual correlation. How expectations are formed and how they affect macroeconomic dynamics remain open questions that I aim to address in this paper.

The novelty of the paper is to investigate the role of expectations in the household credit cycle regardless of credit frictions, heterogeneity, house prices, or credit supply. The study focuses on beliefs on future income and consumer credit, defined as the total credit extended to individuals, excluding loans secured by real estate. The aim is indeed to explain the reasons behind the sharp increase in consumer credit in the last decades beyond credit supply and to understand why during booms consumers increase the demand for debt in the first place (Figure 1.2). My contribution to the literature is threefold. First, I provide some empirical evidence that expectations data on future economic activity are strong predictors of household debt. Survey data on expectations are useful to understand the dynamics of consumer credit and are not meaningless noise. The predictions are in line with the permanent income hypothesis so that optimism on future economic growth predicts an increase in household debt. Second, using survey data, I investigate how beliefs are formed and show that expectations depart from the rational benchmark displaying rigidity at the aggregate level in a way coherent with the theory of natural expectations, as proposed by Fuster et al. (2010). Natural expectations imply that agents have wrong beliefs about the real process of the fundamental and over-estimate the persistence of the process. Finally, I develop a theoretical model that accounts for both pieces of the empirical evidence that allows investigating the im-

[1]The upper panel of Figure 1.1 shows that there is a negative and significant correlation between the Personal saving rate and the Consumer Confidence Index (-0.5***) and a positive one between the Total Consumer Credit flow and the Consumer Confidence Index (0.45***).

[2]The lower panel of Figure 1.1 plots the Consumer Confidence Index with the residuals obtained from estimating a regression of the Personal saving rate and Total Consumer Credit flow respectively on the short and long-term interest rate and a measure of credit supply, the âNet Percentage of Domestic Banks Reporting Increased Willingness to Make Consumer Installment Loansâ by the Senior Loan Officer Opinion Survey on Bank Lending Practices. The correlation between the residuals from the Personal saving rate and the Consumer Confidence Index is negative and significant (-0.45***), while the relationship between the residuals from the Total Consumer Credit and the Consumer Confidence Index is positive (0.4***).

plications of non-rational expectations on the debt dynamics. In this economy, shocks to income affect expectations which in turn drive household debt. A positive shock to income generates a boom-bust cycle in debt as observed in the data. The representative consumer fails to forecast long-run income, gets over-optimistic and over-indebted; eventually, expectations adjust and debt declines. Overall, the model's predictions match well the positive correlation between debt and income and the negative correlation between current debt to real GDP and future income growth that cannot be captured with rational expectations nor with alternative beliefs′ hypotheses.

I begin by documenting that survey forecasts on future growth have substantial predictive power for household debt, savings, and consumption, using data from the Survey of Professional Forecasters (SPF) over the period 1968: Q4-2016: Q2. Although few studies analyse the relationship between households′ expectations and consumption, no other papers focus on the link between households expectations on future income and unsecured debt, using survey data.[3] The predictability is considerable in magnitude: a one standard deviation increase in expected real GDP (RGDP) growth is associated with an increase in credit growth of 2.08 percentage points over the next year, controlling for proxies of the supply of credit; this result is significant considering that, over the spanned period, the annual credit growth is on average 7.7%. Results are coherent with the predictions of the permanent income hypothesis (PIH), according to which, for a given interest rate, household debt is driven by the expected future income growth and the willingness to smooth consumption over time. The findings are consistent using consumers′ expectations on future economic conditions from the Survey of Consumers by the University of Michigan. The relation is robust across time, countries and specifications. Further, results are consistent at the micro level, using the Survey of Consumers Expectations (SCE) for the period 2013-2016. Overall, the 2000s′ boom in household debt and the Great Recession in the USA are not the unique reasons driving the described results, survey data on expectations are a relevant tool to understand the household debt dynamics and are not pure noise.

Since household credit is closely linked to agents′ expectations, it becomes crucial to identify possible models of beliefs formation and subsequently develop a theoretical framework to investigate the implications of miss-specified beliefs on the credit cycle. Do expectations depart from rationality? How? Exploiting survey data from the SPF,

[3]Other papers instead focus on mortgages, the role of collateral and extrapolative expectations on house prices (i.e. Glaeser and Nathanson (2017)).

I show that expectations diverge from the rational benchmark in a way coherent with the natural expectations hypothesis. Precisely, first, following the work by Coibion and Gorodnichenko (2012a) I show that expectations depart from the rational benchmark in favour of models with non-rational beliefs at the aggregate level. Then, I document that the US RGDP exhibits (partial) mean reversion and that agents use a simple autoregressive rule to predict future growth, as assumed by the natural expectations theory. These two facts imply that the forecasts over-estimate the long run persistence of the RGDP, being over-optimist during booms and pessimist during busts. Combining these results, I contribute to the literature by providing empirical evidence in favour of the natural expectations hypothesis leveraging on survey data.

In the last section, I introduce the natural expectations in a standard consumption-saving model and compare the predictions of the model with those obtained assuming rationality. The model accounts for the described empirical factors. First, according to the PIH, the model predicts a positive relationship between expectations on future income and household debt. Second, natural expectations are coherent with the results of non-rational expectations at the aggregate level. Except for the beliefs, the model is a standard open economy with a representative borrower with no financial frictions. Every period the agent updates his forecasts on future income given the realised income and makes his borrowing decisions. In this economy, the agent optimises given the constraints, but he considers his biased beliefs as correct. Following Fuster et al. (2010), the income process has an hump-shaped dynamics, and the agent overestimates the persistence of good (or bad) news. This over-estimation happens because the agent applies a simple autoregressive forecasting rule that does not capture the hump-shaped dynamics of the true process. Positive news on current income induces the agent to over-estimate the long-run income and subsequently to increase the demand for credit leading to over-borrowing relative to the rational case. Eventually, beliefs are disappointed and converge to the rational benchmark, thus debt decreases, generating a boom-bust cycle in debt. The predictions of the model match well the empirical impulse responses from a VAR in 5 variables, which show that following a positive shock to the RGDP growth rate, debt jumps on impact and exhibits a hump-shaped path. The model captures the initial jump and the partial mean reversion in debt observed in the empirical impulse responses; rational expectations, instead, generate a negative correlation between income and debt as an increase in income is associated to an increase in savings. Overall, natural expectations match well some features of the data, such as the positive contemporaneous correlation between income and debt, and the negative correlation between current debt to real GDP and future income growth three years ahead. These features instead cannot be accounted with rational expectations, nor alternative hypothesis expectations, such as noisy signals (Sims (2003)), sticky information (Mankiw and Reis (2002)), and adaptive expectations.[4]

The paper aims at filling the gap on the role of households' expectations in the demand side of the credit cycle by documenting that non-rational expectations have significant implications for the credit dynamics, coherently with the empirical evidence. Exploiting survey data, the paper provides new insights that the theoretical literature on household credit and business cycles should take into account introducing non-rational beliefs in the demand side when studying the credit dynamics.

The remainder of the paper is organised as follows. In Session 2 I briefly discuss the related literature. In Section 3 I describe the expectations, consumer credit and macroeconomic data. Section 4 presents the empirical evidence on the link between beliefs on future output and household debt. Section 5 describes the natural expectations hypothesis and provide empirical evidence in favour of it. Section 6 presents the theoretical model of natural expectations and household debt. Finally, the last section concludes.

1.2 Related literature

The paper is related to several strands of research. First of all, it connects to the natural expectation theory as proposed by Fuster et al. (2010); the study introduces a parsimonious model of quasi-rational expectations, and analyzes its consequences on excess returns in a Lucas tree model. The model is extended by Fuster et al. (2012) which investigates the effects of natural expectations on asset prices showing that they generate empirically serveral observed patterns in the US data. My work adopts their theory but looks at implications of non-rational beliefs on consumer credit. Hence, I contribute to their work by providing a new application for the natural expectations hypothesis and novel empirical evidence in favour of this theory exploiting survey data.[5]

Second, the paper is linked to the empirical literature on survey expectations data. In

[4] In the Appendix, I compare the model to alternative hypotheses on expectations, such as noisy signals, sticky information, and adaptive expectations. These hypotheses on beliefs' formation have substantially different implications on long-term expectations and hence on credit responses.

[5] I focus on beliefs on future income rather than house prices as in Pancrazi and Pietrunti (2014).

macroeconomics, most of the works focus on inflation forecasts (i.e. Coibion and Gorodnichenko (2015a); Malmendier and Nagel (2015)), and only a few extend the analysis to other macroeconomic variables (i.e. Souleles (2004)). Using the Survey of Professional Forecasters (SPF), Coibion and Gorodnichenko (2012a) document the rejection of full-information rational expectations in the direction of sticky information, as proposed by Mankiw and Reis (2002), and noisy information, as Woodford (2001). Agents in these models are perfectly rational but subject to information frictions. Here, I implement the tests proposed by Coibion and Gorodnichenko (2012a) to understand how beliefs evolve. Relative to their work, my contribution is to map the empirical evidence to the natural expectations hypothesis using data on the expected real GDP growth rate.

This study is then partially related to the literature that explores the link between financial markets and the real economy through the debt-driven consumption channel. This strand focuses mainly on the role of credit supply shocks with nominal rigidities (Eggertsson and Krugman (2012); Guerrieri and Lorenzoni (2017)), demand externalities (i.e. Bianchi (2011)) or preference shocks. Quantitative analyses, however, suggest that credit supply and preference shocks are not able to account for the boom and bust in private credit observed during the financial crisis (Justiniano et al. (2015)) suggesting that other mechanisms contribute to the dynamics of credit. Further, those models fail to account for predictable forecast errors as they rely on rational expectations and exogenous financial or preference shocks.

On the other hand, researches in finance revived the old argument (e.g. Minsky (1977)) that investors' sentiment drives the supply of credits (Greenwood et al. (2016); Bordalo et al. (2018); Lopez-Salido et al. (2017)) and discard the rational expectation hypothesis to explain the credit cycle. Similarly to these works, I abandon the rational expectation hypothesis; but I look at the demand side and focus on households' biased beliefs rather than investors'. I introduce a risk-averse agent who makes consumption-savings choices, as opposed to a risk-neutral investor who is asked to finance risky projects. The aim of this paper is indeed to explain how biased beliefs on the future income affect demand for credit, without looking at the supply of funds.

Finally, this work is related to the macroeconomic literature on expectations shocks. This literature maintains the assumption of rational expectations and introduce exogenous shocks to justify shifts in optimistic or pessimistic beliefs. These papers consider either a confidence shocks (Angeletos et al. (2018)), noise shocks (Lorenzoni (2009)) or noise and news shocks (i.e. Barsky and Sims (2012)).[6] Here, instead, I look at shifts in beliefs driven by changes in the fundamentals rather than exogenous sentiment shocks. For this reason, the paper is more closely linked to the literature on behavioural biases and, in particular, to the natural expectations theory.

1.3 Data

The empirical analysis focuses on three categories of data: (1) expectations data at the aggregate level (consumers' and professional forecasters'), (2) data on macroeconomic variables, householdâs debt, and savings, and (3) micro-level data on consumers' expectations and intention to borrow. Data are available at quarterly frequency.

1.3.1 Expectations data

Professional forecasters expectations (SPF)

Data on professional forecasters expectations come from the quarterly U.S. Survey of Professional Forecasters (SPF) provided by the Federal Reserve Bank of Philadelphia. The survey began in 1968: Q4, although few variables were added in 1981: Q3. The forecasters provide projections for five quarters for several macroeconomic variables, including Real GDP growth rate (RGDP), Civilian Unemployment Rate (UNEMP), 3 Months Treasury Bill Rate (3MTBill), Price Index for the GDP (PGDP), and 10 years Treasury Bond (10YTBond). Specifically, respondents are asked to indicate their forecast for the current quarter - the quarter in which the survey is conducted - and the following four quarters. Expectations are aggregated by average across forecasters' responses. Further, I average forecasts over 4 quarters ahead to obtain 4-quarters forecasts; this procedure allows to reduce the noise and make them comparable to the Michigan Survey of Consumers data and the existing literature.[7] In the Appendix, I also use the data on the forecasts for the annual average 10-years real GDP growth rate. The analysis focuses on the expected

[6] Recently, Boz and Mendoza (2014) considered the role of expectations on the ability to borrow, while Kaplan et al. (2017) studied the role of beliefs about future house price growth in the credit cycle. In this paper, instead, I investigate the link between expectations and household debt, without relying on the role of the collateral or house prices.

[7] To get expectations on the RGDP growth rate one year ahead, I use the SPF data on the RGDP âAnnualized Percent Change of Mean Responsesâ which reports the forecast for quarter-over-quarter growth in each quarter over several horizons. Then, I compute the geometric mean of the expected forecast for the next 4 quarters (as in Coibion and Gorodnichenko (2015a)). I adopt a similar procedure for the expected unemployment rate.

RGDP growth rate and unemployment rate; Table 1.1 reports the descriptive statistics for these two measures for the period: 1968:Q4 to 2016:Q2.

Household expectations by the Michigan Survey of Consumers

Data on household expectations are from the Survey of Consumers by the University of Michigan (MSC). The survey asks respondents about the predicted direction of future changes in the unemployment rate, the interest rate, and the business conditions for the following 12 months and collects qualitative answers.[8] At the aggregate level, I use quarterly time series, reporting the relative share of agents who give a particular answer; i.e. for the Expected Change in Business condition in 1 year, the index is computed as the fraction of agents that expects the business condition to be better over the next year minus the fraction of agents that the expects the business condition to be worst over the next year plus 100. Aggregate data are available at the quarterly frequency from 1968:Q4 to 2016:Q2.

Professionals' and households' expectations: correlation

The empirical analysis focuses mainly on data by the SPF for two reasons. First, this data allows making quantitative predictions, which cannot be obtained using the MSC given the qualitative nature of the data.[9] Second, professional forecasters are some of the most informed economic agents, so they can provide a conservative benchmark for assessing potential deviations from full-information rational expectations. Households' and professional forecasters' expectations are highly positively correlated with each other. Table 1.2 reports partial correlations for the variables of interest. Most relationships are positive and significantly different from zero. For instance, the average correlation between the Michigan Expected Change in Business Conditions in 1 Year and the SPF RGDP growth rate over the next year is 63% (Figure 1.3).[10] The high degree of correlation between the time series suggests that survey data contain shared beliefs on the aggregate state of the economy; this evidence goes against a common criticism on survey data on

[8]For instance, respondents are asked: *Now turning to business conditions in the country as a whole – do you think that during the next 12 months we will have good times financially or bad times or what?* Answer: *Good times, Uncertain, Bad times.*

[9]The first best for the analysis is to test the rational expectation hypothesis on households' expectations; however, to my knowledge, no quantitative data on consumers' expectations is available for a sufficiently long series.

[10]The correlation is lower for the expected interest rate and the unemployment rate. The former may be because the MSC asks respondents about the Prime rate, while the SPF asks about the 3 Months Treasury Bill Rate (or 10YTBond). The latter instead may occur because the SPF measures the expected unemployment rate, while the MSC measures the expected change in unemployment.

expectations that are believed to be noisy and meaningless, as described by Greenwood and Shleifer (2014). To confirm that professional forecastersâ expectations are a good proxy for householdsâ expectations, I regress the MSC expectations on the SPF data, controlling for the business cycle (Table 1.3). Specifically, I use as controls (i) the US Business Cycle Expansions and Contractions[11] and (ii) the Real GDP growth rate. The positive relationship between the professional forecastersâ and consumersâ expectations is not driven by the trend but survives within the business cycle. These results suggest that it is reasonable to extend the results obtained with the SPF data to consumers' expectations.

1.3.2 Consumer credit and macroeconomic data

Survey data are merged with a database containing the historical values of the variables object of the forecast and additional macroeconomic variables. I use data from different sources. The Federal Reserve Bank of Philadelphia itself provides third release data for the variables object of the forecast. I collect aggregate data on consumer credit from the Consumer Credit (G.19) database provided by the Federal Reserve Board. Data on Total Consumer Credit Outstanding (hereafter Household credit) in levels are available at a monthly frequency and span the period 1943:Q1-2016:Q2.[12] For the empirical analysis, I reconstruct quarterly data in levels and percentage changes quarter-over-quarter and year-over-year. Total Consumer Credit Outstanding covers most credit extended to individuals, excluding loans secured by real estate. Data on additional macroeconomic variables at quarterly frequency are from the Federal Reserve Economic Data (FRED) by the Federal Reserve Bank of St. Louis. The Personal consumption expenditures measures goods and services purchased by U.S. residents (PCE); while the Personal saving rates (PSAVERT) is the personal saving as a percentage of disposable personal income (DPI). I use the FRED series for the Real Gross Domestic Product in Billions of Chained 2009 Dollars, Quarterly, Seasonally Adjusted (GDPC96). The other macroeconomic variables are the unemployment rate (UNRATE), the 3MTBill (TB3MS), the 10YTBond (GS10) and the consumer price index (CPIAUCSL).

[11] The time series indicates the US Business Cycle Expansions and Contractions as defined by the National Bureau of Economic Research (NBER). This is a dummy variable that is equal to 1 during recessions and 0 during expansions.

[12] I use the series: Total consumer credit owned and securitized, seasonally adjusted level.

1.3.3 Micro level data: Survey of Consumers Expectations

Finally, I use data from the Survey of Consumers Expectations (SCE) for the period 2013: Q4- 2016: Q4. The survey is nationally representative, internet-based, and involves a rotating panel of about 1,300 household heads; for a detailed description, see Armantier et al. (2017). The survey includes information on expectations about various macroeconomic and household level variables and, at the quarterly level, information on access to credit. For the analysis, I build two measures of the intention to apply for a loan. The first one is a discrete variable which indicates the average likelihood of applying for a credit card, an auto loan, an increase in the limit of the credit card over the following year, the larger it is, the larger is the likelihood of applying; the second one is the average percent chance of applying for the same sets of loans over the next year. Then I build two proxies for the expected change in income one year ahead. The first one is a discrete variable, increasing with the expected income, that indicates the expected financial condition in the next 12 months; the second one is the expected percentage increase in household income in the next 12 months.[13] Summary statistics are provided in Table 1.1.

1.4 Expectations data and household debt

In this section, I show that survey data on expectations are good predictors of household debt as predicted by the permanent income hypothesis (PIH). I begin by testing the link between expectations and consumers' credit and I use a forecasting framework to answer the question: do beliefs on future output help predict changes in household debt in the short run? Formally, I estimate variants of the following regression using quarterly U.S. data:

$$D_{t,t+4} = \alpha + \beta E_t(RGDPgr_{t,t+4}) + \delta(L)Z_t + \gamma(L)D_{t,t+4} + u_{t,t+4} \tag{1.1}$$

Where $E_t(RGDPgr_{t,t+4}) = E_t(RGDPgr_{t,t+4}|I_t)$ denotes the SPF average forecast of Real GDP growth rate (RGDP gr) over the next four quarters, made at time t, given the information set. $D_{t,t+4}$ is the percentage change between time t and $t+4$ of the dependent variable, which can be Household credit, Personal savings rate, and Personal consumption expenditures. In the baseline specification, Z_t includes contemporaneous macroeconomic variables known at time t ($Z_t \in I_t$) that may affect both expectations and households

[13] I use the following variables of the SCE: (i) n17_a and n17_b on the likelihood of applying for a loan, and (ii) q2 and q25v2part2 on the expected income.

credit, like unemployment, GDP, Consumer Price Index, long and short-term interest rate.[14] In this specification, the interest rates are proxies for the credit market conditions as we aim to pick out the link between expectations and credit demand for a given supply. All specifications include two lags of the dependent variable: these lags ensure that mean reversion of the dependent variable is not driving our results.[15] β is the primary object of study to test whether expectations are informative. According to the PIH, given the interest rate, there is a positive relationship between the change in debt and the expected income growth. Expectations on future income have opposite effects on savings and credit; agents increase debt when they expect the permanent income to increase and otherwise expand savings when they expect it to decrease. These predictions hold regardless of how expectations are formed.

Table 1.4 (Panel A) reports the baseline results from OLS regressions. In Columns (1) and (4) the dependent variable is Household credit; this measure covers most credit extended to individuals, excluding loans secured by real estate; Column (2) refers to the Personal savings rate, and Column (3) to Personal consumption expenditures. The coefficient β_{debt} is positive and significantly different from zero; while the estimated coefficient $\beta_{savings}$ has a negative sign, as predicted by the theory. Changes in expected future growth have strong forecasting power for household debt: a one standard deviation increase in expected RGDP is associated with an increase in credit growth of 2.08 percentage points, conditional on the interest rates and the macroeconomic variables. This result is significant considering that, over the spanned period, the annual credit growth is on average 7.7%. The effect is greater on credit than consumption; a one standard deviation increase in expected output growth raises consumption growth by 0.46 percentage points, and the average personal consumption expenditure growth rate is 6.89%. Individuals' forecasts on future economic growth have substantial predictive power for household debt and savings. SPF expectations alone explain the 27% of the variability of household debt. Further, when controlling for the macroeconomic variables, the increment to the adjusted R^2 that results from augmenting the baseline regression with the SPF expectations is higher than 17 percent. On the contrary, household credit decreases when agents' expect higher unemployment rate (Table 1.4, Panel A, Column (7)); a one standard deviation increase in the expected unemployment rate is associated

[14] Results are very similar when controlling for lagged variables assuming that agents have information only on the previous quarter when forming their forecasts ($Z_t \notin I_t$).

[15] The lag structure has been defined to keep it reasonable and according to the information criteria (AIC/BIC).

with a reduction in household credit growth of 6.39 percentage points.

Robusteness check. These results are open to a variety of interpretations. One possibility is that expectations reflect pieces of information available at time t, omitted in the regression, that drive household debt. A large number of predictors may be added in the forecasting regression. To address this high-dimensional problem, I adopt a two steps procedure as suggested by Stock and Watson (2002). First, I estimate a time series of the factors from the potential predictors using principal components and 257 macroeconomic variables.[16] Second, I estimate a linear regression including as regressors the estimated factors. Let $D_{t,t+4}$ be the time series of household debt growth rate to be predicted and X_t be an N-dimensional time series of candidate predictors. Let ($D_{t,t+4}$, X_t) admit a factor model representation with r common latent factors F_t,

$$\begin{aligned} X_t &= \lambda F_t + e_t \\ D_{t,t+4} &= \gamma_t F_t + \delta w_t + \epsilon_{t,t+4} \end{aligned}$$

Where e_t is a $nx1$ vector of idiosyncratic disturbances and w_t a $mx1$ vector of observed predictors, including $E_t(RGDPgr_{t,t+4})$. I fix $r = 8$ according to McCracken and Ng (2015). Table 1.4 (Panel A) reports the results of the baseline regression (1.1) augmented with the estimated common factors.[17] The coefficient of interest β is always positive and significant, suggesting that expectations on future RGDP growth do not exclusively reflect omitted information (Columns (5) and (6)). The same is true for expectations on the future unemployment rate (Column (8)). The 257 macroeconomic variables include forward-looking measures such as stock-market indices; therefore, these results confirm that survey data on expectations are fundamentally different from other forward-looking variables.

Expectations or realisations? So far, we have seen that expectations contain meaningful information that helps to predict household debt. Anyhow debt may fluctuate in response to other factors anticipated by the forward-looking beliefs. In particular, expectations may anticipate future realised growth, which in turn may drive household debt. To alleviate any concern that the correlation between expected and future growth might be driving the results, I include the future realised RGDP growth rate as control

[16] For a detailed description, see the FRED-QD quarterly large macroeconomic databases.

[17] The estimated regressions include all the factors.

and estimate the following regression:

$$D_{t,t+4} = \alpha + \beta_1 E_t(RGDPgr_{t,t+4}) + \beta_2 RGDPgr_{t,t+4} + \delta(L)Z_t + \gamma(L)D_{t,t+4} + u_{t,t+4} \quad (1.2)$$

The coefficient of interest β_1 is positive and significant (Table 1.4, Panel B, Columns (1) and (2)). The effect remains large in magnitude: one standard deviation increase in SFP expectations increases debt by 1.32 percentage points when controlling for macroeconomic variables, including the interest rate. For a given realised state of the economy, preceding optimism over future growth is associated with an increase in household debt growth. Expectations appear to be important determinants of household debt and overtake the realised RGDP growth. As a robustness check, I add as controls the factors described in the previous paragraph, over the current (t) and future ($t+4$) quarter. This specification allows controlling for the events that arise at time $t+4$, affect household debt and may be anticipated by the forward-looking beliefs.[18] Columns (3)-(5) confirm the previous results. Results are mostly consistent also for the expected unemployment rate (Columns (6) and (8)).[19]

Hints on the theory. These results go in favour of the PIH according to which, for a given interest rate, household debt is driven by the expected future income growth and the willingness to smooth consumption over time. At the same time, I can reject alternative models in which household borrow to get liquidity anticipating bad times. If this was the case, the relation between the expected future income growth and household debt growth should have the opposite sign. Further, expectations on future income growth are a key driver of the demand for debt as opposed to realised income, for a given interest rate, as predicted by the PIH. Given the realised income, optimism leads to an increase in household debt and pessimism to a reduction in debt. These results can be coherent both with models of rational and non-rational expectations in which beliefs drive choices.

Short or long-term expectations? In the analysis, I consider short-term forecasts, while the PIH defines a relation between the forecasts of the long-run income growth and credit. SPF data do not allow to explore this relation since it provides forecasts on RGDP over the next 10 years only at the annual frequency and from 1992. Although only a short time series is available, the correlation between long and short-term expectations

[18] Given the short sample, I don't control for the factors for all quarters from t to $t+4$. Adding as controls the macroeconomic indicators for t and $t+4$, results do not change.

[19] In Columns (6)-(8) do not control for the realized unemployment rate between time $t-4$ and t as it is strongly correlated with the expected unemployment rate.

is positive and significant (0.5) (see the Appendix). Further, short-term expectations are strongly autocorrelated (0.85); therefore, it is reasonable to approximate the PIH relation using the short-term beliefs.

Additional robustness. Results are consistent with consumers' expectations by the Michigan Survey of Consumers. Further, results are consistent across countries, using a panel of 11 countries, and across time, on a sub-sample excluding post-2000 data. This evidence suggests that the Great Recession is not driving the previous results (see the Appendix for details). Finally, so far I assumed that heterogeneity in beliefs is not key to understand the household credit cycle. In the Appendix, using SPF data, I provide evidence in support of this hypothesis showing that the positive relationship between household debt and expected income growth observed in Table 1.4 is not driven by a small fraction of agents. The results, in fact, are robust when we exclude the most optimistic (or pessimistic) agents.

Micro data. Results are robust when estimating a similar set of regressions on the microdata from the Survey of Consumers Expectations (SCE) for the period 2013: Q4-2016: Q4. Table 1.5 displays the results. The dependent variable is the willingness of applying for a loan in the next year and the independent variable the expected change in the household financial conditions one year ahead. Panel A and B of Table 1.5 adopt different proxies of these variables as described in Section 2.3; the former uses discrete variables while the latter continuous variables. Column (1) shows that there is a positive and significant relationship, also at the individual level. Micro-data allows controlling for time and state fixed effect which soak up all common components of expectations, such as shocks to the credit supply and the business cycles (Column (2)).[20] Further, they allow to include potential determinants of borrowing, which may correlate with expected future income; these are demographic characteristics, expectations on macroeconomic variables, and current financial conditions (Column (3)). When adding these controls, the coefficient of interests always remains positive and significant indicating that, at the individual level, optimism on future income increases the willingness of applying for a loan, given the aggregate state of the economy and market conditions.

[20] Data do not allow to add state-time fixed effect.

1.5 The natural expectations hypothesis

We have seen that expectations on future growth are strong predictors of consumer credit in the short-term, and this result is robust to several specifications. In the light of these findings, it becomes relevant to understand how consumers form beliefs and subsequently develop a theoretical model coherent with the empirical evidence to investigate their impacts on the household credit cycle. In this section, I present the natural expectations hypothesis, as proposed by Fuster et al. (2010) and provide evidence that this hypothesis is coherent with the survey data on expectations.

1.5.1 Definition

The natural expectations theory is based on two main assumptions. First, the time series of interest y_t, in level or logs, has long-horizon hump-shaped dynamics so that it exhibits (partial) mean reversion in the long run. Second, the representative agent forecasts future changes in y_t by estimating an autoregressive process on Δy_t with few lags of historical changes and fails to estimate the medium- long-run properties of y_t. Expectations are not consistent with the true model, but they are empirically disciplined since the agent estimates the parameter using historical data. Let the true data generating process (DGP) for y_t be:

$$y_{t+1} = \phi(L)y_t + \beta(L)\epsilon_t \tag{1.3}$$

The parameter $\phi(L)$ and $\beta(L)$ are such that the process exhibits mean reversion in the long run. The agent does not know the true DGP and bases his beliefs on the following forecasting rule:

$$\widetilde{E_t}(\Delta y_{t+1}) = \rho(L)\Delta y_t + \eta_t \tag{1.4}$$

Here the $\widetilde{E_t}(\ .\ )$ is the natural expectations operator. The rule is such that the agent does not capture the hump-shaped dynamics of the true process in equation (1.3) and the parameters in $\rho(L)$ are estimated using historical data. The agent extrapolates historical data into the future; he knows the true parameter ρ of the AR(1) model and no uncertainty nor learning on the parameter is in place.[21] Expectations deviate from rationality since they fail to forecast the mean reversion, and over-predict the long-run

[21] Fuster et al. (2010) introduce an additional parameter λ that I assume to be equal to zero. Further, they provide an example in which y_{t+1} is an AR(2) and the agent estimates an AR(1) on Δy_{t+1} such that: $y_{t+1} = \phi_1 y_t + \phi_2 y_{t-1} + \epsilon_t$, $\Delta y_{t+1} = \rho\Delta y_t + \eta_t$ with $(\phi_1 + \phi_2) < 1$ and $\rho > 0$.

persistence of positive and negative shocks. Each period, given the realized Δy_t, short and long-term beliefs adjust according to the forecasting rule in equation (1.4).

1.5.2 Why natural expectations? Match with the empirical evidence from survey data

Overall natural expectations are described by few characteristics: (i) expectations are non-rational at the aggregate level, (ii) the true process has an hump-shaped dynamics, (iii) households use simple forecasting rules, (iv) beliefs overestimate the persistence of the true process, and (v) there is a positive relationship between long and short term expectations. In this section, I document that these features are coherent with the empirical evidence.

Testing rationality: forecast error predictability

Natural expectations imply that expectations depart from rationality as they are extrapolative and characterised by frequent updating. The agent extrapolates historical data into the future and assign excessive weights to recent observations. Further, each period expectations adjust according to the observed realisations of Δy_t. Following Coibion and Gorodnichenko (2015a), I investigate whether beliefs are systematically biased in a way coherent with the natural expectations hypothesis, testing the predictability of ex-post forecast errors and revisions. Rational expectations imply that agents have full knowledge of the economy and exploit the information optimally; therefore, errors and revisions are unpredictable and orthogonal to the information set available to agents at the time the prediction is made.

Let $\Delta y_{t,t+4}$ be the real GDP growth rate in the current quarter t and the following three quarters (or the unemployment rate) and consider the following definitions relative to time t:

- Realized value of $\Delta y_{t,t+4}$: $\Delta y_{t,t+4}$
- Average expectations of $\Delta y_{t,t+4}$: $\widetilde{E_t}(\Delta y_{t,t+4})$
- Forecast error on $\Delta y_{t,t+4}$: $\varepsilon_{t,t+4} = Fe_{t,t+4} = \Delta y_{t,t+4} - \widetilde{E_t}(\Delta y_{t,t+4})$
- Forecast revision on $\Delta y_{t,t+4}$: $Fr_{t,t+4} = \widetilde{E_t}(\Delta y_{t,t+4}) - \widetilde{E_{t-1}}(\Delta y_{t,t+4})$
- Forecast revision on $\Delta y_{t-1,t+3}$: $Fr_{t-1,t+3} = \widetilde{E_t}(\Delta y_{t-1,t+3}) - \widetilde{E_{t-1}}(\Delta y_{t-1,t+3})$
- Lagged forecast revision on $\Delta y_{t-1,t+3}$: $Fr_{lag,t-1,t+3} = \widetilde{E_{t-1}}(\Delta y_{t-1,t+3}) - \widetilde{E_{t-2}}(\Delta y_{t-1,t+3})$

Using survey data from the SPF, I compute these measures. Forecast errors are constructed using the third release data, provided by the SPF; this is the best comparison measure since final data may be affected by re-classifications or re-definitions; while first release data are less accurate and are subject to measurement errors. Given these definitions, I run three different tests to check if expectations depart from the full information rational expectations hypothesis (FIRE); results are reported in Table 1.6. Specifically, I study the following properties:

1. The agent extrapolates past growth into the future: recent observations are predictors of future forecast errors. To test this assumption I run the following test:

$$\varepsilon_{t,t+4} = \alpha_1 + \beta_1 \Delta y_{t-4,t} + u_{t,t+4} \tag{1.5}$$

   Under rational expectations, the coefficients α_1 and β_1 are zero. This test is informative of whether expectations are rational or extrapolative. Expectational errors are systematically biased and predictable at the aggregate level, as shown in Column (1). The null of FIRE is rejected at 1% level of statistical significance and the negative coefficient β_1 suggests that agents extrapolate recent observations into the future.

2. Forecast revisions predict ex-post forecast errors. To test this assumption, I estimate the following regression:

$$\varepsilon_{t,t+4} = \alpha_2 + \beta_2 Fr_{t,t+4} + u_{t,t+4} \tag{1.6}$$

   Under rational expectations, the coefficients α_2 and β_2 are zero since the $Fr_{t,t+4}$ is in the information set at time t. The null hypothesis of rational expectations is rejected, as shown in Column (2), and forecast revisions predict future forecast errors. The positive coefficient $\beta_2 < 1$ implies slow reaction to macroeconomic shocks; positive revisions are associated with a rise in the forecast error meaning that expectations adjust but not enough. The positive coefficient is consistent with the natural expectations hypothesis under some assumptions on the parameters. For instance, let $\Delta y_{t+1} = \beta(L)\epsilon_t$ and $\widetilde{E_t}(\Delta y_{t+1}) = \rho \Delta y_t + \eta_{t+1}$. Then $\beta_1 > \rho$ implies a positive relationship between the forecast errors and the forecast revisions as the one observed in the data. Table 1.7 shows the estimates for the ARIMA(1,1,0), ARIMA(0,1,11) and ARIMA(0,1,18) for the quarterly US RGDP; the estimated coefficients are such that $\beta_1 > \rho$.

3. Current forecast revisions predict future forecast revisions. Finally, I run the following test:

$$Fr_{t-1,t+3} = \alpha_3 + \beta_3 Fr_{lag,t-1,t+3} + u_{t-1,t+3} \tag{1.7}$$

Rationality implies α_3 and β_3 equal to zero since the forecast revision should be unpredictable given the information set at time t; while a positive β_3 implies persistent and smooth adjustment. Column (3) shows that forecast revisions are predictable: an upward revision at time $t-1$ predicts a positive update at time t. The positive sign of the coefficient is coherent with the natural expectations hypothesis under some assumptions on the parameters. These conditions are satisfied when considering an ARIMA(1,1,0) and an ARIMA(0,1,18) (or an ARIMA(0,1,11)) for the expected and realized quarterly US RGDP (Table 1.7).[22]

In conclusion, we find evidence in favour of non-rational expectations at the aggregate level. Forecast error predictability reflects extrapolative expectations and slow updating to macroeconomic shocks and results are consistent with the natural expectations hypothesis.

The true process: the hump-shaped dynamics

Second, the natural expectations hypothesis assumes that the true process of y_t exhibits partial mean reversion in the long run. Fuster et al. (2010) show that many macroeconomic series, including the US RGDP and the unemployment rate, have an hump-shaped dynamics with (partial) mean reversion. Further, they document that to capture the short-term properties of macroeconomic time series it is necessary to estimate a highly flexible statistical model as low-order models have difficulty in detecting hump-shaped

[22] Let $\Delta y_{t+1} = \beta(L)\epsilon_t$ and $\widetilde{E_t}(\Delta y_{t+1}) = \rho \Delta y_t + \eta_{t+1}$. Assume $\Delta y_0 = \widetilde{E_0}(\Delta y_t) = 0$ for any $t > 0$ and $\Delta y_1 = e_1 > 0$. At time $t = 1$, the agent observes Δy_1 and $Fe_1 = \Delta y_1 - \widetilde{E_0}(\Delta y_1) = \Delta y_1 = e_1 > 0$ and slowly adjust his forecasts upward such that $Fr_2 = \widetilde{E_1}(\Delta y_2) - \widetilde{E_0}(\Delta y_2) = \rho e_1 > 0$. At time $t = 2$, under the assumption that $\beta_1 > \rho$, $Fe_2 = \Delta y_2 - \widetilde{E_1}(\Delta y_2) = (\beta_1 - \rho)e_1 > 0$ and $Fr_3 = \widetilde{E_2}(\Delta y_3) - \widetilde{E_1}(\Delta y_3) = e_1\rho(\beta_1 - \rho) > 0$. Hence, there is a positive link between the forecast error and the forecast revision, as well as between current and future revisions. At $t = 3$, the agent observes $\Delta y_3 = \beta_2 e_1$ and revises upward such that $Fe_3 = \Delta y_3 - \widetilde{E_2}(\Delta y_3) = (\beta_2 - \rho\beta_1)e_1$ and $Fr_4 = \widetilde{E_3}(\Delta y_4) - \widetilde{E_2}(\Delta y_4) = \rho e_1(\beta_2 - \beta_1)$. Under the chosen calibration, $(\beta_2 - \rho\beta_1) > 0$ and both the forecast error and the forecast revision are positive. A similar reasoning applies for the following quarters.

dynamics.[23] To confirm this pattern in this application, I estimate several ARIMA(0,1,q) for the US RGDP in log, with large q. When q is sufficiently large, this flexible representation implies a hump-shaped pattern and partial mean reversion (Table 1.7 and Figure 1.4).

The estimated process: a simple autoregressive rule

The model assumes a representative household who uses a simple autoregressive model to forecast future changes in y_t. To provide evidence in support of this assumption, I exploit SPF data on expectations and run an out-of-sample forecasting exercise to choose the best order of the following process $\widetilde{E_t}(\Delta y_{t,t+4}) = \rho(L)\Delta y_{t-4,t} + \epsilon_{t,t+4}$. In this exercise, I estimate each model on a sub sample and then predict the future values of $\widetilde{E_t}(\Delta y_{t,t+4})$ beyond the estimation sample using the realized values of the forecasted variable; specifically to predict $\widetilde{E_t}(\Delta y_{t,t+4})$ I use $\Delta y_{t-4,t}, \Delta y_{t-5,t-1}, \Delta y_{t-6,t-2}$ and so on, which are in the information set at time t. Table 1.8 reports the Mean-Square Forecast Error (MSFE) and the Mean Absolute Error (MAE) from this exercise.[24] The MSFE and MAE are two measures of predictions' accuracy such that the fit of the model increases when these measures decrease. The smallest forecast error is associated with the following forecasting rule: $\widetilde{E_t}(\Delta y_{t,t+4}) = \rho\Delta y_{t-4,t} + \epsilon_{t,t+4}$. Results are consistent when changing the out-of-sample window (different columns). This evidence goes in favour of the idea that the representative agent estimates an auto-regressive model AR(1) and extrapolates from the past, giving excessive weight to recent changes.[25] In the Appendix, I show that this is the best process, in terms of out-of-sample performance, also when we allow for moving average components. The forecasting rule with only one recent observation is the one that approximates better the survey data in terms of out-sample performance, hence the hypothesis of simple models associated with natural beliefs seems justifiable.

Over-estimates of the long-run persistence

The hump-shaped dynamics of the fundamental process matched with a simple autoregressive forecasting rule imply that shocks are expected to be more persistent than they

[23] Fuster et al. (2010) provide a Monte Carlo analysis to show that ARIMA(0,1,q), with large q, is better to capture empirically relevant low-frequency mean reversion compared to a lower order ARIMA(p,1,q) with p and q $\geq$ 3. They show that ARIMA(0,1,q) processes are not subject to bias in the estimated persistence when the true data generating process is not an ARIMA(0,1,q).

[24] I use the out-of-sample forecast to reduce the risk of over-fitting.

[25] Otherwise, under alternative hypothesis on the beliefs, the forecast should depend on all, or more than one, past lags of the dependent variable. For instance, adaptive expectations use all historical data and down-weight old data exponentially.

are, so that good (or bad) times are expected to persist; indeed, the order of the model determines the estimated long-run persistence which is equal to (1+ma(L))/(1-ar(L)).[26] The long run persistence is defined as the ultimate impact of the shock on the level of y_t. Table 1.7 shows how the persistence for the US RGDP varies with the estimated model. For instance, an ARIMA(0,1,18) exhibits mean reversion and the associated persistence is equal to (1+ma(L))=0.299, while an ARIMA(1,1,0) implies a persistence equal to $1/(1-\rho)$=1.476 with the estimated ρ equal to 0.3227.[27] This feature characterises natural expectations; the household estimates low order models and overestimate the persistence of the process. This fact implies that beliefs do not capture the long-run mean reversion of the series so that they are over-optimistic during booms, and over-pessimistic during busts.

Short and long term expectations

Natural expectations imply that short and long term forecasts are positively correlated. Let the agent estimates an AR(1) model with $\rho > 0$ such that $\widetilde{E_t}(\Delta y_{t+1}) = \rho \Delta y_t$, and $\widetilde{E_t}(\Delta y_{t+j}) = \rho^j \Delta y_t$ for any $j \geq 0$. Short and long term predictions move in the same direction and their relationship is described by the parameter ρ. Accordingly, the Appendix shows that there is a positive correlations between the expected real GDP growth rate at 1 year (4 quarters ahead) and at 10 years. Further, coherently with the hypothesis of natural expectations, during booms short term expectations are greater than long term expectations; while the opposite is true during busts.

Additional evidence

The hypothesis of natural expectations is coherent with the extensive experimental evidence on extrapolation where recent observations are those that matter the most and, in particular, with Beshears et al. (2013) who show that agents fail to forecast long-term mean-reversion, while agents capture well short-run momentum. Besides, the model relies on the additional assumption that no learning on the true process is in place. The experimental evidence by Landier et al. (2017) supports the hypothesis of no learning in the short-medium run, showing that extrapolation is the primary driver of expectations, and no learning occurs. Finally, the natural expectations hypothesis does not allow for model

[26] Here ma(L) is equal to the sum of the moving average components and ar(L) is the sum of the autoregressive parameters such that ma(L)=$\beta_1 + \beta_2 + \beta_3 + ...$ and ar(L)=$\phi_1 + \phi_2 + \phi_3 + ...$.

[27] An estimated persistence equal to 0.3 means that a one unit shock as an ultimate impact on the level of y_t of 0.3. Using survey data on the expected RGDP growth over the next year the estimated persistence is about 1.17 with $\hat{\rho} = 0.1459$.

uncertainty; the agent does not take into account model and parameter uncertainty when forming his predictions. This assumption is coherent with the idea of overconfidence on predictions (over-precision), according to which agents assigns overly narrow confidence intervals to their predictions. This effect can be explained by anchoring (Tversky and Kahneman (1974)) to the initial estimate which in this case is the natural forecast.

1.6 A model of natural expectations and debt

In the previous section, we have seen two facts. First, expectations data are not pure noise and help to predict household debt in line with the permanent income hypothesis (PIH) (Section 1.4). Second, expectations are non-rational and display aggregate rigidities in a way coherent with natural expectations (Section 1.5). Given this empirical evidence, now I test if non-rational expectations can account for the household credit cycle. Can non-rational expectations generate boom and bust in household debt? To address this question, I propose a model in which consumers form natural expectations, and these beliefs influence their borrowing decisions. The model reflects the cited empirical factors. It predicts a positive relationship between expectations on future income and household debt growth, for a given credit supply, according to the PIH. Further, it accounts for the evidence on non-rational extrapolative expectations with slow adjustment at the aggregate level.

