# A distributional theory of household sentiment[‡]

**Marco Bellifemine** LSE

**Adrien Couturier** LSE

**Seyed M. Hosseini** Harvard

September 23, 2025

**Abstract**

We embed diagnostic expectations into an otherwise standard incomplete-markets model of consumption-saving with idiosyncratic income risk. In this framework, households form beliefs that overweight recent income shocks, a bias we summarize with a sufficient statistic—sentiment—that distorts perceptions of future income. We discipline our model empirically using the Italian Survey of Household Income and Wealth. Taking advantage of a continuous-time formulation, we derive a closed-form characterization of how sentiment dampens the saving motive under optimism and amplifies it under pessimism, causing households to overreact to income shocks in their consumption-saving choices. We then show that the interaction of sentiment with borrowing constraints generates a "diagnostic poverty trap": positive shocks fuel over-consumption rather than asset accumulation, making it harder for constrained households to escape the hand-to-mouth state. This simple behavioral friction rationalizes the persistence of hand-to-mouth households observed in the data and helps match their empirical prevalence without invoking illiquid assets or preference heterogeneity.

[‡]We thank Ben Moll and Ricardo Reis for their encouragements and continuous guidance on this project. We would also like to thank Matthias Doepke, Wouter den Haan, Tomer Ifergane, Cameron Peng, Maarten de Ridder, Walker Ray, and Adi Soenarjo. Oliver Seager provided excellent research assistance.

Bellifemine: m.bellifemine@lse.ac.uk.
Couturier: y.couturier-roguet@lse.ac.uk.
Hosseini: shosseinimaasoum@fas.harvard.edu.

# 1 Introduction

A growing literature documents that economic agents' forecasts are excessively swayed by recent information. The theory of diagnostic expectations formalizes this form of over-extrapolation: agents overweight future states that are representative of recent news (Bordalo et al., 2018). In this paper we ask whether embedding this friction into a standard incomplete-markets model—where households face idiosyncratic income risk and borrowing constraints—can help resolve persistent puzzles about household consumption-saving behavior in the data.

In particular, we show that introducing this deviation from rational expectations helps reconcile two empirical observations that standard models cannot explain. First, past income growth is positively correlated with households' forecast errors about future income, a new fact we document. Second, the hand-to-mouth status is highly persistent, consistent with the evidence in Aguiar et al. (Forthcoming), but difficult to generate under rational expectations.

This paper connects the incomplete-markets literature on consumption–saving with the growing literature on deviations from rational expectations. We embed diagnostic expectations about idiosyncratic income into the workhorse heterogeneous-agents model (Bewley, 1977, Imrohoroglu, 1989, Huggett, 1990, Aiyagari, 1994). A key challenge is that expectations depend on the entire history of shocks. Building on the continuous-time formulation of diagnostic expectations in Maxted (2023a), we capture this history with a single household-level sufficient statistic, which we call *sentiment*. Sentiment distorts beliefs about the drift of future income: positive shocks raise optimism, while negative shocks induce pessimism. and jumps whenever new shocks arrive. As a result, the stationary distribution of the economy depends jointly on income, wealth, and sentiment.

To capture the deviations from rational expectations induced by sentiment, we develop a new tool that we call the rationality wedge. This operator enters the household's Hamilton–Jacobi–Bellman (HJB) equation and represents misperceptions about the evolution of state variables. The wedge formulation allows us to extend the intuition of diagnostic expectations beyond simple AR(1) processes to the richer stochastic environments commonly used in household finance. It also enables us to represent the economy as a mean-field-game fixed point between a behavioral HJB—pinning down households' policy functions under distorted beliefs—and a Kolmogorov forward equation that governs the true evolution of the economy. More generally, our rationality wedge provides a flexible method for embedding a wide range of belief distortions into heterogeneous-agent models.