Except for the expectation process, the model is a standard small open economy with a representative borrower. The representative agent chooses consumption and credit and optimises given the budget constraint and his biased expectations. Beliefs represent the sole source of distortion and non-rationality. I study this tractable model as it reflects the PIH and allows to highlight the effects of non-rational beliefs on the dynamics of credit without the need of additional confounding factors.[28] Following Fuster et al. (2010), the fundamental has an hump-shaped dynamics, and the agent has natural expectations; hence he has wrong beliefs about the dynamics of the fundamental and underestimates the degree of mean-reversion of the income process. This fact happens because the agent applies a simple AR(1) forecasting rule, over-weights recent observations and fails to forecast long-run income. Under this assumption, beliefs on the long-run income are excessively optimistic during booms and pessimist during busts. In this economy, a shock

[28] Although the US economy is not a small open economy; this model is reasonable for our purposes since in recent years the United States has been borrowing heavily on foreign capital markets at low-interest rates according to the ”saving glut” hypothesis by

to income translates into a shift in the expected future income and in turn into a change in debt; every period, the agent observes the realised income, updates his beliefs and his choices for debt. The model implies that the agent separates the forecasting and the choice problem and expectations drive choices.[29] A positive shock to current income induces agents to over-estimate the long-run income and subsequently to increase the demand for credit, over-borrowing compared to the full-information rational equilibrium. This phase is followed by a period of disappointed expectations and gradual adjustment. Overall, natural expectations generate a boom-bust cycle in household credit without relying on heterogeneous beliefs, collateral constraints or supply shocks.

1.6.1 Theoretical model

Consider a representative agent with an infinite life horizon and preferences described by the following utility function for $t = 0, 1, .., \infty$:

$$\widetilde{E_t}\Sigma_{j=0}^{\infty}\beta^t U(c_{t+j}) \tag{1.8}$$

$\widetilde{E_t(\ . \ )}$ denotes the biased expectation operator at date t (given the information set available at time t) and β the discount factor. Each period, the household receives an exogenous and stochastic endowment y_t and can borrow or lend in a one-period bond that pays a constant interest rate. The endowment process represents the unique source of uncertainty. Given y_0, the evolution of the income is given by a ARIMA(0,1,q) process:

$$\Delta y_{t+1} = \beta(L)\epsilon_t \qquad \epsilon_t \sim N(0, \sigma^2) \tag{1.9}$$

The income y_t follows a stochastic process I(1), Δy_t is a stationary process and ϵ_t is a zero mean stochastic variable. For any $t = 0, 1, .., \infty$, the period-by-period budget constraint faced by the agent is:

$$c_t + (1 + r)d_t = y_t + d_{t+1} \tag{1.10}$$

Where d_{t+1} denotes the debt position assumed (chosen) in period t and r is the risk-free interest rate. The model can be considered as a small open economy so that r is the world interest rate, or as a partial equilibrium model of households in a closed economy, where r is the exogenous risk-free rate. The interest rate is kept fixed, assuming that the

[29] Believes do not depend on the context or on their potential consequences. Further, the agent ignores potential model or parameter uncertainty; his rationality is bounded so that he is not able to deal with complicated frameworks but prefer to adopt simple rules.

supply of credit by the banks is (almost perfectly) elastic as long as the agent respects the inter temporal budget constraint.

Every period t, for $t = 0, 1, .., \infty$, given his beliefs the consumer optimizes rationally choosing the processes $\{c_{t+j}, d_{t+j+1}\}_{j=0}^{\infty}$, given d_t and subject to equations (1.9)-(1.10) and the no-Ponzi constraint of the form:

$$\lim_{j\to\infty} \frac{d_{t+j+1}}{(1+r)^j} \leq 0 \tag{1.11}$$

The following Euler condition is obtained by combining the first-order conditions of the agent's optimisation problem:

$$U'(c_t) = \beta(1+r)\widetilde{E_t}U'(c_{t+1}) \tag{1.12}$$

With rational expectation, the Euler equation is defined in the same way, except for the expectations operator. To achieve a greater tractability, I set $\beta(1 + r) = 1$ and a quadratic utility function such that $U(c_t) = c_t - \frac{\gamma}{2}c_t^2$. Under these assumptions, equation (1.12) becomes $c_t = \widetilde{E_t}(c_{t+1})$; using the inter-temporal budget constraints, the transversely condition and the law of iterated expectations:[30]

$$c_t = \frac{r}{1+r}\Sigma_{j=0}^{\infty}(\frac{1}{1+r})^j\widetilde{E_t}(y_{t+j}) - rd_t \tag{1.13}$$

and

$$\Delta d_{t+1} = \Sigma_{j=1}^{\infty}(\frac{1}{1+r})^j\widetilde{E_t}(\Delta y_{t+j}) \tag{1.14}$$

According to equations (1.13)-(1.14), changes in consumption depend on the revision in the expected permanent income, while changes in debt on expected future variations in income. Coherently with the empirical evidence, the model predicts a positive correlation between expected income growth and changes in debt, for a given interest rate: the demand for debt increases when the representative agent expects his permanent income to increase in the future.

[30] To solve the problem we can apply the law of iterated expectations since the agent takes as correct his beliefs, and cannot predict systematic changes in the forecasts. Further, the transversality condition imposes that the no-Ponzi constraint is satisfied with strict equality. Finally, I assume that consumption is always positive.

1.6.2 The income process

Following Fuster et al. (2010), the income process exhibits partial mean reversion in the long run. Considering an hump-shaped process is relevant for two reasons. First, it is the best way to describe the US RGDP medium-term dynamics which is at the core of this study. Second, it allows us to focus on cases in which the agent slowly learns about the income reversion, and recent observations are key drivers of the expected long-run income. To calibrate the true data generating process (DGP), I estimate several ARIMA(0,1,q) for the US RGDP, with large q. When q is sufficiently large, this flexible representation implies a hump-shaped pattern and partial mean reversion.[31] The baseline model is calibrated on an ARIMA(0,1,18) (Table 1.7). Although it may not be the best process to describe the US RGDP, it is a plausible sufficiently flexible model with an hump-shaped dynamics. Further, results do not depend on the chosen order of the process, but the theoretical impulse responses are very similar when considering alternative orders that exhibit partial mean reversion (i.e. ARIMA(0,1,11)).

1.6.3 Natural expectations

Natural expectations imply that the perceived income process differs from the true one. Given the true DGP for Δy_t, the agent estimates an AR(1), such that for any $j \geq 1$:

$$\widetilde{E_t}(\Delta y_{t+1}) = \rho \Delta y_t \tag{1.15}$$

$$\widetilde{E_t}(\Delta y_{t+j}) = \rho^j \Delta y_t \tag{1.16}$$

Each period, the agent observes the realized Δy_t and updates his short and long-term beliefs using the rules in equations (1.15) - (1.16). Overall, the agent infers a mistaken sequence of income shocks and attributes his forecasts errors entirely to the unpredictable shocks. Otherwise, under rational expectations, if no shocks arise, the agent has no reason to update his forecasts since he has complete knowledge of the true DGP. The agent estimates an auto-regressive model AR(1), according to the out-of-sample forecasting exercise, presented in Table 1.8; he extrapolates from the past, giving excessive weight to recent changes.

[31] For this aim the AIC/BIC criteria are not appropriate since these measures tend to prefer models with few parameters that may not fully capture complex dynamics.

1.6.4 Calibration

An equilibrium is a set of processes $\{c_t, d_{t+1}, \Delta y_t, \widetilde{E_t}(\Delta y_{t+j})\}_{j=1}^{\infty}$ that for any t satisfies equations (1.13) e (1.14), given (1.9)-(1.11), (1.15)-(1.16) and d_0. I calibrate the model at quarterly frequencies aiming at replicating some facts of the US credit cycles. I calibrate Δy_t on an ARIMA(0,1,q) with $q = 18$, estimated on the US RGDP (Table 1.7). The discount factor is calibrated on the average Effective Federal Funds Rate (Percent, Quarterly) over the period 1968:Q4- 2016:Q2 such that $\beta = 0.95$. The standard deviation of the income growth rate shock σ equal to 0.008 is chosen to replicate as well as possible the impulse responses from a VAR in 5 variables (Figure 1.7).[32] The process for the expectations is calibrated on an AR(1) using quarterly data for the US RGDP (Table 1.7), such that ρ is equal to 0.3227.[33] Finally, I set the initial debt to income ratio equal to $d_0 = .10$ to match the average Household credit to real GDP ratio over the period 1968: Q4-2016: Q2.

1.6.5 Impulse responses

Figure 1.5 displays the term structure for the expectations on y_{t+j} and Δy_{t+j} to a shock to Δy_t. The term structure reports the long term expectations $\widetilde{E_1}(\Delta y_{t+j})$ computed at time t=1 for any $j \geq 1$. After a positive shock to income, the agent overstates the persistence of the shock since the AR(1) process implies a perceived persistence equal to 1.479, while the true one is 0.299. The agent does not capture the hump-shaped dynamics of y_{t+j}, under-predicts the income over the short term and over-predicts the long run income (Panel A). This is the case because he estimates an AR(1) and forecasts positive (decreasing) values of Δy_{t+j} until convergence to zero (Panel B). Every quarter, given the realized Δy_{t+j}, the agent updates his short and long term expectations revising upward if $\Delta y_{t+j} > \widetilde{E_t}(\Delta y_{t+j})$ and downward if $\Delta y_{t+j} < \widetilde{E_t}(\Delta y_{t+j})$. Eventually, natural beliefs converge to the rational ones when $\Delta y_{t+j} = \widetilde{E_t}(\Delta y_{t+j}) = 0$. Figure 1.6 reports the impulse response functions to a one standard deviation shock to Δy_t. Under rational expectations, consumption permanently increases on impact and then remains constant. Debt initially raises, because the consumer envisages an increase in the permanent income. When the current income gets larger than the future permanent income (at t=3), debt starts decreasing: the rational agent foresees the hump-shaped dynamics and saves the exceeding income to finance future consumption. Finally, when the income reaches its

[32] Results are similar with $\sigma = 0.006$, the estimated standard deviation of the residuals from an ARIMA (0,1,18).

[33] ρ is not calibrated on the survey data because they are related to the expected RGDP gr over the next year, while in the model $\widetilde{E_t}(\Delta y_{t+1})$ is the expected income growth over the next quarter.

new steady state, debt stabilises at a new level. With natural expectations, credit and consumption rise on impact as the consumer forecast a permanent increase in income (Figure 1.5). Each quarter, the agent revises his expectations on Δy_{t+j} and y_{t+j} for any $j \geq 1$. At t=2, expectations adjust upward, and consumption and debt rise accordingly. The same occurs at t=3. As long as the Δy_t is positive, the consumer expects a positive permanent growth in income and increases debt and consumption; in this phase, in fact, both debt and income increase and finance consumption. When Δy_t gets negative, the expectations immediately adjust downward anticipating a reduction in income; thus both debt and consumption reduce. Eventually, when natural beliefs converge to the rational benchmark, consumption and debt reach the new steady-state level. Over the long run, a positive shock leads to a permanent increase in debt and to a consumption level inferior to the rational benchmark since the agent has to pay back the initial extra consumption. After good news, the model generates a credit boom that does not occur with rational expectations. Further, the model implies an endogenous partial reversion for the debt due to forecast revisions. These dynamics arises because miss-specified beliefs drive the demand for debt.

1.6.6 Model evaluation

Conditional correlations

In this section, I explore the qualitative implications of natural expectations on the household credit path. The theoretical framework is too stylized for full quantitative analysis, but it can help answer the question: can biased expectations generate cyclical changes in debt? To address this question, I compare the impulse responses obtained from the model, with those generated by a VAR(1) in 5 variables:[34]

$$[RGDPgr_t, log(Household\_credit_t), Unrate_t, 10YTBond_t, E_t(RGDP_{t,t+4})] \quad (1.17)$$

Figure 1.7 reports the impulse responses to a shock to the real GDP growth rate ($RGDPgr_t$), from the VAR. The shaded areas represent one-standard-error bias-corrected bootstrap confidence bands of Kilian (1998). The shock is identified using a Cholesky identification scheme so that shocks to fundamentals affect expectations contemporaneously, while shocks to expectations do not affect macro variables on impact. This assumption ensures consistency with the theory, in which shocks to the fundamentals affect beliefs and endogenous variables contemporaneously. The VAR includes two ad-

[34] The order of the VAR is chosen according to the information criteria.

ditional variables to reduce the risk of miss-specification: the unemployment rate and the 10-years Treasury Bond rate as a proxy for the long-term interest rate. Otherwise, if relevant variables are omitted, their presence in the VAR residuals may determine our result. Impulses are consistent when using as an alternative measure of expectations the Michigan Consumer Confidence Index, as well as alternative orders (see the Appendix).

Following a positive shock to the Real GDP growth rate, expectations on income growth jump on impact and then decrease gradually (Figure 1.7). The impulse response of Household credit to innovations in the Real GDP growth rate is significant, fast-building, and declines after few periods showing an hump-shaped pattern. The model captures well the partial mean reversion in debt observed in the empirical impulses, that cannot be matched with rational expectations. The empirical impulse response functions are smoother compared to those predicted from the theory since the VAR(1) implies a smoother process for Δy_t compared to the ARIMA(0,1,18) chosen for the calibration; in turn, also the responses for debt and Δy_t are smoother. In Appendix, I compare natural expectations to alternative hypotheses, such as adaptive expectations, sticky information and noisy signals. These hypotheses have substantially different implications on long-term expectations and on credit responses. The alternative models, except for the adaptive theory, predict an increase in savings after a positive shock to income growth since rational agents know the true DGP and correctly foresee the income path.

Unconditional correlations

Finally, I evaluate the model in terms of unconditional correlations. Table 1.9 reports the statistical moments from the US data and the simulated models, see the Appendix for details on the alternative hypothesis. Data are quarterly to match the frequency of the calibration. The models are simulated for 190 periods of quarterly data, to match the length of the series. I run 1000 simulations with $\sigma = 0.008$, compute the moments for each simulation and average the moments over these 1000 independent simulations.
First of all, the natural expectation model matches well the positive correlation between household credit and income. Under natural expectations, indeed, an increase in RGDP is associated to a rise in household debt; rational expectations, with an hump-shaped fundamental, instead generate a negative relationship between current income and debt as an increase in income is associated to a rise in savings. Then the model matches well the negative relationship between current household credit to RGDP and future RGDP growth rate, at $t + 4$, $t + 8$ and $t + 12$. The observed correlation between current debt

to RGDP and future RGDP growth rate at $t + 12$ is -0.147 and the simulated one -0.171; rational expectations instead imply a positive correlation of about 0.1. Natural expectations capture better these moments compared to the other models except for adaptive expectations which match well the sign of the correlations. These results show a predictable reduction in RGDP growth, coherent to the empirical evidence: a rise in debt to RGDP ratio predicts a decrease of the RGDP growth three years ahead (as in Mian et al. (2017)). The natural and the adaptive expectations models get the true sign of these moments because debt increases during the boom phase, while alternative hypotheses imply an increase in savings. Further, the model captures the mean reversion in debt, getting the negative correlation between current and future debt at $t + 8$ and $t+12$. In the data, the correlation of current and future debt growth rate at $t+12$ is -0.241, while the simulated one is -0.15; rational expectations instead generate a correlation close to zero (-0.07). Overall, the model represents an improvement in matching the moments and the empirical impulse responses relative to the rational case and the alternative hypotheses on the beliefs' process.

1.7 Conclusions

The paper examines the effect of non-rational expectations on the household credit dynamics. Introducing biased beliefs in the demand side is a novel way of thinking about the credit cycle and allows to explain why during good times consumers increase debt in the first place, regardless of credit frictions, house prices or supply shocks. First, I show that expectations on future income growth have strong predictive power for the path of household debt. Second, exploiting survey data, I demonstrate that beliefs depart from rationality in a way coherent with the natural expectations hypothesis. Then, I study the demand side by introducing natural expectations in a consumption-saving model and show that they have strong implications on debt. The ability to generate sizeable credit cycles rests crucially on the assumption made about the beliefs' process. Qualitatively natural expectations represent a significant improvement compared to the rational case; indeed, following a positive shock to income growth, they generate an initial increase in debt and a subsequent reduction, as observed in the data. Rational expectations, instead, cannot explain a similar boom-bust cycle in household debt.

In conclusion, the paper shows that households' biased expectations play a large role in the debt dynamics and alternative assumptions on the beliefs have substantially different macroeconomic implications. The presence of natural expectations alone can generate sizable fluctuations in debt in a simple model with no frictions, heterogeneity nor supply side through the amplification of shocks to the fundamentals. A richer model with the supply side and financial frictions may help in getting a better empirical fit. Natural expectations indeed can amplify the effects of expansionary credit supply shocks that generates an increase in income, such as a reduction in the interest rate or a change in the credit constraint. If such shock arises, it generates a boom during which natural consumers over-estimate the long run level of income over-borrowing relative to the rational benchmark. Hence this demand-side bias exacerbates the boom in debt and represent a strong complementary channel to the supply side hypothesis.

1.8 Figures and Tables

Table 1.1: Summary statistics

Variables	(1) Mean	(2) Std. Dev.	(3) Min.	(4) Max.	(5) N
$E_t(RGDPgr_{t,t+4})$ (SPF)	2.936	1.078	-0.097	5.919	186
Forecast error on RGDP	-0.134	1.661	-7.734	4.252	191
Forecast revision on RGDP	-0.121	0.895	-3.397	2.694	186
$E_t(Unrate_{t,t+4})$ (SPF)	6.503	1.522	4.006	10.5	186
Forecast error on Unrate	-0.029	0.428	-0.887	1.803	191
Forecast revision on Unrate	-0.021	0.431	-0.882	1.803	188
Exp. Change in Business Conditions in 1 Year (Michigan)	108.105	12.67	73	146	191
Exp Change in Unemployment during the Next Year (Michigan)	82.4	17.15	37	126	191
Household credit $\text{gr}_{t.t+4}$	7.671	4.939	-3.8	19.2	191
Savings rate $\text{gr}_{t,t+4}$	7.396	22.367	-49.7	122.7	191
Personal consumption $\text{gr}_{t,t+4}$	6.896	2.928	-3.2	12.7	191
Consumer Price Index	137.171	63.541	35.433	239.409	191
Unrate	6.284	1.582	3.4	10.7	191
10YTBond	6.618	2.914	1.64	14.85	191
3MTBill	4.926	3.349	0.01	15.05	191
GDP	10486.094	3998.852	4844.779	17622.486	191
RGDP $\text{gr}_{t,t+1}$ 3th release	2.558	3.301	-11.416	13.385	191
Unrate_t 3th release	6.275	1.575	3.4	10.6	191
Likelihood of applying for loans (SCE)	1.622	0.018	1	5	9,600
Perc. chance of applying for loans (SCE)	10.161	0.344	0	100	8,207
Exp. financial condition in the next 12 months (SCE)	3.2372	0.017	1	5	11,295
Exp. perc. increase in household income in the next 12 months (SCE)	4.869	0.318	-100	100	11,298

Note: Survey data are from the U.S. Survey of Professional Forecasters (SPF) provided by the Federal Reserve Bank of Philadelphia. Aggregate expectations are the average across forecasters'. Data on household expectations come from the Survey of Consumers by the University of Michigan. Data on Consumer Credit Outstanding and macroeconomic variables are from the FRED. The RGDP $\text{gr}_{t,t+1}$ is the third revision used to compute the forecast errors provided by the SPF. Similarly, Unrate_t 3th release. Data are quarterly, 1968:Q4-2016:Q2. Few observation are missing in the SPF data for the following periods: 1969: Q1-Q3, 1970:Q1 and 1974: Q3. Micro level data are from the Survey of Consumers Expectations (SCE), 2013:Q4-2016:Q4. See Section 2.3 for details.

Table 1.2: Professionals' and households' beliefs: correlations

Variables	(1) $E_t(RGDPgr_{t,t+4})$	(2) $E_t(PGDP_{t,t+4})$	(3) $E_t(3MTBill_{t,t+4})$	(4) $E_t(Unrate_{t,t+4})$
Exp Change in Business Conditions in 1 Year (n.obs 186)	0.670 (0.000)			
Business Conditions Exp during the Next Year (n.obs 186)	0.558 (0.000)			
Business Conditions Exp during the Next 5 Years (n.obs 186)	0.418 (0.002)			
Exp Change in Prices in the Next Year (n.obs 191)		0.866 (0.000)		
Exp Change in Interest Rates during the Next Year (n.obs 140)			0.260 (0.019)	
Exp Change in Unemployment during the Next Year (n.obs 186)				0.114 (0.120)
Probability of Losing a Job during the Next 5 Years (n.obs 65)				0.505 (0.000)

Note: SPF forecasts in columns and Michigan data in rows. All data are quarterly. P-values in parenthesis. See Section 2.3 for details. In the Michigan Survey of Consumers, some variables are available for a shorter period.

Table 1.3: Professionals' and households' beliefs: SPF and Michigan data

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables	Exp Change in Business Conditions				Exp Change in Unemployment			
$E_t(RGDPgr_{t,t+4})$	7.941***	7.573***	7.288***					
	(0.648)	(0.675)	(0.710)					
$E_{t-1}(RGDPgr_{t-1,t+3})$				6.224***				
				(0.732)				
$E_t(Unrate_{t,t+4})$					1.285	1.772***	2.569***	
					(0.823)	(0.662)	(0.672)	
$E_t(Unrate_{t-1,t+3})$								3.087***
								(0.793)
US Business cycle		-3.894*				-29.94***		
		(2.128)				(2.945)		
RGDP gr_t			2.247**				12.58***	
			(0.952)				(1.244)	
Constant	84.93***	86.53***	85.45***	90.13***	74.35***	75.21***	56.16***	62.75***
	(2.028)	(2.197)	(2.059)	(2.291)	(5.498)	(4.408)	(4.710)	(5.305)
Observations	186	186	175	185	186	186	175	185
R^2	0.449	0.459	0.471	0.283	0.013	0.369	0.389	0.076
Adjusted R^2	0.446	0.453	0.465	0.279	0.008	0.362	0.382	0.071
Root MSE	9.509	9.449	9.565	10.72	17.05	13.67	13.50	16.37

Note: Table reports results from estimating $Exp_t(Business_{t,t+4}) = \alpha + \beta E_t(RGDPgr_{t,t+4}) + u_{t,t+4}$. In Columns (1)-(4) $Exp_t(Business_{t,t+4})$ is the Michigan Expected Change in Business condition in 1 year, this is the relative share of agents who expects the Business condition to be better over the next year. In Columns (5)-(6) it is the Michigan Expected increase in Unemployment in 1 year, this is the share of agents who expects unemployment's to raise over the next year. $E_t(RGDPgr_{t,t+4})$ is the SPF expected Real GDP growth rate (RGDP gr) over the next four quarters. Similarly, $E_t(unrate_{t,t+4})$ is the SPF expected unemployment rate over the next four quarters. US Business cycle is a dummy equal to 1 during recessions, 0 otherwise. *** p<0.01, ** p<0.05, * p<0.1. For details see Section 2.3.

Table 1.4: SPF Expectations and Household Credit

Panel A: Expected RGDP gr and unemployment rate

Variables	(1) H. Credit	(2) Savings	(3) Cons.	(4) H. Credit	(5) H. Credit	(6) H. Credit	(7) H. Credit	(8) H. Credit
$E_t(RGDP_{t,t+4})$	2.050***	-2.551	0.725***	1.738***	1.912***	1.528***		
	(0.285)	(1.668)	(0.265)	(0.263)	(0.380)	(0.248)		
$E_t(Unrate_{t,t+4})$							-4.225***	-3.297***
							(1.034)	(0.895)
Constant	-2.463**	16.96***	-0.400	-3.916	0.933	-13.12	3.429	-18.76*
	(1.144)	(6.051)	(0.989)	(4.972)	(5.368)	(10.65)	(5.434)	(10.56)
Baseline controls$_t$	NO	NO	NO	NO	YES	YES	YES	YES
Factors$_t$	NO	NO	NO	YES	NO	YES	NO	YES
Observations	181	181	181	180	181	180	181	180
Adj. R^2	0.660	0.079	0.618	0.795	0.688	0.802	0.639	0.783
Root MSE	2.921	21.70	1.821	2.272	2.797	2.233	3.008	2.338

Panel B: Expected and realized RGDP gr and unemployment rate

Variables	(1) H. Credit	(2) H. Credit	(3) H. Credit	(4) H. Credit	(5) H. Credit	(6) H. Credit	(7) H. Credit	(8) H. Credit
$E_t(RGDPgr_{t,t+4})$	1.197***	1.219***	1.517***	1.720***	1.485***			
	(0.325)	(0.378)	(0.285)	(0.242)	(0.270)			
RGDP gr$_{t,t+4}$	-0.108	-0.179	-0.133	0.131	-0.100			
	(0.0998)	(0.123)	(0.138)	(0.141)	(0.147)			
RGDP gr$_{t-4,t}$	0.753***	0.714***	0.481***	0.692***	0.441***			
	(0.145)	(0.127)	(0.137)	(0.151)	(0.139)			
$E_t(Unrate_{t,t+4})$						-2.198*	0.513	-3.316***
						(1.135)	(0.406)	(0.918)
Unrate$_{t,t+4}$						-1.325***	-0.0740	0.0851
						(0.230)	(0.529)	(0.547)
Constant	-1.857*	-1.214	4.923	11.23*	-0.791	1.853	-19.36*	-19.56*
	(1.076)	(4.985)	(4.758)	(6.530)	(10.79)	(5.186)	(11.21)	(10.60)
Baseline controls$_t$	NO	YES	NO	NO	YES	YES	NO	YES
Factors$_t$	NO	NO	YES	YES	YES	NO	YES	YES
Factors$_{t+4}$	NO	NO	NO	YES	NO	NO	NO	NO
Adj. R^2	0.744	0.762	0.815	0.850	0.815	0.712	0.734	0.782
Root MSE	2.535	2.444	2.158	1.968	2.159	2.690	2.591	2.344

Note: Table reports results from estimating $D_{t,t+4} = \alpha + \beta E_t(RGDPgr_{t,t+4}) + \delta(L)Z_t + \gamma(L)D_{t,t+4} + u_{t,t+4}$ and $D_{t,t+4} = \alpha + \beta E_t(Unrate_{t,t+4}) + \delta(L)Z_t + \gamma(L)D_{t,t+4} + u_{t,t+4}$. $E_t(RGDPgr_{t,t+4})$ is the SPF average forecast of Real GDP growth rate (RGDP gr) over the next four quarters. $E_t(Unrate_{t,t+4})$ is the SPF average forecast of the unemployment rate over the next year. $D_{t,t+4}$ is the percentage change between time t and t+4 of the dependent variable, which can be Household credit, Personal savings, and Personal consumption expenditures. All regressions include two lags of the dependent variable. Baseline controls are: Consumer Price Index, 10YTBond, 3MTBill, GDP, and unemployment rate. Newey-west standard errors are reported in parentheses with 4 lags. *** p<0.01, ** p<0.05, * p<0.1. Few observations are missing in the SPF data. For details see section 1.4.

Table 1.5: SCE Expectations and Household Credit (Micro data)

Panel A: Likelihood of applying for loans (discrete variables)

	(1)	(2)	(3)
Variables	Likelihood of applying for loans		
Expected financial condition in the next 12 months (discrete)	0.179***	0.179***	0.0641***
	(0.0213)	(0.0209)	(0.0234)
Date fixed effect	NO	YES	YES
State fixed effect	NO	YES	YES
Baseline controls$_{i,t}$	NO	NO	YES
Observations	9585	9585	8682

Panel B: Percent chance of applying for loans (continuous variables)

	(1)	(2)	(3)
Variables	Percent chance of applying for loans		
Expected percentage increase in household income in the next 12 months (continuous)	0.0775***	0.0714***	0.0374**
	(0.0166)	(0.0165)	(0.0187)
Date fixed effect	NO	YES	YES
State fixed effect	NO	YES	YES
Baseline controls$_{i,t}$	NO	NO	YES
Observations	8195	8195	7432

Note: Table reports results from estimating $D_{i,t,t+4} = \alpha + \beta E_{i,t}(income_{i,t,t+4}) + \delta Z_{i,t} + u_{i,t,t+4}$. In Panel A, $D_{i,t,t+4}$ is the likelihood of applying for loans (credit cards, auto loans, request an increase in the limit of the credit card), the higher it is the higher is the likelihood. $E_{i,t}(income_{i,t,t+4})$ is the expected financial conditions over the next 12 months, the higher it is the higher is the optimism. Both are discrete variables, and the table shows the results from the ordered probit estimation. In Panel B, $D_{i,t,t+4}$ is the percentage chance of applying for loans (credit cards, auto loans, request an increase in the limit of the credit card). $E_{i,t}(income_{i,t,t+4})$ is the expected percentage change increase in household income in the next 12 months. Both are continuous variables, and the table shows the OLS estimation with robust standard errors in parenthesis. Baseline controls are expected house price, expected credit conditions, expected interest rate, expected stock prices, expected inflation, expected unemployment rate, demographic variables, outstanding debt, the amount requested last year, expected probability not being able to repay in the next 3 months, last year late repayment. *** p<0.01, ** p<0.05, * p<0.1. Few observations are missing in the SCE data.

Table 1.6: Testing ex-post forecast error predictability at aggregate level

Panel A: Expected Real GDP growth (RGDP gr)

Variables	(1) Forecast errors	(2) Forecast errors	(3) Forecast revisions
Lag RGDP gr	-0.0933		
	(0.0597)		
Forecast revision		0.746**	
		(0.288)	
Lag forecast revision			0.421***
			(0.138)
Constant	0.136	0.0267	-0.0204
	(0.256)	(0.182)	(0.0601)
Observations	187	185	184
R^2	0.017	0.067	0.083
F-statistic	0.274	0.036	0.005

Panel B: Expected unemployment rate (unrate)

Variables	(1) Forecast errors	(2) Forecast errors	(3) Forecast revisions
Lag unrate	-0.0852***		
	(0.0303)		
Forecast revision		0.459***	
		(0.155)	
Lag forecast revision			0.356***
			(0.0766)
Constant	0.525**	-0.0506	-0.0109
	(0.214)	(0.0416)	(0.0168)
Observations	187	185	184
R^2	0.100	0.123	0.183
F-statistic	0.012	0.000	0.000

Note: In Panel A, Column (1): $Fe_{t,t+4} = \alpha_1 + \beta_1 RGDP_{t-4,t} + u_{t,t+4}$; Column (2): $Fe_{t,t+4} = \alpha_2 + \beta_2 Fr_{t,t+4} + u_{t,t+4}$; Column (3): $Fr_{t-1,t+3} = \alpha_3 + \beta_3 Fr_{lag,t-1,t+3} + u_{t-1,t+3}$. The Forecast error is defined as $Fe_{t,t+4} = RGDP_{t,t+4} - E_t(RGDP_{t,t+4})$, the Forecast revision on $RGDP_{t,t+4}$ as $Fr_{t,t+4} = E_t(RGDP_{t,t+4}) - E_{t-1}(RGDP_{t,t+4})$, the Forecast revision on $RGDP_{t-1,t+3}$ as $Fr_{t-1,t+3} = E_t(RGDP_{t-1,t+3}) - E_{t-1}(RGDP_{t-1,t+3})$, the Lag forecast revision on $RGDP_{t-1,t+3}$ as $Fr_{lag,t-1,t+3} = E_{t-1}(RGDP_{t-1,t+3}) - E_{t-2}(RGDP_{t-1,t+3})$. The same definitions apply for the expected unemployment rate (Panel B). In Panel A, there are few missing observations for the Forecast error due to missing SPF data. The F-statistics test the null hypothesis of rational expectations: $(\alpha, \beta) = (0, 0)$. Newey-west standard errors are reported in parentheses with 4 lags. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.7: The RGDP process ARIMA(p,1,q)

Variables	(1) ARIMA(1,1,0)	(2) ARIMA(0,1,11)	(3) ARIMA(0,1,18)
$RGDP_{t-1}$	0.3227***	-	-
	(0.598)	-	-
Ma_1	-	0.323***	0.339***
	-	(0.0783)	(0.103)
Ma_2	-	0.227***	0.288***
		(0.0738)	(0.0768)
Ma_3	-	0.156*	0.117
	-	(0.0903)	(0.1)
Ma_4	-	0.153**	0.162
	-	(0.0626)	(0.000)
Ma_5	-	-0.0615	-0.0497
	-	(0.0799)	(-0.131)
Ma_6	-	0.047	0.0807
	-	(0.0818)	(0.0803)
Ma_7	-	-0.0829	-0.0429
	-	(0.0954)	(0.0995)
Ma_8	-	-0.197**	-0.245**
	-	(0.093)	(0.0955)
Ma_9	-	0.0261	0.0851
	-	(0.0943)	(0.106)
Ma_{10}	-	-0.0222	-0.00211
	-	(0.0855)	(0.0847)
Ma_{11}	-	0.168*	0.0755
	-	(0.0935)	(0.105)
Ma_{12}	-	-	-0.172**
	-	-	(0.0837)
Ma_{13}	-	-	-0.0157
	-	-	(0.101)
Ma_{14}	-	-	-0.00953
	-	-	(0.0901)
Ma_{15}	-	-	-0.221**
	-	-	(0.0998)
Ma_{16}	-	-	0.0596
	-	-	(0.0841)
Ma_{17}	-	-	0.00705
	-	-	(0.0866)
Ma_{18}	-	-	-0.157*
	-	-	(0.0881)
Constant	0.0067***	0.00725***	0.00685***
	(0.0008)	(0.0003)	(0.0045)
Observations	190	190	190
Persistence	1.476	0.736	0.299

Note: Quarterly data from 1968:Q4 to 2016:Q2. The dependent variable is $\ln(RGDP_t)$. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.8: Out of sample performance - AR process

Model	MSFE (obs 122)	MAE (obs 122)	MSFE (obs 142)	MAE (obs 142)	MSFE (obs 162)	MAE (obs 162)
L(0)	.1144	.1469	.0902	.1018	.0315	.0407
L(0/1)	.1319	.1671	.0953	.1139	.0342	.0472
L(0/2)	.1392	.1740	.1020	.1213	.0402	.0537
L(0/3)	.1420	.1767	.1077	.1243	.0431	.0555
L(0/4)	.1465	.1790	.1107	.1266	.0432	.0560
L(0/5)	.1496	.1817	.1110	.1285	.0385	.0552
L(0/6)	.1537	.1815	.1169	.1314	.0369	.0543
L(0/7)	.1665	.1868	.1262	.1353	.0357	.0529
L(0/8)	.1765	.1931	.1327	.1404	.0401	.0571
L(0/9)	.1990	.2091	.1409	.1440	.1409	.0622
L(0/10)	.2293	.2272	.1556	.1530	.0646	.0681

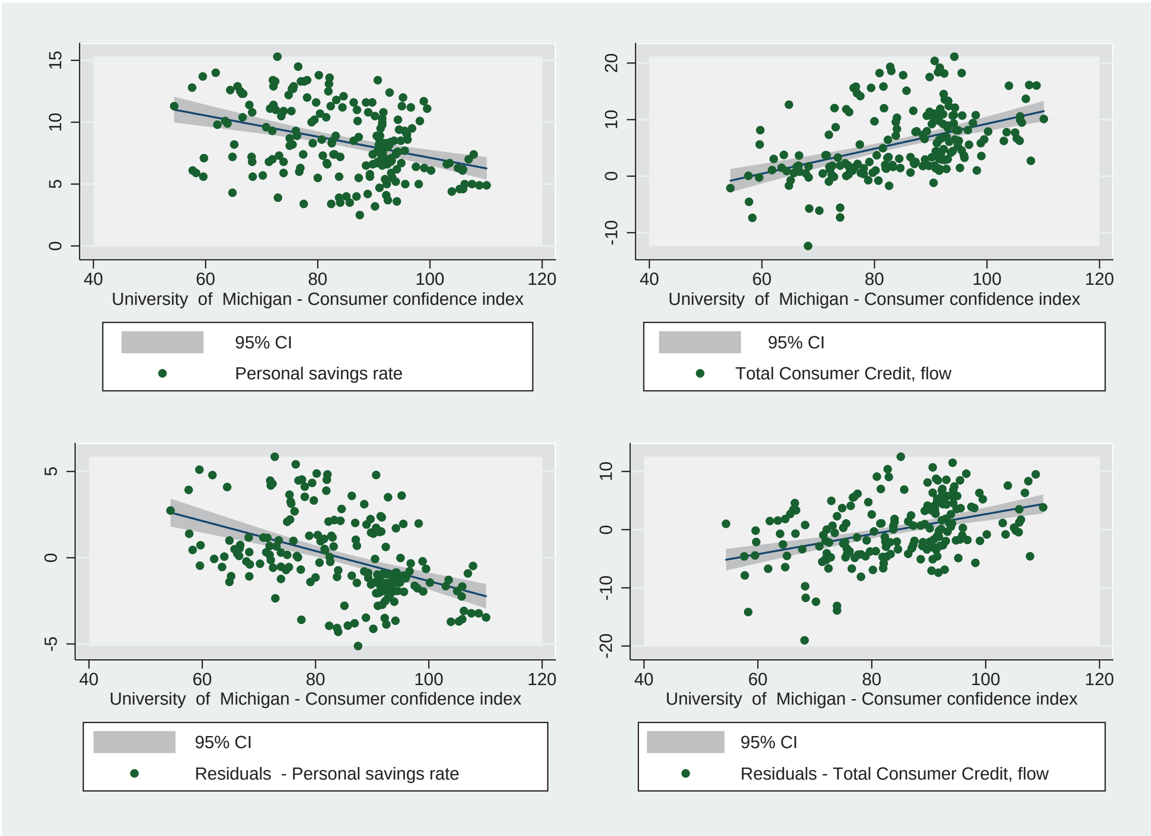
Note: Mean Square Forecast error (MSFE) and Mean Absolute Error (MAE) from an out of sample forecasting exercises used to choose the correct order of the following process: $\widetilde{E_t}(\Delta y_{t,t+4}) = \rho(L)\Delta y_{t-4,t} + \epsilon_{t,t+4}$. For instance, L(0) uses only $\Delta y_{t-4,t}$, while L(0/1) both $\Delta y_{t-4,t}$ and $\Delta y_{t-5,t-1}$ and so on. Different columns indicate different out-of-sample windows. I use SPF survey data on expectations and US ln(RGDPt)$data at quarterly frequency$.

Table 1.9: Unconditional correlations - model evaluation

Correlations	US Data	Rational $\sigma = 0.008$	Natural $\rho = 0.3227$	Sticky info $\lambda = 0.438$	Adaptive $\phi = 0.33$	Noisy signals $\sigma_\eta^2 = 0.00623^2$
$\rho(debt_t,\, rgdp_t)$	0.997***	-0.6794	1.000	-0.6547	0.9646	-0.6960
	(0.996 - 0.997)					
$\rho(\frac{debt_t}{rgdp_t},\, rgdp\_gr_t)$	-0.115	0.2280	0.841	0.2106	-0.0724	0.2173
	(-0.271 - 0.042)					
$\rho(\frac{debt_t}{rgdp_t},\, rgdp\_gr_{t+4})$	-0.109	0.1485	-0.2008	0.1395	-0.212	0.1509
	(-0.265 - 0.047)					
$\rho(\frac{debt_t}{rgdp_t},\, rgdp\_gr_{t+8})$	-0.144*	0.1082	-0.2000	0.1030	-0.1848	0.1117
	(-0.317 - 0.029)					
$\rho(\frac{debt_t}{rgdp_t},\, rgdp\_gr_{t+12})$	-0.147*	0.0747	-0.1711	0.0733	-0.1468	0.0789
	(-0.318 - 0.025)					
$\rho(debt\_gr_t,\, debt\_gr_{t+1})$	0.803***	0.7979	0.2999	0.9128	0.3847	0.8983
	(0.732 - 0.874)					
$\rho(debt\_gr_t,\, debt\_gr_{t+4})$	0.426***	0.3947	0.0847	0.5164	0.15	0.5179
	(0.300 - 0.552)					
$\rho(debt\_gr_t,\, debt\_gr_{t+8})$	-0.046	0.1828	-0.1719	0.1915	-0.0352	0.1945
	(-0.194 - 0.103)					
$\rho(debt\_gr_t,\, debt\_gr_{t+12})$	-0.241***	-0.0682	-0.1494	-0.0740	-0.0761	-0.0810
	(-0.194 - 0.103)					

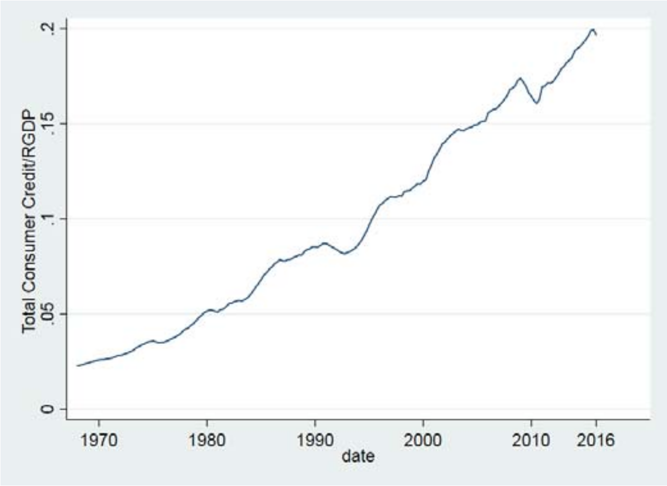
Note: The table compares the correlations computed on the US data in log, at quarterly frequency from 1968:Q4 to 2016:Q2, and the moments from the simulations with alternative assumptions: rational expectations, natural expectations, sticky information, adaptive expectations, and noisy signals. For details on the alternative models see the Appendix. I run 1000 simulations, compute the moments for each simulation and average the moments over the 1000 independent simulations. I set $\sigma = 0.008$ to match the impulse of the VAR in 5 variables and calibrate the model on US quarterly data in log (see section 1.6.4). Confidence intervals in parenthesis are obtained bootstrapping the statistics by re-sampling observations (with replacement). *** p<0.01, ** p<0.05, * p<0.1.

Figure 1.1: Consumer Confidence Index - Michigan Survey



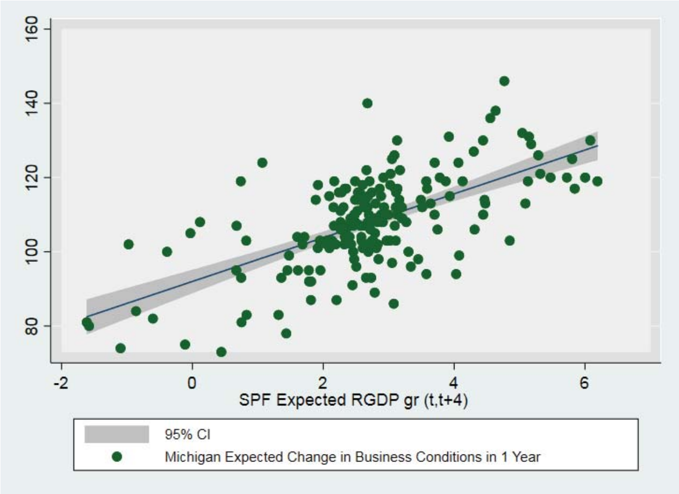
Note: In the upper left hand side panel, the Personal savings rate is reported on the y axis, this is the personal saving as a percentage of disposable personal income (PSAVERT); the University of Michigan Consumer confidence index is on the x axis. The two series are negatively correlated (-0.5***), and the correlation is negative also when considering Personal savings in level rather than the personal savings rate. In the upper right hand side panel, the Total Consumer Credit Owned and Securitized, in flow, is reported on the y axis; the flow data represent changes in the level of credit due to economic and financial activity (FLTOTALSL). The University of Michigan Consumer confidence index on the x axis. The correlation is positive and significant (0.45***). All data are quarterly, from 1968:Q4 to 2016:Q2. The lower panels on the y axis display respectively the residuals obtained from regressing the Personal saving rate and the Total Consumer Credit on the short (3MTbill) and long (10YTBond) term interest rate and the "Net Percentage of Domestic Banks Reporting Increased Willingness to Make Consumer Installment Loans" by the Senior Loan Officer Opinion Survey on Bank Lending Practices (DRIWCIL). Those variables are used as proxies for credit supply. The correlations between the residuals from the Personal saving rate and the Consumer confidence index is negative and significant (-0.45***), while the one between the residuals from the Total Consumer Credit and the Consumer confidence index is positive (0.4***).

Figure 1.2: Total Consumer Credit, outstanding



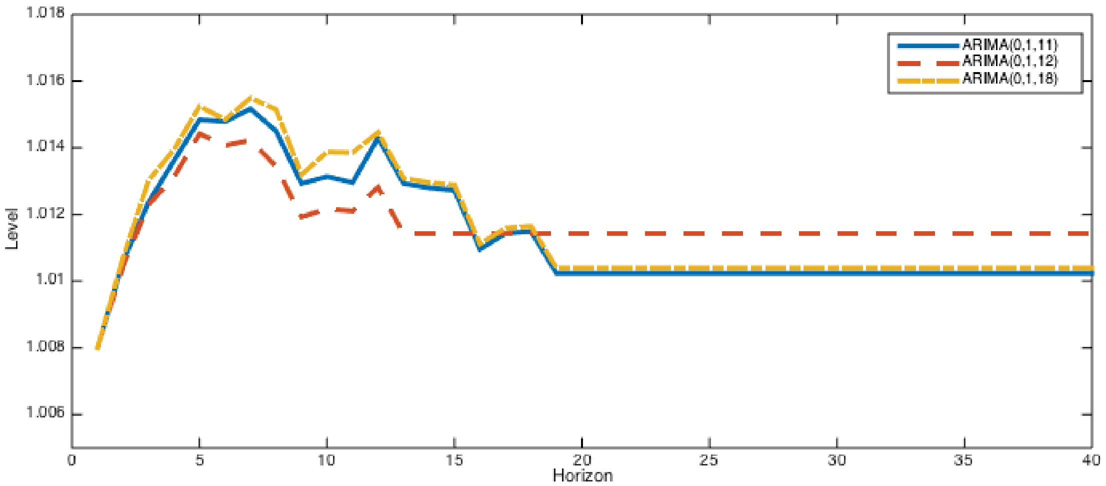
Note: The figure displays the ratio between the Total Consumer Credit Owned and Securitized, outstanding (TOTALSL) in Billions and the Real GDP in Billions of dollars Chn. 2012 \$. All data are quarterly, from 1968:Q4 to 2016:Q2.

Figure 1.3: Professionals' and households' expectations: SPF and Michigan



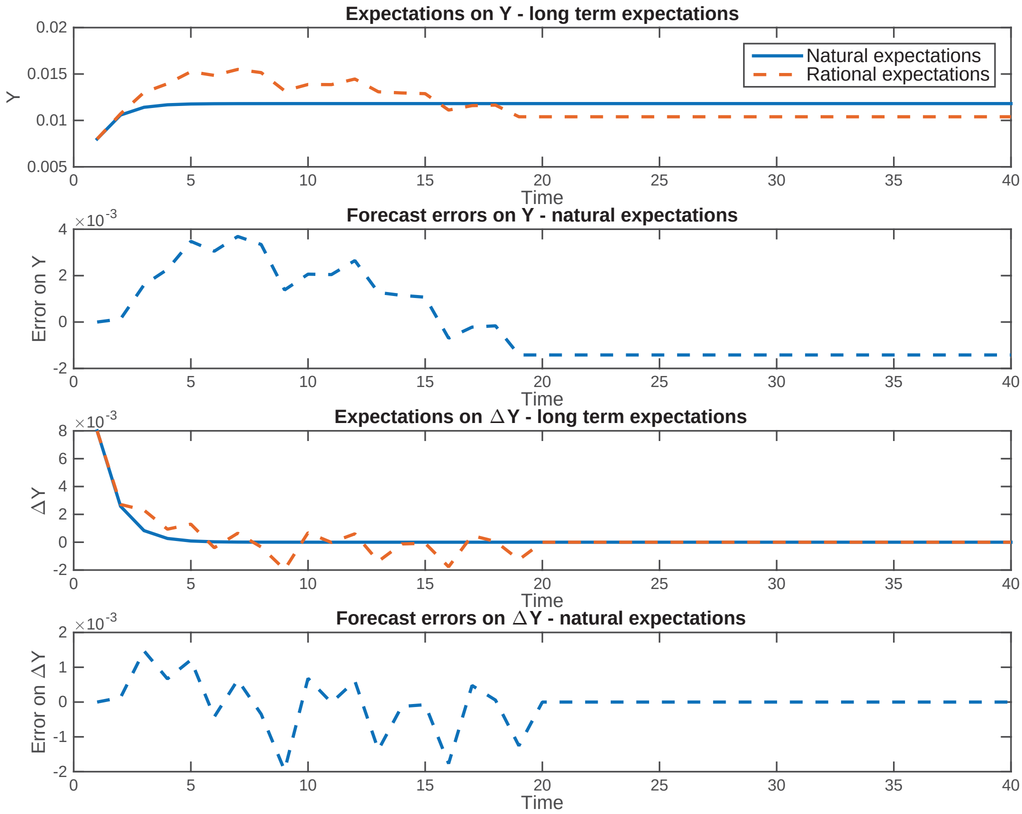
Note: The Expected Change in Business condition in 1 year, on the y axis, is the relative share of agents who expects the Business condition to be better over the next year. The SPF Expected $RGDPgr_{t,t+4}$, on the x axis, is the SPF expected Real GDP growth rate (RGDP gr) over the next 4 quarters. Data are quarterly, 1968:Q4 - 2016:Q2. The correlation is positive and significant (0.6***).

Figure 1.4: Hump-shaped dynamics under flexible ARIMA(p,1,q) models



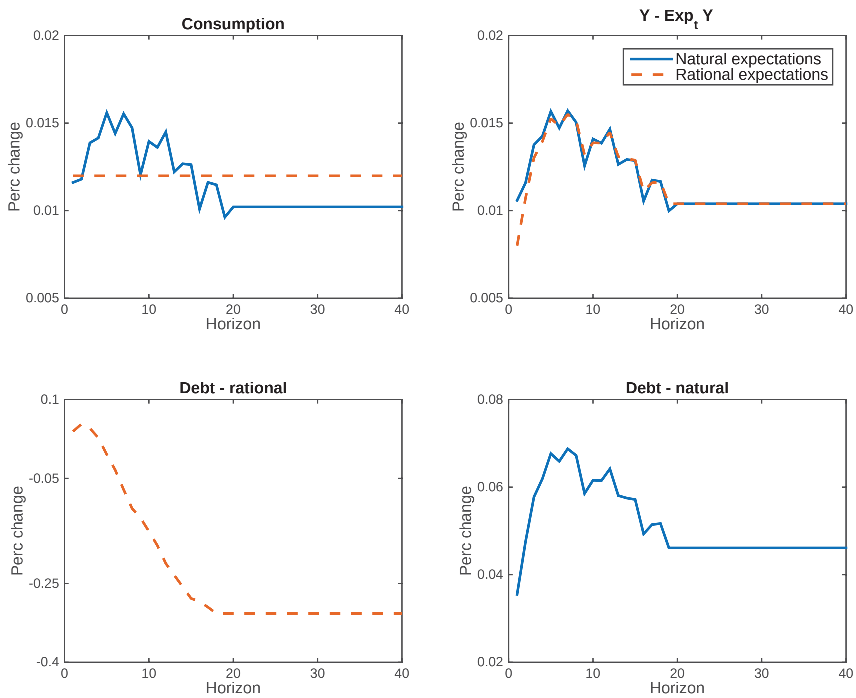
Note: Impulse responses of y_t to a one standard deviation shock to Δy_t, under alternative calibrations. The blue line represents an ARIMA(0,1,11), the orange dashed line an ARIMA(0,1,12) and the yellow dashed an ARIMA(0,1,18). The process for Δy_t is estimated using quarterly data from 1968:Q4 to 2016:Q2 for the US $\ln(\mathrm{RGDP}_t)$.

Figure 1.5: Beliefs term structure following one positive shock



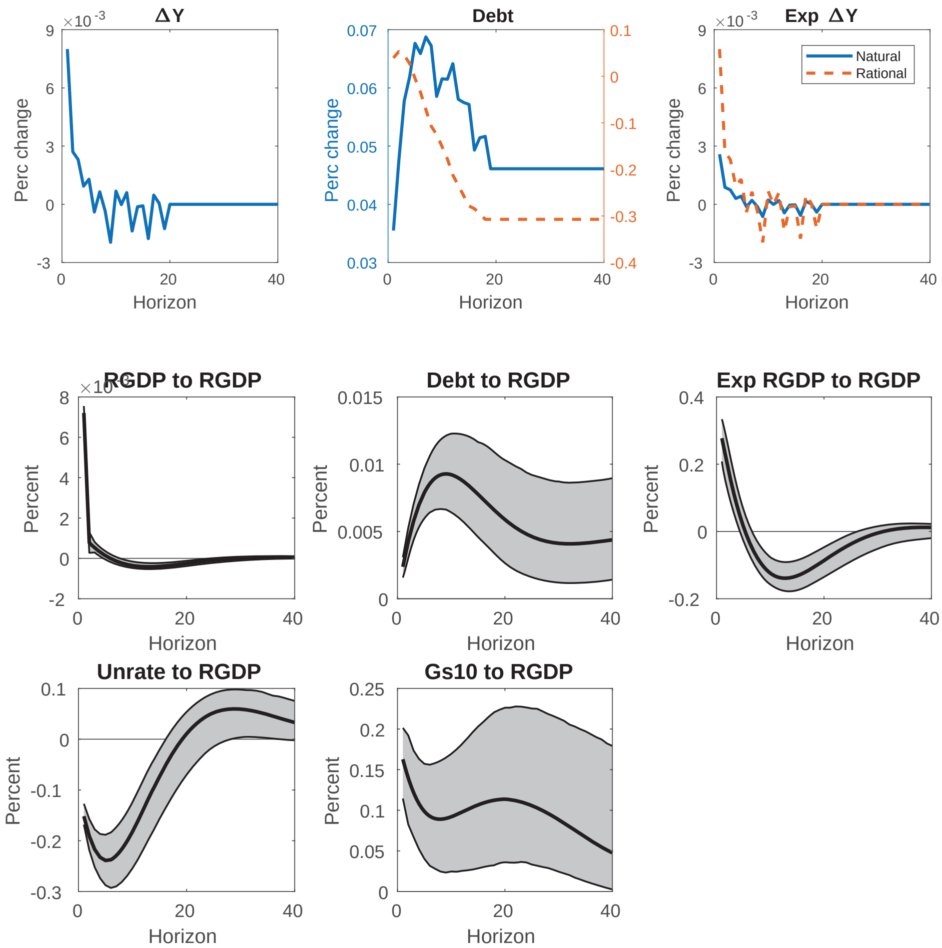
Note: Beliefs term structure (in percentage change) following one standard deviation shock to Δy_t. Dashed orange lines are the rational expectations case. Solid blue lines the natural expectations, with $\rho = 0.3227$. The upper panel displays $\widetilde{E_{t=1}}(y_{t+j})$ and the relative forecast error $(y_{t+j} - \widetilde{E_{t=1}}(y_{t+j}))$ for $j \geq 1$. The lower panel displays $\widetilde{E_{t=1}}(\Delta y_{t+j})$ and the relative forecast error $(\Delta y_{t+j} - \widetilde{E_{t=1}}(\Delta y_{t+j}))$ for $j \geq 1$. Time on the horizontal axis is the forecast horizon in quarters.

Figure 1.6: Impulse responses following one positive shock



Note: Impulse responses following one standard deviation shock to Δy_t. Dashed orange lines are the rational expectations case. Solid blue lines the natural expectations, with $\rho = 0.3227$. In the upper-right panel, the solid blue line is $\widetilde{E}_t(y_{t+1})$ for $t \geq 1$; the dashed orange line instead indicate the realized values at time t, y_t.

Figure 1.7: Empirical and theoretical responses to a shock to the fundamental



Note: The upper graph displays the theoretical impulses to a one standard deviation shock to Δy_t for the variables of interest. The lower one reports the IRFs to a one standard deviation innovation to the Real GDP growth rate ($RGDPgr_t$), from a VAR in 5 variables. Shocks are identified with a Cholesky scheme and $RGDPgr_t$ is ordered first. Shaded areas are the confidence bands. Debt is $ln(Household\_credit)$, Gs10 is the 10 Year Treasury Bond rate (10YTBond), Unrate is the unemployment rate and Exp RGDP is the expected RGDP growth rate over next year ($E_t(RGDPgr_{t,t+4})$). US quarterly data: 1968:Q4-2016:Q2.

Chapter 2

Heterogeneity in inflation expectations and personal experience

(co-authored with Federica Di Giacomo, PhD Candidate at Tor Vergata University)

2.1 Introduction

Inflation expectations play a crucial role in macroeconomic dynamics and monetary policy effectiveness. Monetary policy announcements indeed aim at boosting consumption influencing householdsâ expectations on the future real interest rate, and in turn affecting financial choices. Despite the extensive empirical literature, however, there is little understanding of how beliefs are formed and the reasons behind the sizable disagreement about inflation among households.

Using data from the Michigan Survey of Consumers (MSC), we observe a strong heterogeneity in the expected inflation across income classes. Figure 2.1 plots the inflation expectations one year ahead for four income groups as deviation from the cross-sectional mean in each period for the period 1978-2018. At each point in time, high-income agents expect lower inflation than low-income ones and this is true along the entire time series. The dispersion is economically large; for instance, during the period 2004-2014, the disagreement between the bottom and the top class is on average above 1.3 percentage points. This fact remains unexplained using existing models (i.e. Malmendier and Nagel (2015)) suggesting that more work is needed to gain a deep understanding of the disagreement on future inflation. At the same time, households have different grocery experiences and face heterogeneous inflation rates (Kaplan and Schulhofer-Wohl (2017)). Using data on actual purchases by the Kilts-Nielsen Consumer Panel (KNCP), we calculate inflation rates at the household level and find that they are highly dispersed across income classes over the period 2004-2014 (Figure 2.2). Low-income households experience a higher inflation rate relative to high-income agents. The dispersion is significant, the quarterly inflation rate is on average 0.4 percentage points higher for low-income agents relative to high-income consumers. Given this evidence, in this paper, we aim at understanding if the heterogeneity in the shopping experience can explain the disagreement in the expected inflation.

We contribute to the literature by suggesting a new channel to explain the dispersion in inflation expectations among income classes, using inflation rates computed at the household level (hereafter individual inflation rate). Exploiting data on real purchases, we empirically investigate if the heterogeneity in the personal experience can explain the disagreement in inflation expectations across income classes. The idea is that agents observe prices while shopping and use these pieces of information to predict the future path of prices. Households indeed may use several sources of information beyond the aggregate US CPI when forming their expectations, and one of these is the signal received while shopping. We sketch a model of intuitive thinking in which the agents use both private and public signals to predict the future inflation rate. We introduce a public signal (the US CPI) common to all agents and a private one (the individual inflation rate) which is household specific. The latter is a false signal of the aggregate level of prices that rational agents disregard. We assume instead that agents are non-rational, they adopt an intuitive approach to combine the two signals and differ in how they process them. Heterogeneity hence can derive from differences in the private signals, in the sources of information used, and in the way agents handle the various pieces of information. The intuitive thinking derives from the fact that when asked about their beliefs agents answer the question quickly by naturally leveraging on their personal and vivid memories without a full awareness of their informativeness. Using this framework, we address the following research questions: Do consumers use intuitive reasoning by exploiting personal experience to form inflation expectations? Can heterogeneity in individual inflation rate explain heterogeneity in inflation expectations across income groups? Do agents focus on salient goods when predicting future inflation?

Investigating the relationship between inflation expectations and realised inflation at micro level requires the knowledge of both a householdâs actual purchases and beliefs in the corresponding period. Unfortunately, to the best of our knowledge, none of the existing datasets meets these conditions. Since we cannot run the ideal test, we use data from the Kilts-Nielsen Consumer Panel (KNCP) and the University of Michigan Survey of Consumers (MSC), and we adopt an imperfect matching based on the householdsâ characteristics with a particular focus on the income class. The KNCP is a dataset records the prices, quantities and goodsâ characteristics in 500 million transactions by about 50,000 U.S households from 2004 through 2014. This dataset provides us with detailed information on consumersâ actual shopping and allows us to compute the inflation rates at households level. The MSC dataset instead provides micro-level data on households′ inflation expectations one year ahead. Exploiting these two datasets, we first build measures of realised and expected inflation rates at the individual level and then aggregate them at the income group level. We look only at income groups as the MSC does not provide enough observations to allow for a multi-dimension matching also based on other characteristics such as age, education and geography. Further, the data show that there is sharp disagreement about the expected inflation among income classes, which is larger than the one observed between cohorts and education classes for the period spanned, as explained below.

We address the research questions in a series of steps. First, by employing data from the KNCP, we compute the inflation rate at households level (individual inflation rate), allowing for both households specific consumption bundle, price paid and quantity. Then, we aggregate the individual inflation rates across consumers within the same income class to get an income specific inflation rate. As the second step, using survey data from the MSC, we compute income specific inflation expectations by aggregating the individual beliefs across consumers within the same income class. By looking at the descriptive statistics, we find substantial heterogeneity in the experienced and expected inflation across income groups: those groups who experience higher inflation rates are the same that expect higher inflation. Given this evidence, motivated by the theoretical framework, we empirically investigate whether the shopping experience is a relevant determinant of inflation expectations at the group level. We find that both private (individual inflation) and public (US CPI) signals affect households′ expectations. Further, the groups differ in the weights assigned to the two components. High-income households prefer the aggregate CPI inflation, while low-income classes focus more on their personal experience. This result is robust to several specifications in which we introduce additional sources of information, such as the oil price, house prices and news from the media. Results validate the theoretical framework suggesting that agents adopt an intuitive approach to predict future inflation; households indeed use both signals although the private one has no predictive power for future prices, beyond the US CPI. All income groups form non-rational expectations; overall, high-income agents make more accurate forecasts compared to low-income consumers because they assign a larger weight to the correct piece of information (the US CPI) although they do not process it rationally. We conclude that the heterogeneity in inflation expectations derives from heterogeneity in the individual inflation rate and in the way agents treat the multiple pieces of information.

Then, we investigate if agents overweight prices of salient goods when predicting the inflation rate. Up to this point, we assumed that the shopping experience provides the agents with a unique signal and that the agents recall the prices of all the purchased goods in the same way. We relax this hypothesis and assume people′ s attention is drawn by specific products that come to mind when predicting future inflation. For instance, an agent might be shocked when observing a much higher price for water than the *normal* one he usually pays and hence may have a vivid memory of such a high cost. Alternatively, he may notice any small variations in the price for milk since he buys it very frequently. In particular, we test if agents focus on the prices of the products that: (i) the agent buys more frequently; (ii) require a significant expenditure; and (iii) exhibit large price variations from the previous period. We build our theoretical framework on Bordalo et al. (2013) such that a product is salient in one attribute (i.e. frequency) if it stands out relative to that attributeâs average level in the consumption bundle. For instance, a good may be salient in frequency if it is purchased every day, while the other goods are purchased less frequently. Exploiting the theory, we compute few alternative private signals reflecting price changes of the products salient in these characteristics and test if agents take them into account when forming expectations. Rational agents do not consider their personal experience nor focus on striking goods when predicting future inflation. A salient thinker instead exploits his own shopping experience and is particularly sensitive to the price variations of products that are salient in some attributes. We find that consumers are affected by striking goods which represent an additional source of heterogeneity in the agents′ information set. Agents focus mainly on products that exhibit a significant positive price variation, rather than on products outstanding in terms of frequency of purchase. Overall, salience depicts a channel through which personal experience affects households′ expectations and in turn an additional source of heterogeneity.

Lastly, we show that the intuitive (experience-based) inflation expectations are consistent with consumers' choices by looking at the relationship with the readiness to spend. Groups whose experience leads them to predict a higher inflation rate are more willing to buy durable goods compared to the others. To sum up, we provide evidence that personal shopping experiences play a significant role in casting inflation expectations as consumers recall the observed prices when forming their expectations in different ways. People extrapolate their expertise to the aggregate level and heterogeneous consumption experiences translate into diverse information sets and beliefs. Overall, these findings shed some lights on how expectations are formed and the reasons behind the disagreement on future inflation that policymakers should take into account for monetary policy purposes.

Related literature. Among the others, most importantly, our work connects to the seminal paper by Malmendier and Nagel (2015) which proposes a learning from experience model to explain disagreement about future inflation between young and old individuals. The paper proposes a model in which individuals use historical data to forecast future aggregate inflation rates, but overweight data realised during their lifetimes, so that young individuals assign a larger weight to recent observation compared to old individuals. This mechanism helps explaining disagreement between young and old individuals in periods of high surprise inflation but does not look at disagreement across the income distribution. In our paper, we suggest an alternative channel and look at the personal experience on the (individual) inflation rate across income groups as opposed to experience on (aggregate) inflation at cohort level. For the period 2004-2014, experience on the aggregate inflation rate cannot fully explain differences in expectations across income groups; as shown in Figure 2.3, there is unexplained variation within cohorts, suggesting the presence of alternative explanations.[1] Further, excluding the high-inflation years of the 1970s-1980s, in normal times the disagreement between income groups (Figure 2.1) is substantially larger than the one observed between cohorts, as shown in Figure 2.4.

The paper also connects to the growing literature that studies how personal or local experience influence individuals′ macroeconomic expectations. Within this strand, Kuchler and Zafar (2015) show that individuals systematically extrapolate from the local home prices when asked of their expectations about U.S. house price changes; moreover experiencing unemployment leads respondents to be pessimistic about future nationwide unemployment. Das et al. (2017) instead study how the socioeconomic status influences consumers′ macroeconomic expectations and find that people with higher income are more optimistic about future business conditions, unemployment and stock markets returns. Similarly to these works, we study the link between individuals′ experiences and expectations on aggregate variables, yet here we focus on expectations of future inflation rather than house price, unemployment or economic conditions. Further, we assume that the socioeconomic status does not have a direct effect on people′ beliefs; individuals extrapolate their grocery experience, and the socioeconomic status may have only an indirect effect on beliefs through its impact on consumption choices.

The paper contributes to the extensive macroeconomic literature that looks at heterogeneity in inflation expectations and provides several explanations such as sticky information (Mankiw et al. (2004)), financial literacy (Bruine de Bruin et al. (2010)), inattention to news reports on the views of professional forecasters (Carroll (2003)), information frictions (Coibion and Gorodnichenko (2012b), Armantier et al. (2016)), learning from experience (Madeira and Zafar (2015)) and rational inattention combined with cognitive limitations (Cavallo et al. (2017)). Among the others, Bruine de Bruin et al. (2010) shows that higher inflation expectations are reported by individuals who are female, poorer, single and less educated, claiming that differences in financial literacy may partially explain these demographic differences in inflation expectations. Cavallo et al. (2017) instead use a novel survey experiment to present evidence in favour of rational inattention and cognitive limitations. They show that even if information about inflation statistics is available, individuals use inaccurate sources of information, such as the memories

[1]In the Appendix, we show that there is not much variation at cohorts (or education) level within income groups.

of the prices of the products purchased. These results are in line with our findings, although we use actual purchases data rather than survey data and adopt a different framework to test the theory.

The paper is also closely linked to Johannsen (2014), who analyses the dispersion in expectations within demographic groups. Using microdata from the Consumer Expenditure Survey (CEX) and sub-indexes of the consumer price index (CPI), the author shows that households with low levels of income (and education) have greater dispersion in the experienced and expected inflation compared to those with high levels of income (education). Compared to his work, first, we use actual shopping data and do not assume that all agents pay the same price for a given good. This feature is particularly important since most of the dispersion in individual inflation rates derives from differences in the price paid for a specific good (Kaplan and Schulhofer-Wohl (2017)). Second, we look at between groups dispersion as the variation across classes is substantially larger than the within group one, both for the expected and the realised inflation rates, as shown in Table 2.2.

Further, our paper is related to the literature on individual inflation rates that exploits scanner data. These works study how the inflation rates vary across demographic groups but do not look at the relationship with expectations. For instance, Kaplan and Schulhofer-Wohl (2017) use the KNCP scanner data to document that inflation rates at the household level are highly heterogeneous and have an annual inter-quartile range of 6.2-9.0 percentage points. Jaravel (2019) and Argente and Lee (2015) instead look at the dispersion in the inflation rate over the income distribution studying the reasons and the implications in terms of inequality. Coherently with our estimates, Jaravel (2019) finds that annual inflation for retail products was 0.661 percentage points higher for the bottom quintile of the income distribution relative to the top quintile. Similar patterns are shown by Argente and Lee (2015). Relative to this literature, this article refines the methodology according to our purpose and investigate the link with the households$'$ expectations.

Overall, this Chapter contributes to the existing literature by suggesting a new channel to explain the disagreement about future inflation by linking the strands that investigate how inflation expectations are formed, the dispersion in the inflation rate computed at the household level and how personal experience affects agents$'$ macroeconomic expectations.

Layout. The remainder of the paper is organised as follows. Section 2 presents a model of intuitive inflation expectations. Section 3 describes the data, the empirical methodology used to compute the income specific inflation rates and the descriptive statistics. In Section 4, we test the theoretical framework and comment on the primary results. Section 5 presents an extension of the model and explores the role of salience goods in the intuitive thinking. Section 6 investigates the link between the intuitive expectations model and consumers′ choices. Finally, Section 7 reports the main conclusions.

2.2 A model of intuitive inflation expectations

Consider a model with heterogeneous agents; each agent j forms expectations on the future inflation rate at each point in time, given the information set available. We make two main assumptions. First, all agents know the current level of inflation π_t, which is a public signal, common to all individuals. Second, each agent j goes shopping and evaluates $N_{j,t}$ goods in a choice set $C_{j,t} = \{q_{j,k,t}, p_{j,k,t}\}_{k=1,2,...,N_{j,t}}$. The choice set is composed of all the goods that the agent j buys in a given period t, and two attributes describe each good k: the quantity bought $q_{j,k,t}$ and the price paid $p_{j,k,t}$. Both attributes are individual and time specific. The shopping experience provides the agent with a private signal on the change in the price level of goods $s_{j,t}(\pi_t)$.[2] This is however only a pseudo signal, which does not convey additional information content on π_{t+1} beyond π_t since it reflects only a tiny fragment of the overall change in prices. Rational agents, hence, do not take into account the private signal when forming expectations on the aggregate inflation rate and use the following rule:

$$E_t(\pi_{t+1}) = f(\pi_t) \tag{2.1}$$

We assume instead that the agents mistakenly believe that the signal is representative of the general level of prices. Now, when asked to form expectations on the future level of the inflation rate, the agents recall both the public and the private signal received during the shopping experience:

$$\tilde{E}_{j,t}(\pi_{t+1}) = g_i(s_{j,t}(\pi_t), \pi_t) \tag{2.2}$$

This approach can be explained by the fact that agents form expectations and evaluate the problem based on what first comes to their mind (Tversky and Kahneman (1974), Gennaioli and Shleifer (2010)); people may be able to form more accurate expectations but choose to answer the question quickly recalling both memories from their personal experience and the collective knowledge. Intuitive agents develop their expectations by processing the various pieces of information, without complete awareness of their interpretation or informativeness. People leverage on the memories that can be easily retrieved and personal experiences are readily accessible (Kahneman (2003)) compared to other pieces of information such as economic news. We say

[2] $s_{j,t}(\pi_t)$ indicates that the signal is believed to be informative about future inflation π_t. Notice that the signal is not a function of π_t.

that people adopt an intuitive approach in predicting future inflation using all signals received.[3]

Assuming a linear function, the rational agent uses the following rule:

$$E_t(\pi_{t+1}) = \gamma * \pi_t \tag{2.3}$$

The intuitive agent, instead, exploits both signals:

$$\tilde{E}_{j,t}(\pi_{t+1}) = \beta_j * s_{j,t}(\pi_t) + \gamma_j * \pi_t \tag{2.4}$$

The parameters β_j and γ_j are individual specific and determine the gap between the final assessment and the rational benchmark. We assume they are time-invariant so that each agent uses the same forecasting rule over time. Heterogeneity in the weights assigned to the private and the public signals depends on how easily agents recall them; we allow agents to differ in the facility with which they retrieve these pieces of information so that some agents might find easier to remember their own experiences or the news from the media than others. Heterogeneity in expectations hence derives from heterogeneity in the private signals and the weights assigned to the two components. In the next sections, we test this hypothesis and investigate these two sources of heterogeneity.

2.3 Data and descriptive statistics

To test the proposed model, we need to build a proxy of the private signal and a measure of the future expected inflation rate. For the former, we use the individual inflation rate computed by exploiting data on actual purchases from the Kilts-Nielsen Consumer Panel (KNCP). For the latter, we use the Michigan Survey of Consumers (MSC). See the Appendix for the summary statistics on the demographic characteristics.

2.3.1 Kilts-Nielsen Consumer Panel (KNCP)

To get a proxy of the private signal from the shopping experience (individual inflation rate), we use data provided by Kilts-Nielsen Data Center at the University of Chicago Booth School of Business. The KNCP is a panel dataset that tracks the shopping behaviour of 40,000 to 60,000 US households drawn from geographically dispersed markets, known as Scantrack markets, corresponding to 54 Metropolitan Statistical Areas (MSA) over the period 2004: m1 to 2014: m12. Individuals in the panel provide information about each of their shopping trips using a home

[3] Alternatively, we can think that the signal affects households perceptions on the current level of inflation which in turn determine the expectations.

scanner to record the details of their purchases: the items purchased, the number of units for each item, the date and store where the transactions were made. If the panellist buys the good at a store covered by Nielsen, the price is set automatically equal to the average price of the item at that store during the week when the purchase occurred. Otherwise, the panellist enters the price. Households also record whether the item was acquired using store features, coupons, or other deals and in case input the value of the ticket. A 12-digit Universal Product Code (UPC) identifies the items the panellists′ purchase at the barecode level. For each UPC, the data provide information on the brand, size, packaging, and other product characteristics. KNCP defines a product hierarchy so that UPCs are classified in departments (i.e. dry grocery, general merchandise, health and beauty car, fresh produce, frozen food, alcohol, etc.), product groups (i.e. soup, beer, cosmetics, frozen pizza, detergent, candles, etc.) and detailed product modules (i.e. mozzarella cheese, soft-drinks carbonated, soft-drinks low-calorie, etc.). Overall, the data contain around 1.5 million distinct UPCs grouped into 1,075 product modules, about 124 product groups and 10 departments. For instance, *Six-Pack of Coca-Cola Classic* is a UPC, linked with the product module *Soft drinks-carbonated*, the product group *Carbonated beverages* and the department *Dry Grocery.* Data cover the transactions made by all the components of the households. Demographic data on household members are collected at the entry into the panel and are updated annually through a written survey during the fourth quarter of each year. The available information includes age, education, marital status, employment, type of residence, and race. See the Appendix for the summary statistics.

By using the appropriate weights provided by the KNCP, the dataset is representative of US expenditures. To check the representatives of our data, we compare the evolution of spending in the Consumer Expenditure Survery (CEX) and Nielsen finding that the latter represents well the US total spending (Figure 2.5). Still, one limitation of the dataset is that it does not contain all of the transactions that take place in the market. The data indeed cover around 40% of all of the expenditures on goods in the CPI, it over-represents some categories such as food, under-represents others (i.e. transportation or housing) and excludes few classes as gasoline or medical care. For further details on the goods included in the KNCP, see Kaplan and Schulhofer-Wohl (2017).

2.3.2 Individual inflation rate

Using the KNCP data, for each household in the panel, we can compute the individual inflation rate given his experienced consumption. The inflation rate calculated at the household level differs from the aggregate inflation rate reported by the authorities since the latter is calculated on a representative consumption bundle with average prices for goods and services which may not correspond to the consumption bundle and prices experienced by a given household. This index indeed does not allow consumer choices of products to vary across socio-demographic groups nor in response to relative price changes (Boskin et al. (1996)). The reasons why the individual inflation rate deviates from the aggregate statistics are three: (i) each family can buy a different coterie of goods compared to the representative bundle; (ii) the weight that each good plays in the personal basket can vary across households; (iii) the price paid for the same good can be different across families. Our data allow us to deal with all these sources of heterogeneity, enabling us to build an inflation index that takes into account the heterogeneity in the set of the purchased products, prices, and weights. Considering the actual costs is an essential advantage of our analysis, in fact, as a pointed out by Kaplan and Menzio (2015) in the U.S. there is a pervasive dispersion in prices for an identical good over markets and time periods; moreover, the variation in the individual inflation rate derives mostly from differences in the amount paid for a given product as shown by Kaplan and Schulhofer-Wohl (2017). Note that KNCP data allow us to overcome these limitations of the aggregate statistics; however, it meets another restriction as it does not consider all the products consumed by the households and included in the US CPI but only a subset of it.

For each household, we compute individual inflation rates at the monthly frequency and product module level. The monthly basket for the household j is composed of all the products modules that the household j purchases in three consecutive periods (months), the current one and the previous two months. We adopt a rolling scheme so that when computing the inflation rate, the price indexes refer to the same set of goods, which are those purchased in the current and the previous two months. Items not purchased in three consecutive months are omitted from the computation. To each product k we assign a weight equal to the quantity purchased expressed in terms of volume (i.e. grams or litres) and the price paid by household j at time t. If a household buys the same product more than once in a month, we set the household$'$ s monthly price for that item equal to the volume-weighted average of prices that the household paid.

We define the goods according to the product module classification as we believe this is the appropriate measure for our purposes. The classification at the UPC level would be too detailed for computing the inflation expectations and would reduce substantially the average dimension of the baskets of goods on which we calculate the inflation rates. For instance, *Six-Pack of Coca-Cola Classic* and *Two-pack of Coca-Cola Classic* correspond to two different UPCs; therefore, adopting the UPC classification, if a household buys a *Six-Pack of Coca-Cola Classic* in month t, and one (or more) *Two-Pack* in month $t + 1$, the goods are excluded from the computation of the inflation rate. Otherwise, both products are linked with the product module *Soft drinks-carbonated*, so when using the product modules classification, both transactions are considered in the estimate. Our procedure differs from Kaplan and Schulhofer-Wohl (2017) as the latter, define the items by barcodes (UPCs) and calculate the quarterly householdâs inflation rate, considering only goods that the household bought two consecutive quarters.

For each household j, we compute three standard price indexes: the Laspeyres index, which weights the price changes between two dates by basket at the initial date, the Paasche index, which weights the price changes by the basket at the final date and the Fisher index which is the geometric mean of the previous two.

1. Laspeyres index: $L_{j,t} = \frac{\sum_k P_{j,k,t} Q_{j,kt-1}}{\sum_k P_{j,k,t-t} Q_{j,k,t-1}}$
2. Paschee Index: $P_{j,t} = \frac{\sum_k P_{j,k,t} Q_{j,kt}}{\sum_k P_{j,k,t-1} Q_{j,k,t}}$
3. Fisher Index: $F_{j,t} = \sqrt{(}L_{j,t} \times P_{j,t})$

Where $P_{j,k,t}$ is the price paid by the household j for item k at time (month) t, and $Q_{j,k,t}$ is the quantity consumed by the household j for item k at time t. Using these indexes, we compute three measures of inflation at household level:

1. Laspeyres inflation rate: $\pi^L_{j,t} = (\frac{L_{j,t}}{L_{j,t-1}} - 1)$
2. Paschee inflation rate: $\pi^P_{j,t} = (\frac{P_{j,t}}{P_{j,t-1}} - 1)$
3. Fisher inflation rate: $\pi^F_{j,t} = (\frac{F_{j,t}}{F_{j,t-1}} - 1)$

In the paper, we report results obtained using the Laspeyres index as the three indexes are strongly correlated. On average the individual inflation rate is about 0.356 (Table 2.3). The CPI inflation rate and our index are not directly comparable because they consider different sets of products and our index over-weights food and beverage and omit other products such as energy and transportation. For these reasons, the two measures differ in the average level and dynamics, and a light correlation links them (about 0.2, see Appendix). This evidence is in line with the relationship between the aggregate CPI inflation and the one computed only on food and beverage, which is about 0.2 on a monthly basis, as shown in Appendix.[4] The detected low correlation between the individual inflation rates and the aggregate CPI inflation rate suggests that the two measures capture different pieces of information that may be reflected in the experience-based expectations.

[4] The low correlation is observed because the CPI components tend to have different patterns.

2.3.3 Michigan Survey of Consumers (MSC)

To retrieve information about households inflation expectations, we use microdata from the Survey of Consumers conducted by the Survey Research Center at the University of Michigan. The samples for the Surveys of Consumers are statistically designed to be representative of all American households, excluding those in Alaska and Hawaii. Each month, a minimum of 500 interviews are conducted by telephone. The survey also collects demographic information for the surveyed households. In particular, they provide information about age, gender, education, marital status, income and race. These data are available at a monthly frequency and cover the period of our analysis 2004:01 to 2014:12. Consumers are asked two related questions about inflation expectations. The first question asks:

Q1. "During the next 12 months, do you think that prices, in general, will go up, or down, or stay where they are now?"

Q2. "By about what percent per year do you expect prices to go (up/ down) on the average, during the next 12 months?"

The MSC combines the two answers to derive the direction of the expectations (up or down) and get a numerical value of the expected change (Curtin (1996)). Following Bachmann et al. (2015), we set the expectations equal to 20% if (in the data) the expected inflation is higher than 20% and equal to -20% if the expected inflation is lower than -20%.[5] These values may be due to typos, lack of attention or misunderstandings. We set a threshold so high because as explained in Curtin (1996) at each step of the data collection extreme values are verified, interviewers probe unusually large responses, asking respondents to confirm their answers and check that the values were accurately transcribed, so remaining extreme values are assumed to reflect the respondentâs answers correctly.

2.3.4 Descriptive statistics at income class level

As the first step, we look at the descriptive statistics on the experienced and expected inflation rates to understand how they differ between income groups. To study the relationship between personal spending and inflation expectations we aggregate the individual data over income groups, distinguishing between four groups according to their annual income: (i) less than 25,000, (ii) 25,000 to 50,000, (iii) 50,000 to 100,000, and (iv) more than 100,000. These

[5]Overall, we drop a few observations most of which because above 20 per cent and relative measures shows that a large share of the dropped observations is associated with low-income respondents. When computing the $ratio_i$ % as the Number of dropped observations for the income class i/ Total numbers of observations for income class i, we find that the deleted observations involve mainly low-income household.

groups match the quartiles of the income distribution in the KNCP data and are representative of the US income distribution. Using these classes, we build four time series of the individual inflation rate, one for each income class. Similarly, we adopt these classes for the MSC data and get four inflation expectations time series.

We find substantial heterogeneity in the inflation rate experienced across income groups. Figure 2.7 reports the mean and the median of the inflation rate (Laspeyres index) across income classes. The households in the highest income class experiment lower inflation than the others and the individual inflation rate is decreasing with income. This finding is consistent with Jaravel (2019) who show that during the period 2004-2015, the annual inflation was 0.65 percentage points lower for households earning above $100,000 a year relative to those earning less than $30,000 a year. Jaravel (2019) explains this fact by showing that product innovation disproportionately benefited high-income households; firms innovated more in the products targeted to high-income consumers as a result prices in this market segment lowered due to the competitive pressure. Further, we document considerable disagreement on expected inflation: the forecasted inflation is decreasing with income and those groups who expect higher inflation are those experiencing higher inflation (Figure 2.8). A massive wedge arises between the consumers′ inflation expectations and the CPI inflation rate. Households mistakenly report the inflation in level, although they capture the dynamics of the inflation rate (Figure 2.6); consumers′ inflation expectations indeed are positively and strongly correlated with the US CPI inflation (about 0.58).[6]

2.4 Testing the theoretical framework

In this section, we test the theoretical framework, discussed in Section 2.2. In the empirical analysis, we consider four representative agents i one for each income group, assuming that agents within the same income group receive similar signals and similarly process the pieces of information available. As explained in the introduction, this assumption is necessary given the data limitation; however, it appears reasonable considering that the within-group variation is limited compared to the between-class dispersion (Table 2.2).[7] For each group, we adopt the average inflation rate and the average expected rate described in Section 2.3.4, and we use the former as a proxy for the private signal. Adopting this approach we make three additional implicit assumptions. First, agents recall only the prices of the purchased goods; agents are likely to observe prices of other products on the shelf, but they remember better the paid costs

[6] Consumers' inflation expectations indeed are also positively and strongly correlated with professional forecasters' beliefs.

[7] Likewise, the within-class variation in the expected inflation is lower compared to the across-class variation (see Appendix).

rather than the observed one. Second, households recall the actual amount spent, without considering the effect of potential sales or coupons. We make this assumption because it's not clear how consumers evaluate the discounts, they might interpret them as a sign of decreasing prices or as an exceptional event (i.e. black Friday). Third, we assume that these pieces of information are shared within the household as the NNCP data covers the transactions of all individuals within the family.

2.4.1 Baseline model

We test the theoretical model by estimating the equation (2.5) using the the panel data dimension. We regress the average expected inflation rate formulated at time t by the income group i on a proxy of the private signal and the US aggregate inflation, such that:

$$\bar{E}^i_{i,t}(\pi_{t,t+12}) = \alpha + \beta\bar{\pi}^i_{i,t-12,t} + \gamma\pi_{t-12,t} + \epsilon_{i,t} \tag{2.5}$$

Where t is the month, and $\bar{E^i}_{i,t}(\pi_{t,t+12})$ is the average expected inflation rate over the next twelve months of income group i at time t. $\bar{\pi}^i_{t-12,t}$ is the average monthly realized individual inflation of the same income group i over the last year. $\pi_{t-12,t}$ instead is the average monthly aggregate CPI inflation rate over the last year.[8] We use these averages, computed between time $t-12$ and time t, to capture the idea that agents retrieve memories received in a long time horizon, not only recently in the current month t. For the same reason, we do not merely use the percentage change from the previous year as it would not to summarise the entire information set received in the last 12 months. Equation (2.5) is our baseline regression; then we augment it with income fixed effect and interactions terms to allow for income-specific weights to $\bar{\pi}^i_{t-12,t}$ and $\pi_{t-12,t}$. Table 2.4 displays the estimated results. The specification in Column (1) considers only the public common signal. Column (2) introduces the private signal and imposes that the weights assigned to the two components are invariant across groups. Column (3) instead allows for heterogeneity in the weight assigned to the public signal (and the intercept), while Column (4) to the private signals (and the intercept). Finally, Column (5) allows for heterogeneity in the weights assigned to both the private and public signals (and the intercept). Results show that people indeed use both private and public signals to form their expectations. Income groups differ in the relative importance granted to the two sources of information; low-income groups assign a significantly larger weight to the private component compared to high-income groups (Column (4)). Further, we investigate if heterogeneity in expectations derives mainly form heterogeneity in the individual inflation rates or the weights assigned to the two signals.

[8] $\bar{\pi}^i_{t-12,t}$ and $\pi_{t-12,t}$ are monthly averages computed as geometric means of monthly inflation rates, individual or aggregate, between time $t-12$ and time t. For instance, $\pi_{t-12,t} = x$ indicates that prices increased monthly by $x\%$ on average over the last 12 months. Alternatively, results do not change when using the annualized inflation rate (see Appendix).

Comparing Columns (1) and (2), we observe that the heterogeneity in the private signal helps to explain the variation in the dependent variable, as the adjusted R-sq increases from about 0.09 to 0.18. Comparing Columns (2) and (5), instead, we observe that the adjusted R-sq increases considerably when we add the interaction terms suggesting that income-specific weights are relevant to explain heterogeneity in expectations and income specific weight to $\bar{\pi}^i_{t-12,t}$ are particularly powerful (Column (4)). Overall, allowing for income-specific weights, the public and the private signals explain more than 74% per cent of the variation in the expected inflation rate across income classes.[9] The unexplained variation may be due to additional pieces of information relevant to the forecasts and omitted in the model.

Then we test the theoretical model by estimating equation (2.6) for each income group at aggregate level to exploit the time series dimension. Specifically, we run 4 separate regressions, one for each income group, in which we regress the average expected inflation rate over the next year formulated at time t by the income group i on the a proxy of the private signal and the US aggregate inflation, such that:

$$\bar{E}^i_t(\pi_{t,t+12}) = \alpha + \beta\bar{\pi}^i_{t-12,t} + \gamma\pi_{t-12,t} + \epsilon^i_t \tag{2.6}$$

Here t is the month, and $\bar{E}^i_t(\pi_{t,t+12})$ is the average expected inflation rate over the next twelve months of income group i at time t. $\bar{\pi}^i_{t-12,t}$ is the average monthly realized inflation of the same income group i over the last year. $\pi_{t-12,t}$ is the average monthly US CPI inflation rate over the last year, between time $t-12$ and time t. Table 2.5 confirms that individual inflation affects expectations for most of the income groups. Most coefficients are positive and statistically significant, suggesting that people extrapolate their personal experience to the aggregate inflation rate. The effect is stronger for the lowest income class for which the increase of 1 percentage in the average individual inflation rate is associated to a rise of 2.13 percentage points in the average expected inflation; this is not a negligible increase considering that the average individual inflation rate is for this class is 4.5. On the other hand, for the highest income class, an increase of 1 percentage in the average personal inflation rate is associated to a rise of 0.55 percentage points in the expected inflation; however, estimating four separate models results in loss of statistical power and the coefficient is not significant. The adjusted R-sq for the highest income groups, in fact, is lower compared to those observed for the other groups suggesting that high income agents also exploit different sources of information. In Section 2.4.3 we test this hypothesis. The coefficient associated with US CPI inflation is statistically significant for the entire income distributions, meaning that all households use public information when predicting future inflation. Results are robust when we replace the average expected inflation for each income group with the median for each group (see Appendix).