We first document new evidence on households' expectations about idiosyncratic income. Using the Italian Survey of Household Income and Wealth (SHIW), we show that households systematically overreact to recent income shocks: those who recently experienced positive income growth are overly optimistic about the future, while those with negative shocks are overly pessimistic. We also confirm a pattern from the Michigan Survey that high-income households tend to be optimistic and low-income households pessimistic (Rozsypal and Schlafmann, 2023).

Our framework reproduces these systematic deviations from rational expectations. Because forecast errors in the model depend on households' recent income history, agents with a sequence of positive shocks accumulate high sentiment and become overconfident, while those with negative shocks become overly pessimistic. This mechanism generates dispersed and predictable forecast errors, consistent with the survey evidence and difficult to capture in standard incomplete-markets models (Broer et al., 2021, 2022, Rozsypal and Schlafmann, 2023).

We then turn to the persistence of the hand-to-mouth state. Exploiting the long panel dimension of the SHIW, we estimate that roughly 60 percent of households classified as HtM remain so after two years, about 55 percent after four years, and nearly 40 percent even after fourteen years. These findings closely mirror and extend the evidence in Aguiar et al. (Forthcoming).

Our framework naturally accounts for this persistence. Under diagnostic expectations, positive income shocks raise sentiment and induce constrained households to overconsume rather than save, slowing their transition out of the borrowing limit. Negative shocks, by contrast, cannot trigger additional saving at the constraint. This asymmetry generates a "diagnostic poverty trap," consistent with the long-run stickiness observed in the SHIW. Quantitatively, the model matches the prevalence and persistence of hand-to-mouth households without relying on illiquid assets or preference heterogeneity, and implies welfare losses averaging 3.3 percent of lifetime consumption. Moreover, because sentiment shifts marginal propensities to consume, the model provides a behavioral channel for the heterogeneity in consumption responses documented in Lewis et al. (2019), Arellano et al. (2023).

**Related Literature** This paper relates to three different strands of literature. First, we build upon the recent literature embedding behavioral frictions in incomplete market models.[1] A growing strand of literature has been incorporating present bias of the Harris

[1] See also Pappa et al. (2023), who study the interaction between incomplete markets and expectational shocks in a macroeconomic model with labor market frictions.

and Laibson (2013) type to heterogeneous agents models, see Laibson and Maxted (2023) and Maxted (2023b). Laibson et al. (2021) show that present bias amplifies the household balance-sheet channels of macroeconomic policy when agents can conduct cash-out refinancing. We differ from this strand by introducing behavioral frictions on beliefs rather than preferences, in this sense we are closer to Rozsypal and Schlafmann (2023) who introduce over-persistence bias to an Ayiagari-Bewley-Hugett economy. We show that introducing sentiment can potentially explain the persistence in the hand-to-mouth status documented in Aguiar et al. (Forthcoming) and latent heterogeneity in consumption response to shocks discussed in Lewis et al. (2019) and Arellano et al. (2023).

Second, we extend the literature on diagnostic expectations pioneered by Bordalo et al. (2018). We rely and extend the continuous-time representation of diagnostic expectations proposed in Maxted (2023a). Unlike this paper, we apply diagnostic expectations to idiosyncratic shocks to income, instead of aggregate shocks, and develop a methodology to apply them to jump-drift processes instead of Brownian motions. We provide novel evidence and quantitative estimates on the degree of diagnosticity in households' expectations about their own future income, complementing the results in Gennaioli et al. (2016), Bordalo et al. (2020), and Ma et al. (2020). Our approach differs from that of Bianchi et al. (2024) and L'Huillier et al. (2023), who integrate diagnostic expectations into standard real business cycle and new Keynesian models respectively, but assume a representative agent. We focus instead on an incomplete market framework.

Finally, our research also contributes to the rich body of theoretical work on psychology-driven poverty traps. The studies by Banerjee and Mullainathan (2010) and Bernheim et al. (2015) underscore the influence of present bias and temptation in driving poverty. Studies that relate to our work include Thakral and Tô (2021), which posit an alternate psychology-oriented theory of poverty traps, proposing that consumers might be disinclined to save, knowing their future selves could squander savings, or Sergeyev et al. (2023), in which poverty traps result from financial stress in the vicinity of the borrowing limit. Similarly, the study by Dalton et al. (2016) adds a valuable dimension to this discussion by examining the impact of reference dependence and aspirations on poverty traps, through a different psychological lens. We contribute to this strand of literature by showing that a well accepted form of deviation from rational expectations can generate poverty traps.

**Outline** Section 2 lays out our model of diagnostic expectations in an incomplete market setting. Section 3 characterizes the effect of sentiment on household consumption-saving behavior. Section 4 provides suggestive evidence of diagnostic expectations with respect to income and discusses calibration. Section 6 present steady state implications of household's sentiment. Section 7 concludes. Appendix A characterizes the effect of sentiment on households' MPCs.

# 2 A partial equilibrium model of sentiment

This section introduces a simple incomplete market model where agents feature diagnostic expectations for their idiosyncratic income. We then derive analytically and numerically the main implications for household's consumption-saving behavior in partial equilibrium.

## 2.1 Model set up

**Earnings dynamics** Time $t$ is continuous. Households are endowed with an idiosyncratic flow of productivity $e^{y_t}$ that they supply to firms inelastically against a wage $w$, earning a total labor income $we^{y_t}$. As is standard in the quantitative heterogeneous agents literature, individual productivity follows a jump-drift process in logs. Jumps arrive at a Poisson rate $\lambda$. Conditional on a jump, a new log-productivity state $y'$ is drawn from a Normal distribution with mean zero and variance $\sigma^2$, i.e., $y' \sim \mathcal{N}(0, \sigma^2)$. Between jumps, the process mean-reverts exponentially at rate $\mu y_t$ . Formally, the process for $y_t$ can be represented as[2]

$$dy_t = -\mu y_t dt + dN_t \tag{1}$$

where the shock $dN_t$ captures the income change from $y$ to $y'$. The reason why we don't rely on a more simple two-state Markov chain is that diagnostic expectations, described momentarily, is more amenable to a setting in which the support for productivity is continuous.

**Diagnostic expectations and Sentiment** We extend the definition of sentiment presented in Maxted (2023a) to the jump-drift process in (1). Households' perception of the drift of log productivity is biased by a "sentiment" variable $\mathcal{S}_t$ which captures recent income shocks. Intuitively, household who received a negative sequence of income shocks will feature negative sentiment and will thus underestimate their future income (and vice-

[2] More formally, the infinitesimal generator of this process is given by $\mathcal{I}v(y) = -\mu y \partial_y v(y) + \lambda \int v(y') - v(y) d\Phi\left(y'/\sigma\right)$, where $\Phi(\cdot)$ is the CDF of the normal distribution.

versa). Formally, households' perceived log-productivity process is given by

$$\widetilde{dy_t} = \Big( -\mu y_t + \theta \mathcal{S}_t \Big) dt + dN_t, \quad \mathcal{S}_t \equiv \int_{-\infty}^{t} e^{-\kappa(t-s)} dN_s. \tag{2}$$

where $\widetilde{dy_t}$ denotes the *perceived* law of motion for $y_t$, as opposed to the real one. Sentiment is defined as a weighted sum of past log-productivity shocks, with an exponential discounting parameter $\kappa$.[3] In this sense, $\kappa$ can be understood as a "traumatic memory" parameter which captures the extent to which past shocks persistently affect households' psychology. The parameter $\theta$ captures the degree of diagnosticity, with $\theta = 0$ nesting rational expectations.

A convenient property of this formulation is that the law of motion for sentiment can be expressed recursively and hence we can get the joint law of motion for log productivity and sentiment:

$$dy_t = -\mu y_t dt + dN_t, \quad d\mathcal{S}_t = -\kappa \mathcal{S}_t dt + dN_t.$$

From the equations above, it is clear that the processes for income and sentiment are correlated as they are hit with the same shocks $dN_t$. In particular, the infinitesimal generator for the joint process $(y_t, \mathcal{S}_t)$ is given by

$$\mathcal{B}v(y, \mathcal{S}) = -\mu y \partial_y v(y, \mathcal{S}) - \kappa \mathcal{S} \partial_{\mathcal{S}} v(y, \mathcal{S}) + \lambda \int v(y', y' - y + \mathcal{S}) - v(y, S) d\Phi \left( y'/\sigma \right).$$

Where the argument $y' - y + \mathcal{S}$ captures the fact that $\mathcal{S}$ is shocked by the change in income $y' - y$.