[9] The improvement in the adjusted R-sq is due to income fixed effects as well as the interaction terms.

2.4.2 Is it rational to use the individual inflation?

So far, we have shown that agents use an intuitive forecasting rule as the one described in equation (2.4). People exploit their personal experience to form their expectations and the weights assigned to the private and public information differ across income groups. Now we test if a rational agent should use the model described in equation (2.1) and if it is rational for agents to use their individual inflation to predict aggregate inflation. Precisely, we check whether the private signal ($\bar{\pi}^i_{t-12,t}$) has additional predictive content relative to the public one ($\pi_{t-12,t}$) to forecast aggregate inflation $\pi_{t,t+12}$ over the next year. The intuition is the following. We assume that agents correctly understand the question and predict (think of) the aggregate US inflation when reporting their answer; if entirely rational, the respondents should use the best predictors to form their expectations as they would have complete knowledge of the true generating process of the inflation rate and they would fully exploit the available information set. Therefore, rational agents should use their private signals only if they convey some useful predicting content. We run the following forecasting equation for each income group and compare it with the previous results from the regression (2.6):

$$\pi_{t,t+12} = \alpha + \beta\bar{\pi}^i_{t-12,t} + \gamma\pi_{t-12,t} + e_t \tag{2.7}$$

Where i is the income group, and t is the month. We find that the coefficient γ is positive and significant suggesting that past US inflation is a strong predictor of future US inflation (Table 2.6). On the contrary, the coefficient β is not significant, suggesting that $\bar{\pi}^i_{t-12,t}$ does not convey useful information to predict $\pi_{t,t+12}$, conditional on $\pi_{t-12,t}$. Rational agents hence should look exclusively at the US CPI as the private information is a false signal, conditional on the public signal, as assumed in our theoretical framework. Comparing Tables 2.6 and 2.5, we conclude that agents are not rational, but they have biased expectations as they miss-use both the private and public inflation to predict future US aggregate inflation. The bias is strong for all income classes, although they differ in the weights assigned to the two signals. The lowest income class over-weights the personal inflation relative to the optimal benchmark as well as the aggregate US inflation; the higher-income groups instead over-weight the US aggregate inflation compared to the benchmark. High-income consumers make more accurate predictions compared to low-income households because they focus on the correct piece of information $\pi_{t-12,t}$; however they do not have rational expectations nor process the data in a better way.

As an additional test, we define the ex-post forecast error and check whether this is predictable given the personal and aggregate signals. Under the hypothesis of rational expectations, the forecast error should be unpredictable given the information set available when the forecast is made. For each income group i, we define the forecast error and run the test as follows:

$$fe_t^i = (\pi_{t,t+12}) - \bar{E}_t^i(\pi_{t,t+12}) \tag{2.8}$$

$$fe_t^i = \tilde{\alpha} + \tilde{\beta}\bar{\pi}_{t-12,t}^i + \tilde{\gamma}\pi_{t-12,t} + \epsilon_t^i \tag{2.9}$$

Under the hypothesis of rational expectations the coefficients $(\tilde{\alpha}, \tilde{\beta}, \tilde{\gamma}) = (0, 0, 0)$ for each income group. Table 2.7 displays the results and confirms our previous findings. Consumers extrapolate the personal inflation rate relative to the rational benchmark, as $\bar{\pi}_{t-12,t}^i$ predict future forecast errors. An increase in the index translates into a rise in the expected inflation rate and hence into a reduction in the forecast error (as defined above). This finding is true for all income groups except the top one. On the other hand, all groups over-weight the US aggregate inflation rate relative to the rational benchmark: $\pi_{t-12,t}$ predicts the ex-post forecast errors such that an increase in the US aggregate inflation translates into an over-prediction of the future inflation rate.

Robustness. What if they are predicting their personal inflation rather than the aggregate US inflation? The wording of the question may affect the reported expectations leading respondents to think of their personal price experiences (Bruine de Bruin et al. (2012)). To address this issue, we test if the personal and aggregate inflation rates have predictive power for the future individual inflation rate. The intuition is the following. If agents are rationally reporting (thinking of) their individual inflation, then they should optimally use their information set. For each group, we estimate the following forecasting equation and compare it with the previous results:

$$\bar{\pi}_{t,t+12}^i = \alpha + \beta\bar{\pi}_{t-12,t}^i + \gamma\pi_{t-12,t} + e_t^i \tag{2.10}$$

Where i is the income group, and t is the month. Consumers$'$ expectations are non-rational, also under the hypothesis that respondents report their experienced inflation rate (Table 2.8). Even under this hypothesis, agents use uninformative signals. Indeed, first, $\bar{\pi}_{t-12,t}^i$ does not have predictive power for $\bar{\pi}_{t,t+12}^i$, aside the lowest income group; second, $\pi_{t-12,t}$ is never informative of future individual inflation. This result suggests that agents understand the question and report the aggregate inflation rate; however, their intuitive approach leads them to use their private signals mistakenly and overweight the pieces of information available.

2.4.3 Robustness: new sources of information

In the baseline model, we assume that the expectations are affected by both personal and US aggregate inflation rates. In this section, we relax the hypothesis that people receive only two signals on future price variations and assume that they use other sources of information to form their beliefs. Explicitly, we allow agents to gain extra public knowledge, such as oil price, house prices and economic news. We test this hypothesis, augmenting the baseline regression (2.6) with additional controls.

During the past decade, an increasing number of studies has established a link between oil prices and inflation expectations. Starting with the onset of the crisis, inflation expectations reacted quite strongly oil shocks (Coibion and Gorodnichenko (2015b)). The intuition is that gasoline prices are among the most visible costs to consumers; thus households may pay more attention to them when forming expectations on other prices. To test this mechanism, we add as a control in our regression the growth rate in oil price, which represent a signal common to all individuals.[10]

$$\bar{E}_t^i(\pi_{t,t+12}) = \alpha + \beta\bar{\pi}_{t-12,t}^i + \gamma\pi_{t-12,t} + \delta oil_t + \epsilon_t^i \tag{2.11}$$

Table 2.9 reports the estimates of equation (2.11) and shows that our results are robust to the introduction of the growth rate of oil prices. The coefficients associated with the individual inflation rate remain positive and significant. Result suggest that the heterogeneity in inflation expectations cannot be fully explained by differences in the attention paid to the gasoline price; besides, the relevance of oil price varies substantially across groups as it seems particularly important for high-income agents.[11] A test for rationality, similar to the one presented in Section 2.4.2, shows that oil price does not have additional predictive power for the aggregate inflation rate, beyond the US CPI inflation; hence people should not focus on it (see the Appendix). Further, heterogeneity in inflation expectations does not appear to be driven by the expected oil price since high and low-income consumers have similar oil price expectations (Figure 2.10).[12]

Results are also confirmed when adding house prices, and the cost of energy (Tables 2.10 - 2.11).[13] These are two crucial components of the households' overall expenditure that are omitted from our indexes, which mainly reflects food and beverages. Although we do not know how the four income groups differ in the price paid and the quantities purchased of houses and

[10] We use data on "Price of West Texas Intermediate (WTI) crude" (https://www.eia.gov/opendata/qb.php?sdid=PET.RWTC.D) as in Coibion and Gorodnichenko (2015b). Further, please notice that oil is not within the set of goods in the KNCP dataset.

[11] For high-income groups, the adjusted R-sq increases significantly from the baseline regression; the coefficient associated to the individual inflation rate gets significant suggesting that conditional on oil price the personal experience helps in explaining the changes in the expected inflation rate over time.

[12] Respondents are asked: *Now thinking only about the next twelve months, do you think that the price of gasoline will go up during the next twelve months, will gasoline prices go down, or will they stay about the same as they are now?*

[13] We use the SP/Case-Shiller U.S. National Home Price Index (CSUSHPINSA) and the CPI series for energy from the by the U.S. Bureau of Labor Statistics (CUSR0000SA0E) and look at their first differences.

energy, we can test if they assign different weights to these components when forming inflation expectations. We find that house and energy prices do not enter in the expectation formation process and do not provide additional insight beyond the individual and the aggregate inflation rates.

The recent literature shows that economic news, especially in times of downturn, affect the agent′s expectations of future economic prospects and inflation expectations change in response to incoming news so that well-informed households have better forecasts. In light of these pieces of evidence, we check if income groups differ in the intensity of news heard. In the Appendix, Figure B.2 shows that high-income consumers declare to hear more economic news compared to low-income ones.[14] All the series follow the same pattern and the news coverage increases with the Financial crisis. On the other hand, there are no substantial differences across income groups when looking at news about inflation, and only a few respondents get this kind of news. We augment the baseline regression with the indicator $Inf\_News^i_t$, which measures the fraction of agents in a given class that read/heard economic news. Table 2.12 displays results for the model (2.12).

$$\bar{E}^i_t(\pi_{t,t+12}) = \alpha + \beta\bar{\pi}^i_{t-12,t} + \gamma\pi_{t-12,t} + \delta Inf\_\bar{N}ews^i_t + \phi Inf\_\bar{N}ews^i_t * \pi_{t-12,t} + \epsilon^i_t \quad (2.12)$$

Where i is the income group, and t is the month. The coefficients associated with the individual inflation rates are still positive and statistically significant for all the income classes, as well as those related to the US CPI inflation. Households overweight both the private and public signals when forming their expectations, relative to the rational benchmark (in Appendix). The coefficients associated with the indicator and the interaction terms instead are not significant, suggesting that heterogeneity in the news heard or in their interpretation is not driving heterogeneity in expectations. This evidence goes against the hypothesis that a substantial portion of the dispersion in expectations is driven by economic literacy (Burke and Micheal (2014)), although $Inf\_News^i_t$ is only an imperfect proxy for financial literacy. Finally, we take a broader approach by considering that agents look at several sources of information that may not be observable or available. Therefore, we summarize all the information received during the month t using the difference between the expected inflation rate at time t, $E^i_t(\pi_{t,t+12})$, and the one predicted in the previous period $E^i_{t-1}(\pi_{t-1,t+11})$. Results are consistent when controlling for this additional measure (Table 2.12) and the adjusted R-sq mostly increase suggesting that additional sources of information are relevant in the expectations formation process.

[14]In the Michigan survey, respondents are asked if they heard the economic news and in particular if they heard news about inflation. They provide dichotomy answers to this questions (yes/no), and we use these pieces of information to construct two income specific indexes.

2.5 Extension: the role of salient memories

2.5.1 Theoretical framework: salient goods and expectations

To refine our understanding of the transmission channel of the personal experience to the expectations we extend our baseline model. Up to this point, we assumed that the shopping experience provides the agents with a unique signal, that we approximate with the individual inflation rate; hence we presumed that the agent recalls the prices of all the goods purchased in the same way. Now we relax this hypothesis and assume that while shopping the agent observes several prices, but processes them in different ways according to their relevance and characteristics. People's attention is drawn by specific goods that come to mind when predicting future inflation; these goods are outstanding in some dimension so that agents strongly remember their prices. In particular, we assume that an agent may focus on the price variations of the products that: (i) he buys more frequently; (ii) require a significant expenditure; and (iii) exhibit large price variations from the previous period. We introduce a few additional private signals reflecting price changes in these sets of products and address the following questions: do people select memories of salient goods' prices when forecasting future inflation? Which are the goods people focus on?

The intuition is the following. People may focus on salient goods' prices because they recall better unusual events, that are surprising relative to the standard or expected case. For instance, an agent might be shocked when observing a much higher price for water than the *normal* one he usually pays and might have a vivid memory of such a high price. Alternatively, the same agent may not notice a tiny increase in the amount paid for the same bottle of water. Overall, people select some memories from their personal experience and focus on these in the spontaneous recall process. The intuition generalises what we believe happens in people's mind when predicting future prices: consumers exploit their personal experience and overweight few outstanding memories. In the model, we look at three goods' characteristics that may draw people's attention:

1. Frequency of purchases
   Frequently encountered prices come to mind more readily when asked about price changes since there is a positive effect of recency and repetition on subsequent recall. For instance, an agent may have vivid memories of the price of the frequently purchased milk and therefore notice any change in price, although small.

2. Expenditure
   People may recall better prices of goods that require a notable expenditure in a given period. For instance, the electricity bill is paid infrequently but usually absorbs a large

share of the consumers' monthly budget; therefore price changes in the electricity bill may be easily recognisable, even when relatively small.

3. Price variation from the preceding purchase
People recall better large price changes rather than small price variations. For instance, an agent might be shocked when seeing a much higher price for a given good than the expected one and record a clear memory of the amount paid.

We build our theoretical framework on the Salience theory of attention and choice proposed by Bordalo, Gennaioli, and Shleifer (hereafter BGS, 2013, 2012) who propose a theory of context-dependent choice in which a consumerâs attention is drawn to salient attributes of goods, such as prices or quantities. According to their model, a characteristic (attribute) is salient for a good when it stands out among the goodâs attributes relative to that attributeâs average level in the choice set. Salience distorts the valuation (utility) of a good so that the salient thinker enhances the relative utility weight attached to the salient attribute. Here instead we do not look at consumption choices but at the expectations formation process, and we talk about context-dependent memories as opposed to decisions. Specifically, we adapt the BGS definition of salience to test whether households attach an excessive weight to the prices of goods that are salient in the frequency of purchases, the expenditure and the price variations when forming their inflation expectations. We assume that memories are context dependent and each agent$'$ s attention focuses on goods that are very different from a reference point in his consumption basket. In our framework, the context is the consumer$'$ s choice set; the reference point is the average good in the consumption basket, and the degree of salience of an item in a given attribute depends on its distance from the reference good. In the next paragraph, we try to formalise the intuition behind such reasoning.

Let each consumer j, each month t, evaluate $N_{j,t} > 1$ goods in a choice set $C_{j,t} = \{p_{j,k,t}, q_{j,k,t}, a_{j,k,t}\}_{k=1,2,...,N_{j,t}}$. As in the baseline model, the choice set is composed of all the goods that the agent buys in a month t. Each good k is described by an additional generic attribute $a_{j,k,t}$ in which a good can be salient; in our application, $a_{j,k,t}$ can be the frequency of purchase ($f_{j,k,t}$), the overall expenditure ($exp_{j,k,t}$) and the price variation from the preceding period ($\Delta p_{j,k,t}$). These are the attributes that may draw the consumers$\hat{a}$ attention.[15] At each point in time, each agent predicts future inflation given his information set, which is composed by a public signal (US CPI inflation) and his memories from the shopping experiences synthesised in few private signals. The signals summarise the observed price changes over-weighting

[15] Each good k is described by three attributes that are all involved the private signals that the consumers receive while shopping. The first two, price and quantities, are needed to compute the private signals, that we approximate with the individual inflation rate; while the other one is informative of the relevance of the products in the consumers$\hat{a}$ mind and habits.

goods that are outstanding in a given attribute because, in the recall process, the consumerâs attention is drawn to the most salient goods. Each month t, each consumer j evaluates the salience of each purchased good k in the attribute $a_{j,k,t}$ relative to the choice set, comparing it with a reference point that is the good defined as having the average level of each attribute, where the average is taken over the products in the choice set. For each consumer j and month t, we set a reference good as the one described by the average attributes $(\bar{p}_{j,t}, \bar{q}_{j,t}, \bar{a}_{j,t})$ in $C_{j,t}$ where $\bar{p}_{j,t} = \frac{1}{N_{j,t}} \sum_{k}^{N_{j,t}} p_{j,k,t}$, and similar averages are defined for the other attributes. Following BGS, we define a function $\sigma(.,.)$ which indicates the degree of salience of a good in an attribute given the choice set and whose value depends on the level of the attribute itself for good k and its average level in $C_{j,t}$. Specifically, we define the following salience function:

$$\sigma(a_{j,k,t}, \bar{a}_{j,t}) = \frac{|a_{j,k,t} - \bar{a}_{j,t}|}{|a_{j,k,t} + \bar{a}_{j,t}|} \tag{2.13}$$

where $\sigma(0,0) = 0$ and $a_{j,k,t}, \bar{a}_{j,t} \neq 0$ for any j, k, t. Hereafter, we simplify the notation by indicating $\sigma(a_{j,k,t}, \bar{a}_{j,t})$ with $\sigma(a_{j,k,t})$. The function is continuous, symmetric and has the properties of ordering, diminishing sensitivity and homogeneity of degree zero. The property of ordering implies that the salience increases in the distance from the reference point; diminishing sensitivity implies that salience decreases as the value of an attribute uniformly increases for all good. Homogeneity of degree zero means that given a change in $a_{j,k,t}$, ordering dominates diminishing sensitivity if and only if the change in $a_{j,k,t}$ is larger than one in the average level $\bar{a}_{j,t}$. A detailed description of the properties of the salience function is provided by Bordalo et al. (2012). This function indicates whether the good k is outstanding concerning the attribute $a_{j,k,t}$, such that the larger is the value of $\sigma_{j,k,t}(a_{j,k,t})$ the higher is the degree of salience of product k for $a_{j,k,t}$. Overall, the degree of salience of k in the attribute $a_{j,k,t}$ is high when $a_{j,k,t}$ is far from the average value of the same attribute in the choice set $\bar{a}_{j,t}$. The function is such that a good k can display a large degree of salience in an attribute if (i) $a_{j,k,t}$ is much larger than its average level in the choice set $\bar{a}_{j,t}$, or (ii) $a_{j,k,t}$ is much smaller than $\bar{a}_{j,t}$. For instance, a product may be salient in frequency because it is acquired daily or only once a month. We distinguish between the two cases separating the goods that exhibit $(a_{j,k,t} - \bar{a}_{j,t}) > 0$ from those with $(a_{j,k,t} - \bar{a}_{j,t}) < 0$. Further, the degree of salience of an attribute is time-varying. This feature, for instance, captures the fact that strawberries, or ice-cream, may be salient in price variation during the winter, and the frequency of purchases in summer. Besides, salience is time varying as a given event may be striking the first time it arises, but over time, if agents get used to it, it will not be surprising any more.

We assume that, in the recall process, each household focus on goods characterized by a high degree of salience which reflect the relevance of the product k in the consumersâ mind and habits. When forecasting inflation, hence the memories evoked are aggregated in signals by weighting the goods according to the degree of salience in a given attribute. For instance, a salient thinker retrieves memories of the prices paid to predict the future inflation rate and combine them by over-weighting the prices of the goods that he acquires frequently, such as milk. Overall the shopping experience provides the agents with few signals that overweight the price changes of those goods that are salient in the frequency of purchases, in the expenditure or in the price variation ($s^{SAL}_{j,t}(\pi_t)$). Without salience distortions, rational consumers are unaffected by the individual purchases nor salient products and estimate the model in equation (2.1). Salient thinkers instead look at their personal experience and, within the observed goods, focus on those that "stand out" among the others in a particular attribute relative to the reference good. Price changes of these products engrave in memory so, when asked to forecast future prices, salient thinkers recall those prices for which they have vivid memories.

Assuming linearity, salient thinkers adopt the following intuitive rule:

$$\tilde{E}_{j,t}(\pi_{t+1}) = \beta_j * s_{j,t}(\pi_t) + \gamma_j * \pi_t + \delta_j * s^{SAL}_{j,t}(\pi_t) \tag{2.14}$$

Where the private signal $s^{SAL}_{j,t}(\pi_t)$ reflects the price changes of the products that are salient in one (or more) dimension (i.e. frequency of purchases, expenditure or price variation). Relative to equation (2.4), we introduce two new sources of heterogeneity in expectations: the salience private signals and the weights assigned to them. In the next sections, we describe how we build the proxies for these additional signals exploiting actual purchases data and test the theory.

2.5.2 Individual inflation rates on salient goods

Given the theoretical framework described in the previous section, in the empirical analysis, we build the proxies of the additional private signals as follow. First, we compute the three measures of interest ($f_{j,k,t}, exp_{j,k,t}, \Delta p_{j,k,t}$), for each household j, month t and product k. KNCP data provide the frequency of purchase within a month, the quantity bought and the price paid, and the price for each transaction made by each household. For each good k, consumer j and time t, we compute a monthly index of the frequency of purchase equal to the number of trips in which household j buys the product k over the total number of trips of j in month t. This index lies between (0-1] and reaches its maximum for the items that the household buys on any trip. The expenditure is the quantity (volume) of the product times the price paid for one unit of it. The price variation from the preceding month is the difference in the price paid for a given good k by household j, in the current month and the previous one. Second, for each consumer j and month t, we compute the average attributes within the choice set $C_{j,t}$ which are described by:

$$\bar{f}_{j,t} = \frac{1}{N_{j,t}} \sum_{k}^{N_{j,t}} f_{j,k,t}, \quad e\bar{x}p_{j,t} = \frac{1}{N_{j,t}} \sum_{k}^{N_{j,t}} exp_{j,k,t},$$
$$\bar{\Delta p}^{>0}_{j,t} = \frac{1}{N_{j,t}} \sum_{k}^{N_{j,t}} \Delta p_{j,k,t}|_{\Delta p_{j,k,t}>0}, \quad \bar{\Delta p}^{<0}_{j,t} = \frac{1}{N_{j,t}} \sum_{k}^{N_{j,t}} \Delta p_{j,k,t}|_{\Delta p_{j,k,t}<0}.$$

We distinguish between positive and negative price variations, to take into account their potential asymmetric effect on expectations. Then, for each consumer j, month t, we build the following salience functions:

$$\begin{aligned}\sigma(f_{j,k,t}) &= \frac{|f_{j,k,t} - \bar{f}_{j,t}|}{f_{j,k,t} + \bar{f}_{j,t}} \quad if \quad (f_{j,k,t} - \bar{f}_{j,t}) > 0 \\ &= 0 \quad otherwise\end{aligned} \tag{2.15}$$

$$\begin{aligned}\sigma(exp_{j,k,t}) &= \frac{|exp_{j,k,t} - e\bar{x}p_{j,t}|}{exp_{j,k,t} + e\bar{x}p_{j,t}} \quad if \quad (exp_{j,k,t} - e\bar{x}p_{j,t}) > 0 \\ &= 0 \quad otherwise\end{aligned} \tag{2.16}$$

$$\begin{aligned}\sigma(\Delta p^{>0}_{j,k,t}) &= \frac{|\Delta p_{j,k,t} - \bar{\Delta p}^{>0}_{j,t}|}{\Delta p_{j,k,t} + \bar{\Delta p}^{>0}_{j,t}} \quad if \quad (\Delta p_{j,k,t} - \bar{\Delta p}_{j,t}) > 0 \\ &= 0 \quad otherwise\end{aligned} \tag{2.17}$$

$$\begin{aligned}\sigma(\Delta p^{<0}_{j,k,t}) &= \frac{|\Delta p_{j,k,t} - \bar{\Delta p}^{<0}_{j,t}|}{|\Delta p_{j,k,t} + \bar{\Delta p}^{<0}_{j,t}|} \quad if \quad (\Delta p_{j,k,t} - \bar{\Delta p}_{j,t}) < 0 \\ &= 0 \quad otherwise\end{aligned} \tag{2.18}$$

Equation (2.15) is positive only for goods with $(f_{j,k,t} - \bar{f}_{j,t}) > 0$ since we aim to build an index that over-weights the price changes of those goods that are purchased with high frequency. We say that a good is salient in the frequency of purchases if $(f_{j,k,t} - \bar{f}_{j,t}) > 0$, and we apply a similar definition for the other attributes. Uniformly, equation (2.16) captures the degree of salience of those goods that require a large expenditure relative to the reference goods, rather than an exceptionally low one; equation (2.17) measures the degree of salience of positive price variations, while equation (2.18) of negative changes. Exploiting these measures, for each consumer j and month t, we compute four Laysperes indexes varying the weight assigned to each product according to its degree of salience in a given attribute such that, for a generic attribute $a_{j,k,t}$, the index is computed as:

$$L^a_{j,t} = \frac{\sum_k P_{j,k,t} * \sigma(a_{j,k,t}) * Q_{j,k,0}}{\sum_i P_{j,k,0} * \sigma(a_{j,k,t}) * Q_{j,k,0}} \tag{2.19}$$

And $a_{j,k,t} \in (f_{j,k,t}, exp_{j,k,t}, \Delta p^{>0}_{j,k,t}, \Delta p^{<0}_{j,k,t})$. The Laysperes index is defined in Section 2.3.2, the main difference is that now the index overweights the goods that are salient in a given attribute $a_{j,k,t}$ (i.e. $f_{j,k,t}$) relative to the others in the choice set. Using these new Laysperes indexes, we build four monthly inflation rates at the household level: $\pi^f_{j,t}, \pi^{exp}_{j,t}, \pi^{\Delta p>0}_{j,t}$, and $\pi^{\Delta p<0}_{j,t}$ Finally, we aggregate each measure by income group to get income-specific salient signals, following the approach described in Section 2.3.

The new rates are substantially different from the individual inflation rates and display heterogeneity across income classes (Table 2.14). The differences between the baseline individual inflation rate and these measures derive from the differences in weights involved in each index. By looking at the distribution of the degree of salience, we find that consumersâ baskets are heterogeneous in terms of positive price variations and several goods record high value of $\sigma(\Delta p^{>0}_{j,k,t})$. The consumersâ choice sets instead are more homogeneous in terms of expenditure, as only a few items record a high degree of salience in this attribute. Finally, only a few items display a high degree of salience for frequency and these are varying within and between classes. On average $\bar{\pi}^{i,\Delta p>0}_{t-12,t}$ is higher than the other measures since it over-weights those goods that exhibit a significant price variation, while $\bar{\pi}^{i,\Delta p<0}_{t-12,t}$ is negative as it considers only the items that display a negative price variation. Further, $\bar{\pi}^{i,\Delta p>0}_{t-12,t}$ is substantially larger for agents at the top of the income distribution possibly because low-income agents postpone (or substitute) the purchase when observing a significant price increase while high-income consumers do not. The inflation rate that over-weights the goods salient in the frequency of purchase instead is substantially lower than the others as only a few products are purchased very frequently and they tend to have stable prices.

2.5.3 Testing the salience theory

We test if consumers adopt the forecasting rule in equation (2.14), by considering the new private signals one at the time. Once again we consider four representative agents i, one for each income group, and estimate the following regression:

$$\bar{E}^i_t(\pi_{t,t+12}) = \alpha + \beta\bar{\pi}^i_{t-12,t} + \gamma\pi_{t-12,t} + \delta\bar{\pi}^{i,SAL}_{t-12,t} + \epsilon^i_t \tag{2.20}$$

Where $\bar{\pi}^{i,SAL}_{t-12,t} \in \{\bar{\pi}^{i,f}_{t-12,t}, \bar{\pi}^{i,exp}_{t-12,t}, \bar{\pi}^{i,\Delta p>0}_{t-12,t}, \bar{\pi}^{i,\Delta p<0}_{t-12,t}\}$. For these measures, we compute the monthly average over the last 12 months, between time $t-12$ and t, to capture the signals received by consumers in a long time horizon. Tables 2.15-2.17 suggest that consumers are affected by salient attributes in different ways and overall high-income agents are more sensitive to outstanding products relative to low-income ones. The adjusted R-sq are mostly higher than those observed in the baseline regression (Table 2.5), suggesting that these signals convey additional information content on the expectations beyond the individual inflation rates. The increase in the adjusted R-sq is larger for the groups at the top of the income distribution. The coefficients associated to $\bar{\pi}^i_{t-12,t}$ are mostly positive and significant indicating that agents look at the entire bundle of purchased goods, although they pay particular attention to some specific products. Goods that exhibit large positive price variations are relevant for all the classes, while goods that display a reduction in price are not involved in the expectation formation process (Table 2.15). Agents overweight goods salient in expenditure, especially high-income consumers (Table 2.16). Finally, only low-income individuals focus on goods frequently purchased when predicting future inflation (Table 2.17). In Appendix, we test few additional hypotheses on the salient bias, to understand the effect of the interactions between salient attributes (i.e. if agents particularly overweight goods that are both notable in price variations and frequency) and of the items that instead are purchased with low frequency or require little spending. Overall, the results confirm that agents mostly focus on goods that record a substantial increase in price and a significant expenditure and overweight the products that have both these characteristics. In conclusion, this section shows that all groups overweight salient goods relative to the rational benchmark (see the Appendix) and depicts an additional dimension in which expectations systematically depart from rationality and a further source of dispersion in beliefs.

2.6 Inflation expectations and consumption choices

So far we have shown that personal experiences help explaining differences in beliefs about future inflation. Now, we wonder if these beliefs are consistent with consumersâ choices and if our model of inflation expectations explain heterogeneity in consumption decisions. According to the Euler and the Fisher equations, higher expected inflation should stimulate spending since expenditure is inversely related to the real interest rate. Households expecting an increase in the inflation rate should be willing to consume more, both durables and nondurables goods; this argument is well explained by Bachmann et al. (2015) who test the relationship between expected inflation rate and readiness to spend using the Michigan survey at individual level and do not find a significant association between householdsâ inflation expectations and their readiness to spend on durable consumption goods. Other studies instead show that survey measures are associated with agents′ behaviours such as buying, borrowing or investment (i.e. Malmendier and Nagel (2015)). The objective of this section is to provide some evidence on the relationship between the intuitive (experience-based) inflation expectations described in Section 2.2 and consumersâ spending. We look at readiness to spend elicited from the following

Michigan questions:

Q3. "About the big things people buy for their homes - such as furniture, a refrigerator, stove, television, etc. Generally speaking, do you think now is a good or a bad time for people to buy major household items?"

Q4. "Speaking of the automobile market – do you think the next 12 months or so will be a good time or a bad time to buy a car?"

Responses take on three different categories: good time, bad time, or uncertain. To estimate the effect of intuitive experience-based forecasts on householdsâ decisions, we estimate the following regression separately for each income group:

$$\bar{C_t^i} = \alpha + \beta \tilde{E}_t^i(\pi_{t,t+12}) + \gamma \bar{X}_t^i + \epsilon_t^i \tag{2.21}$$

Where $\bar{C_t^i}$ is the readiness to spend for income group i at time t, which is the share of agents that believe it is a good time to buy durable goods or cars within the income class i. $\tilde{E}_t^i(\pi_{t,t+12})$ is the intuitive forecasts for the income group i, computed as fitted values from our baseline regression (2.6). These measures reflect the variation in the expectations that is explained by our model with private and public signals. $\bar{X}_t^i$ are additional variables at income class level that may affect spending and may correlate with inflation expectations, such as the expected interest rate, the expected future income or business conditions. Adding these controls allows us to avoid spurious correlation.

Tables 2.18 present the estimation results. We observe that the intuitive expectations are positively related to the readiness to spend both in durable goods and cars. As predicted by the theory, households who expect a higher inflation rate and, in turn, a lower real interest rate, take on more spending.[16] The effect is heterogeneous across classes; the magnitude is larger for cars and mostly increasing in income; for instance, for the higher income class a one standard deviation increase in the expected inflation is associated to an increase of 0.58 standard deviation of the dependent variable, while it is 0.22 for the lowest income class. Further, for all income groups, the most common reason for households to have a positive view of buying conditions is that *prices are low.* Overall this section confirms that personal experience affects individualsâ expectations and their consumption decisions.

[16]The interpretation of the magnitude of these coefficients is not straightforward given the qualitative nature of the data. The coefficients are very similar when looking at the actual inflation expectations by income groups rather than the fitted values. See Appendix.

2.7 Conclusion

The paper suggests personal shopping experience as a new channel to explain the heterogeneity in inflation expectations with a particular focus on differences across income classes. By employing data from Kilts-Nielsen Consumer Panel (KNCP) and Michigan Survey of Consumers (MSC), we test a model of intuitive inflation expectations in which agents use both private (shopping experience) and public (US CPI) signals. We use the inflation rate computed at the household level as a proxy for the private signal and show that agents adopt a spontaneous recall process and take into account shopping memories when predicting inflation. Finally, we sketch a model of a salient thinker and depict a particular channel through which the personal expertise can affect expectations. We conclude that heterogeneity in expectations derives from differences in the private signals and in how people combine private and public information.

These findings stand in contrast with the rational (full information) expectation hypothesis, according to which consumersâ should form homogeneous expectations without leveraging on their personal experience. These results may have substantial implications concerning monetary policy effectiveness and we believe they might open an avenue for further research. In particular, they may help to explain the forward guidance puzzle (Del Negro et al. (2012)) as with biased heterogeneous expectations, announcements have a weaker a impact on consumersâ beliefs compared to what is predicted by standard macroeconomic models with rational agents. Overall, allowing for heterogeneity in inflation and expectations might help to explain how people react to monetary policy announcements although further research is needed to fully characterize the implications of biased expectations in macroeconomic models. Further, these results have relevant policy implications as they suggest that providing individuals with accurate information on the official statistics may not be sufficient to reduce the bias and disagreement in the expectations. Agents indeed observe multiple pieces of information, including the aggregate CPI inflation, but do not know their informativeness nor how to interpret them. Policymakers, therefore, should inform agents by making easily accessible the correct sources of information (i.e. CPI inflation) but also facilitate the publicâ s understanding and processing of the various pieces of information. Finally, policymakers should fully understand the mechanisms behind the expectations formation process if they want to manipulate expectations using targeted announcements, as proposed by Coibion et al. (2018). A full understating of the reasons behind the disagreement about future inflation may allow policymakers to build tailored messages to specific groups of individuals that are less sensitive to official announcements; this targeted approach may help in increasing the effectiveness of monetary policy announcements and generating stronger shifts in expectations. Overall, this paper provides some empirical evidence in this direction and future researches may build on it to shed some lights on how expectations are formed and how they differ across demographic groups using micro-level data.

2.8 Figures and Tables

	Age class		Income class	
	All sample	After 1990	All sample	After 1990
$\frac{Exp(Low)-Exp(High)}{mean}$	0.079	-0.023	0.2320	0.3069
$\frac{\lvert Exp(Low)-Exp(High)\rvert}{mean}$	0.1795	0.1220	0.2519	0.3104
SD between classes	0.4096	0.2626	0.5102	0.5375

MSC data at monthly frequency for the period 2004-2014. The following three age classes are considered: Age < 40, Age from 40 to 60, Age > 60. The following four income classes are considered: Income $<$ \$25.000, Income between \$25,000 and \$50,000, Income between \$50,000 and \$100,000, Income $> \$100,000$. Exp(Low) and Exp(High) are the average expected inflation rate for the class at the bottom and at the top of the (age or income) distribution, respectively. Mean instead is the average expected inflation rate on the entire sample. SD is the standard deviation computed across classes.

Table 2.1: Dispersion between income and age classes

Table 2.2: Within and between variability by income group

Sample	Individual inflation	Expected inflation
Within income class 1	0.4072	4.71
Within income class 2	0.4190	4.26
Within income class 3	0.4376	3.91
Within income class 4	0.4068	3.76
Between income classes	0.8767	5.73

Note: The table displays the standard deviations computed on the realized individual inflation rates and the expected inflation one year ahead within each income classes and the one between income classes, the latter is calculated on the data aggregated by income groups.

Table 2.3: Cyclical inflation (month on month)

Index	Mean	Stand. Dev.
$\pi_{j,t}^{L}$	0.3564	0.1133
$\pi_{j,t}^{P}$	0.3634	0.1189
$\pi_{j,t}^{F}$	0.3699	0.1201
US inflation π_t	0.2443	

Note: $\pi_{j,t}^{L}$ is the monthly Laspeyres inflation rate computed at the household level using the KNCP, $\pi_{j,t}^{P}$ the Paschee rate, $\pi_{j,t}^{F}$ the Fisher rate. π_t is the monthly US inflation rate. Data are at monthly frequency, 2004-2014.

Table 2.4: Expected and realized inflation - panel data analysis

	(1) $\bar{E}^i_t(\pi_{t,t+12})$	(2) $\bar{E}^i_t(\pi_{t,t+12})$	(3) $\bar{E}^i_t(\pi_{t,t+12})$	(4) $\bar{E}^i_t(\pi_{t,t+12})$	(5) $\bar{E}^i_t(\pi_{t,t+12})$
$\bar{\pi}^i_{t-12,t}$		1.229**	1.261***	2.138***	2.133***
		(0.019)	(0.000)	(0.000)	(0.000)
$\pi_{t-12,t}$	1.795**	1.685**	1.848**	1.667***	1.805***
	(0.015)	(0.012)	(0.032)	(0.000)	(0.000)
Income 2*$\bar{\pi}^i_{t-12,t}$				-0.878*	-0.901*
				(0.069)	(0.071)
Income 3*$\bar{\pi}^i_{t-12,t}$				-1.240	-1.241*
				(0.125)	(0.076)
Income 4*$\bar{\pi}^i_{t-12,t}$				-1.529**	-1.579**
				(0.020)	(0.011)
Income 2*$\pi_{t-12,t}$			-0.471		0.0136
			(0.870)		(0.610)
Income 3*$\pi_{t-12,t}$			-0.0362		-0.129
			(0.950)		(0.865)
Income 4*$\pi_{t-12,t}$			-0.349		-0.586
			(0.689)		(0.503)
Cons	4.414***	4.888***	5.612***	5.870***	5.909***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Income dummy	NO	NO	YES	YES	YES
N	432	432	432	432	432
Adj. R-sq	0.094	0.182	0.726	0.740	0.741

p-values in parentheses* p<0.10, ** p<0.05, *** p<0.01. $\bar{\pi}^i_{t-12,t}$ is the average realized inflation of the income group i. $\pi_{t-12,t}$ is the US inflation. Newey-west standard errors with 12 lags. The baseline group in Columns (3)-(5) is the income group 1.

Table 2.5: Expected and realized inflation by income classes

	Income 1	Income 2	Income 3	Income 4
	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$
$\bar{\pi}_{t-12,t}^i$	2.133***	1.232***	0.893*	0.554
	(0.000)	(0.000)	(0.093)	(0.255)
$\pi_{t-12,t}$	1.805***	1.818***	1.676**	1.218**
	(0.000)	(0.000)	(0.018)	(0.043)
Cons	5.909***	4.561***	4.133***	4.784***
	(0.000)	(0.000)	(0.000)	(0.000)
N	108	108	108	108
Adj. R-sq	0.661	0.553	0.298	0.130

Note: p-values in parentheses* p<0.10, ** p<0.05, *** p<0.01. $\bar{\pi}_{t-12,t}^i$ is the average realized inflation of the income group i. $\pi_{t-12,t}$ is the US inflation. Newey-west standard errors with 12 lags.

Table 2.6: Test for rationality

	Income 1	Income 2	Income 3	Income 4
	$\pi_{t,t+12}$	$\pi_{t,t+12}$	$\pi_{t,t+12}$	$\pi_{t,t+12}$
$\bar{\pi}_{t-12,t}^i$	0.388	-0.241	-0.0488	-0.165
	(0.287)	(0.297)	(0.652)	(0.355)
$\pi_{t-12,t}$	0.182***	0.0873**	0.144***	0.247***
	(0.000)	(0.045)	(0.000)	(0.000)
Cons	0.440***	0.250***	0.279***	0.241***
	(0.000)	(0.003)	(0.000)	(0.000)
N	108	108	108	108
Adj. R-sq	0.341	0.108	0.169	0.290

Note: p-values in parentheses* p<0.10, ** p<0.05, *** p<0.01. $\bar{\pi}_{t-12,t}^i$ is the average realized inflation of income group i. $\pi_{t-12,t}$ and $\pi_{t,t+12}$ are the US inflation. Newey-west standard errors with 12 lags.

Table 2.7: Testing forecast errors predictability

	Income 1	Income 2	Income 3	Income 4
	fe_t^i	fe_t^i	fe_t^i	fe_t^i
$\bar{\pi}^i_{t-12,t}$	-1.948***	-0.655**	-1.339***	-0.827
	(0.000)	(0.016)	(0.000)	(0.324)
$\pi_{t-12,t}$	-3.254***	-2.055***	-2.971***	-2.526***
	(0.000)	(0.009)	(0.000)	(0.000)
Cons	-5.855***	-4.758***	-4.561***	-4.043***
	(0.000)	(0.000)	(0.000)	(0.000)
N	108	108	108	108
Adj. R-sq	0.641	0.245	0.637	0.349

Note: p-values in parentheses* p<0.10, ** p<0.05, *** p<0.01. $\bar{\pi}^i_{t-12,t}$ is the average realized inflation of income group i. $\pi_{t-12,t}$ is the US inflation. fe_t is the forecast error for income group i. Newey-west standard errors with 12 lags.

Table 2.8: Test for rationality, robustness

	Income 1	Income 2	Income 3	Income 4
	$\bar{\pi}^i_{t,t+12}$	$\bar{\pi}^i_{t,t+12}$	$\bar{\pi}^i_{t,t+12}$	$\bar{\pi}^i_{t,t+12}$
$\bar{\pi}^i_{t-12,t}$	0.355*	0.0721	0.0758	0.0433
	(0.077)	(0.755)	(0.185)	(0.933)
$\pi_{t-12,t}$	-0.0338	-0.284	0.279	0.642
	(0.897)	(0.268)	(0.204)	(0.203)
Cons	0.225**	0.307***	0.223***	0.220
	(0.012)	(0.000)	(0.000)	(0.392)
N	108	108	108	108
Adj. R-sq	0.134	0.059	0.240	0.213

Note: p-values in parentheses* p<0.10, ** p<0.05, *** p<0.01. $\bar{\pi}^i_{t-12,t}$ and $\bar{\pi}^i_{t,t+12}$ are the average realized inflation of the income group i. $\pi_{t-12,t}$ is the US inflation. Newey-west standard errors with 12 lags.

Table 2.9: Expected and realized inflation with oil price

	Income 1	Income 2	Income 3	Income 4
	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$
$\bar{\pi}^i_{t-12,t}$	2.145***	0.605**	1.182***	1.303**
	(0.000)	(0.048)	(0.000)	(0.044)
$\pi_{t-12,t}$	1.573***	1.191**	1.743***	1.156**
	(0.000)	(0.017)	(0.000)	(0.015)
Growth rate oil price	1.236**	0.995	0.976	2.836***
	(0.031)	(0.231)	(0.145)	(0.000)
Cons	5.837***	4.793***	4.526***	4.143***
	(0.000)	(0.000)	(0.000)	(0.000)
N	108	108	108	108
Adj. R-sq	0.703	0.161	0.587	0.517

Note: p-values in parentheses* p<0.10, ** p<0.05, *** p<0.01. $\bar{\pi}^i_{t-12,t}$ is the average realized inflation of the income group i. $\pi_{t-12,t}$ is the US inflation. Growth rate oil price is the monthly growth rate. Newey-west standard errors with 12 lags.

Table 2.10: Expected and realized inflation with house prices

	Income 1	Income 2	Income 3	Income 4
	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$
$\bar{\pi}^i_{t-12,t}$	2.139***	0.493*	1.207***	0.844*
	(0.000)	(0.083)	(0.000)	(0.085)
$\pi_{t-12,t}$	1.841***	1.331**	1.871***	1.639***
	(0.000)	(0.019)	(0.000)	(0.009)
House price	-2.876	-11.79	-3.916	6.617
	(0.659)	(0.403)	(0.575)	(0.550)
Cons	5.922***	4.823***	4.569***	4.108***
	(0.000)	(0.000)	(0.000)	(0.000)
N	108	108	108	108
Adj. R-sq	0.664	0.179	0.558	0.311

Note: p-values in parentheses* p<0.10, ** p<0.05, *** p<0.01. $\bar{\pi}^i_{t-12,t}$ is the average realized inflation of the income group i. $\pi_{t-12,t}$ is the US inflation. House price is the change from the preceding period. Newey-west standard errors with 12 lags.

Table 2.11: Expected and realized inflation with energy price

	Income 1	Income 2	Income 3	Income 4
	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$
$\bar{\pi}^i_{t-12,t}$	1.930***	0.574*	1.091***	0.882*
	(0.000)	(0.089)	(0.000)	(0.086)
$\pi_{t-12,t}$	1.894***	1.199**	1.836***	1.679**
	(0.000)	(0.044)	(0.000)	(0.019)
Energy price	0.0798	-0.0290	0.0755*	-0.00323
	(0.325)	(0.710)	(0.082)	(0.955)
Cons	5.896***	4.780***	4.525***	4.129***
	(0.000)	(0.000)	(0.000)	(0.000)
N	108	108	108	108
Adj. R-sq	0.682	0.136	0.575	0.298

Note: p-values in parentheses* p<0.10, ** p<0.05, *** p<0.01. $\bar{\pi}^i_{t-12,t}$ is the average realized inflation of the income group i. $\pi_{t-12,t}$ is the US inflation. Energy price is the change from the preceding period. Newey-west standard errors with 12 lags.

Table 2.12: Expected and realized inflation with news

	Income 1	Income 2	Income 3	Income 4
	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$
$\bar{\pi}^i_{t-12,t}$	1.933***	0.759**	1.199***	0.956**
	(0.000)	(0.030)	(0.000)	(0.049)
$\pi_{t-12,t}$	1.752***	1.379**	1.730***	1.702***
	(0.000)	(0.012)	(0.000)	(0.008)
$Inf\_news^i_t$	2.437	9.055	6.233	0.893
	(0.140)	(0.117)	(0.129)	(0.688)
$\pi_{t-12,t} * Inf\_news^i_t$	3.640	16.37	11.57	1.016
	(0.399)	(0.189)	(0.154)	(0.884)
Cons	5.747***	4.703***	4.398***	4.127***
	(0.000)	(0.000)	(0.000)	(0.000)
N	108	108	108	108
Adj. R-sq	0.682	0.289	0.629	0.303

Note: p-values in parentheses* p<0.10, ** p<0.05, *** p<0.01. $\bar{\pi}^i_{t-12,t}$ is the average realized inflation of the income group i. $\pi_{t-12,t}$ is the US inflation. $Infl\_news^i_t$ is the Economic news index. Newey-west standard errors with 12 lags.

Table 2.13: Expected and realized inflation with additional news

	Income 1	Income 2	Income 3	Income 4
	$\bar{E}_t^i(\pi_{t,t+12})$	$\bar{E}_t^i(\pi_{t,t+12})$	$\bar{E}_t^i(\pi_{t,t+12})$	$\bar{E}_t^i(\pi_{t,t+12})$
$\bar{\pi}_{t-12,t}^i$	2.132***	0.447**	1.201***	0.827
	(0.000)	(0.044)	(0.000)	(0.155)
$\pi_{t-12,t}$	1.591***	1255**	1.713***	1.671**
	(0.000)	(0.040)	(0.000)	(0.013)
$\bar{E}_t^i(\pi_{t,t+12}) - \bar{E}_{t-1}^i(\pi_{t-1,t+11})$	0.0967**	0.126	0.143*	0.112*
	(0.030)	(0.132)	(0.055)	(0.054)
Cons	5.834***	4.749***	4.536***	4.112***
	(0.000)	(0.000)	(0.000)	(0.000)
N	108	108	108	108
Adj. R-sq	0.690	0.181	0.581	0.315

Note: p-values in parentheses* p<0.10, ** p<0.05, *** p<0.01. $\bar{\pi}_{t-12,t}^i$ is the average realized inflation of the income group i. $\pi_{t-12,t}$ is the US inflation. Newey-west standard errors with 12 lags.

Table 2.14: Individual inflation rates on salient goods by income classes

		Income 1	Income 2	Income 3	Income 4
$\bar{\pi}_{t-12,t}^i$	- Laysperes index	0.441	0.387	0.364	0.355
$\bar{\pi}_{t-12,t}^{i,f}$	- Salience in frequency	0.176	0.223	0.215	0.259
$\bar{\pi}_{t-12,t}^{i,exp}$	- Salience in expenditure	2.728	2.954	3.212	3.241
$\bar{\pi}_{t-12,t}^{i,\Delta p>0}$	- Salience in positive price variation	1.997	2.273	2.472	3.116
$\bar{\pi}_{t-12,t}^{i,\Delta p<0}$	- Salience in negative price variation	-2.646	-2.129	-2.118	-2.110

Note: $\bar{\pi}_{t-12,t}^i$ is the individual inflation rate. $\bar{\pi}_{t-12,t}^{i,f}$ is the income specific inflation rate computed by over-weighting goods that are salient in frequency. Similarly, $\bar{\pi}_{t-12,t}^{i,exp}$ over-weighting goods that are salient in expenditure. $\bar{\pi}_{t-12,t}^{i,\Delta p>0}$ over-weighting goods that are salient in positive price variations and $\bar{\pi}_{t-12,t}^{i,\Delta p<0}$ in negative price variations. Data are at monthly frequency, 2004-2014.

Table 2.15: Expected and realized inflation with salience in price variation

	Income 1	Income 2	Income 3	Income 4
	$\bar{E}_t^i(\pi_{t,t+12})$	$\bar{E}_t^i(\pi_{t,t+12})$	$\bar{E}_t^i(\pi_{t,t+12})$	$\bar{E}_t^i(\pi_{t,t+12})$
$\bar{\pi}_{t-12,t}^i$	1.873***	0.410**	1.174***	1.122**
	(0.000)	(0.038)	(0.001)	(0.032)
$\pi_{t-12,t}$	2.482***	2.460**	1.623**	1.440*
	(0.000)	(0.027)	(0.026)	(0.077)
$\bar{\pi}_{t-12,t}^{i,\Delta p>0}$	0.0898*	0.169**	0.131*	0.104*
	(0.076)	(0.015)	(0.055)	(0.083)
$\bar{\pi}_{t-12,t}^{i,\Delta p<0}$	-0.0597	-0.0455	0.0985	0.0550
	(0.463)	(0.603)	(0.133)	(0.643)
Cons	9.053***	9.459***	5.054***	5123***
	(0.000)	(0.000)	(0.000)	(0.000)
N	108	108	108	108
Adj. R-sq	0.696	0.400	0.679	0.393

Note: p-values in parentheses* $p<0.10$, ** $p<0.05$, *** $p<0.01$. $\pi_{t-12,t}$ is the US inflation. $\bar{\pi}_{t-12,t}^{i,\Delta p>0}$ is the income specific inflation rate computed by over-weighting goods that are salient in positive price variations. Similarly, $\bar{\pi}_{t-12,t}^{i,\Delta p<0}$ for negative price variations. Newey-west standard errors with 12 lags.

Table 2.16: Expected and realized inflation with salience in expenditure

	Income 1	Income 2	Income 3	Income 4
	$\bar{E}_t^i(\pi_{t,t+12})$	$\bar{E}_t^i(\pi_{t,t+12})$	$\bar{E}_t^i(\pi_{t,t+12})$	$\bar{E}_t^i(\pi_{t,t+12})$
$\bar{\pi}_{t-12,t}^i$	1.956***	0.373*	1.018***	1.559**
	(0.000)	(0.094)	(0.001)	(0.015)
$\pi_{t-12,t}$	2.110***	2.494**	2.302***	2.423***
	(0.000)	(0.016)	(0.000)	(0.000)
$\bar{\pi}_{t-12,t}^{i,exp}$	0.353*	0.720***	0.285*	1.110***
	(0.080)	(0.003)	(0.068)	(0.005)
Cons	6.188***	5.575***	4.840***	5.422***
	(0.000)	(0.000)	(0.000)	(0.000)
N	108	108	108	108
Adj. R-sq	0.671	0.287	0.565	0.523

Note: p-values in parentheses* $p<0.10$, ** $p<0.05$, *** $p<0.01$. $\pi_{t-12,t}$ is the US inflation. $\bar{\pi}_{t-12,t}^{i,exp}$ is the income specific inflation rate computed by over-weighting goods that are salient in expenditure. Newey-west standard errors with 12 lags.

Table 2.17: Expected and realized inflation with salience in frequency

	Income 1	Income 2	Income 3	Income 4
	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$
$\bar{\pi}^i_{t-12,t}$	2.093***	0.560*	1.143***	0.776
	(0.000)	(0.089)	(0.000)	(0.155)
$\pi_{t-12,t}$	1.854***	1213**	2.062***	1.764***
	(0.000)	(0.045)	(0.000)	(0.008)
$\bar{\pi}^{i,f}_{t-12,t}$	0.922**	-0.0178	0.334	-0.559
	(0.010)	(0.981)	(0.365)	(0.601)
Cons	5.765***	4.789***	4.554***	4.240***
	(0.000)	(0.000)	(0.000)	(0.000)
N	108	108	108	108
Adj. R-sq	0.706	0.130	0.562	0.306

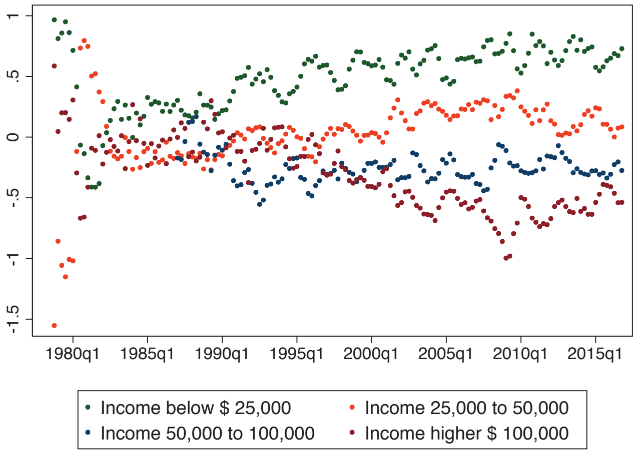
Note: p-values in parentheses* p<0.10, ** p<0.05, *** p<0.01. $\pi_{t-12,t}$ is the US inflation. $\bar{\pi}^{i,f}_{t-12,t}$ is the income specific inflation rate computed by over-weighting the goods that are salient in frequency. Newey-west standard errors with 12 lags.

Table 2.18: Readiness to spend and inflation expectations

	Income 1	Income 2	Income 3	Income 4
		Durable goods		
$\tilde{E}_t^i(\pi_{t,t+12})$ - Fitted values	0.0381**	0.0348	0.0599***	0.0206
	(0.040)	(0.636)	(0.008)	(0.545)
N	108	108	108	108
Adj. R-sq	0.786	0.716	0.844	0.816
		Cars		
$\tilde{E}_t^i(\pi_{t,t+12})$ - Fitted values	0.0243**	0.136***	0.0583***	0.0954***
	(0.030)	(0.000)	(0.000)	0.000)
N	108	108	108	108
Adj. R-sq	0.757	0.777	0.826	0.810

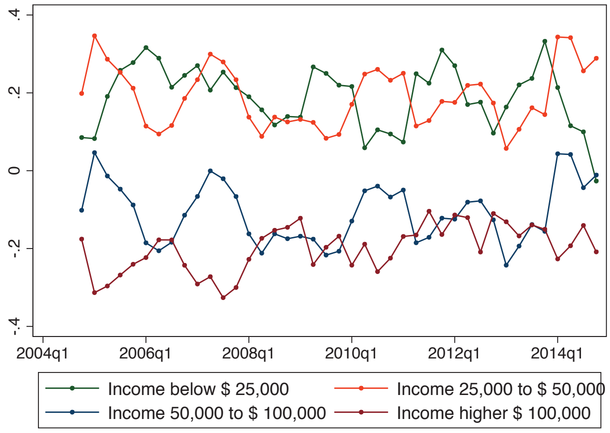
Note: p-values in parentheses * p<0.10, ** p<0.05, *** p<0.01. Newey-west standard errors with 12 lags. $\tilde{E}_t^i(\pi_{t,t+12})$ are the fitted values from the baseline model of intuitive expectations in equation (2.6). All columns include the following controls: shares of respondents that expect better business condition, higher interest rates and greater income. See Appendix.

Figure 2.1: Inflation expectations by income group relative to cross-sectional mean



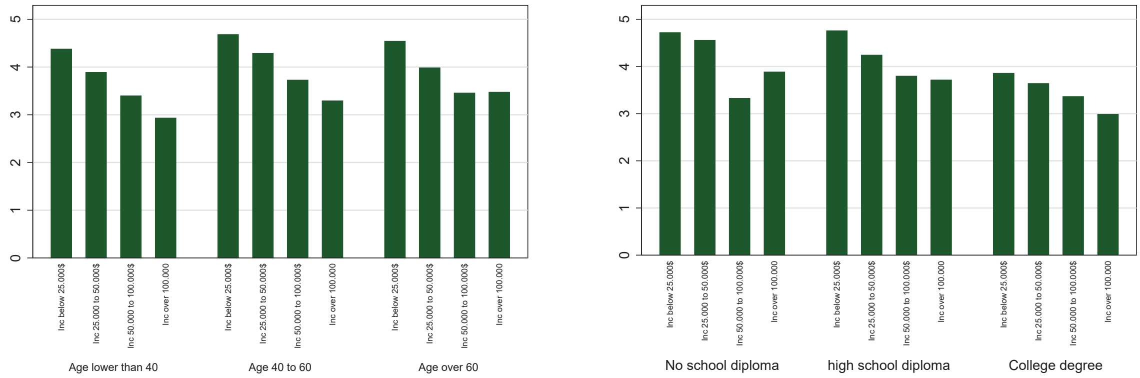
Note: Four-quarter moving average of the mean one-year inflation expectations by income groups, shown as deviation from the cross-sectional mean expectation. MSC data at quarterly frequency, 1978-2018.

Figure 2.2: Individual inflation by income group relative to cross-sectional mean



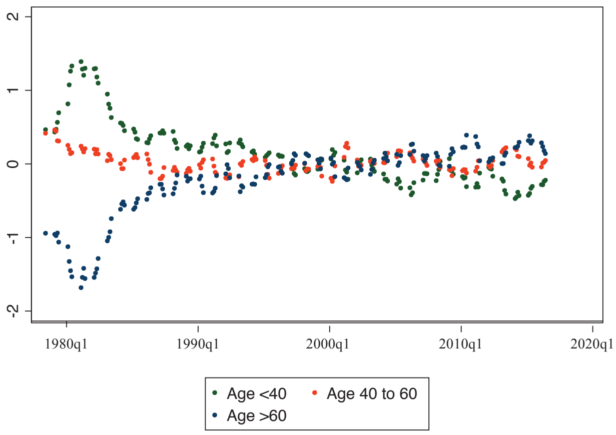
Note: Four-quarter moving average of the mean quarterly individual inflation rate by income groups, shown as deviation from the cross-sectional mean expectation. KNCP data at quarterly frequency, 2004-2014.

Figure 2.3: Inflation expectations by age, education and income group



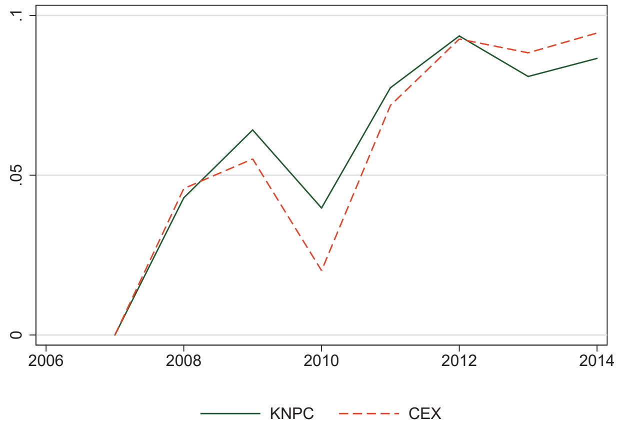
Note: Data are at monthly frequency, 2004-2014. The plot reports the mean expected inflation one year ahead for each group.

Figure 2.4: Inflation expectations by age group relative to cross-sectional mean



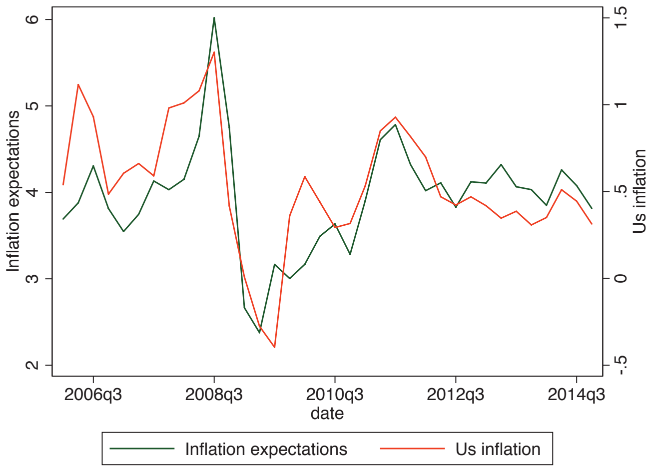
Note: Four-quarter moving average of the mean one-year inflation expectations by age groups, shown as deviation from the cross-sectional mean expectation. MSC data are at quarterly frequency, 1978-2018.

Figure 2.5: Aggregate expenditure growth rate relative to 2007



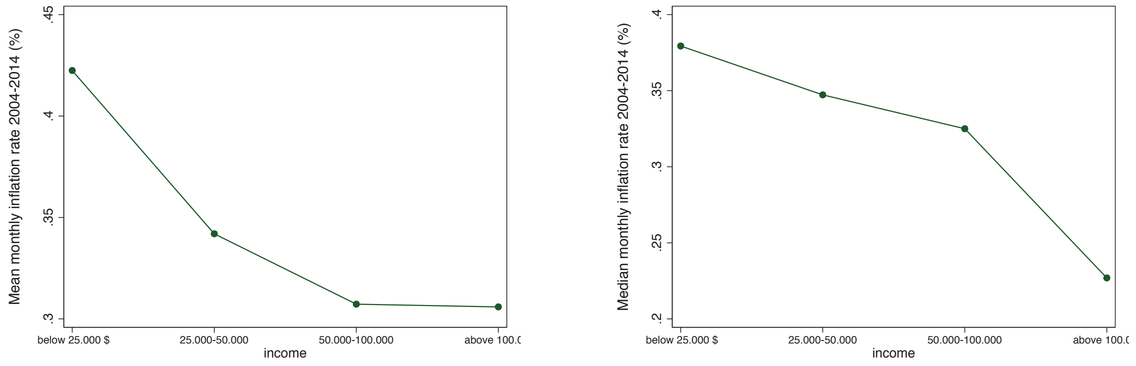
Note: Using KNCP and Consumer Expenditure Survey (CEX) data from 2007-2014, we compute the annual aggregate expenditure, and the growth rate relative to 2007: log(expenditure$_t$)/log(expenditure$_{2007}$). As explained in the text, the KNCP covers only a subset of goods, while the CEX survey covers the entire consumption basket.

Figure 2.6: US CPI inflation and inflation expectations



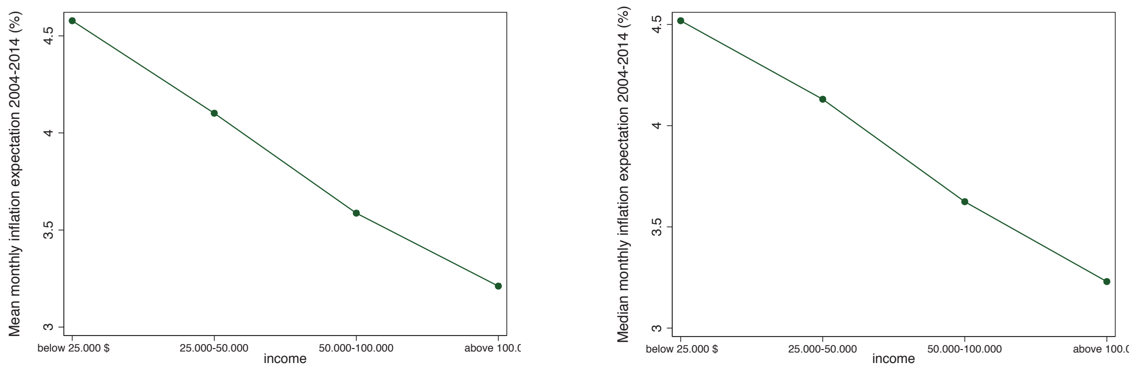
Note: Four-quarter moving average of the mean one-year inflation expectations and monthly US CPI inflation. Data are at quarterly frequency, 1978-2018.

Figure 2.7: Inflation rate (Laspeyres index) across income classes



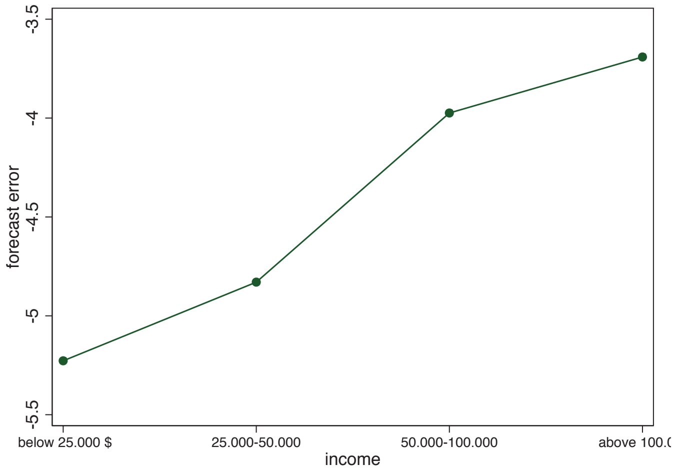
Note: The right hand side panel displays the mean individual inflation rate by income group; the left hand side the median. Individual data at monthly frequency are aggregated by income class and then collapsed over the period 2004-2014. The inflation rate (Laspeyres index) at the household level is computed as described in Section 2.3.2.

Figure 2.8: Inflation expectations across income classes



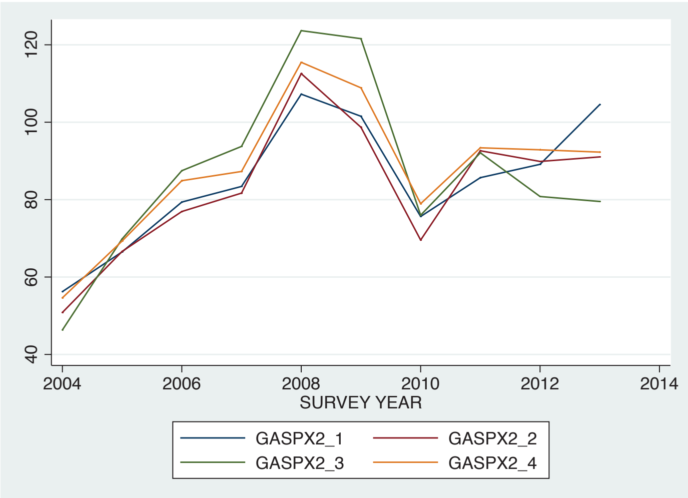
Note: The right hand side panel displays the mean expected inflation rate over the next year for each income group; the left hand side the median. Individual data at monthly frequency on expectations are aggregated by income class and then collapsed over the period 2004-2014. Data on inflation expectations at the household level are provided by the MSC (Section 2.3.3).

Figure 2.9: Forecast error by income class



Note: Note: For each income group, the forecast error is defined as $fe_t^i = (\pi_{t,t+12}) - \bar{E}_t^i(\pi_{t,t+12})$. Monthly data are collapsed over the period 2004-2014.

Figure 2.10: Expected gasoline price by income groups



Note: The figure plots the fraction of agents who expect oil price to increase during the next twelve months, provided by the MSC. Data are at annual frequency, for the period 2004-2014.

Chapter 3

Parenting Style and Women Empowerment

3.1 Introduction

The project investigates how economic development, rights, and social norms affect parenting styles, female education and in turn women empowerment, defined as the ability of women to access education, earning opportunities and rights of choosing their career. Equal access to education is at the core of gender equality, since it permits to gain skills, independence and participate in knowledge-intensive economic activities. At the same time, parents′ expectations for their daughters determine girls′ aspirations, investment in human capital as well as future career choices. In this framework, I address the following research question: How do economic and social incentives drive parenting styles towards daughters? Do parenting styles affect women education and future success?

To address this question I use an overlapping generation model, that builds on Doepke and Zilibotti (2017), in which socioeconomic variables determine the equilibrium parenting style, which in turns feeds back into daughters′ preferences for education, human capital and future career success. The model reflects the idea that parents′ expectations for their daughters determine girls′ aspirations, self-confidence and future choices. Children develop their beliefs and dreams within the family and therefore are strongly affected by their parents′ views and behaviours. *Parenting style* refers to the strategy that parents adopt to raise their children (Doepke et al. (2019)). Parents have altruistic and paternalistic motives and can affect their daughter by shaping her preferences (persuasion) and through restricting her choices (coercion). Parents can choose between three different styles: authoritarian, authoritative, and permissive (Baumrind (1967)). The selected style depends on both parental preferences and socioeconomic environment and feeds back into the girlâs ambition (preferences), education and occupation.

Parenting styles are driven by (i) economic incentives, such as return to schooling, and (ii) the institutional setting (social and legal norms) which determines the perceived cost of emancipation; for instance, the cost is high when there are cultural stereotypes against women abilities or frictions in the labour market. In equilibrium, if the perceived cost of emancipation is high, parents are authoritarian and exclude their daughter from the formal labour market. If the cost of emancipation is low and return to schooling is high, instead authoritative parents boost the girl′s ambitions and willingness to study so that women education and participation in the labour market increase. Finally, if the cost of emancipation and return to schooling are low, permissive parents have no incentive to exclude the girl from the labour market nor to enhance her aspirations, so they leave her free to choose the level of education and occupation according to her preferences. In this framework, both societal and economic institutions are key elements to explain how parenting styles and women empowerment vary across time and countries and the model's predictions are consistent with the cross-country empirical evidence.

The predictions of the model are coherent with the empirical evidence on the link between economic development, social and legal norms, and gender equality. As explained by Duflo (2012) economic development, return to education, and job opportunities, increase empowerment of women. At the same time, there is evidence that growth is not enough to overcome discrimination; for instance, the gap between girls and boys in terms of schooling, salary and legal rights have not improved in Asian countries, (i.e. China and India), despite rapid economic growth. The model reflects both these facts. On the one hand, a rise in return to education helps women′s empowerment; on the other side, economic growth alone is insufficient to ensure significant progress since legal rights, social norms and stereotypes against womenâs ability play a major role in driving gender equality. The model is also coherent with the evolution of women rights and education in most advanced economies in the last 200 years, where the technological changes were accompanied by a drastic change in women's legal rights, as well as in their perceived role in the society and their education and occupations. As argued by Doepke and Zilibotti (2018), indeed, the evolution of gender roles and child-rearing practices, over time and across countries, has been shaped by technological and cultural changes; the socio-economic environment is a key driver of parents′ attitudes towards girls and parents choose how to raise girls according to the observed economic role of women to prepare them for the adult life.

Overall, the model connects to the idea that culture, defined as *beliefs and values that ethnic, religious, and social groups transmit fairly unchanged from generation to generation*, has an impact on economic outcomes (Guiso et al. (2006)). In this framework, indeed, gender norms affect parents' behaviour and in turn daughters' economic preferences and choices. Further, gender-role attitudes, social and legal norms, are persistent overtime since they are transmitted from parents to children and determine women′s inherited ambition. This transmission generate a feedback loop between gender norms, women education and labour force participation that reinforce each other. Finally, the model is consistent with the evidence that in advanced economies women views′ on the benefit from working evolve over time, as argued by Fernandez (2013). In this framework, in fact, when the barriers to emancipation are low, women themselves decide their occupation and education and their attitude towards working evolve with the socioeconomic environment.

The paper builds on the work by Doepke and Zilibotti (2017), who propose an economic theory that rationalises the choice between alternative parenting styles. In the present project, I adopt their framework to study how economic background, local culture, social norms and institutions shape parents′ aspirations for their daughters and subsequently women empowerment. The contribution of this work is to suggest a new application of the economic theory of parenting style proposed by Doepke and Zilibotti (2017), to adapt it to explain the wide variation in female labour force participation and education worldwide and to provide empirical evidence in favour of it (see Doepke et al. (2019) for a survey of the growing economic literature on parenting). The study also connects to the discussion on how parents raise boys and girls differently by Doepke and Zilibotti (2018). In Chapter 6, the authors argue that gender differences in child-rearing depend on the expectations about what roles women and men should play in the society and explain how differences in gender roles affect the incentive that parents face in raising their daughters. Child-rearing practices are rooted in the socioeconomic environment in which parents themselves grew up and in which they expect their children to leave; in societies where the traditional role models prevail, parents seek to prepare their daughters for a similar environment where women and men carry out different tasks. Relative to their discussion, here I formalise the intuition that economic and cultural incentives drive parenting styles which in turn shape girls′ behaviours in a framework where parents are motivated by altruism. Specifically, I focus on the role of return to education, social norms and institutions to explain differences in parents attitudes towards girls and women empowerment across countries. The paper also connects to the model by Glaeser and Ma (2014) in which parents deceive their daughters about their ability to affect their choices in terms of childbearing. In their model parents know the daughters' ability but daughters themselves do not and infer the ability from the parental investment in their human capital. The idea is that parents care about their grandchildren and attempt to increase their daughters′ childbearing by persuading them that the returns from working are lower than those from bearing children. Similarly, in the proposed model parents shape daughters′ beliefs to affect their future choices; however the objective pursued by the parents and their incentives are different. First, parents shape girls' ambitious and self-confidence aiming at affecting their investment education, while in Glaeser and Ma (2014) parents choose the level of human capital to drive fertility choices. Second, parents' attitudes are driven by socioeconomic incentives as opposed to the desire for grandchildren as in Glaeser and Ma (2014).

Finally, the study also contributes to the literature on the determinants of women empowerment (i.e. Guiso et al. (2008); Alesina et al. (2013)) providing a new framework to study its evolution and in particular to the strand of the literature that investigate the link between culture and women$slabor force participation (i.e. Fernandez$ (2013)$, Fernandez et\ al.$ (2004)$, Fernandez$ (2007)).

The remainder of the article proceeds as follows. The next section describes the theoretical framework and summarizes the testable prediction from the model. Section 3 describes the data used to test the mentioned prediction and Section 4 displays the empirical evidence. Section 5 concludes.

3.2 A model of preference transmission

In this section, I provide a sketch of the theoretical model that builds on the one proposed by Doepke and Zilibotti (2017) and applies it to the transmission of "ambition" to daughters, features that I assume to be important for women empowerment as well as human capital and wealth accumulation.

Framework. I consider an overlapping generation model in which people live for two periods, childhood (c) and parenthood (p), and each parent has one daughter. Daughters (children) and parents have different preferences: $U^c(c|a) \neq U^p(c|a)$. Parents have altruistic and paternalistic motives; the latter derives from the fact that children and parents have different utility functions. Daughters choose $x = [x^e, x^\mu]$, where $x^e \in [0, 1]$ is the investment in education and $x^\mu \in X = \{x^{HOME}, x^{WORK}\}$ is the occupation. $x^\mu = x^{HOME}$ implies that the daughter is excluded from the labour market (or otherwise is employed in an unskilled job); $x^\mu = x^{WORK}$ instead implies that she works in a skill-intensive activity that she is free to choose. The daughter$'$s choices are affected by the preference vector (ambition or self-confidence) $a' \in [1, \bar{a}]$, which is defined during childhood and influences her utility in both periods as it determines the level of education. A lower level of a' corresponds to a higher ambition and willingness to study. The investment in education represents an initial cost in terms of consumption for the child, but increases her utility in the following period given the positive return to schooling. Each parent can affect child's choices through shaping her ambition $a' \in [1, \bar{a}]$ (persuasion) or restricting the set X imposing her not to work (coercion), such that $X = \{x^{HOME}\}$. Children naturally set $a' = \bar{a}$ to maximise their childhood utility, but the parents can have the incentive to lower it to boost the child's ambition and preference for education. The level of education chosen by the daughter x^e is decreasing in a'; a lower level of a' indeed reduces the daughter's instant utility but increases her investment in education and future utility.

Following Doepke and Zilibotti (2017), each parent maximizes an objective function that depends on the future choices made by the daughter. The value function of the parents is given by:

$$v^p(a,h) = \max_{c^p,a',X} \{U^p(c^p|a) + \delta w(X,a'|a)\} \tag{3.1}$$

Where $w(X,a'|a)$ is the utility that parents derive form the children experience, δ measures the degree of altruism and

$$w(X,a'|a) = \{(1-\lambda)U^c(c^c|a') + \lambda U^p(c|a) + \beta v^p(a',h')\} \tag{3.2}$$

Here λ is the relative weight of the paternalistic component and $(1-\lambda)$ the one associated to the altruistic component. With $\lambda = 1$, the altruistic component disappears and parents take into account only their point of view. Alternatively, when $\lambda = 0$ there's no paternalistic motive. β is the discount factor. The value function of the daughter (children) is given by:

$$v^c(X,a') = \max_{c^c,x^e,h'} \{U^c(c^c|a') + \beta v^p(a',h')\} \tag{3.3}$$

Let $R > 0$ be the return from education, $y(x^\mu)$ the output obtained for a given occupation x^μ, and $0 < \sigma < 1$, we set:

$$U^p(c|a) = \frac{(c^p)^{1-\sigma}}{(1-\sigma)}$$

$$U^c(c|a) = a\frac{(c^c)^{1-\sigma}}{(1-\sigma)}$$

$$c^c = (1-x^e)y(x^\mu)$$

$$c^p = h = (1+Rx^e)y(x^\mu)$$

$$y(x^{HOME}) = \mu\frac{y_H}{(1+Rx^e)}$$

$$y(x^{WORK}) = y_W \text{ with } y_W > y_H$$

The main difference from the original model is $y(x^{HOME})$; here it is such that the future consumption does not depend on the education level when the daughter stays at home and does not work. In this case, the girl has no incentive to invest in education as it represents an immediate cost and it's not associated with any benefit in the future. The inverse relationship between $y(x^{HOME})$ and x^e can be explained by the fact that while studying girls may lose some skills useful for housework; these are innate skills that can be developed without studying and may be helpful in the marriage market. This hypothesis is consistent with the evidence provided by Bertrand et al. (2016) that show how the marriage market is unfavorable for highly educated women. College-educated women, in fact, marry at a lower rate then less educated women; the gap has been decreasing in several countries, mostly in North America and Nordic countries,

but it remains large in others such as East Asian countries. The investment in human capital instead is profitable in the case of work, as it pays a positive return.
The relative advantage of working depends on the return to education R, the differences in the outcome between the two occupations $(y_W - y_H)$ and the parameter $\mu > 1$ which reflects the *cost of emancipation* that rely on the institutional setting and the local culture; for instance, it captures the religion and social norms, or the barriers/frictions of the labour market. The benefit from being at home indeed depend both on economic and social factors, such as the difference between the productivity in the workplace and in the household, the setting of the labour market (i.e. if formal laws impose no gender discrimination) and the societal beliefs on women's ability and their role in the economy. The larger is μ, the larger is the cost of emancipation and the benefit from not working. The interpretation of this parameter is a key novelty as it crucially affects the equilibrium parenting style and female education. The cost can exist both in developed and developing countries. In developing countries, the cost of emancipation reflects mainly gender or religious norms and barriers to the labour market, for instance if women have fewer job opportunities or lower salaries compared to men. In more advanced economies, these frictions are smaller, but present although the formal legal norms tend to reduce gender discrimination; the costs, in fact, may incorporate religious and cultural norms as well as a "stereotype threat" on women abilities that may affect womenâs confidence (as well as confidence in women), participation in the labour market and investments in human capital. Further, μ reflects the societal bias and does not necessarily reflect the parents' view on gender stereotypes; parents may find convenient for girls to be involved in housework although they do not share the local culture because they aim at preparing the child for adult life in the observed environment (Doepke and Zilibotti (2018)). Overall, this cost may be widely heterogeneous across and within countries and capture different features in different areas according to the societies' characteristics.

Parenting styles. Each parent can choose among three parenting styles (Baumrind (1967)), authoritarian, authoritative, and permissive, defined as follows:

1. A parent is authoritarian if he restricts the choice set of his daughter to $X = \{x^{HOME}\}$, so that excludes her from the labour market. An authoritarian parent teaches obedience to his children and imposes her the best occupation according to his point of view.

2. A parent is authoritative if he increases the child's ambition $(a' < \bar{a})$ so that she is more willing to invest in education. An authoritative parent does not impose obedience but persuades the child to be more ambitious to influence her choices.

3. A parent is permissive if he is not authoritarian nor authoritative and grants independence to the child.

Authoritarian. Parents first choose if they want to be authoritarian or not by solving the maximization problem. This choice hinges on the value of μ. If authoritarian the parents restrict the choice set of the daughter to $X = \{x^{HOME}\}$ and sets $a' = \bar{a}$ and $x^e = 0$. In this case, indeed, parents have no incentive to reduce a' to boost education since by reducing a' they would reduce the daughter instant utility. The idea here is that parents may prefer to have their daughters not working, or working in some specific occupations (i.e. unskilled jobs or housework) given the socioeconomic background. For instance, this may happen in traditional societies where the appropriate place for women is within the home, and there's a negative social attitude towards working women so that parents discourage their daughters from participating in non-domestic activities. The cost μ reflects the overall economic and social cost that parents and girls would pay and can be expressed as a function of social, religious and legal norms. This cost is small in economies characterised by low gender inequality where women are allowed to participate freely in employment outside the home; while it is larger in countries where the marriage market is unfavourable for educated women or with strong inequalities.

Non-authoritarian. If non-authoritarian, parents choose between authoritative and permissive parenting. A non-authoritarian parent does not restrict the choice set of the daughter so that $X = X^{FREE} = \{x^{HOME}, x^{WORK}\}$. If μ is very high, the child prefers to stay at home although she is free to choose a different activity. This is the case in which the social or economic costs of emancipation are very high, so the girl herself does not want to participate in the labour market. The threshold level of μ above which the girl decides not to study is increasing in R. Otherwise, if μ is sufficiently low, she will choose to study and work. In this case, for a given a', the education level selected by the daughter is:

$$x^e(a') = \underset{x^e}{argmax}\{a'\tfrac{(1-x^e)^{1-\sigma}}{(1-\sigma)} + \beta\tfrac{(1+Rx^e)^{1-\sigma}}{(1-\sigma)}\}$$

The non-authoritarian problem remains equal to the one described by Doepke and Zilibotti (2017). The non-authoritarian parents cannot impose a particular level of x^e but can induce it modifying a'. The idea is that parents can persuade their child for instance by choosing their experiences, friends, or readings. The choice hinges on the parameters R and λ. The authors explain that there is a marginal benefit from lowering a', since education increases, and a marginal cost that derives from the utility loss suffered by the child. The marginal benefit and cost depend on the level of paternalism λ and return to human capital R. If $\lambda = 0$ the parents are fully altruist, the marginal benefit from lowering a' vanishes, parents are permissive and $a' = \bar{a}$; if $\lambda = 1$, instead, the marginal cost is null, parents evaluate the child's utility exclusively from their point of view, they choose to be authoritative and set $a' = 1$. For intermediate values of λ, R plays a crucial role in the equilibrium parenting style as the larger it is, the larger is the benefit from education. Hence, for a given level of $\lambda \in (0, 1)$, parents prefer to be authoritarian for large values of R, and permissive for low returns from education.

Equilibrium parenting style. The choice between authoritarian and non-authoritarian depends on the level of μ. For a given level of λ, there exists a unique threshold $\hat{\mu}(R)$, such that for $\mu \geq \hat{\mu}(R)$, parents choose to be authoritarian. The choice between permissive and authoritative parenting instead depends on the level of R and λ. A detailed solution to this problem is provided by Doepke and Zilibotti (2017), Proposition 1.

Let $\lambda^* = \frac{\sigma\bar{a}}{\bar{a}-(1-\sigma)}$ and $\bar{R} = (\frac{\sigma\bar{a}(1-\lambda)}{(\lambda(\bar{a}-(1-\sigma)-\sigma\bar{a}})^{\frac{\sigma}{1-\sigma}}(\frac{\bar{a}}{\beta})^{\frac{1}{1-\sigma}}$. There exists a function $\hat{\mu}(R) > 0$ such that:

(1) If $\lambda > \lambda^*$ then:
1. If $\mu > \hat{\mu}(R)$, parents choose an authoritarian style ($a' = \bar{a}, x^e = 0$).
2. If $\mu \leq \hat{\mu}(R)$ and $R > \bar{R}$, parents choose an authoritative style ($a < \bar{a}, x^e > 0$).
3. If $\mu \leq \hat{\mu}(R)$ and $R \leq \bar{R}$, parents choose a permissive style ($a' = \bar{a}, x^e \geq 0$).

(2) If $\lambda \leq \lambda^*$ then:
1. If $\mu > \hat{\mu}(R)$, parents choose an authoritarian style ($a' = \bar{a}, x^e = 0$).
2. If $\mu \leq \hat{\mu}(R)$, parents choose a permissive style ($a' = \bar{a}, x^e \geq 0$).

This result is similar to the one reported by Doepke and Zilibotti (2017), Proposition 2. There are two main differences. First, the threshold $\hat{\mu}(R)$ does not have a lower bound independent of R. Second, in equilibrium when parents are authoritarian x^e is equal to zero, rather than being an increasing function in R. Alternatively, we may impose a minimum legally required level of investment in education such as $x^e \in [x^e_{MIN}, 1]$ with $x^e_{MIN} > 0$; in this case, with authoritarian parents the investment in education would be positive equal to x^e_{MIN}, while with authoritative and permissive parents it would be larger (or equal) than x^e_{MIN}.

Testable predictions. The theoretical model gives several testable predictions that I check in the next sections. First of all, the socioeconomic background shapes parenting styles, so that:
1. When the cost of emancipation is high, parents choose to be authoritarian; otherwise, parents choose a non-authoritarian behaviour.
2. When the return from education is high, and the cost of emancipation is low, parents prefer an authoritative parenting style.
Second, parenting styles have long-run implication for women education and future career:
3. Authoritarian parents exclude their daughters from the formal labour market (or restrict the set of possible occupations), and girls do not invest in education.
4. Authoritative parenting is associated with high investment in education and participation in the labour force.

3.3 Data

To test the model's predictions I use several sources of data:

1. World Value Survey (WVS)

   Data on parenting style comes from the World Value Survey (WVS) Waves 1-6 conducted during the period 1981-2014 (Inglehart et al. (2014)). [1] Each wave involves a large number of respondents in different countries, for instance the last wave includes interviews with about 85,000 individuals in 60 countries. In the survey people are asked *Here is a list of qualities that children can be encouraged to learn at home. Which, if any, do you consider to be especially important? Choose up to five qualities listed above.*[2] I use answers to this question to build measures of parenting style considering only respondents that declare to have at least one child. Following Doepke and Zilibotti (2017), I label as authoritarian parents that choose *obedience* as the desired quality, authoritative those who list *hard work* and are not authoritarian, and permissive those who are not authoritarian nor authoritative and list *independence*. The three styles are mutually exclusive but not exhaustive. Data are then aggregated at the country-wave level. Table 3.1 reports the summary statistics on the parenting styles for the final sample. The WVS also provides information on local culture, religion and gender-related questions that provides information on social and religious norms and the perceived role of women in the society; these pieces of information are used to build proxies of the cost of emancipation as described in Table 3.3.

2. Gender Equality Database (GED)

   Data on gender inequality comes from the Gender Equality Database (GED) provided by the World Bank, which contains the time series of a wide range of indicators for over 260 countries for the period 1981-2014.[3] Specifically, it provides information on female education, school enrollment, and occupation as well as labour market legislation regarding gender inequality. For these indicators, for each Wave of the WVS, I compute the average value over the years in which the survey WVS is conducted (i.e. for the first wave, I average over the period 1981-1984). These data are used to compute the cost of emancipation deriving from the formal institutions and labour-market laws (Table 3.3) and the socioeconomic variables, such as female education and labour force participation (Table 3.2).

[1] The waves are for the periods: 1981-1984, 1989-1993, 1994-1998, 1999-2004, 2005-2009, 2010-2014.

[2] The qualities are *Independence, Hard work, Feeling of responsibility, Imagination, Tolerance and respect for other people, Thrift saving money and things, Determination, perseverance, Religious faith, Unselfishness, Unselfishness, Obedience, Self-expression*

[3] The GED provides data for a large number of countries, but with missing data since some indicators are not always available.

3. World Bank Open Data

   Finally, I use macroeconomic variables from World Bank Open Data on income inequality to build proxies of the return from education. This proxy seems appropriate, since as described by Doepke and Zilibotti (2017), the advantage associated to schooling is lower in countries where there's low inequality. Definitions and summary statistics for these variables are provided in Table 3.2.