**Household's problem** To study the implications of sentiment on the consumption-saving behavior of households, we embed the joint process for log productivity and sentiment into an otherwise standard incomplete markets framework. Throughout the rest of the paper, we focus on partial equilibrium behavior. Households solve the income fluctuation problem by choosing how much to consume and save in a single risk-free asset $a$ yielding a return of $r$. Agents face a standard borrowing limit $\underline{a}$. The household's problem is:

$$\max_{\{c_t\}_{t \geq 0}} \ \widetilde{\mathbb{E}}_0 \ \int_0^{\infty} e^{-\rho t} u(c_t) dt \quad s.t. \quad \dot{a}_t = r a_t + w e^{y_t} - c_t, \quad a \geq \underline{a}$$

[3] Since $dN_s$ is a jump process, the integral boils down to a simple sum of shocks. In particular let $\{\ldots, t_2, t_1\}$ be the timing of past realized shocks and $\{\ldots, dN_2, dN_1\}$ the corresponding income changes. Sentiment then simply reads $\mathcal{S}_t = \sum_{s=1}^{\infty} e^{-\kappa t_s} dN_s$.

where the only deviation from the standard formulation is in the expectation operator $\widetilde{\mathbb{E}}_0$, which captures households' deviations from rational expectations.

## 2.2 Recursive representation: the rationality wedge

We now present a general recursive representation of non-rational expectations, captured by a linear operator which we call the "rationality wedge". This wedge is able to handle a wide range of deviations from rational expectations in heterogeneous agents models with an arbitrary number of states in a tractable way. In particular, by leveraging the rationality wedge, we can extend diagnostic expectations to general classes of stochastic processes, thus going beyond the AR(1) case which the literature has typically focused on.

**Hamilton-Jacobi-Bellman (HJB) equation** In addition to assets $a$ and log-productivity $y$, we include sentiment $\mathcal{S}$ as a third state variable. We then derive a representation of deviations from rational expectations as a wedge in the agent's HJB equation. We refer to this wedge as the "rationality wedge" and denote it by $\Psi(\mathcal{S})$. This formalization of the deviations makes computations more transparent and can be generalized to a wide range of deviations from rational expectations.

As is standard in the behavioral economics literature, we can distinguish between two types of agents: sophisticates and naïve. Sophisticates understand that they have sentiment and internalize its true law of motion. Naïve agents, on the other hand, are not aware that they misperceive the true law of motion for their income process, and as a result also ignore the law of motion for sentiment. The value function of the sophisticated household solves the following HJB equation:[4]

$$\rho V(a,y,\mathcal{S}) = \max_c u(c) + \underbrace{(ra + we^y - c)\partial_a V(a,y,\mathcal{S})}_{\text{propagation in } a} + \underbrace{\mathcal{B}V(a,y,\mathcal{S})}_{\text{propagation in } (y,\mathcal{S})} + \underbrace{\theta\mathcal{S}\partial_y V(a,y,\mathcal{S})}_{\text{rationality wedge}} \tag{3}$$

The rationality wedge $\Psi(\mathcal{S})V(a,y,\mathcal{S}) = \theta\mathcal{S}\partial_y V(a,y,\mathcal{S})$ captures the fact that households misperceive the drift of their income by a factor $\theta\mathcal{S}$. Apart from this misperception of the income drift, and because of sophistication, they understand the true joint law of motion of income and sentiment, captured by $\mathcal{B}$.