Merging the three sources of data, I get a final dataset with 93 countries, that I observe on average 2.4 times, and 227 observations. The final dataset is an unbalanced panel data with data aggregated at country-wave level.[4] The variables are defined as follows (Tables 3.1-3.3).

3.4 Empirical evidence

In this section, I first show that the popularity of the different parenting styles varies across countries that differ in social and legal norms, and benefits from education. Second, I document that the three parenting styles are associated with different outcomes concerning female education and occupation.

Emancipation cost and parenting style. I start by testing the first prediction of the model, according to which when the cost of emancipation is high, parents choose an authoritarian style. The emancipation cost represents the cost of being engaged in a skilled occupation, instead than being a housewife (or being involved in an unskilled job). The costs and benefits of working depend on several distinct aspects, such as the time for the family life, the potential prejudices against working women derived from religiosity or traditional culture, as well as the remuneration from working and the functioning of the labour market. Here, I consider three main components: the religious norms, the social norms and the labour market institutions. The religious norms reflect the role that religiosity plays in shaping the involvement of women in society. I build a proxy of the religion-related costs as the fraction of religious agents in a given country using three different questions of the WVS, as described in Table 3.3. The social norms, instead, reflect the potential prejudice toward women participation in the labour market in societies where the dominant belief is that women should be involved in domestic activities.

[4]The sample includes: Albania, Andorra, Azerbaijan, Argentina, Australia, Bangladesh, Armenia, Bosnia Herzegovina, Brazil, Bulgaria, Belarus, Canada, Chile, China, Colombia, Croatia, Cyprus, Czech Rep., Dominican Rep., Ecuador, El Salvador, Ethiopia, Estonia, Finland, France, Georgia, Germany, Ghana, Haiti, Hong Kong, Hungary, India, Indonesia, Iran, Iraq, Israel, Italy, Japan, Kazakhstan, Jordan, South Korea, Kuwait, Kyrgyzstan, Lebanon, Latvia, Libya, Lithuania, Malaysia, Mali, Mexico, Moldova, Montenegro, Morocco, Netherlands, New Zealand, Nigeria, Norway, Pakistan, Peru, Philippines, Poland, Puerto Rico, Qatar, Russia, Rwanda, Saudi Arabia, Serbia, Singapore, Slovakia, VietNam, Slovenia, South Africa, Zimbabwe, Spain, Sweden, Switzerland, Thailand, Trinidad and Tobago, Tunisia, Turkey, Uganda, Ukraine, Macedonia, Egypt, United Kingdom, Tanzania, United States, Burkina Faso, Uruguay, Uzbekistan, Venezuela, Yemen, Zambia.

A proxy for this cost is built as the fraction of agents who believe that women have less right to study and work than men, should earn less than their husband, or that when women work their children suffer, as described in Table 3.3. Finally, the benefits from working derive also from the degree of legal protection against gender discrimination in the labour market and the maternity leave. The degree of protection varies substantially across countries. A proxy of this cost is computed as described in Table 3.3, the larger is the value of the proxy the larger are the barriers women face to enter the labour market. As predicted by the model, there is a positive and significant correlation between the popularity of the authoritarian parenting and the religion and social costs; while there is a negative relationship between these costs and the presence of non-authoritarian parenting (Table 3.4). Similarly, when the discrimination in the labour market is low, non-authoritarian parenting increase. Religiosity appears to be a primary determinant of the parenting style (Figure 3.1); religious norms explain up to 40% of the variation in the authoritarian style, and an increase of one unit in religiosity is associated with an increase of 0.5 in the fraction of authoritarian parents. This result is confirmed in Table 3.5 which show that adding the other indexes does not improve the adjusted r-squared; further, the evidence is consistent at the micro level data (see Appendix).

Return from education and parenting style. Second, the model predicts that when the return from education is high, and the cost of emancipation is low, parents prefer an authoritative parenting style. The intuition is that when the profits at stake are high the benefit from increasing daughters' ambition is larger. To test this prediction, following Doepke and Zilibotti (2017), I use measure of income inequality as proxies for the gain from education. The idea is that in more re-distributive societies the benefits from schooling and earnings are lower compared to more unequal societies. Precisely, I look at the Gini coefficient and the differences between the income share held by the top 10th (20th) percent and the lowest 10th (20th) percent. The larger are these indexes the larger is the degree of income inequality within a country. As predicted by the theory, for a given cost of emancipation, there is a negative correlation between the popularity of permissive parenting and inequality (Table 3.6, Figure 3.2), this result is consistent with Doepke and Zilibotti (2017) which provide a detailed analysis of the link between return to education and parenting style. When looking at authoritative parents, the correlations are not significant and the adjusted R-squared reduces substantially, suggesting that alternative motivations for being authoritative are in place. Social norms may play a role here, indeed, the coefficients associated to the social norms (Columns (1)-(3)-(5)) are positive and significant, while the theory predicts a negative correlation. This result may be because when stereotypes against women are strong parents may have the incentive to be authoritative to make their daughters' working even harder to cope with the prejudices and have success in their future career. This might be an additional mechanism not captured by the sketched model. Finally, the coefficients associated with the religious norms are negative as predicted by the theory, while institutions are not relevant.

Parenting style and women education and occupation. Finally, the model predicts that parenting styles affect women education and participation in the labour market. To test these predictions, I look at the relationship between parenting styles and the percentage of women who completed at least the upper secondary school (Educational attainment, at least complete upper secondary, female (%)) and the female labor force participation rate (% of total labour force). I consider the secondary school since, as pointed out by Duflo (2012), in recent years primary school enrollment has become almost universal for both boys and girls, while the gap in terms of secondary school enrollment is still present in many countries. As expected authoritarian parenting is linked with lower levels of education and women participating in the labour force; the correlations are negative and significant (Table 3.7, Figure 3.3). On the other hand, authoritative parenting is associated with higher investment in education and higher labour force participation rate. The negative correlation between authoritarian parents and educational attainment is confirmed when looking at the percentage of women that obtains a bachelor degree or complete tertiary education (Table 3.8), confirming that parents' attitudes have a long-term effect on human capital. Further, on average in countries with a significant fraction of authoritarian parents, the expected years of schooling for girls are lower compared to countries where the alternative parenting styles are more popular. The link between authoritarian parenting and education is particularly strong, while the one with the labour force participation is lower; this evidence can be rationalised considering that the labour force also includes unskilled jobs for which investment in human capital is not needed.

3.5 Conclusion

The paper analyses how economic development, rights, and cultural norms affect parenting attitudes towards girls and in turn gender inequality. I suggest an overlapping generation model in which the socio-economic environment defines the equilibrium parenting style, which feeds back into daughters' ambitions, investment in schooling, and future economic success. The model can be used to describe variations across countries, and the predictions are in line with the cross-country empirical evidence. In this framework, women's beliefs are developed within the family sphere where gender stereotypes are implicitly transmitted across generations and have a long-term impact on women educational and occupational choices. The model reflect the evidence that in some occurrences, girls grow up with low-confidence in their abilities and expectations for their future career prospects in others instead parents foster daughters ambition and confidence and encourage them to fulfil their career goals. Returns to education, social and religious norms and institutions are key elements to explain parents attitudes towards girls and hence women empowerment.

3.6 Figures and Tables

Table 3.1: Parenting style

Parenting style	Question	Answer	Mean and Sd	Obs	Sourcs
Authoritarian	Here is a list of qualities that children can be encouraged to learn at home.	Obedience	40.10 (18.40)	227	WVS
Authoritative	Which, if any, do you consider to be especially important?	Hard work (if not Authoritarian)	34.00 (17.56)	227	WVS
Permissive		Independence (if not Authoritative and Authoritarian)	15.63 (13.64)	227	WVS

Note: Data on parenting style comes from the World Value Survey (WVS) Waves 1-6 conducted during the period 1981-2014. Microdata from the WVS is aggregate at the wave-country level. Standard deviations in parenthesis.

Table 3.2: Socioeconomic variables

Variables	Definition	Mean and Sd	Obs	Source
Women education	Educational attainment, at least complete upper secondary, female (%)	48.09 (24.93)	111	GED
Labour force participation	Labor force participation rate, female (% of total labour force)	40.07 (9.01)	217	GED
Bachelor degree	Educational attainment, at least Bachelor's or equivalent, female (%)	15.16 (10.32)	115	GED
Terziary education	Tertiary education, gross completion ratio, female	30.7 (20.03)	108	GED
Expected schooling	Expected Years of School, Female	13.3 (2.9)	167	GED
Gini Index	GINI index (World Bank estimate)	39.6 (9.30)	150	World Bank
Income top 10% - lowest 10 %	Total share held by highest 10% minus the share held by the lowers 10%	27.44 (9.58)	155	World Bank
Income top 20% - lowest 20 %	Total share held by highest 20% minus the share held by the lowers 20%	38.58 (12.28)	155	World Bank

Note: Data on gender inequality, for the period 1981-2014, comes from the Gender Equality Database (GED) provided by the World Bank. Other inequality indicator are also form the World Bank Open Data. Standard deviations in parenthesis. *Women Education* is the percentage of population ages 25 and over that attained or completed upper secondary education (SE.SEC.CUAT.UP.FE.ZS). *Labour force participation* is the female labor force as a percentage of the total. Labor force comprises people ages 15 and older who supply labor for the production of goods and services during a specified period (SL.TLF.TOTL.FE.ZS). *Bachelor degree* is the percentage of population ages 25 and over that attained or completed Bachelor's or equivalent as the highest level of education (SE.TER.HIAT.BA.FE.ZS). *Terziary education* is number of graduates from first degree programmes expressed as a percentage of the population of the theoretical graduation age of the most common first degree programme (SE.TER.CMPL.FE.ZS). *Expected Schooling* is the the number of years a child of school entrance age is expected to spend at school, or university, including years spent on repetition (SE.SCH.LIFE.FE). The GINI Index measures the extent to which the distribution of income (or, in some cases, consumption expenditure) among individuals or households within an economy deviates from a perfectly equal distribution (SI.POV.GINI). *Income share top 10% - lower 10%* is the income share held by highest 10% minus the one held by the lowers 10%. Similarly, *Income share top 20% - lower 20%* considers the income share held by highest 20% and the lowers 20%.

Table 3.3: Emancipation cost (Cost of not staying at home)

Index	Question	Answer/Value	Mean and Sd	Obs	Source
{Religious norms}	How important is in life: Religion?	Very important	46.69 (29.82)	215	WVS
	How often do you attend religious services?	At least once a week	32.66 (23.81)	218	WVS
	Are you a religious person	Yes	70.79 (20.94)	217	WVS
Religious norms	Computed as the average of the three indexes above		50.19 (23.04)	225	WVS
{Social norms}	If a woman earns more money than her husband, it's almost certain to cause problems	Agree	31.82 (13.85)	56	WVS
	When a mother works for pay, the children suffer	Agree	54.3 (20.43)	68	WVS
	A university education is more important for a boy than for a girl	Agree	24.43 (13.25)	199	WVS
	When job are scarce, man should have more right to a job than women	Agree	39.16 (10.89)	215	WVS
Social norms	Computed as the average of the three indexes above		34.35 (16.04)	216	WVS
{Institutions}	Law mandates equal remuneration for females and males for work of equal value (Yes/No)	Yes	0.25 (0.44)	51	GED
	Law mandates nondiscrimination based on gender in hiring (Yes/No)	Yes	0.58 (0.500)	102	GED
	Law mandates paid or unpaid maternity leave (Yes/No)	Yes	0.96 (0.19)	102	GED
Institutions	Computed as the average of the three indexes above * (-1)		-0.69 (0.26)	102	GED

Note: Data on the religion related cost comes from the World Value Survey (WVS) Waves 1-6 conducted during the period 1981-2014. Microdata from the WVS is aggregate at the wave-country level. Data on gender inequality, for the period 1981-2014, comes from the Gender Equality Database (GED) provided by the World Bank, I use the following series: SG.LAW.EQRM.WK, SG.LAW.NODC.HR, and SG.LAW.LEVE.PU. Standard deviations in parenthesis. All data are at country level.

Table 3.4: Emancipation cost and parenting style

Variables	(1) Authoritarian	(2) No-Authoritar.	(3) Authoritarian	(4) No-Authoritar.	(5) Authoritarian	(6) No-Authoritar.
Religious norms	0.525***	-0.525***				
	(0.0421)	(0.0421)				
Social norms			0.211***	-0.211***		
			(0.0702)	(0.0702)		
Institutions					17.09**	-17.09**
					(6.876)	(6.875)
Constant	13.83***	86.17***	33.64***	66.36***	52.61***	47.39***
	(2.150)	(2.150)	(2.510)	(2.510)	(5.204)	(5.202)
Observations	225	225	216	216	102	102
Adj. R-squared	0.429	0.429	0.029	0.009	0.0434	0.0434

Note: Table displays results from estimating univariate regressions $Parenting_{i,t} = \alpha + \beta Indicator_{i,t} + \varepsilon_{i,t}$, where i is the country and t is the time (wave). The dependent variable is the fraction of respondents in the WVS that are authoritarian or non-authoritarian. No-authoritarian are those parents that are either authoritative or permissive. The independent variables instead are computed as described in Table 3.3. Heteroskedasticity consistent standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.5: Emancipation cost and parenting style

Variables	(1) Authoritarian	(2) Non Authoritarian
Religious norms	0.561***	-0.561***
	(0.0762)	(0.0762)
Social norms	-0.0857	0.0858
	(0.120)	(0.120)
Institutions	-2.626	2.635
	(7.029)	(7.027)
Constant	13.35*	86.67***
	(7.862)	(7.861)
Observations	102	102
Adj. R-squared	0.396	0.396

Note: Table displays results from estimating multivariate regressions $Parenting_{i,t} = \alpha + \beta Indicator_{i,t} + \varepsilon_{i,t}$, where i is the country and t is the time (wave). See note in Table 3.4 for details.

Table 3.6: Income inequality and parenting style

VARIABLES	(1) Authoritative	(2) Permissive	(3) Authoritative	(4) Permissive	(5) Authoritative	(6) Permissive
GINI INDEX	0.162	-0.582***				
	(0.220)	(0.158)				
Income top 10% - lowest 10%			-0.152	-0.410***		
			(0.207)	(0.146)		
Income top 20% - lowest 20%					-0.0800	-0.319***
					(0.158)	(0.122)
Religious norms	-0.386***	-0.0173	-0.304***	-0.0519	-0.318***	-0.0578
	(0.117)	(0.0775)	(0.106)	(0.0892)	(0.103)	(0.0890)
Social norms	0.497***	-0.393***	0.438***	-0.375***	0.448***	-0.375***
	(0.165)	(0.113)	(0.167)	(0.124)	(0.165)	(0.124)
Institutons	0.805	-4.163	0.116	-4.054	0.304	-3.828
	(8.777)	(5.284)	(8.584)	(5.462)	(8.627)	(5.420)
Constant	31.91***	48.91***	39.34***	38.87***	38.87***	40.41***
	(11.61)	(9.316)	(10.64)	(7.803)	(10.82)	(8.302)
Observations	85	85	86	86	86	86
Adj. R-squared	0.0914	0.322	0.089	0.277	0.089	0.277

Table displays results from estimating univariate regressions $Parenting_{i,t} = \alpha + \beta Inequality_{i,t} + \varepsilon_{i,t}$ where i is the country and t is the time (Wave). The dependent variable is the fraction of respondents in the WVS that are authoritative or permissive. The GINI Index measures the extent to which the distribution of income among individuals or households within an economy deviates from a perfectly equal distribution. *Income top 10% - lower 10%* is the income share held by highest 10% minus the one held by the lowers 10%. Similarly, *Income top 20% - lower 20%* considers the income share held by highest 20% and the lowers 20%. The other independent variables are computed as described in Table 3.3. Heteroskedasticity consistent standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.7: Parenting style, education and labour force participation

VARIABLES	(1) W. Education	(2) W. Education	(3) W. Education	(4) Labour force	(5) Labour force	(6) Labour force
Authoritarian	-0.751***			-0.128***		
	(0.103)			(0.0345)		
Authoritative		0.330**			0.121***	
		(0.139)			(0.0289)	
Non Authoritative			-0.330**			-0.121***
			(0.139)			(0.0289)
Constant	79.42***	37.39***	70.40***	45.26***	35.88***	48.01***
	(4.841)	(5.099)	(9.591)	(1.256)	(1.356)	(1.735)
Observations	111	111	111	217	217	217
Adj. R-squared	0.309	0.041	0.041	0.064	0.052	0.056

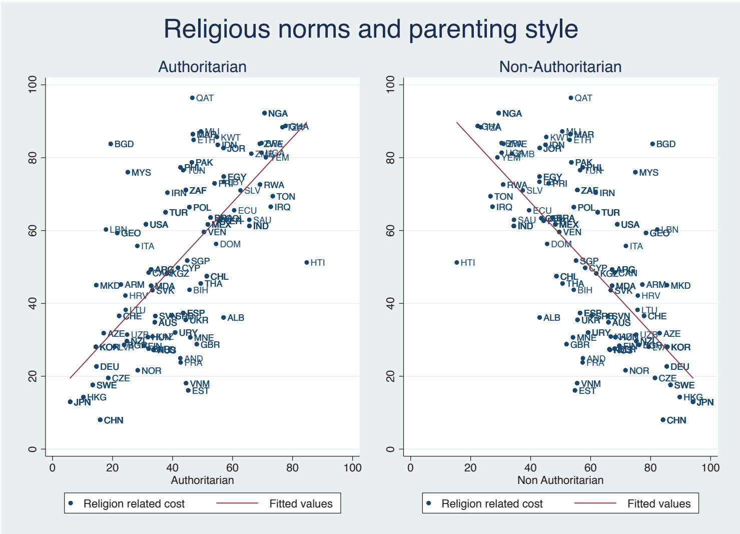
Table displays results from estimating univariate regressions $Education_{i,t} = \alpha + \beta Parenting_{i,t} + \varepsilon_{i,t}$ where i is the country and t is time (wave). *Parenting* is the fraction of respondents in the WVS that are authoritarian, authoritative or non-authoritative. *Women Education* is the percentage of population ages 25 and over that attained or completed upper secondary education. *Labour force participation* is the female labor force as a percentage of the total. Labor force comprises people ages 15 and older who supply labor for the production of goods and services during a specified period. Heteroskedasticity consistent standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.8: Parenting style and education

VARIABLES	(1) Bachelor	(2) Terziary education	(3) Expected schooling
Authoritarian	-0.201***	-0.357***	-0.0616***
	(0.0395)	(0.0833)	(0.0112)
Constant	23.34***	44.57***	15.68***
	(1.967)	(3.934)	(0.473)
Observations	115	108	167
Adj.R-squared	0.137	0.114	0.145

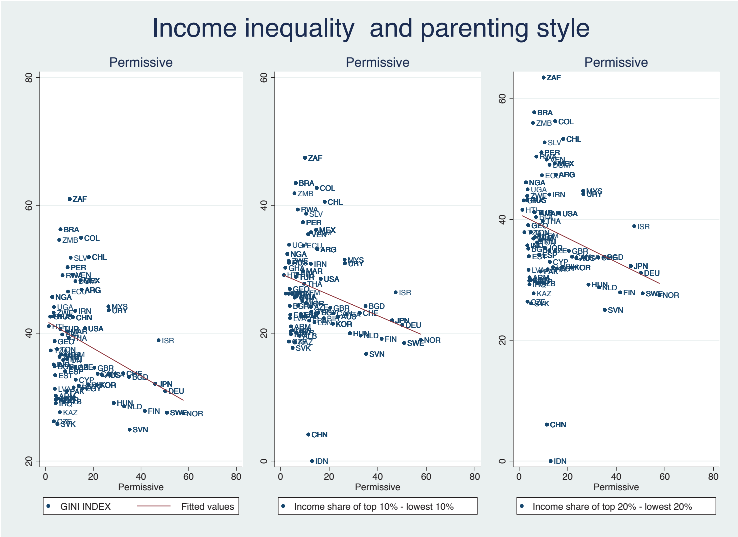
Table displays results from estimating univariate regressions $Education_{i,t} = \alpha + \beta Parenting_{i,t} + \varepsilon_{i,t}$, where i is the country and t is the time (wave). *Parenting* is the fraction of respondents in the WVS that are authoritative or non-authoritative. *Bachelor* percentage of population ages 25 and over that attained or completed Bachelor's or equivalent. *Terziary education* is number of graduates from first degree program expressed as a percentage of the population of the theoretical graduation age of the most common first degree program. *Expected Schooling* is the expected years of schooling. Heteroskedasticity consistent standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Figure 3.1: Religious norms and parenting style



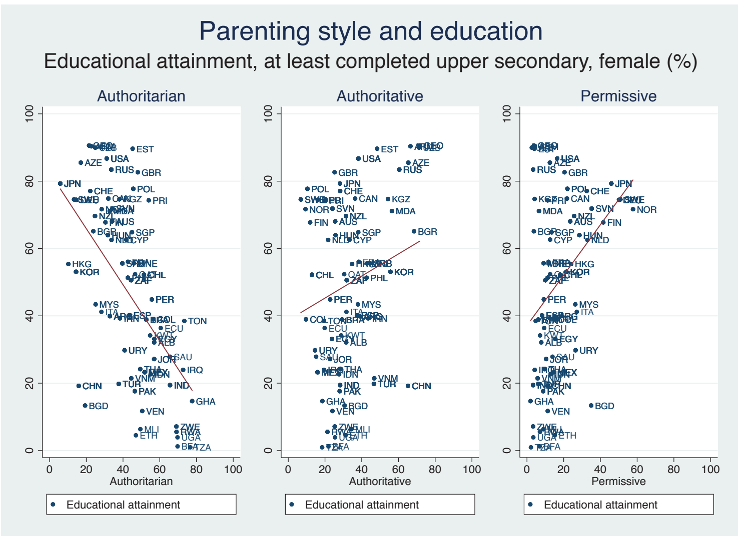
Note: Data are collapsed at country level. Figures display the fraction of respondents in the WVS that are authoritarian or non-authoritarian and the religious norms computed as described in Table 3.3. Non-authoritarian are those parents that are either authoritative or permissive.

Figure 3.2: Income inequality and parenting style



Note: Data are collapsed at country level. Figures display the fraction of respondents in the WVS that are permissive and income inequality indexes described in Table 3.2.

Figure 3.3: Parenting style and education



Note: Data are collapsed at country level. Figures display the fraction of respondents in the WVS that are authoritarian, authoritative and permissive and female education (Table 3.2).

Bibliography

Alesina, A., P. Giuliano, and N. Nunn (2013). On the Origins of Gender Roles: Women and the Plough. *The Quarterly Journal of Economics 128*(2), 469–530.

Angeletos, G.-M., F. Collard, and H. Dellas (2018). Quantifying confidence. *Econometrica 86*(5), 1689–1726.

Argente, D. and M. Lee (2015). Cost of living inequality during the great recession. *Nielsen Dataset Paper Series Working Paper*, Working Paper.

Armantier, O., S. Nelson, G. Topa, W. van der Klaauw, and B. Zafar (2016). The price is right: Updating inflation expectations in a randomized price information experiment. *The Review of Economics and Statistics 98*(3), 503–523.

Armantier, O., G. Topa, W. Van der Klaauw, and B. Zafar (2017). An overview of the Survey of Consumer Expectations. *Economic Policy Review* (23-2), 51–72.

Bachmann, R., T. O. Berg, and E. R. Sims (2015). Inflation expectations and readiness to spend: Cross-sectional evidence. *American Economic Journal: Economic Policy 7*(1), 1–35.

Barsky, R. B. and E. R. Sims (2012). Information, animal spirits, and the meaning of innovations in consumer confidence. *American Economic Review 102*(4), 1343–77.

Baumrind, D. (1967). Child care practices anteceding three patterns of preschool behaviour. *General Psychology Monograph 75*, 43–88.

Bertrand, M., P. Cortes, C. Olivetti, and J. Pan (2016). Social norms, labor market opportunities, and the marriage gap for skilled women. Working Paper 22015, National Bureau of Economic Research.

Beshears, J., J. J. Choi, A. Fuster, D. Laibson, and B. C. Madrian (2013). What goes up must come down? experimental evidence on intuitive forecasting. *The American Economic Review 103*(3), 570–574.

Bianchi, J. (2011). Overborrowing and systemic externalities in the business cycle. *American Economic Review 101*(7), 3400–3426.

Bordalo, P., N. Gennaioli, and A. Shleifer (2012). Salience theory of choice under risk. *Quarterly Journal of Economics 127*(3), 1243–1285.

Bordalo, P., N. Gennaioli, and A. Shleifer (2013). Salience and consumer choice. *Journal of Political Economy 121*(5), 803–843.

Bordalo, P., N. Gennaioli, and A. Shleifer (2018). Diagnostic expectations and credit cycles. *The Journal of Finance 73*(1), 199–227.

Boskin, M., E. Dulberger, R. Gordon, Z. Griliches, and D. Jorgenson (1996). Toward a more accurate measure of the cost of living. final report to the senate finance committee from the advisory commission to study the consumer price index. Technical report.

Boz, E. and E. G. Mendoza (2014). Financial innovation, the discovery of risk, and the U.S. credit crisis. *Journal of Monetary Economics 62*(C), 1–22.

Bruine de Bruin, W., W. van der Klaauw, G. Topa, J. S. Downs, B. Fischhoff, and O. Armantier (2012). The effect of question wording on consumersâ reported inflation expectations. *Journal of Economic Psychology 33*(4), 749 – 757.

Bruine de Bruin, W., W. Vanderklaauw, J. S. Downs, T. G. Fischhoff, B., and O. Armantier (2010). Expectations of inflation: The role of demographic variables, expectation formation, and financial literacy. *Journal of Consumer Affairs 44*(2), 381–402.

Burke, M. A. and M. Micheal (2014). Economic literacy and inflation expectations: Evidence from a laboratory experiment. *Journal of Money, Credit and Banking 46*(7), 1421–1456.

Carroll, C. D. (2003). Macroeconomic Expectations of Households and Professional Forecasters*. *The Quarterly Journal of Economics 118*(1), 269–298.

Cavallo, A., G. Cruces, and R. Perez-Truglia (2017). Inflation expectations, learning, and supermarket prices: Evidence from survey experiments. *American Economic Journal: Macroeconomics 9*(3).

Coibion, O. and Y. Gorodnichenko (2012a). What can survey forecasts tell us about information rigidities? *Journal of Political Economy 120*(1), 116–159.

Coibion, O. and Y. Gorodnichenko (2012b). What can survey forecasts tell us about information rigidities? *Journal of Political Economy 120*(1), 116–159.

Coibion, O. and Y. Gorodnichenko (2015a). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review 105*(8), 2644–78.

Coibion, O. and Y. Gorodnichenko (2015b). Is the phillips curve alive and well after all? inflation expectations and the missing disinflation. *American Economic Journal: Macroeconomics 7 (1)*, 197–232.

Coibion, O., Y. Gorodnichenko, S. Kumar, and M. Pedemonte (2018). Inflation expectations as a policy tool? Working Paper 24788, National Bureau of Economic Research.

Curtin, Richard, T. (1996). Procedure to estimate price expectations. Technical report, Surveys of Consumers.

Das, S., C. M. Kuhnen, and S. Nagel (2017). Socioeconomic status and macroeconomic expectations. Working Paper 24045, National Bureau of Economic Research.

Del Negro, M., M. Giannoni, and C. Patterson (2012). The forward guidance puzzle. *Federal Reserve Bank of New York Staff Reports No. 574*.

Doepke, M., S. Giuseppe, and F. Zilibotti (2019). The economics of parenting. Working Paper 25533, National Bureau of Economic Research.

Doepke, M. and F. Zilibotti (2017). Parenting with style: Altruism and paternalism in intergenerational preference transmission. *Econometrica 85*(5), 1331–1371.

Doepke, M. and F. Zilibotti (2018). *Love, Money, and Parenting*. Princeton University Press.

Duflo, E. (2012). Women empowerment and economic development. *Journal of Economic Literature 50*(4), 1051–79.

Eggertsson, G. B. and P. Krugman (2012). Debt, deleveraging, and the liquidity trap: A fisher-minsky-koo approach. *The Quarterly Journal of Economics 127*(3), 1469–1513.

Fernandez, R. (2007). Alfred marshall lecture women, work, and culture. *Journal of the European Economic Association 5*(2â3), 305–332.

Fernandez, R. (2013). Cultural change as learning: The evolution of female labor force participation over a century. *American Economic Review 103*(1), 472–500.

Fernandez, R., A. Fogli, and C. Olivetti (2004). Mothers and sons: Preference formation and female labor force dynamics. *The Quarterly Journal of Economics 119*(4), 1249–1299.

Fuster, A., B. Hebert, and D. Laibson (2012). Natural expectations, macroeconomic dynamics, and asset pricing. *NBER Macroeconomics Annual 26*(1), 1–48.

Fuster, A., D. Laibson, and B. Mendel (2010, December). Natural expectations and macroeconomic fluctuations. *Journal of Economic Perspectives 24*(4), 67–84.

Gennaioli, N. and A. Shleifer (2010). What comes to mind. *Quarterly Journal of Economics 125 (4)*, 1399–1433.

Glaeser, E. L. and Y. Ma (2014). *The Supply of Gender Stereotypes and Discriminatory Beliefs*, pp. 355–389. University of Chicago Press.

Glaeser, E. L. and C. G. Nathanson (2017). An extrapolative model of house price dynamics. *Journal of Financial Economics 126*(1), 147 – 170.

Greenwood, R., S. G. Hanson, and L. J. Jin (2016, August). A model of credit market sentiment. Working Paper 17-015, Harvard Business School.

Greenwood, R. and A. Shleifer (2014, 01). Expectations of Returns and Expected Returns. *The Review of Financial Studies 27*(3), 714–746.

Guerrieri, V. and G. Lorenzoni (2017). Credit crises, precautionary savings, and the liquidity trap. *The Quarterly Journal of Economics 132*(3), 1427–1467.

Guiso, L., F. Monte, P. Sapienza, and L. Zingales (2008). Culture, gender, and math. *Science 5880*(320), 1164–1165.

Guiso, L., P. Sapienza, and L. Zingales (2006). Does culture affect economic outcomes? *Journal of Economic Perspectives 20*(2), 23–48.

Inglehart, R., C. Haerpfer, A. Moreno, C. Welzel, K. Kizilova, J. Diez-Medrano, M. Lagos, P. Norris, E. Ponarin, B. Puranen, and et al (2014). World values survey: Round six - country-pooled datafile version. http://www.worldvaluessurvey.org/wvsdocumentationwv6.jsp. Technical report.

Jaravel, X. (2019). The unequal gains from product innovations: Evidence from the u.s. retail sector. *Quarterly Journal of Economics*, 1â69.

Johannsen, B. K. (2014). Inflation experience and inflation expectations: Dispersion and disagreement within demographic groups. *Working Paper*.

Jorda, O., M. Schularick, and A. M. Taylor (2013). When credit bites back. *Journal of Money, Credit and Banking 45*(s2), 3–28.

Justiniano, A., G. Primiceri, and A. Tambalotti (2015). Household leveraging and deleveraging. *Review of Economic Dynamics 18*(1), 3–20.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *The American Economic Review 93*(5), 1449–1475.

Kaplan, G. and G. Menzio (2015). The morphology of price dispersion. *International Economic Review 56*(4), 1165–1206.

Kaplan, G., K. Mitman, and G. L. Violante (2017). The housing boom and bust: Model meets evidence. Working Paper 23694, National Bureau of Economic Research.

Kaplan, G. and S. Schulhofer-Wohl (2017). Inflation at the household level. *Journal of Monetary Economics 19-38*, 19–38.

Kilian, L. (1998). Small-sample confidence intervals for impulse response functions. *The Review of Economics and Statistics 80*(2), 218–230.

Kuchler, T. and B. Zafar (2015). Personal experiences and expectations about aggregate outcomes. Staff Reports 748, National Bureau of Economic Research.

Landier, A., Y. Ma, and D. Thesmar (2017). New experimental evidence on expectations formation. Working paper, SSRN Electronic Journal. 10.2139/ssrn.3046955.

Lopez-Salido, D., J. C. Stein, and E. Zakrajsek (2017, 05). Credit-market sentiment and the business cycle. *The Quarterly Journal of Economics 132*(3), 1373–1426.

Lorenzoni, G. (2009). A theory of demand shocks. *American Economic Review 99*(5), 2050–84.

Madeira, C. and B. Zafar (2015). Heterogeneous inflation expectations and learning. *Journal of Money, Credit and Banking 47*(5), 867–896.

Malmendier, U. and S. Nagel (2015). Learning from inflation experiences. *The Quarterly Journal of Economics 131*(1), 53–87.

Mankiw, N. G. and R. Reis (2002). Sticky information versus sticky prices: A proposal to replace the new keynesian phillips curve. *The Quarterly Journal of Economics 117*(4), 1295–1328.

Mankiw, N. G., R. Reis, and J. Wolfers (2004). Disagreement about Inflation Expectations. In *NBER Macroeconomics Annual 2003, Volume 18*, NBER Chapters, pp. 209–270. National Bureau of Economic Research, Inc.

McCracken, M. W. and S. Ng (2015, June). Fred-md: A monthly database for macroeconomic research. Working papers, Federal Reserve Bank of St. Louis.

Mian, A., A. Sufi, and E. Verner (2017). Household debt and business cycles worldwide. *The Quarterly Journal of Economics 132*(4), 1755–1817.

Minsky, H. P. (1977). The financial instability hypothesis: An interpretation of keynes and an alternative to "standard" theory. *Nebraska Journal of Economics and Business 16*(1), 5–16.

Pancrazi, R. and M. Pietrunti (2014). Natural expectations and home equity extraction. Working Paper 984, Bank of Italy.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics 50*(3), 665 – 690.

Souleles, N. (2004). Expectations, heterogeneous forecast errors, and consumption: Micro evidence from the michigan consumer sentiment surveys. *Journal of Money, Credit and Banking 36*(1), 39–72.

Stock, J. H. and M. W. Watson (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association 97*(460), 1167–1179.

Tversky, A. and D. Kahneman (1974). Judgment under uncertainty: Heuristics and biases. *Science 185*(4157), 1124–1131.

Woodford, M. (2001). Imperfect common knowledge and the effects of monetary policy. In *Knowledge, Information and Expectations in Modern Macroeconomics, in Honour of Edmund S. Phelps*, pp. 25–58. Princeton University Press.

Appendix A

Chapter 1

A.1 Alternative beliefs' models

Adaptive expectations

Adaptive expectations are obtained by weighting past observations exponentially. They derive from the following model:

$$\widetilde{E_t^A}(\Delta y_{t+1}) = \phi \Sigma_{j=0}^{\infty} (1-\phi)^j \Delta y_{t-j} \tag{A.1}$$

With $0 \leq \phi \leq 1$. This is equivalent to:

$$\widetilde{E_t^A}(\Delta y_{t+1}) = \widetilde{E_{t-1}^A}(\Delta y_t) + \phi(\Delta y_t - \widetilde{E_{t-1}^A}(\Delta y_t)) \tag{A.2}$$

The agent adjusts his prior expectations according to the last observed error by a factor ϕ. This one is a parsimonious model with only one free-parameter. In this economy, the agent optimises knowing his preferences, constraints, beliefs, and the exogenous realised income. Differently, he does not know the true income process. Each quarter, the agent forms adaptive expectations on Δy_{t+j} for any $j \geq 1$, and in turn for the income process y_{t+j}, according to the equation (A.1). He expects the growth rate to be constant forever and consequently the income to grow over time such that for any $j \geq 1$:

$$\widetilde{E_t^A}(\Delta y_{t+1}) = \widetilde{E_t^A}(\Delta y_{t+j}) \tag{A.3}$$

Each period, based on the observables, the agent updates his short and long-term beliefs using these rules.

Match with the empirical evidence. Adaptive expectations are a particular case of extrapolative expectations coherently to Table 1.6 (Column (1)). Then, let Δy_t be an ARIMA(0,1,q) such that $\Delta y_{t+1} = \beta(L)\epsilon_t$; $\phi < \beta_1$ implies slow reaction to shocks as predicted by the empirical evidence in Table 1.6 (Column (2)). After a shock, expectations adjust gradually and forecast revision at time t predicts the revision at time $t+1$ as in Table 1.6 (Column (3)). Given the time series for the US RGDP growth rate, I compute the adaptive expectations for several values of the parameter ϕ. I set the parameter ϕ equal to 0.33, to minimize the absolute difference between the computed adaptive expectations and the SPF data. The parameter ϕ defines how far the agent extrapolates into the past; a low value for ϕ implies that the agent assigns a smaller weight to new information relative to past expectations.

Impulse responses. Figure A.1 shows the term structure for the expectations on y_{t+j} and Δy_{t+j} to a shock to Δy_{t+j}. The term structure reports the long term expectations $\widetilde{E_1^A}(\Delta y_{t+j})$ and $\widetilde{E_1^A}(y_{t+j})$ computed at time $t = 1$ for any $j \geq 1$. Rational expectations coincide with the true data generating process. Following a positive shock, the adaptive agent expects y_t to grow at a positive constant rate, although small; the forecast error, is negative and gets larger (in absolute value) with the forecasting horizon j. Any quarter, given the realized Δy_{t+j}, expectations are updated and eventually converge to the rational ones. Until reaching convergence, the agent substantially fails to forecast long-run income and these over-optimistic beliefs drive his debt choices.

Figure A.4 displays the impulse responses following a one standard deviation shock to Δy_t. A positive shock to the fundamentals implies an immediate jump in consumption and debt since the consumer expects his income to growth permanently at a positive growth rate and strongly over-predict the long run income. Each period, the agent gets disappointed and revises his forecasts towards rational ones; accordingly, consumption declines over time. Debt increases as long as the consumer expects a positive increase in income; when expectations get pessimistic with $\widetilde{E_t^A}(\Delta y_{t+j}) < 0$, debt starts decreasing. The level of debt remains high as it is needed to finance consumption and to pay back the previous debt. Over the long run, a positive shock leads to a permanent increase in debt and to a consumption level inferior to the rational benchmark since the agent needs to pay back the initial extra consumption. Further, the model generates a credit boom disconnected from the true dynamics of the fundamentals.

Sticky information

In this section, I consider a model of inattentive agents as proposed by Mankiw and Reis (2002). Let the economy be populated by a continuum of identical agents. The agents update their information sets each period with probability $(1 - \lambda)$ and acquire no new information with probability λ. Agents indeed may rationally choose not to revise every period when they face a fixed cost to updating their information (Reis, 2006). $1/(1 - \lambda)$ is the average duration of information updates, and it is a measure of information rigidity. When agents update their information sets, they acquire full-information and have rational expectations. At time t, the average expectations across agents is a weighted average of current and past full-information rational expectations:

$$\widetilde{E_t^S}(\Delta y_{t+k}) = (1 - \lambda)\Sigma_{j=0}^{\infty}\lambda^j E_{t-j}(\Delta y_{t+k}) \tag{A.4}$$

This is equivalent to:

$$\widetilde{E_t^S}(\Delta y_{t+k}) = (1 - \lambda)E_t(\Delta y_{t+k}) + \lambda\widetilde{E_{t-1}^S}(\Delta y_{t+k}) \tag{A.5}$$

Where $E(.)$ is the rational expectations operator. As in Reis (2006), I consider agents to be *consumption planners*. This means that they fix their level of consumption and let their savings adjust one to one with their income when they don't update. After a positive shock, a fraction $(1-\lambda)$ of agents updates the beliefs rationally and chooses consumption and savings coherently. A fraction λ instead doesn't update, nor change the consumption plan; this portion of agents receives a larger income and therefore let savings increase. Every quarter, a fraction $(1-\lambda)$ of consumers updates and chooses consumption rationally. Overall, consumption is a weighted average of current and past rational consumption plans and debt changes accordingly.[1]

Match with the empirical evidence. As proved by Coibion and Gorodnichenko (2015), there is a perfect map between the coefficient β_2 obtained from regressing the forecast error on the forecast revision (Column (2), Table 1.6) and the parameter λ. A positive value of β_2 is coherent with the hypothesis of sticky information and $\beta_2 = \frac{\lambda}{(1-\lambda)}$. Therefore, I set $\lambda = 0.438$.

Impulse responses. Figure A.2 shows the term structure for the expectations on y_{t+j} and Δy_{t+j} to a shock to Δy_{t+j}. The term structure reports the long term expectations $\widetilde{E_1^S}(\Delta y_{t+j})$ and $\widetilde{E_1^S}(y_{t+j})$ computed at time t=1 for any $j \geq 1$. Rational expectations coincide with the true data generating process. Following a positive shock to Δy_t, only a fraction of agents immediately updates rationally. Hence, the average forecast made at $t = 1$ on the long run captures the correct path but is below the realized value; while the average forecast error is positive for all horizons (Panel A). Every quarter, a positive fraction of agents updates the beliefs, so the average forecast increases and gets closer to the rational one. Eventually expectations convergence, but in the medium run, biased expectations drive the demand for debt.

Figure A.5 displays the impulse responses following a one standard deviation shock to Δy_t. After a positive shock, consumption and savings gradually increase. A portion $(1-\lambda)$ of consumers correctly anticipate the income dynamics and rationally increases savings and consumption soon after the shock. Savings increases more than rationally because some agents do not revise the consumption plan immediately and over-save the extra-income. Each quarter, more agents revise their beliefs and consumption plans, the average forecast slowly moves upward and consumption gradually increases. Overall, the model yields to a level of consumption and savings larger than under rational expectations, because consumers over-save in the first few quarters and exploit their savings to finance the extra consumption afterwards.

[1] Aggregate consumption is given by $c_t^S = (1-\lambda)c_t^{rat} + \lambda c_{t-1}^S$; while debt can be derived from the budget constraint.

Noisy signals

In this section, I consider a model with noisy signals as in Sims (2003). The representative agent continuously receives information on Δy_t, but observes only a noisy signal s_t such that:[2]

$$\Delta y_t = \beta(L)\epsilon_t, \epsilon_t \sim N(0, \sigma^2_\epsilon) \tag{A.6}$$

and:

$$s_t = \Delta y_t + \eta_t, \quad \eta_t \sim N(0, \sigma^2_\eta) \tag{A.7}$$

The noise term η_t generates an independent source of variation in the expectations and prevents from correctly identifying the permanent component ϵ_t. After observing an increase of s_t, the agent needs time to learn about the nature of the change. Both shocks indeed imply an immediate increase in s_t: however, the noise shock generates only a one-period change in the signal, while ϵ_t has a permanent effect on the income in level. Since the agent does not fully observe the true state (Δy_t), he needs to solve a signal extraction problem via Kalman filtering. Thus the forecasts are a weighted average of agent's prior beliefs and an error correction term:

$$\widetilde{E^N_t}(\Delta y_t) = \widetilde{E^N_{t-1}}(\Delta y_t) + G(s_t - \widetilde{E^N_{t-1}}(\Delta y_t)) \tag{A.8}$$

Where G is the Kalman gain and depends on the signal to noise ratio. Except for the information frictions, the agent is fully rational and computes $E^N_t(\Delta y_{t+j})$ for any $j > 1$ according to the true DGP. [3] The presence of noisy signals generates expectational errors, both when ϵ_t and η_t arrive. The consumer has only limited information regarding the long-term income and these expectations drive the demand for debt.

Match with the empirical evidence. As shown by Coibion and Gorodnichenko (2015), there is a perfect map between the coefficient β_2 obtained from regressing the forecast error on the forecast revision, and the Kalman gain G (Column (2), Table 1.6). A positive value for the parameter β_2 is coherent to the hypothesis of noisy signals and $\beta_2 = \frac{1-G}{G}$. The Kalman gain depends on the signal to noise ratio s.t. $G = \frac{\sigma_{\Delta y}}{\sigma_{\Delta y} + \sigma_\eta}$. Without information frictions G is equal to 1; $G \leq 1$ instead implies information frictions and $(1 - G)$ can be interpreted as the degree of information rigidity. I calibrate the model using the estimated coefficient and the estimated variance for Δy_t. Given these two values, I recover the variance of the noise shock and set $\sigma^2_\eta = 0.0063^2$ (and $\sigma^2_{\Delta y} = 0.008^2$).

[2]The representative agent observes a change in income equal to s_t but doesn't know whether this is equal to the true Δy_t, or not.

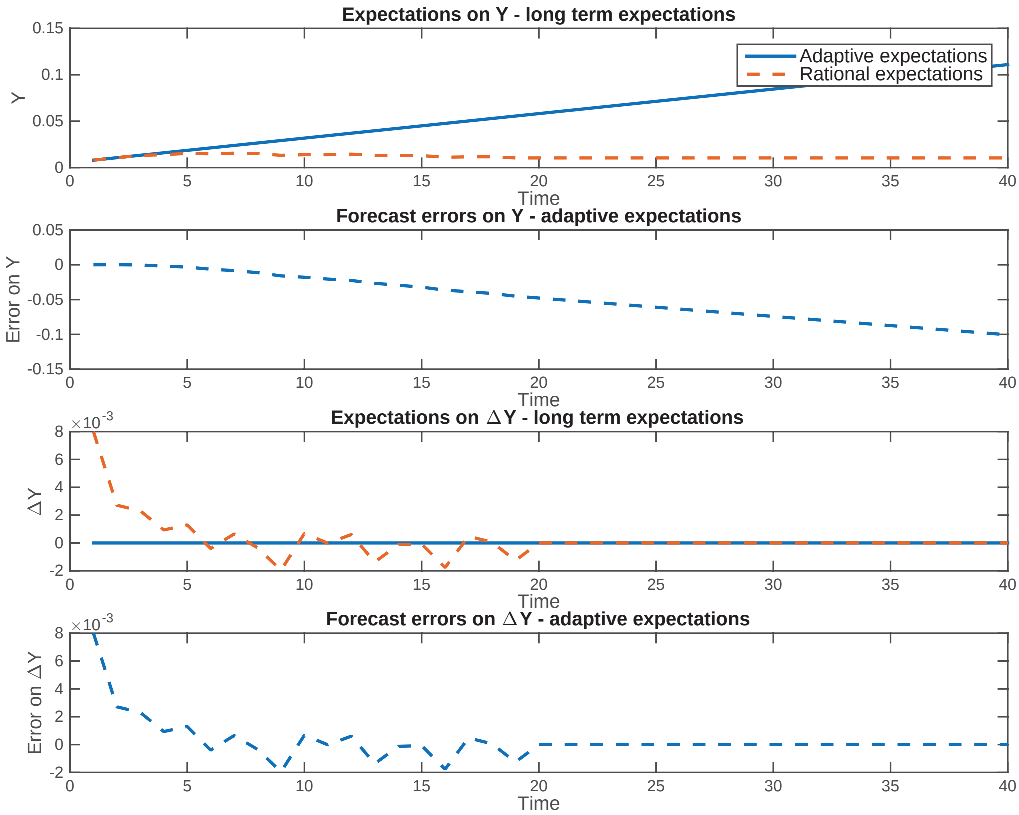
[3]I solve the signal extraction problem, considering a state space representation: (1) $\alpha_t = T\alpha_{t-1} + e_t$ (state equation), and (2) $y_t = Z\alpha_t + \zeta_t$ (measurement equation). α_t is a vector of unobservable state variable and y_t a vector of observables. The error terms have zero mean, are jointly normally distributed and uncorrelated s.t. e_t iid N(0,H) and ζ_t iid N(0,S).

Impulse responses. Figure A.3 shows the term structure for the expectations on y_{t+j} and Δy_{t+j} to a shock to Δy_{t+j}. The term structure reports the long-term expectations $\widetilde{E_1^N}(\Delta y_{t+j})$ and $\widetilde{E_1^N}(y_{t+j})$ computed at time $t = 1$ for any $j \geq 1$. Rational expectations coincide with the true data generating process. Following a positive shock to the fundamental, the representative agent updates his beliefs filtering out the noise and knowing the true DGP. The forecasts rationally capture the hump-shaped pattern but fail to match the long run income in level and generate a positive forecast error on long-term expectations both for y_{t+j} (Panel A) and Δy_{t+j} (Panel B) for any $j \geq 1$. Every quarter, the forecasts move in the direction of the detected error; slow adjustment implies that the forecasts are too low after a positive change, while they are too high after a negative change. Expectations get very close to the rational ones, but convergence occurs after more than 40 quarters. The speed of convergence depends on the signal to noise ratio $\frac{\sigma_{\Delta y}}{\sigma_\eta}$; convergence is achieved faster when the ratio reduces as the forecasts are less sensitive to the observed forecast error. With a sufficiently low ratio, the forecasts instead converge once Δy_{t+j} reaches zero. Overall, in the short run, the agent fails to predict the long-term income, and these biased beliefs affect his consumption/savings choices.

Figure A.6 displays the impulse responses following a one standard deviation shock to Δy_t. After a positive shock to the fundamentals, consumption and savings gradually increase. Initially, the agent foresees a permanent increase in income and hence raises consumption. Consumption grows less than rationally because the agent underestimates the long run income; further, he anticipates the hump-shaped dynamics and rationally increases savings.[4] Each quarter, expectations are revised in the direction of the observed error. Consumption gradually increases towards the rational level as the forecasts move towards the rational benchmark and savings rationally increase. Since expectations are very sensitive to the detected errors, consumption frequently changes over time. After a positive Δy_t, the agent revises his forecast upward on future income and increases consumption; on the contrary, after a negative Δy_t consumption decreases.The learning process and the impulse for consumption depend strongly on the signal to noise ratio $\frac{\sigma_{\Delta y}}{\sigma_\eta}$. However, following a positive shock to the income growth rate, the model always predicts an increase in savings, and never a growth in debt.

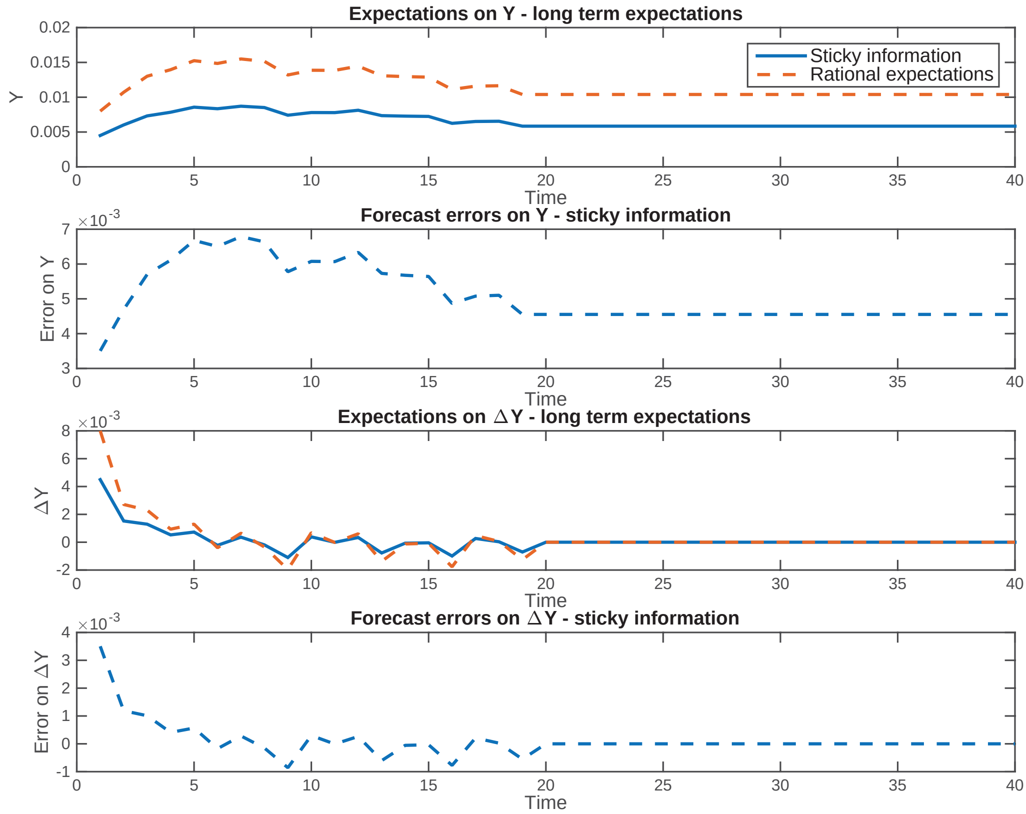
[4]In the first few quarters, debt increases but less than rationally. From t=3, debt starts decreasing and follows a path similar to the rational case.

Figure A.1: Beliefs term structure - Adaptive expectations



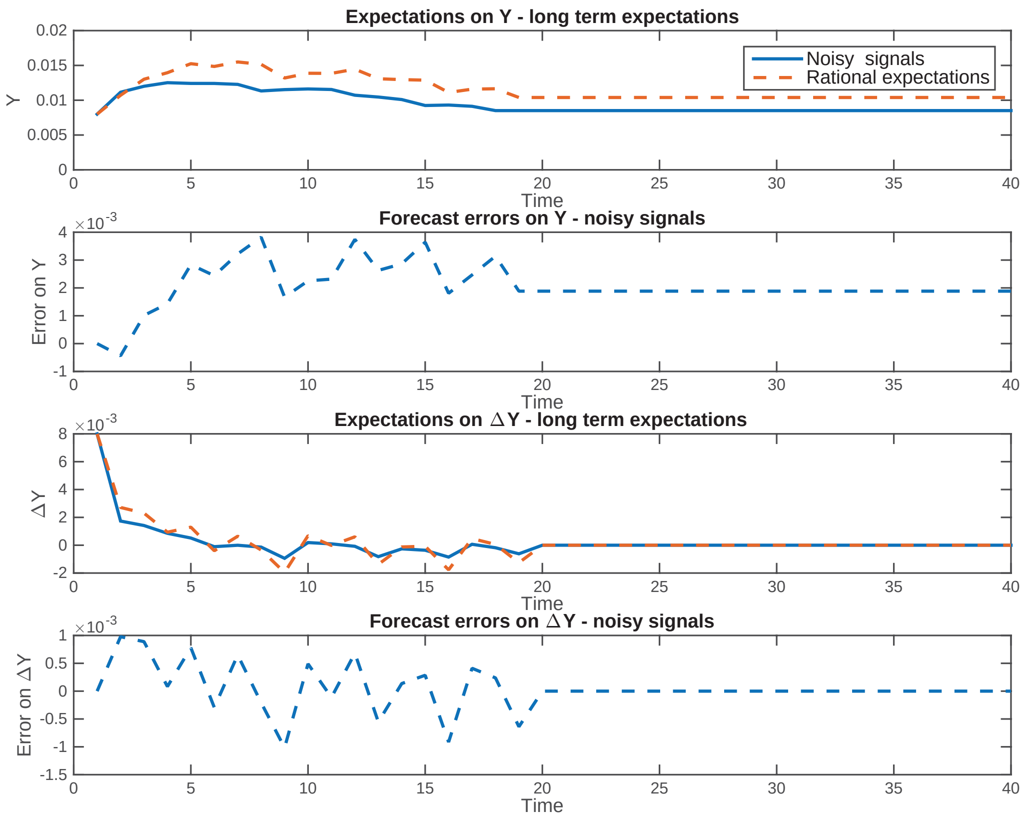
Note: Beliefs term structure (in percentage change) following one standard deviation shock to Δy_t. Dashed orange lines are the rational expectations case. Solid blue lines the adaptive expectations, with $\phi = 0.33$. The upper panel displays $\widetilde{E^A_{t=1}}(y_{t+j})$ and the relative forecast error $(y_{t+j} - \widetilde{E^A_{t=1}}(y_{t+j}))$ for $j \geq 1$. The lower panel displays $\widetilde{E^A_{t=1}}(\Delta y_{t+j})$ and the relative forecast error $(\Delta y_{t+j} - \widetilde{E^A_{t=1}}(\Delta y_{t+j}))$ for $j \geq 1$. Time on the horizontal axis is the forecast horizon in quarters.

Figure A.2: Beliefs term structure - Sticky information



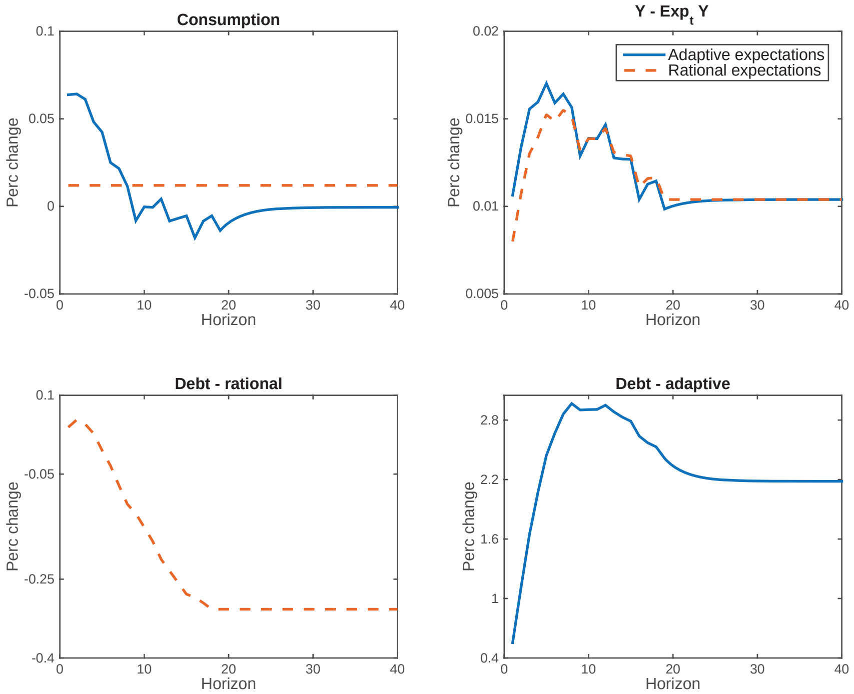
Note: Beliefs term structure (in percentage change) following one standard deviation shock to Δy_t. Dashed orange lines are the rational expectations case. Solid blue lines the sticky information case, with $\lambda = 0.438$. The upper panel displays $\widetilde{E^S_{t=1}}(y_{t+j})$ and the relative forecast error $(y_{t+j} - \widetilde{E^S_{t=1}}(y_{t+j}))$ for $j \geq 1$. The lower panel displays $\widetilde{E^S_{t=1}}(\Delta y_{t+j})$ and the relative forecast error $(\Delta y_{t+j} - \widetilde{E^S_{t=1}}(\Delta y_{t+j}))$ for $j \geq 1$. Time on the horizontal axis is the forecast horizon in quarters.

Figure A.3: Beliefs term structure - Noisy signals



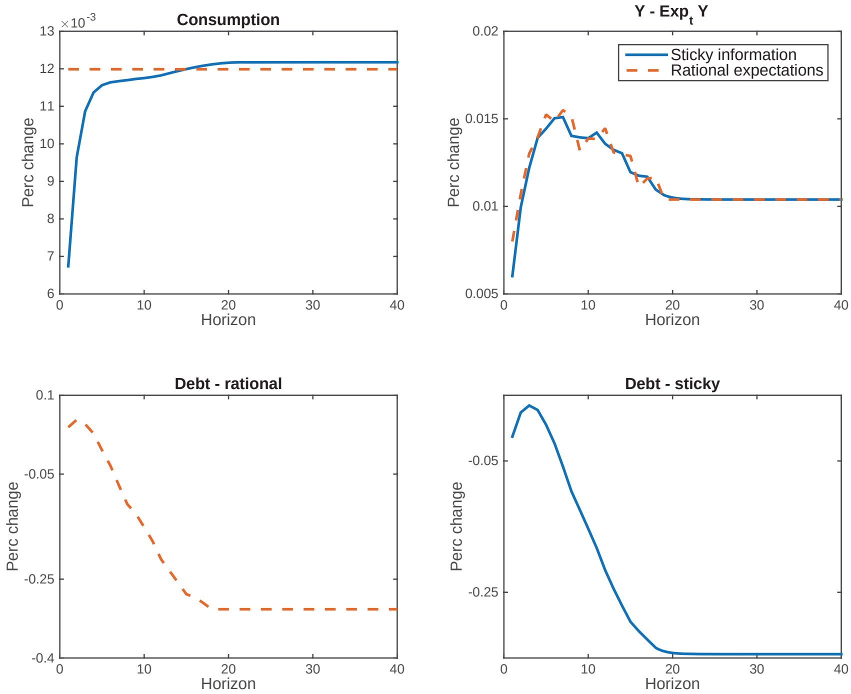
Note: Beliefs term structure (in percentage change) following one standard deviation shock to Δy_t. Dashed orange lines are the rational expectations case. Solid blue lines the noisy signal case with $\sigma_\eta = 0.00623$. The upper panel displays $\widetilde{E^N_{t=1}}(y_{t+j})$ and the relative forecast error $(y_{t+j} - \widetilde{E^N_{t=1}}(y_{t+j}))$ for $j \geq 1$. The lower panel displays $\widetilde{E^N_{t=1}}(\Delta y_{t+j})$ and the relative forecast error $(\Delta y_{t+j} - \widetilde{E^N_{t=1}}(\Delta y_{t+j}))$ for $j \geq 1$. Time on the horizontal axis is the forecast horizon in quarters.

Figure A.4: Impulse responses - Adaptive expectations



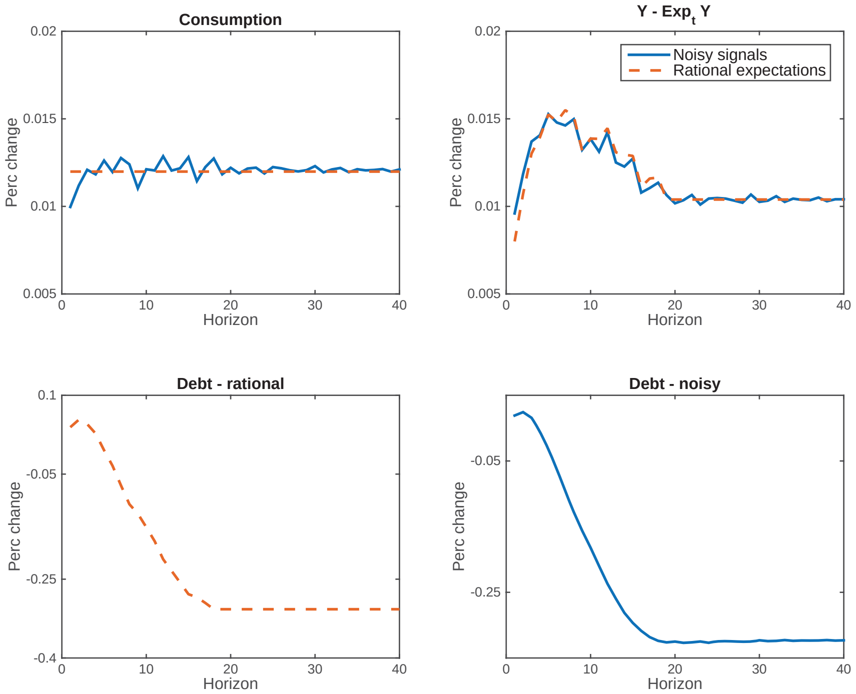
Note: Impulse responses following one standard deviation shock to Δy_t. Dashed lines are the rational expectations case. Solid blue lines the adaptive expectations, with $\phi = 0.33$. In the lower panels, expectations are $\widetilde{E_t^A}(\Delta y_{t+1})$ and $\widetilde{E_t^A}(y_{t+1})$ for $t \geq 1$, respectively.

Figure A.5: Impulse responses - Sticky information



Note: Impulse responses following one standard deviation shock to Δy_t. Dashed orange lines are the rational expectations case. Solid blue lines the sticky information case, with $\lambda = 0.438$. In the lower panels, expectations are $\widetilde{E_t^S}(\Delta y_{t+1})$ and $\widetilde{E_t^S}(y_{t+1})$ for $t \geq 1$, respectively.

Figure A.6: Impulse responses - Noisy signals



Note: Impulse responses following one standard deviation shock to Δy_t. Dashed orange lines are the rational expectations case. Solid blue lines the noisy signal case with $\sigma_\eta = 0.00623$. In the lower panels, expectations are $\widetilde{E_t^N}(\Delta y_{t+1})$ and $\widetilde{E_t^N}(y_{t+1})$ for $t \geq 1$, respectively.

A.2 Robustness checks

Do beliefs on future output help predict changes in household debt in the short run and is it robust to different specifications, countries, and measures of expectations? Do few optimistic agents drive this relationship?

Testing the role of heterogeneity

In this section, I test if a small fraction of optimistic agents drives the relationship observed in Table 1.4. If this is the case, heterogeneity in expectations may be a relevant factor to take into account when studying the household credit cycle. To test this, hypothesis, I use micro-data on forecasters' expectations and check if the most optimistic agents (in the top 5/10/20.. percentiles) are driving the positive relationship observed in Table 1.4. Table A.1 shows that this is not the case. Panel B reports the results of the forecasting regression estimated considering the average expectations of the most 5% pessimistic agents (Column 1), the 10% (Column 2) and so on. The relationship is positive and significant also when considering only the most pessimistic agents; thus, the positive correlation is not exclusively due to a small fraction of optimistic agents.

Michigan expectations

The empirical findings are robust when using consumers' expectations by the Michigan Survey (Table A.2). In Columns (1) and (4), the explanatory variable is the average percentage of respondents who, in a given quarter, believe that there will be better times in the next year. Similarly, I estimate the equation considering the average percentage of respondents who, in a given quarter, believe that there will be higher unemployment (Columns (5) and (8)) over the next year. Results are in line with previous findings. Householdsâ expectations are relevant predictors of credit growth rate: debt increases when agents foresee good times, and it decreases when consumers expect higher unemployment rate. On the contrary, savings contract when consumers are optimistic and rise otherwise. Consumers' expectations on future business conditions alone explain 25% of the variability of debt and the increase obtained from augmenting the baseline regression with the Michigan expectations is greater 12% of the variation in next year debt. Further, when controlling for the common factors (Table A.2, Panel B), the coefficient of interest β_1 remains positive and significant.

Cross-country analysis

The work takes a broader view and test whether results are robust across countries. I use an international survey of professional forecasters by the Consensus Economics, for the period from 2001: Q1 to 2016: Q1, which reports forecast of the real GDP (RGDP) growth rate over the next year for 11 countries: USA, UK, Italy, Germany, Norway, Netherlands, Swiss, Spain, France, Japan, Sweden and Canada. Aggregate expectations are the average across forecasters' responses.[5] Besides, I use data on household debt provided by the BIS. These are quarterly break-adjusted series and capture the outstanding amount of credit to households and non-profit institutions serving households at the end of reference quarter; credit is provided by domestic banks, other institutions and non-residents; it covers the core debt, defined as loans, debt securities and currency and deposits. Results are robust across countries: the coefficient of interest is always positive and significant, even when adding time and country fixed effect and additional controls (Table A.3, Panel A). To avoid Nickell bias, I do not include lags of the dependent variables. Time fixed effects are highly significant suggesting that there is a common time component driving the credit cycle across countries. This result is consistent with the recent empirical evidence on household debt and business cycles word-wide (Mian et al. (2017)). A one standard deviation increase in expectations leads to a 0.48 standard deviation increase in household debt (2.56 percentage points), with the other variables held constant and without fixed effect; results are very similar with time fixed effect. Adding country fixed effect, debt increases by 0.31 standard deviation (1.67 percentage points) and results are very similar including both time and fixed effect. Controlling for the realized RGDP growth rate the effect is reduced, but it is still positive and significant: a one standard deviation increase in expectations leads to a 0.28 standard deviation increase in household debt (1.38 percentage points), with the other variables held constant and without fixed effect (Panel B).

Is it robust across time?

I run the same set of regressions on a sub-sample, excluding post-2000 data. Results are consistent, the coefficient of interest is reduced but remains positive and significant suggesting that the Great Recession is not driving the previous result (Table A.4).

[5] I only have a short time series of these data; however, it is still relevant to look at it since world-wide most of the variation in the growth rate of household debt took place in the last 20 years. The growth rate of household debt follows a similar pattern for most of the advanced economies over this 8 years in terms of household growth rate (and as % of GDP); Japan, Swiss and Germany instead follow a different path.

Table A.1: Household Credit and expected RGDP growth by percentiles

Panel A: Household Credit and expected RGDP growth by optimism

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Household Credit				
$E_t(RGDPgr_{t,t+4})$_p5_opt	1.781***								
	(0.252)								
$E_t(RGDPgr_{t,t+4})$_p10_opt		1.780***							
		(0.254)							
$E_t(RGDPgr_{t,t+4})$_p20_opt			1.796***						
			(0.256)						
$E_t(RGDPgr_{t,t+4})$_p25_opt				1.792***					
				(0.257)					
$E_t(RGDPgr_{t,t+4})$_p50_opt					1.803***				
					(0.270)				
$E_t(RGDPgr_{t,t+4})$_p75_opt						1.819***			
						(0.294)			
$E_t(RGDPgr_{t,t+4})$_p80_opt							1.756***		
							(0.307)		
$E_t(RGDPgr_{t,t+4})$_p90_opt								1.559***	
								(0.333)	
$E_t(RGDPgr_{t,t+4})$_p95_opt									1.966***
									(0.326)
Constant	-0.218	-0.278	-0.293	-0.303	-0.568	-0.759	-0.795	-0.570	1.624
	(4.961)	(4.969)	(4.953)	(4.960)	(5.006)	(5.059)	(5.055)	(4.998)	(4.800)
Baseline controls	YES	YES	YES	YES	YES	YES	YES	YES	YES
R^2	0.725	0.724	0.724	0.723	0.718	0.710	0.704	0.684	0.726

Panel B: Household Credit and expected RGDP growth by pessimism

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Household Credit				
$E_t(RGDPgr_{t,t+4})$_p5	1.288***								
	(0.241)								
$E_t(RGDPgr_{t,t+4})$_p10		1.305***							
		(0.229)							
$E_t(RGDPgr_{t,t+4})$_p20			1.485***						
			(0.237)						
$E_t(RGDPgr_{t,t+4})$_p25				1.529***					
				(0.238)					
$E_t(RGDPgr_{t,t+4})$_p50					1.641***				
					(0.238)				
$E_t(RGDPgr_{t,t+4})$_p75						1.691***			
						(0.243)			
$E_t(RGDPgr_{t,t+4})$_p80							1.693***		
							(0.243)		
$E_t(RGDPgr_{t,t+4})$_p90								1.714***	
								(0.247)	
$E_t(RGDPgr_{t,t+4})$_p95									1.722***
									(0.248)
Constant	0.681	0.300	0.145	0.000870	0.120	-0.0199	-0.0375	-0.0845	-0.155
	(5.460)	(5.067)	(5.016)	(5.017)	(4.952)	(4.943)	(4.952)	(4.957)	(4.970)
Baseline controls	YES	YES	YES	YES	YES	YES	YES	YES	YES
R^2	0.703	0.703	0.718	0.719	0.725	0.726	0.726	0.725	0.724

Note: Results from estimating $D_{t,t+4} = \alpha + \beta E_t(RGDP_{t,t+4}) + \delta(L)Z_t + \gamma(L)D_{t,t+4} + u_{t,t+4}$. In the upper panel, $E_t(RGDPgr_{t.t+4})$_p5_opt indicates the average expectations in the top 5% of optimist agents and so on for the other percentiles. Otherwise, in the lower panel $E_t(RGDPgr_{t.t+4})$_p5 indicates the average expectations in the top 5% of pessimist agents and so on for the other percentiles. The percentiles are computed for each quarter. $D_{t,t+4}$ is the percentage change between time t and t+4 of the Household credit. All regressions include two lags of the dependent variable. Baseline controls are: inflation, 10YTBond, 3MTBill, GDP, and unemployment rate. Newey-west standard errors are reported in parentheses with 4 lags. *** p<0.01, ** p<0.05, * p<0.1.

Table A.2: Household Credit and Michigan Survey of Consumers' expectations

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	H. Credit	Savings	Consumption	H. Credit	Savings	Consumption
Exp Change in Business	0.124***	-0.458*	0.0297			
Conditions in 1 Year	(0.0319)	(0.268)	(0.0201)			
Exp increase in Unempl.				-0.165***	0.508**	-0.0744***
during the Next Year				(0.0323)	(0.224)	(0.0243)
Constant	-9.486	23.20	-0.0561	10.67**	-43.50	6.717**
	(6.722)	(31.79)	(2.835)	(5.106)	(36.29)	(2.868)
Baseline controls	YES	YES	YES	YES	YES	YES
Observations	186	186	186	186	186	186
R^2	0.670	0.208	0.726	0.708	0.213	0.788

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	H. Credit	Savings	Consumption	H. Credit	Savings	Consumption
Exp Change in Business	0.0828***	-0.218	0.0167			
Conditions in 1 Year	(0.0290)	(0.207)	(0.0129)			
Exp increase in Unempl.				-0.0975***	0.239	-0.0331*
during the Next Year				(0.0308)	(0.179)	(0.0195)
Constant	-16.71***	52.01	-1.514	-2.266	18.40	3.269
	(4.952)	(36.35)	(3.806)	(4.760)	(34.02)	(4.040)
Factors_t	YES	YES	YES	YES	YES	YES
Observations	185	185	185	185	185	185
R^2	0.789	0.445	0.592	0.789	0.439	0.858

Note: Table reports results from estimating $D_{t,t+4} = \alpha + \beta Exp_t + \delta(L)Z_t + \gamma(L)D_{t,t+4} + u_{t,t+4}$. Exp_t is the Exp Change in Business in 1 Year or the Exp increase in Unemployment during the next year. The former is the relative share of agents that expect good times in 1 year, the latter the fraction of respondents who expect an increase in unemployment during the next year, in a given quarter, in the Michigan Survey of Consumers survey. $D_{t,t+4}$ is the percentage change between time t and t+4 of the dependent variable, which can be Household credit, Mortgages, Personal savings, and Personal consumption expenditures. All regressions include two lags of the dependent variable. Baseline controls are: Consumer Price Index, 10YTBond, 3MTBill, GDP, and unemployment rate. Newey-west standard errors are reported in parentheses with 4 lags. *** p<0.01, ** p<0.05, * p<0.1.

Table A.3: Household Credit and Consensus Economics

Variables	(1) H. Credit	(2) H. Credit	(3) H. Credit	(4) H.Credit
$E_t(RGDPgr_{t,t+4})$	3.019***	1.967***	3.248***	1.903***
	(0.268)	(0.234)	(0.222)	(0.191)
inflation$_t$	0.304	-0.552***	0.788***	0.0468
	(0.221)	(0.152)	(0.196)	(0.121)
rgdp$_t$	-0.0350	0.0874	0.0549	0.00272
	(0.0718)	(0.0680)	(0.0790)	(0.0785)
3M interbank rate$_t$	-0.0841	-0.199	0.586***	0.593***
	(0.228)	(0.213)	(0.165)	(0.174)
10Y Gov bond$_t$	1.249***	0.835***	0.701***	-0.627**
	(0.253)	(0.225)	(0.190)	(0.245)
unrate$_t$	-0.228***	-1.110***	-0.157***	-0.828***
	(0.030)	(0.059)	(0.026)	(0.054)
Time fixed effect	NO	NO	YES	YES
Country fixed effect	NO	YES	NO	YES
Observations	658	658	658	658
R^2	0.486	0.732	0.645	0.854

Variables	(1) H. Credit	(2) H. Credit	(3) H. Credit	(4) H.Credit
$E_t(RGDPgr_{t,t+4})$	2.132***	1.267***	2.952***	1.708***
	(0.249)	(0.209)	(0.217)	(0.205)
$RGDP\_gr_{t,t+4}$	0.593***	0.609***	0.239***	0.236***
	(0.0666)	(0.0704)	(0.0599)	(0.0478)
inflation$_t$	0.742***	0.00246	0.881***	0.120
	(0.220)	(0.128)	(0.214)	(0.127)
$RGDP_t$	-0.0108	0.0305	0.0628	-0.0437
	(0.0629)	(0.0544)	(0.0832)	(0.0799)
3M interbank rate$_t$	0.308*	0.378**	0.558***	0.736***
	(0.179)	(0.185)	(0.154)	(0.169)
10Y Gov bond$_t$	0.789***	0.299	0.633***	-0.640***
	(0.192)	(0.184)	(0.177)	(0.228)
$Unrate_t$	-0.169***	-1.006***	-0.139***	-0.833***
	(0.0349)	(0.0579)	(0.0310)	(0.0572)
Time fixed effect	NO	NO	YES	YES
Country fixed effect	NO	YES	NO	YES
Observations	641	641	641	641
R^2	0.552	0.784	0.652	0.859

Note: Table reports results from estimating $D_{i,t,t+4} = \alpha + \beta E_{i,t}(RGDPgr_{i,t,t+4}) + \delta(L)Z_{i,t} + u_{i,t,t+4}$. $E_{i,t}(RGDPgr_{i,t,t+4})$ is the SPF average forecast of Real GDP growth rate (RGDP gr) over the next four quarters in the given country i. $D_{i,t,t+4}$ is the percentage change between time t and t+4 of the Household credit. Standard errors in parenthesis are country level. *** p<0.01, ** p<0.05, * p<0.1. Few observations are missing in the SPF data.

Table A.4: Household credit and expected RGDP, excluding post- 2000 data

Panel A: Expected RGDP gr and unemployment rate

Variables	(1) H. Credit	(2) Savings	(3) Cons.	(4) H. Credit	(5) H. Credit	(6) H. Credit	(7) H. Credit	(8) H. Credit
$\mathrm{E}_t(RGDPgr_{t,t+4})$	1.884***	-0.914	0.529***	1.148***	1.223***	0.913***		
	(0.286)	(1.569)	(0.191)	(0.280)	(0.314)	(0.271)		
$\mathrm{E}_t(Unrate_{t,t+4})$							-1.874	-1.731
							(1.408)	(1.208)
Constant	-1.069	9.897*	0.924	29.18***	-14.34	21.00*	-13.11	20.69*
	(1.429)	(5.307)	(0.928)	(6.785)	(10.04)	(12.51)	(9.506)	(12.31)
Baseline controls$_t$	NO	NO	NO	NO	YES	YES	YES	YES
Factors$_t$	NO	NO	NO	YES	NO	YES	NO	YES
Observations	116	116	116	116	116	116	116	116

Panel B: Expected and realized RGDP gr and unemployment rate

Variables	(1) H. Credit	(2) H. Credit	(3) H. Credit	(4) H. Credit	(5) H. Credit	(6) H. Credit	(7) H. Credit	(8) H. Credit
$\mathrm{E}_t(RGDPgr_{t,t+4})$	1.147***	0.875**	0.881***	1.004***	0.917***			
	(0.337)	(0.401)	(0.226)	(0.214)	(0.232)			
$RGDPgr_{t,t+4}$	-0.0853	-0.107	-0.0977	0.110	-0.0714			
	(0.141)	(0.148)	(0.113)	(0.149)	(0.140)			
$\mathrm{E}_t(Unrate_{t,t+4})$						-0.952	0.780**	-1.650
						(1.589)	(0.337)	(1.163)
$Unrate_{t,t+4}$						-1.354**	-0.534	-0.553
						(0.627)	(0.453)	(0.704)
Constant	-0.942	-6.883	36.39***	35.87***	36.11***	-3.320	27.48**	25.19*
	(1.468)	(9.512)	(6.998)	(8.131)	(13.68)	(8.849)	(12.74)	(14.01)
Baseline controls$_t$	NO	YES	NO	NO	YES	YES	NO	YES
Factors$_t$	NO	NO	YES	YES	YES	NO	YES	YES
Factors$_{t+4}$	NO	NO	NO	YES	NO	NO	NO	NO
Observations	116	116	116	116	116	116	116	116

Note: Table reports results from estimating $D_{t,t+4} = \alpha + \beta E_t(RGDPgr_{t,t+4}) + \delta(L)Z_t + \gamma(L)D_{t,t+4} + u_{t,t+4}$ and $D_{t,t+4} = \alpha + \beta E_t(Unrate_{t,t+4}) + \delta(L)Z_t + \gamma(L)D_{t,t+4} + u_{t,t+4}$. $E_t(RGDPgr_{t,t+4})$ is the SPF average forecast of Real GDP growth rate (RGDP gr) over the next four quarters. $E_t(Unrate_{t,t+4})$ is the SPF average forecast of the unemployment rate over the next year. $D_{t,t+4}$ is the percentage change between time t and t+4 of the dependent variable, which can be Household credit, Personal savings, and Personal consumption expenditures. All regressions include two lags of the dependent variable. Baseline controls are: Consumer Price Index, 10YTBond, 3MTBill, GDP, and unemployment rate. Newey-west standard errors are reported in parentheses with 4 lags. *** p<0.01, ** p<0.05, * p<0.1. Few observations are missing in the SPF data. For details see section 1.4.

A.3 Natural expectations

Figure A.7: Short and long term expected RGDP growth rate

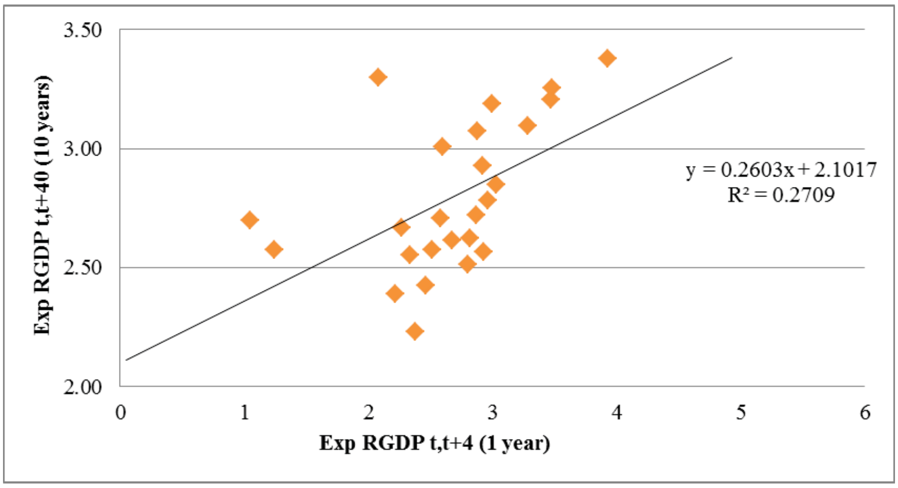
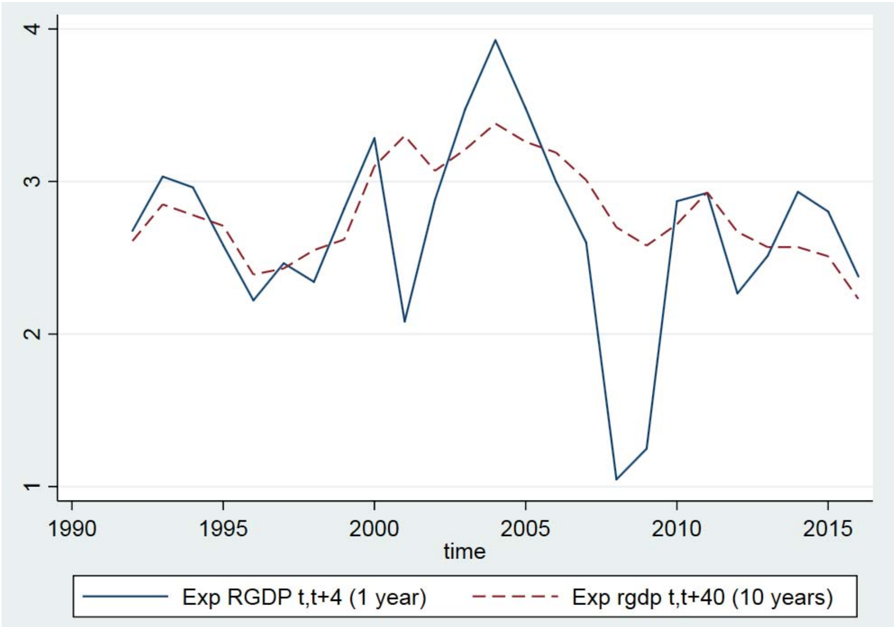
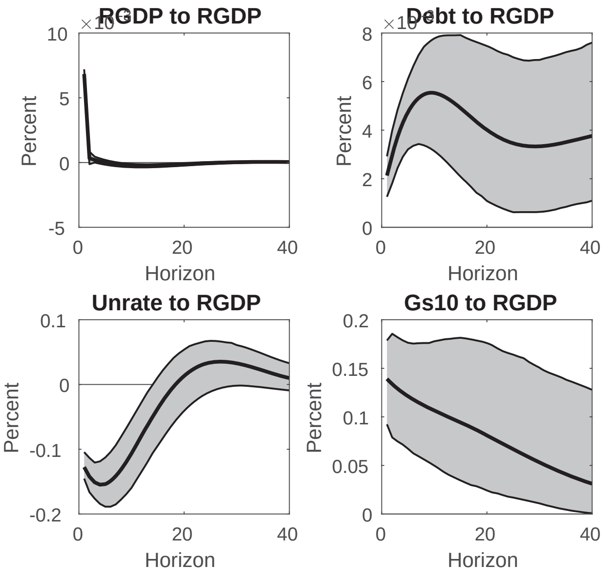


Figure A.8: Short and long term expected RGDP growth rate over time



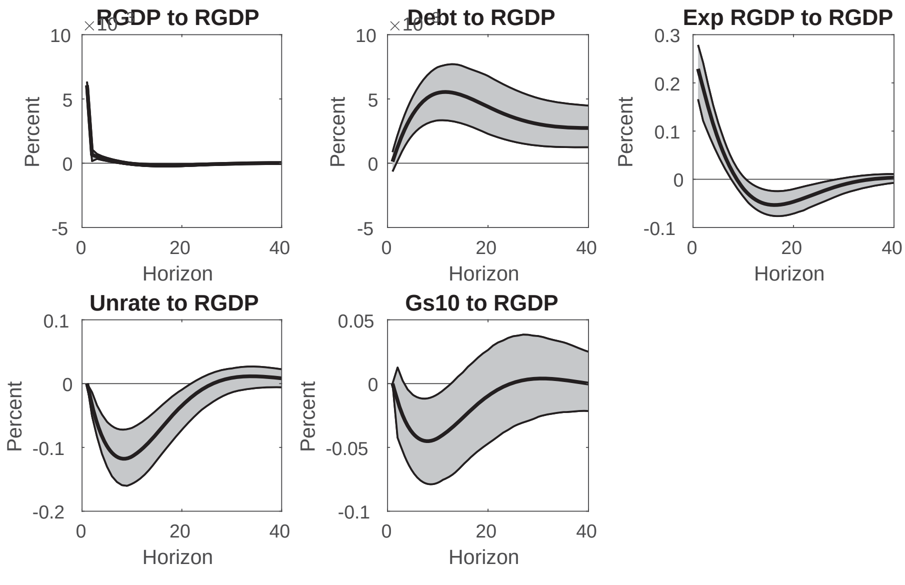
Note: The upper panel reports the correlation between $E_t(RGDPgr_{t,t+4})$ and $E_t(RGDPgr_{t,t+40})$ in percentage, using SPF data. The lower panel reports the time series for these two, where the blu solid line is the expected growth rate at 1 year ($E_t(RGDPgr_{t,t+4})$) and the red dashed line is the expected growth rate at 10 years ($E_t(RGDPgr_{t,t+40})$).

Figure A.9: Empirical impulses with Consumer Confidence



Note: IRFs to a one standard deviation innovation to the Real GDP growth rate ($RGDPgr_t$), from a VAR in 5 variables. Shocks are identified with a Cholesky scheme and $RGDPgr_t$ is ordered first. Shaded areas are the confidence bands. Debt is $ln(Household\_credit)$, Gs10 is the 10 Year Treasury Bond rate (10YTBond), unrate is the unemployment rate and Confidence is the Michigan Consumer Confidence index.

Figure A.10: Empirical impulses with alternative order



Note: IRFs to a one standard deviation innovation to the Real GDP growth rate ($RGDPgr_t$), from a VAR in 5 variables. Shocks are identified with a Cholesky scheme with the following order: $[Unrate_t, 10YTBond_t, RGDPgr_t, log(Household\_credit_t), E_t(RGDP_{t,t+4})]$. A shock to the $RGDPgr_t$ affects expectations and debt on impact as predicted in the model. Shaded areas are the confidence bands. Debt is $ln(Household\_credit)$, Gs10 is the 10 Year Treasury Bond rate (10YTBond), unrate is the unemployment rate and Exp RGDP is the SPF expected real GDP growth rate $E_t(RGDP_{t,t+4})$.

Table A.5: Out of sample performance - ARIMA process

Model	MSFE (obs 122)	MAE (obs 122)	MSFE (obs 142)	MAE (obs 142)	MSFE (obs 162)	MAE (obs 162)
ARIMA(1,1)	2902.171	417.6714	2043.036	287.2562	1043.925	147.7166
ARIMA(1,2)	2903.328	417.6452	2044.118	287.2793	1040.463	147.441
ARIMA(1,3)	2904.568	417.6935	2044.622	287.3007	1042.035	147.5429
ARIMA(1,4)	-	-	2044.944	287.3149	1042.013	147.5437
ARIMA(1,5)	2904.146	417.5428	2044.943	287.3158	1040.727	147.4306
ARIMA(2,1)	2894.203	417.1162	2033.29	286.5477	1014.593	145.5601
ARIMA(2,2)	2904.306	417.6803	-	-	1018.745	145.8748
ARIMA(2,3)	2898.011	417.3764	2032.49	286.5116	1019.326	145.9108
ARIMA(2,4)	2893.419	416.9449	2036.39	286.7803	1025.05	146.2391
ARIMA(2,5)	2892.866	416.9112	2035.603	286.7203	1022.503	146.0669
ARIMA(3,1)	2899.632	417.4858	2034.866	286.6559	1018.949	145.8878
ARIMA(3,2)	2898.971	417.4414	2033.399	286.5583	1018.98	145.8898
ARIMA(3,3)	2892.681	417.0205	2036.266	286.8452	1024.384	146.2286
ARIMA(3,4)	2893.13	416.9283	2035.961	286.749	1017.904	145.8041
ARIMA(3,5)	2916.659	418.5683	2045.846	287.4236	1019.649	145.8959
ARIMA(4,1)	2894.532	417.1412	2032.761	286.5308	1019.018	145.8922
ARIMA(4,2)	2890.647	416.8591	2036.519	286.8562	1021.477	146.0233
ARIMA(4,3)	2894.906	417.1389	2038.286	286.8747	-	-
ARIMA(4,4)	2899.079	417.3546	2040.338	287.0716	-	-
ARIMA(4,5)	2906.384	417.8774	2040.859	287.0692	-	-
ARIMA(5,1)	2892.297	416.8589	2034.206	286.5971	1020.098	145.9021
ARIMA(5,2)	2892.515	416.8685	2035.212	286.6559	1021.071	145.9716
ARIMA(5,3)	2892.549	416.8703	2043.345	287.2954	1020.73	145.9471
ARIMA(5,4)	2892.489	416.8643	2053.379	287.9034	1014.706	145.5558
ARIMA(5,5)	2920.165	418.7772	2045.887	287.4002	1014.706	145.5558

Note: Mean Square Forecast error (MSFE) and Mean Absolute Error (MAE) from an out of sample forecasting exercises used to choose the correct order of the following process: $\widetilde{E_t}(\Delta y_{t+1}) = \rho(L)\Delta y_t + \delta(L)\epsilon_{t+1}$. For instance, ARIMA(1,1) means that the agent estimates an ARIMA(1,1) on Δy_{t+1} such that the estimates model is $\widetilde{E_t}(\Delta y_{t+1}) = \rho\Delta y_t + \delta\epsilon_t + \epsilon_{t+1}$. Different columns indicate different out-of-sample windows. I use SPF survey data for expectations. Data are missing when no convergence is achieved.

Table A.6: Out of sample performance - MA process

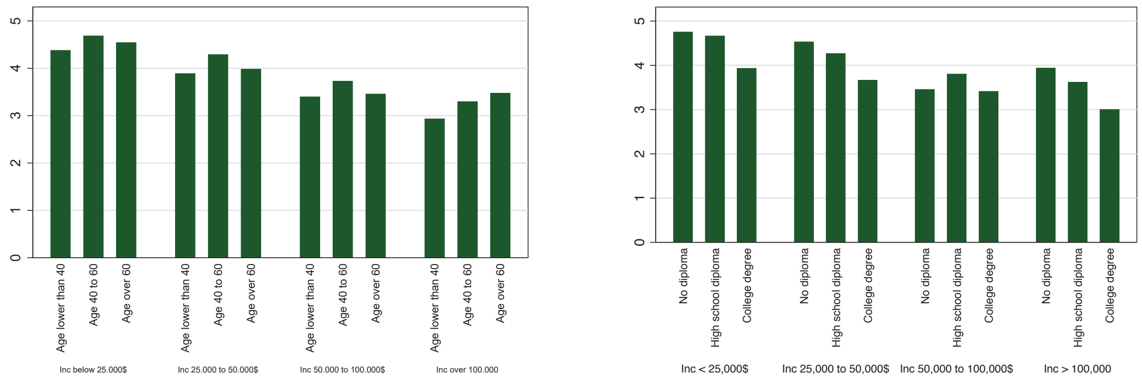
Model	MSFE (obs 122)	MAE (obs 122)	MSFE (obs 142)	MAE (obs 142)	MSFE (obs 162)	MAE (obs 162)
MA(1)	2857.218	416.433	1010.351	145.52	1974.884	283.919
MA(2)	2883.115	417.113	1024.418	146.3314	2011.156	285.7181
MA(3)	2891.454	417.262	1028.069	146.5439	2024.328	286.283
MA(4)	2894.259	417.403	1030.656	146.7556	2029.874	286.6186
MA(5)	2896.748	417.4474	1033.006	146.9121	2033.367	286.7679
MA(6)	2898.358	417.3814	1035.401	147.0763	2036.295	286.8211
MA(7)	2900.837	417.4269	1036.854	147.1714	2042.124	287.1114
MA(8)	2904.807	417.5564	1038.253	147.2594	2048.135	287.4251
MA(9)	2905.21	417.5674	2048.697	147.2934	2048.697	287.4514
MA(10)	2904.516	417.58	1038.802	147.2928	2047.637	287.4113

Note: Mean Square Forecast error (MSFE) and Mean Absolute Error (MAE) from an out of sample forecasting exercises used to choose the correct order of the following process: $\widetilde{E_t}(\Delta y_{t+1}) = \delta(L)\epsilon_{t+1}$. For instance, MA(1) means that the agent estimates an MA(1) on Δy_{t+1} such that the estimated model is $\widetilde{E_t}(\Delta y_{t+1}) = \delta\epsilon_t + \epsilon_{t+1}$. Different columns indicate different out-of-sample windows. I use SPF survey data for expectations. Data are missing when no convergence is achieved.

Appendix B

Chapter 2

B.1 Descriptive statistics



Data are at monthly frequency, 2004-2014. The plot reports the mean expected inflation one year ahead for each group.

Figure B.1: Inflation expectations by age, education and income group

Table B.1: Niesen vs Michigan: summary statistics

	Niesen (Mean)	Michigan (Mean)
Age	55	54
Family size	2,5	2,5
Income	65.000	71.000
Grade School	0.02	0.02
Graduated High School	0.33	0.28
Graduated College	0.20	0.23
Income (p50)	54.000	46.000
One child	0.22	0.14
More than 1 child	0.11	0.13
No child	0.66	0.63

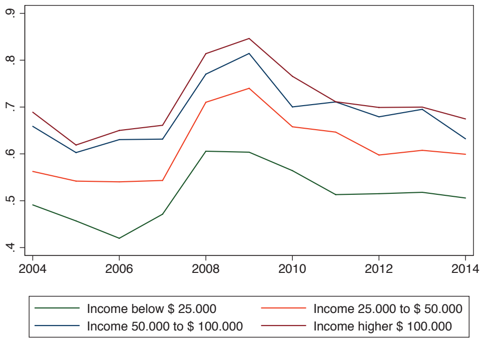
Table B.2: Descriptive statistics - monthly inflation rates

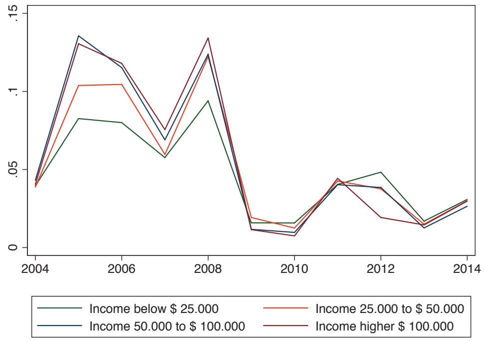
Variables	Income 1	Income 2	Income 3	Income 4
$\bar{\pi}^i_{t-12,t}$	0.441	0.387	0.364	0.355
US CPI $\pi_{t-12,t}$	0.286			

Table B.3: Inflation expectations by income classes

Income class	Mean	Median
Income class 1	4.500	4
Income class 2	4.059	3
Income class 3	3.550	3
Income class 4	3.193	3

Figure B.2: Economic and inflation news heard across income classes





Note: Economic news in the left hand side figure. $News_t^i$ is the fraction of households with income i who heard economic news, at time t. Inflation news in the right hand side figure. $Inf\_news_t^i$ is the fraction of households with income i who heard news about inflation, at time t.

Table B.4: CPI components correlations - 1 month perc. change

Variables	Food and beverages	All items
Food and beverages	1.000	0.182*
All item less energy	0.551***	-
Energy	0.068	-
Transportation services	0.154*	-
Education	-0.019	-

Note: Correlations between the different US CPI components, in 1 month percentage changes. Data are at monthly frequency, 2004-2014.

Table B.5: CPI components correlations - 12 month perc. change

Variables	Food and beverages	All items
Food and beverages	1.000	0.434***
All item less energy	0.763***	-
Energy	0.167*	-
Transportation services	0.177*	-
Education	-0.114	-

Note: Correlations between the different US CPI components, in 12 months percentage changes. Data are at monthly frequency, 2004-2014.

Table B.6: Laspeyres and US CPI correlation - 1 month perc. change

Variables	$\bar{\pi}_{j,t}^i$	US CPI π_t (All items)
$\bar{\pi}_{j,t}^i$	1.000	-
US CPI π_t (All items)	0.225	1.000

Note: Correlations between the individual inflation rate $\bar{\pi}_{j,t}^i$ and the US CPI inflation π_t in 1 month percentage changes. Data are at monthly frequency, 2004-2014.

B.2 Personal experience and inflation expectations

Table B.7: Expected and realized inflation - annualized rate

	(1) $\bar{E}_t^i(\pi_{t,t+12})$	(2) $\bar{E}_t^i(\pi_{t,t+12})$	(3) $\bar{E}_t^i(\pi_{t,t+12})$	(4) $\bar{E}_t^i(\pi_{t,t+12})$	(5) $\bar{E}_t^i(\pi_{t,t+12})$
$\bar{\pi}_{t-12,t}^i$		0.108**	0.187***	0.111***	0.186***
		(0.019)	(0.000)	(0.000)	(0.000)
$\pi_{t-12,t}$	0.154**	0.144**	0.143***	0.159***	0.156***
	(0.016)	(0.013)	(0.000)	(0.000)	(0.000)
Income 2*$\bar{\pi}_{t-12,t}^i$			-0.108		-0.108*
			(0.128)		(0.077)
Income 3*$\bar{\pi}_{t-12,t}^i$			-0.0761*		-0.0781*
			(0.072)		(0.073)
Income 4*$\bar{\pi}_{t-12,t}^i$			-0.134**		-0.139**
			(0.020)		(0.012)
Income 2*$\pi_{t-12,t}$				-0.0419	-0.0522
				(0.599)	(0.490)
Income 3*$\pi_{t-12,t}$				-0.00338	0.000495
				(0.946)	(0.990)
Income 4*$\pi_{t-12,t}$				-0.0304	-0.0117
				(0.689)	(0.859)
Cons	4.420***	4.907***	5.895***	5.632***	5.936***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Income dummy	NO	NO	YES	YES	YES
N	432	432	432	432	432
Adj. R-sq	0.093	0.181	0.739	0.725	0.74

Note: p-values in parentheses* p<0.10, ** p<0.05, *** p<0.01. $\bar{\pi}_{t-12,t}^i$ is the annualized average realized inflation of the income group i. $\pi_{t-12,t}$ is the annualized US inflation. Newey-west standard errors with 12 lags.

Table B.8: Expected and realized inflation - median

	Income 1	Income 2	Income 3	Income 4
	$\bar{E}_t^i(\pi_{t,t+12})$	$\bar{E}_t^i(\pi_{t,t+12})$	$\bar{E}_t^i(\pi_{t,t+12})$	$\bar{E}_t^i(\pi_{t,t+12})$
$\bar{\pi}^i_{t-12,t}$	2.069***	0.634***	1116**	0.119
	(0.000)	(0.003)	(0.024)	(0.785)
$\pi_{t-12,t}$	1.319***	1.114***	1328**	0.388**
	(0.000)	(0.000)	(0.028)	(0.025)
Cons	5.097***	3.619***	3.680***	3.550***
	(0.000)	(0.000)	(0.000)	(0.000)
N	108	108	108	108
Adj. R-sq	0.573	0.441	0.230	0.114

Note: p-values in parentheses* p<0.10, ** p<0.05, *** p<0.01. $\bar{E}_t^i(\pi_{t,t+12})$ is the media expected inflation rate for income group i. $\bar{\pi}^i_{t-12,t}$ is the income specific inflation rate of income group i. $\pi_{t-12,t}$ is the annualized US inflation. Newey-west standard errors with 12 lags.

Table B.9: Testing the information content of the private signal

	Income 1	Income 2	Income 3	Income 4
	$\pi_{t,t+12}$	$\pi_{t,t+12}$	$\pi_{t,t+12}$	$\pi_{t,t+12}$
$\bar{\pi}^i_{t-12,t}$	0.412***	-0.241	-0.0946	-0.270
	(0.008)	(0.313)	(0.435)	(0.173)
Cons	0.524***	0.276***	0.316***	0.278***
	(0.000)	(0.002)	(0.000)	(0.000)
N	108	108	108	108
Adj. R-sq	0.157	0.085	0.016	0.054

Note: p-values in parentheses* p<0.10, ** p<0.05, *** p<0.01. $\bar{\pi}^i_{t-12,t}$ is the average realized inflation of income group i. $\pi_{t,t+12}$ is the US inflation. Newey-west standard errors with 12 lags.

Table B.10: Test for rationality with oil price

	Income 1	Income 2	Income 3	Income 4
	$\pi_{t,t+12}$	$\pi_{t,t+12}$	$\pi_{t,t+12}$	$\pi_{t,t+12}$
$\bar{\pi}^i_{t-12,t}$	0.387	-0.244	-0.0482	-0.150
	(0.187)	(0.294)	(0.674)	(0.382)
$\pi_{t-12,t}$	0.187***	0.0904**	0.143***	0.248***
	(0.000)	(0.033)	(0.001)	(0.000)
Growth rate oil	0.157	-0.0495	-0.0131	0.379
	(0.501)	(0.823)	(0.948)	(0.134)
Cons	0.437***	0.248***	0.280***	0.247***
	(0.000)	(0.005)	(0.000)	(0.000)
N	108	108	108	108
Adj. R-sq	0.348	0.109	0.169	0.351

Note: p-values in parentheses* p<0.10, ** p<0.05, *** p<0.01. $\bar{\pi}^i_{t-12,t}$ is the average realized inflation on the income group i. $\pi_{t-12,t}$ is the US inflation. Growth oil price is the monthly growth rate. Newey-west standard errors with 12 lags.

Table B.11: Test for rationality with house prices

	Income 1	Income 2	Income 3	Income 4
	$\pi_{t,t+12}$	$\pi_{t,t+12}$	$\pi_{t,t+12}$	$\pi_{t,t+12}$
$\bar{\pi}^i_{t-12,t}$	0.395	-0.282	-0.0559	-0.148
	(0.109)	(0.159)	(0.593)	(0.427)
$\pi_{t-12,t}$	0.175***	0.0987*	0.138***	0.248***
	(0.001)	(0.078)	(0.001)	(0.000)
House price	-2.335	-6.407	-1.241	-1.592
	(0.359)	(0.132)	(0.599)	(0.562)
Cons	0.446***	0.247***	0.280***	0.246***
	(0.000)	(0.001)	(0.000)	(0.001)
N	108	108	108	108
Adj. R-sq	0.357	0.265	0.177	0.302

Note: p-values in parentheses* p<0.10, ** p<0.05, *** p<0.01. $\bar{\pi}^i_{t-12,t}$ is the average realized inflation on the income group i. $\pi_{t-12,t}$ is the US inflation. House price is the change from the preceding period. Newey-west standard errors with 12 lags.

Table B.12: Test for rationality with energy price

	Income 1	Income 2	Income 3	Income 4
	$\pi_{t,t+12}$	$\pi_{t,t+12}$	$\pi_{t,t+12}$	$\pi_{t,t+12}$
$\bar{\pi}^i_{t-12,t}$	0.347	-0.234	-0.0676	-0.194
	(0.130)	(0.317)	(0.591)	(0.256)
$\pi_{t-12,t}$	0.180***	0.0978**	0.147***	0.249***
	(0.000)	(0.034)	(0.000)	(0.000)
Energy price	0.0165	-0.00952	0.0108	-0.0113
	(0.528)	(0.337)	(0.571)	(0.342)
Cons	0.433***	0.247***	0.273***	0.226***
	(0.000)	(0.003)	(0.000)	(0.000)
N	108	108	108	108
Adj. R-sq	0.350	0.115	0.176	0.302

Note: p-values in parentheses* p<0.10, ** p<0.05, *** p<0.01. $\bar{\pi}^i_{t-12,t}$ is the average realized inflation on the income group i. $\pi_{t-12,t}$ is the US inflation. Energy price is the change from the preceding period. Newey-west standard errors with 12 lags.

Table B.13: Test for rationality with news

	Income 1	Income 2	Income 3	Income 4
	$\pi_{t,t+12}$	$\pi_{t,t+12}$	$\pi_{t,t+12}$	$\pi_{t,t+12}$
$\bar{\pi}^i_{t-12,t}$	0.212	-0.180	-0.0690	-0.0926
	(0.119)	(0.211)	(0.231)	(0.455)
$\pi_{t-12,t}$	0.120***	0.0363**	0.105***	0.134***
	(0.000)	(0.019)	(0.003)	(0.001)
$Infl\_new^i_t$	1.204	1.742*	1.540	1.576
	(0.302)	(0.055)	(0.199)	(0.159)
$\bar{\pi}^i_{t-12,t}*Infl\_new^i_t$	-0.220	0.959	1.105	1.975
	(0.944)	(0.644)	(0.700)	(0.427)
Cons	0.332***	0.210***	0.222***	0.242***
	(0.000)	(0.000)	(0.000)	(0.000)
N	108	108	108	108
Adj R-sq	0.525	0.384	0.496	0.466

Note: p-values in parentheses* p<0.10, ** p<0.05, *** p<0.01. $\bar{\pi}^i_{t-12,t}$ is the average realized inflation on the income group i. $\pi_{t-12,t}$ is the US inflation. $Infl\_new^i_t$ is the Economic index news (Section 2.4.3). Newey-west standard errors with 12 lags.

B.3 Salience

Table B.14: Expected and realized inflation with salience in frequency (bis)

	Income 1	Income 2	Income 3	Income 4
	$\bar{E}_t^i(\pi_{t,t+12})$	$\bar{E}_t^i(\pi_{t,t+12})$	$\bar{E}_t^i(\pi_{t,t+12})$	$\bar{E}_t^i(\pi_{t,t+12})$
$\bar{\pi}_{t-12,t}^i$	2.138***	0.565*	1.127***	1.402***
	(0.000)	(0.086)	(0.001)	(0.006)
$\pi_{t-12,t}$	1.807***	1.208*	1.955***	0.909**
	(0.000)	(0.095)	(0.000)	(0.011)
$\bar{\pi}_{t-12,t}^{i,f*}$	0.112	-0.0185	0.210	3.145***
	(0.845)	(0.976)	(0.671)	(0.000)
Cons	5.892***	4.790***	4.518***	3.272***
	(0.000)	(0.000)	(0.000)	(0.000)
N	108	108	108	108
Adj. R-sq	0.662	0.130	0.555	0.581

Note: p-values in parentheses* $p<0.10$, ** $p<0.05$, *** $p<0.01$. $\pi_{t-12,t}$ is the US inflation. $\bar{\pi}_{t-12,t}^{i,f*}$ is the income specific inflation rate computed by over-weighting goods that are salient in the frequency of purchase where the degree of salience is computed on all goods with $\sigma(f_{j,k,t}) = \frac{|f_{j,k,t}-\bar{f}_{j,t}|}{f_{j,k,t}+\bar{f}_{j,t}}$. Newey-west standard errors with 12 lags.

Table B.15: Expected and realized inflation with salience in expenditures (bis)

	Income 1	Income 2	Income 3	Income 4
	$\bar{E}_t^i(\pi_{t,t+12})$	$\bar{E}_t^i(\pi_{t,t+12})$	$\bar{E}_t^i(\pi_{t,t+12})$	$\bar{E}_t^i(\pi_{t,t+12})$
$\bar{\pi}_{t-12,t}^i$	2.168***	0.305*	0.968***	1.172**
	(0.000)	(0.094)	(0.001)	(0.030)
$\pi_{t-12,t}$	1.826***	1.346*	2.129***	0.536
	(0.000)	(0.091)	(0.000)	(0.342)
$\bar{\pi}_{t-12,t}^{i,exp*}$	0.299	0.365	0.485	2.807***
	(0.587)	(0.376)	(0.538)	(0.000)
Cons	5.812***	4.552***	4.321***	2.272***
	(0.000)	(0.000)	(0.000)	(0.000)
N	108	108	108	108
Adj. R-sq	0.664	0.143	0.563	0.629

Note: p-values in parentheses* $p<0.10$, ** $p<0.05$, *** $p<0.01$. $\pi_{t-12,t}$ is the US inflation. $\bar{\pi}_{t-12,t}^{i,exp*}$ is the income specific inflation rate computed by over-weighting goods that are salient in expenditures where the degree of salience is computed for all goods with $\sigma(exp_{j,k,t}) = \frac{|exp_{j,k,t}-e\bar{x}p_{j,t}|}{exp_{j,k,t}+e\bar{x}p_{j,t}}$. Newey-west standard errors with 12 lags.

Table B.16: Expected and realized inflation with salience in frequency and price

	Income 1	Income 2	Income 3	Income 4
	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$
$\bar{\pi}_{t-12,t}^{i}$	1.720***	0.235*	0.790***	1.290**
	(0.000)	(0.068)	(0.003)	(0.040)
$\pi_{t-12,t}$	2.247***	2.077**	2.493***	1.832***
	(0.000)	(0.011)	(0.000)	(0.003)
$\bar{\pi}_{t-12,t}^{i,\Delta p^{>0},f}$	0.123***	0.201***	0.186**	0.123**
	(0.004)	(0.005)	(0.012)	(0.031)
Cons	7.854***	8.149***	7.715***	6.431***
	(0.000)	(0.000)	(0.000)	(0.000)
N	108	108	108	108
Adj. R-sq	0.726	0.436	0.695	0.398

Note: p-values in parentheses* p<0.10, ** p<0.05, *** p<0.01. $\pi_{t-12,t}$ is the US inflation. $\bar{\pi}_{t-12,t}^{i,\Delta p^{>0},f}$ is the income specific inflation rate computed by over-weighting goods that are salient in price variation and frequency; in the Laspeyres index each good is weighed by $\sigma(\Delta p_{j,k,t}^{>0}) + \sigma(f_{j,k,t})$. Newey-west standard errors with 12 lags.

Table B.17: Expected and realized inflation with salience in expenditure and price

	Income 1	Income 2	Income 3	Income 4
	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$
$\bar{\pi}_{t-12,t}^{i}$	1.689***	0.241**	0.771***	1.350**
	(0.000)	(0.045)	(0.010)	(0.029)
$\pi_{t-12,t}$	2.235***	2.426***	2.644***	1.976***
	(0.000)	(0.005)	(0.000)	(0.002)
$\bar{\pi}_{t-12,t}^{i,\Delta p^{>0},exp}$	0.0966**	0.195***	0.169**	0.135**
	(0.050)	(0.002)	(0.033)	(0.039)
Cons	7.616***	8.512***	7.759***	6.954***
	(0.000)	(0.000)	(0.000)	(0.000)
N	108	108	108	108
Adj. R-sq	0.704	0.460	0.666	0.433

p-values in parentheses* p<0.10, ** p<0.05, *** p<0.01. $\pi_{t-12,t}$ is the US inflation. $\bar{\pi}_{t-12,t}^{i,\Delta p^{>0},exp}$ is the income specific inflation rate computed by over-weighting goods that are salient in price variation and expenditure; in the Laspeyres index each good is weighed by $\sigma(\Delta p_{j,k,t}^{>0}) + \sigma(exp_{j,k,t})$. Newey-west standard errors with 12 lags.

Table B.18: Expected and realized inflation, salience in expenditure and frequency

	Income 1	Income 2	Income 3	Income 4
	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$	$E_t^i(\pi_{t,t+12})$
$\bar{\pi}_{t-12,t}^{i}$	2.042***	0.392**	1.163***	1.303**
	(0.000)	(0.032)	(0.000)	(0.019)
$\pi_{t-12,t}$	1.899***	1.591*	2.213***	1.792**
	(0.000)	(0.087)	(0.000)	(0.012)
$\bar{\pi}_{t-12,t}^{i,exp,f}$	0.375	0.466	0.354	0.919
	(0.323)	(0.146)	(0.398)	(0.152)
Cons	5.986***	4.890***	4.735***	4.458***
	(0.000)	(0.000)	(0.000)	(0.000)
N	108	108	108	108
Adj. R-sq	0.671	0.155	0.568	0.328

Note: p-values in parentheses* p<0.10, ** p<0.05, *** p<0.01. $\pi_{t-12,t}$ is the US inflation. $\bar{\pi}_{t-12,t}^{i,exp,f}$ is the income specific inflation rate computed by over-weighting goods that are salient in frequency and expenditure; in the Laspeyres index each good is weighed by $\sigma(exp_{j,k,t}) + \sigma(f_{j,k,t})$. Newey-west standard errors, 12 lags.

Table B.19: Test for rationality with salience in price

	Income 1	Income 2	Income 3	Income 4
	$\pi_{t,t+12}$	$\pi_{t,t+12}$	$\pi_{t,t+12}$	$\pi_{t,t+12}$
$\bar{\pi}_{t-12,t}^{i}$	0.224	-0.331	-0.0961	0.0474
	(0.187)	(0.009)	(0.397)	(0.698)
$\pi_{t-12,t}$	0.134***	0.0263**	0.110**	0.179***
	(0.006)	(0.034)	(0.037)	(0.001)
$\bar{\pi}_{t-12,t}^{i,\Delta p<0}$	0.0237	-0.0156	0.0159	-0.00536
	(0.272)	(0.532)	(0.524)	(0.802)
$\bar{\pi}_{t-12,t}^{i,\Delta p<0}$	0.0596	0.0476	0.0421	0.0489
	(0.162)	(0.104)	(0.124)	(0.190)
Cons	1.081	1.524***	0.797	1.481***
	(0.147)	(0.001)	(0.148)	(0.001)
N	1.081	1.524***	0.797	1.481***
Adj. R-sq	(0.147)	(0.001)	(0.148)	(0.001)

Note: p-values in parentheses* p<0.10, ** p<0.05, *** p<0.01. $\pi_{t-12,t}$ is the US inflation. $\bar{\pi}_{t-12,t}^{i,\Delta>0}$ is the income specific inflation rate computed by over-weighting goods that are salient in positive price variations. Similarly, $\bar{\pi}_{t-12,t}^{i,\Delta<0}$ for negative price variations. Newey-west standard errors with 12 lags.

Table B.20: Test for rationality with salience in expenditure

	Income 1	Income 2	Income 3	Income 4
	$\pi_{t,t+12}$	$\pi_{t,t+12}$	$\pi_{t,t+12}$	$\pi_{t,t+12}$
$\bar{\pi}^{i}_{t-12,t}$	0.468	-0.160	0.00352	0.0766
	(0.121)	(0.474)	(0.965)	(0.631)
$\pi_{t-12,t}$	0.184***	0.0451**	0.118***	0.190**
	(0.000)	(0.040)	(0.000)	(0.012)
$\bar{\pi}^{i,exp}_{t-12,t}$	0.0631	-0.113	-0.102	-0.156
	(0.425)	(0.335)	(0.142)	(0.310)
Cons	0.295**	0.637***	0.646**	0.868***
	(0.048)	(0.008)	(0.014)	(0.000)
N	108	108	108	108
Adj. R-sq	0.351	0.213	0.236	0.425

Note: p-values in parentheses* $p<0.10$, ** $p<0.05$, *** $p<0.01$. $\pi_{t-12,t}$ is the US inflation. $\bar{\pi}^{i,exp}_{t-12,t}$ is the income specific inflation rate computed by over-weighting goods that are salient in expenditures. Newey-west standard errors with 12 lags.

Table B.21: Test for rationality with salience in frequency

	Income 1	Income 2	Income 3	Income 4
	$\pi_{t,t+12}$	$\pi_{t,t+12}$	$\pi_{t,t+12}$	$\pi_{t,t+12}$
$\bar{\pi}^{i}_{t-12,t}$	0.433	-0.277	-0.111	-0.185
	(0.007)	(0.120)	(0.362)	(0.229)
$\pi_{t-12,t}$	0.0886*	0.0829*	0.121***	0.252***
	(0.087)	(0.060)	(0.000)	(0.000)
$\bar{\pi}^{i,f}_{t-12,t}$	0.529	0.231	0.168	0.170
	(0.107)	(0.129)	(0.117)	(0.185)
Cons	0.364***	0.184***	0.206***	0.159**
	(0.000)	(0.004)	(0.000)	(0.023)
N	41	37	41	39
Adj. R-sq	0.575	0.173	0.214	0.345

Note: p-values in parentheses* $p<0.10$, ** $p<0.05$, *** $p<0.01$. $\pi_{t-12,t}$ is the US inflation. $\bar{\pi}^{i,f}_{t-12,t}$ is the income specific inflation rate computed by over-weighting goods that are salient in frequency. Newey-west standard errors with 12 lags.

B.4 Inflation expectations and consumption choices

Table B.22: Readiness to spend and inflation expectations - Durable goods

	Income 1	Income 2	Income 3	Income 4
	$\bar{C}_t^i$	$\bar{C}_t^i$	$\bar{C}_t^i$	$\bar{C}_t^i$
$\tilde{E}_t^i(\pi_{t,t+12})$ - Fitted values	0.0381**	0.0348	0.0599***	0.0206
	(0.040)	(0.636)	(0.008)	(0.545)
Exp better buss cond.	0.319**	0.408***	0.456***	0.233**
	(0.018)	(0.002)	(0.000)	(0.047)
Exp higher int. rate	0.0236*	0.0502***	0.0341***	0.0559***
	(0.065)	(0.004)	(0.000)	(0.001)
Exp greater income	0.625***	-0.0452	0.0957	0.106
	(0.000)	(0.730)	(0.285)	(0.283)
Cons	-0.0482	0.317	0.198**	0.387***
	(0.562)	(0.335)	(0.027)	(0.000)
N	108	108	108	108
Adj. R-sq	0.786	0.716	0.844	0.816

p-values in parentheses * p<0.10, ** p<0.05, *** p<0.01. Newey-west standard errors with 12 lags. $\tilde{E}_t^i(\pi_{t,t+12})$ are fitted values from the baseline model of intuitive expectations in eq. (2.6). All columns include following controls: shares of respondents that expect better business condition, higher interest rates and greater income

Table B.23: Readiness to spend and inflation expectations - Cars

	Income 1	Income 2	Income 3	Income 4
	$\bar{C}_t^i$	$\bar{C}_t^i$	$\bar{C}_t^i$	$\bar{C}_t^i$
$\tilde{E}_t^i(\pi_{t,t+12})$ - Fitted values	0.0243**	0.136***	0.0583***	0.0954***
	(0.030)	(0.000)	(0.000)	0.000)
Exp better buss cond.	-0.0228	0.139***	0.0899***	0.0259
	(0.633)	(0.002)	(0.007)	(0.590)
Exp higher int. rate	0.00609	0.0106	0.00250	0.0226***
	(0.277)	(0.193)	(0.570)	(0.000)
Expect greater income	0.380***	0.163**	0.173***	0.0139
	(0.000)	(0.015)	(0.000)	(0.751)
Cons	0.223***	-0.106	0.340***	0.314***
	(0.000)	(0.315)	(0.000)	(0.000)
N	108	108	108	108
Adj. R-sq	0.757	0.777	0.826	0.810

p-values in parentheses * p<0.10, ** p<0.05, *** p<0.01. Newey-west standard errors with 12 lags. $\tilde{E}_t^i(\pi_{t,t+12})$ are fitted values from the baseline model of intuitive expectations in eq. (2.6). All columns include following controls: shares of respondents that expect better business condition, higher interest rates and greater income.

Table B.24: Readiness to spend and inflation expectations - Durable goods

	Income 1	Income 2	Income 3	Income 4
	$\bar{C}_t^i$	$\bar{C}_t^i$	$\bar{C}_t^i$	$\bar{C}_t^i$
$\bar{E}_t^i(\pi_{t,t+12})$ - Exp inflation	0.0378***	0.0593***	0.0316***	0.0433***
	(0.004)	(0.000)	(0.008)	(0.002)
Exp better buss cond.	0.165*	0.327***	0.455***	0.383***
	(0.051)	(0.000)	(0.000)	(0.000)
Exp higher int. rate	0.0315***	0.0367***	0.0324***	0.0291***
	(0.001)	(0.000)	(0.000)	(0.000)
Exp greater income	0.748***	0.151**	0.144**	0.140**
	(0.000)	(0.047)	(0.014)	(0.025)
Cons	-0.0875	0.160**	0.280***	0.285***
	(0.165)	(0.016)	(0.000)	(0.000)
N	108	108	108	108
Adj. R-sq	0.776	0.801	0.848	0.844

p-values in parentheses * $p<0.10$, ** $p<0.05$, *** $p<0.01$. Newey-west standard errors with 12 lags. $\bar{E}_t^i(\pi_{t,t+12})$ are actual inflation expectations. All columns include following controls: shares of respondents that expect better business condition, higher interest rates and greater income.

Table B.25: Readiness to spend and inflation expectations - Cars

	Income 1	Income 2	Income 3	Income 4
	$\bar{C}_t^i$	$\bar{C}_t^i$	$\bar{C}_t^i$	$\bar{C}_t^i$
$\bar{E}_t^i(\pi_{t,t+12})$ - Exp inflation	0.0149***	0.0556***	0.0381***	0.0550***
	(0.008)	(0.000)	(0.000)	(0.000)
Exp better buss cond.	0.0105	0.138***	0.138***	0.115***
	(0.751)	(0.000)	(0.000)	(0.000)
Exp higher int. rate	0.00565	-0.00452	-0.00965***	-0.0134***
	(0.196)	(0.420)	(0.000)	(0.002)
Exp greater income	0.405***	0.240***	0.203***	0.214***
	(0.000)	(0.000)	(0.000)	(0.000)
Cons	0.246***	0.219***	0.396***	0.378***
	(0.000)	(0.000)	(0.000)	(0.000)
N	108	108	108	108
Adj. R-sq	0.773	0.784	0.763	0.707

p-values in parentheses * $p<0.10$, ** $p<0.05$, *** $p<0.01$. Newey-west standard errors with 12 lags. $\bar{E}_t^i(\pi_{t,t+12})$ are actual inflation expectations. All columns include following controls: shares of respondents that expect better business condition, higher interest rates and greater income.

Table B.26: Negative readiness to spend - Durable goods

	Inc 1	Inc 2	Inc 3	Inc 4
	C_t^i	C_t^i	C_t^i	C_t^i
$\tilde{E}_t^i(\pi_{t,t+12})$ - Fitted values	-0.0552**	-0.0886	-0.0682***	-0.0364
	(0.038)	(0.130)	(0.004)	(0.317)
Exp worst buss cond.	-0.0965	0.210	0.393***	0.359**
	(0.679)	(0.234)	(0.005)	(0.031)
Exp higher int. rate	-0.0804***	-0.0937***	-0.0776***	-0.0814***
	(0.000)	(0.000)	(0.000)	(0.000)
Exp lower income	0.390	-0.522***	0.00192	-0.296***
	(0.172)	(0.003)	(0.986)	(0.005)
Cons	0.723***	0.870***	0.559***	0.503***
	(0.000)	(0.002)	(0.000)	(0.001)
N	108	108	108	108
Adj. R-sq	0.490	0.757	0.791	0.830

p-values in parentheses * p<0.10, ** p<0.05, *** p<0.01. C_t^i is the fraction of agents who believe is a bad time to buy durable goods in income class i. Newey-west standard errors with 12 lags. $\tilde{E}_t^i(\pi_{t,t+12})$ are fitted values from the baseline model of intuitive expectations in eq. (2.6). All columns include following controls: shares of respondents that expect better business condition, higher interest rates and lower income.

Table B.27: Negative readiness to spend - Cars

	Income 1	Income 2	Income 3	Income 4
	C_t^i	C_t^i	C_t^i	C_t^i
$\tilde{E}_{j,t}(\pi_{t,t+12})$ - Fitted values	-0.0372***	-0.128***	-0.0479***	-0.0981***
	(0.008)	(0.000)	(0.000)	(0.000)
Exp worst buss cond.	-0.315**	-0.153	0.0437	0.0147
	(0.017)	(0.129)	(0.601)	(0.776)
Exp higer int. rate	-0.0226***	-0.0400***	-0.0286***	-0.0324***
	(0.002)	(0.000)	(0.000)	(0.000)
Exp lower income	0.370***	-0.0182	0.166***	-0.141***
	(0.008)	(0.769)	(0.002)	(0.001)
Cons	0.660***	1.014***	0.483***	0.691***
	(0.000)	(0.000)	(0.000)	(0.000)
N	108	108	108	108
Adj. R-sq	0.500	0.675	0.695	0.854

p-values in parentheses * p<0.10, ** p<0.05, *** p<0.01. C_t^i is the fraction of agents who believe is a bad time to buy cars in income class i. Newey-west standard errors with 12 lags. $\tilde{E}_{j,t}(\pi_{t,t+12})$ are fitted values from the baseline model of intuitive expectations in eq. (2.6). All columns include following controls: shares of respondents that expect better business condition, higher interest rates and lower income.

B.5 Michigan Survey of Consumers questions

Q1. *During the next 12 months, do you think that prices, in general, will go up, or down, or stay where they are now?*

Q2. *By about what percent per year do you expect prices to go (up/ down) on the average, during the next 12 months?*

Q3. *About the big things people buy for their homes - such as furniture, a refrigerator, stove, television, etc. Generally speaking, do you think now is a good or a bad time for people to buy major household items?*

Q4. *Now thinking only about the next twelve months, do you think that the price of gasoline will go up during the next twelve months, will gasoline prices go down, or will they stay about the same as they are now?*

Q5. *No one can say for sure, but what do you think will happen to interest rates for borrowing money during the next 12 months - will they go up, stay the same, or go down?*

Q6. *During the next 12 months, do you expect your income to be higher or lower than during the past year?*

Q7. *During the last few months, have you heard of any favorable or unfavorable changes in business conditions? What did you hear?* Economic news =1 if answer favorable or unfavorable, =0 otherwise. Inflation news =1 if answer news about inflation (favorable or unfavorable), =0 otherwise.

Q8. *Now turning to business conditions in the country as a wholeâdo you think that during the next 12 months weâll have good times financially, or bad times, or what?*

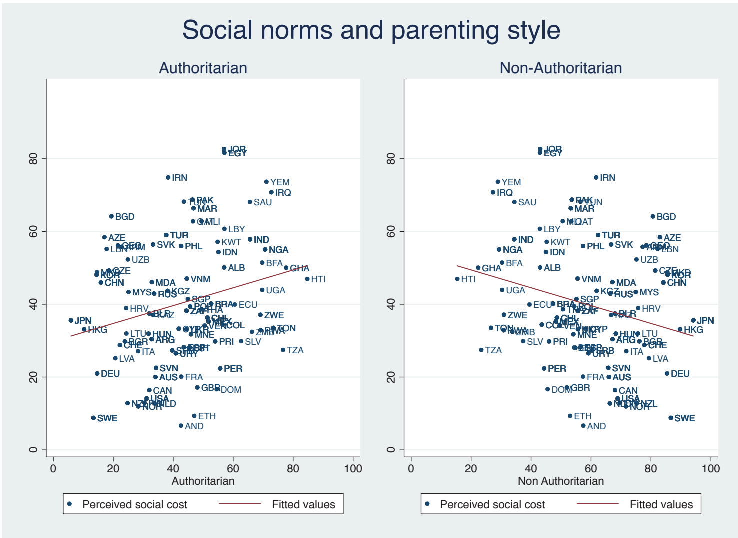
Q9. *How about people out of work during the coming 12 months - do you think there will be more unemployment than now, about the same, or less?*

Appendix C

Chapter 3

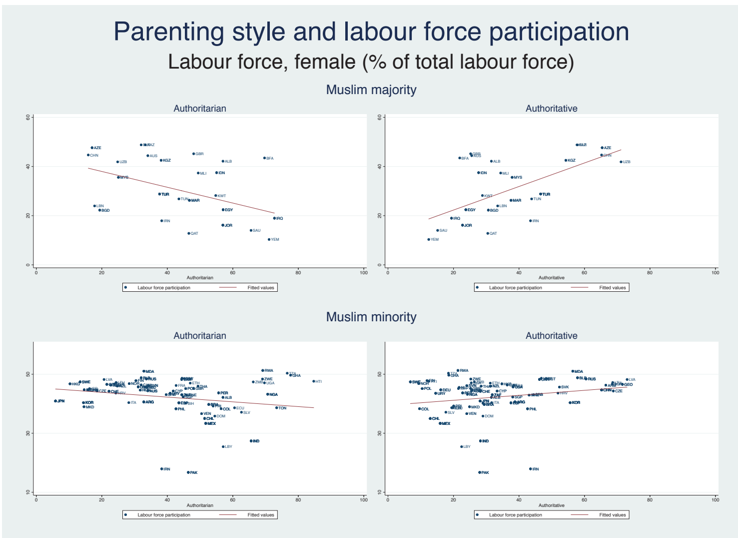
C.1 Figures and Tables

Figure C.1: Social norms and parenting style



Data are collapsed at country level. Figures display the fraction of respondents in the WVS that are authoritarian or non-authoritariaand the social norms computed as described in Table 3.3. Non-authoritarian are those parents that are either authoritative or permissive.

Figure C.2: Parenting style and Education



Data are collapsed at country level. Figures display the fraction of respondents in the WVS that are authoritarian and authoritative and women labour force participation defined in Table 3.2.

C.2 Individual-level data

Table C.1: The determinants of parenting style

Variables	(1) Authoritarian	(2) Authoritative	(3) Permissive	(4) Authoritarian	(5) Authoritative	(6) Permissive
Religious norms	0.822***	-0.550**	0.0233	0.752***	-0.435***	0.0229
	(0.192)	(0.222)	(0.225)	(0.143)	(0.153)	(0.213)
Social norms	0.185	0.303**	-0.269*	0.196	0.284**	-0.258
	(0.133)	(0.128)	(0.162)	(0.131)	(0.126)	(0.162)
Institutions	-0.0194	0.477	-0.836**	0.0778	0.394	-0.743*
	(0.360)	(0.418)	(0.389)	(0.361)	(0.437)	(0.406)
GINI INDEX	0.0116	0.00296	-0.0235*			
	(0.0101)	(0.0126)	(0.0124)			
Income top - lowest 10%				0.0202*	-0.0131	-0.0154
				(0.0120)	(0.0161)	(0.0117)
Constant	-0.662	-0.848	-0.647	-0.554	-0.412	-0.973
	(1.161)	(1.089)	(1.219)	(1.230)	(1.152)	(1.270)
Observations	83,487	83,497	83,489	84,711	84,721	84,713

Table displays results from estimating the following model: $Parenting_{j,i,t} = \alpha + \beta X_{j,i,t} + \delta Indicators_{j,t} + \varepsilon_{j,i,t}$, where j is the individual, i is the country and t is the time (wave). *Parenting* is the fraction of respondents in the WVS that are authoritarian, authoritative and permissive. The GINI Index measures the extent to which the distribution of income (or, in some cases, consumption expenditure) among individuals or households within an economy deviates from a perfectly equal distribution. *Income top 10% - lower 10%* is the income share held by highest 10% minus the one held by the lowers 10%. The independent variables instead are computed as described in Table 3.3. All specifications control for Wave, age, education, sex; standard errors are clustered at the country level.