Naïve agents do not internalize the movement of $\mathcal{S}$ when solving their consumption-

[4] The HJB (4) is subject to the standard boundary condition $\partial_a V(\underline{a},y,\mathcal{S}) \geq u'(r\underline{a} + we^y)$.

saving problem. Their HJB equation is then given by:

$$\rho V(a,y,\mathcal{S}) = \max_c \quad u(c) + \underbrace{(ra + we^y - c)\partial_a V(a,y,\mathcal{S})}_{\text{propagation in } a} + \underbrace{\mathcal{B}V(a,y,\mathcal{S})}_{\text{propagation in } (y,\mathcal{S})}$$
$$+ \underbrace{\mathcal{I}V(a,y,\mathcal{S}) - \mathcal{B}V(a,y,\mathcal{S}) + \theta\mathcal{S}\partial_y V(a,y,\mathcal{S})}_{\text{rationality wedge}} \tag{4}$$

The rationality wedge of the naïve agent additionally captures the fact that they don't perceive the true joint motion of $y$ and $\mathcal{S}$ captured by $\mathcal{B}$, which has to be removed, and instead only perceive the evolution of $y$, captured by its generator $\mathcal{I}$.

In this paper, we will focus the analysis on the naïve case. There are two reasons for this. First, naïveté is more intuitive: agents simply misunderstand the evolution of their income. Second, the naïve policy functions have the nice feature that when sentiment is zero they exactly boil down to the rational expectations case. This is instead not true for the sophisticated agent who realizes that they have sentiment, which distorts their behavior even when sentiment is zero. We performed all the analysis under sophistication and the results are very similar in our calibration.

**Distributional dynamics** Just like income and wealth are distributed in standard incomplete market models, sentiment is unequally distributed between agents and the state of the economy will be captured by a joint measure $G(da, dy, d\mathcal{S})$ over these three states. To simplify notation, let $x \equiv (a, y, \mathcal{S})$ denote the vector of state variables indexing a household. Now denoting by $g(x)$ the joint density over the three states, we can represent the economy as a stationary mean-field-game in the following way:

$$\rho V(x) = \max_c \quad u(c) + (ra + we^y - c)\partial_a V(x) + \mathcal{B}V(x) + \underbrace{\Psi(\mathcal{S})V(x)}_{\text{rationality wedge}} \tag{5}$$

$$0 = -\partial_a[s(x)g(x)] + \mathcal{B}^* g(x) \tag{6}$$

where $s(x) \equiv ra + we^y - c(x)$ denotes optimal savings solving equation (5) and $\mathcal{B}^*$ is the adjoint of the operator $\mathcal{B}$.[5] This system differs from standard rational expectations mean-field-games insofar as the differential operator in the HJB equation is usually the infinitesimal generator of the true stochastic process and hence the differential operator in the Kolmogorov Forward equation is the adjoint of the differential operator in the HJB equation. The rationality wedge violates this duality. Note also that deviations from ratio-

[5]The adjoint $\mathcal{B}^*$ is the continuous-state continuous-time equivalent of a discrete-state transition matrix, hence the adjoint of an operator is akin to the transpose of a matrix, see Achdou et al. (2022).

nal expectations can impact the true state dynamics only via the optimal decisions made by households $s^\star(x)$.

**Belief distortions with sentiment shocks** The rationality wedge can also be used to compute the *perceived* statistical evolution of the agent's states, which allows us to understand the way individuals' belief distortions evolve over different forecast horizons. Suppose the agent starts with some states $x_0$. Their perceived evolution of future states, which we capture with the evolution of the perceived density of $x_{t+\tau}$ denoted by $\widetilde{f}(\cdot|x_0)$, is obtained from a Kolmogorov Forward equation featuring the adjoint of the rationality wedge :

$$\partial_t \widetilde{f}_t(x|x_0) = -\partial_a[s^\star(x)\widetilde{f}_t(x|x_0)] + \mathcal{B}^* \widetilde{f}_t(x|x_0) + \Psi(\mathcal{S})^* \widetilde{f}_t(x|x_0) \tag{7}$$

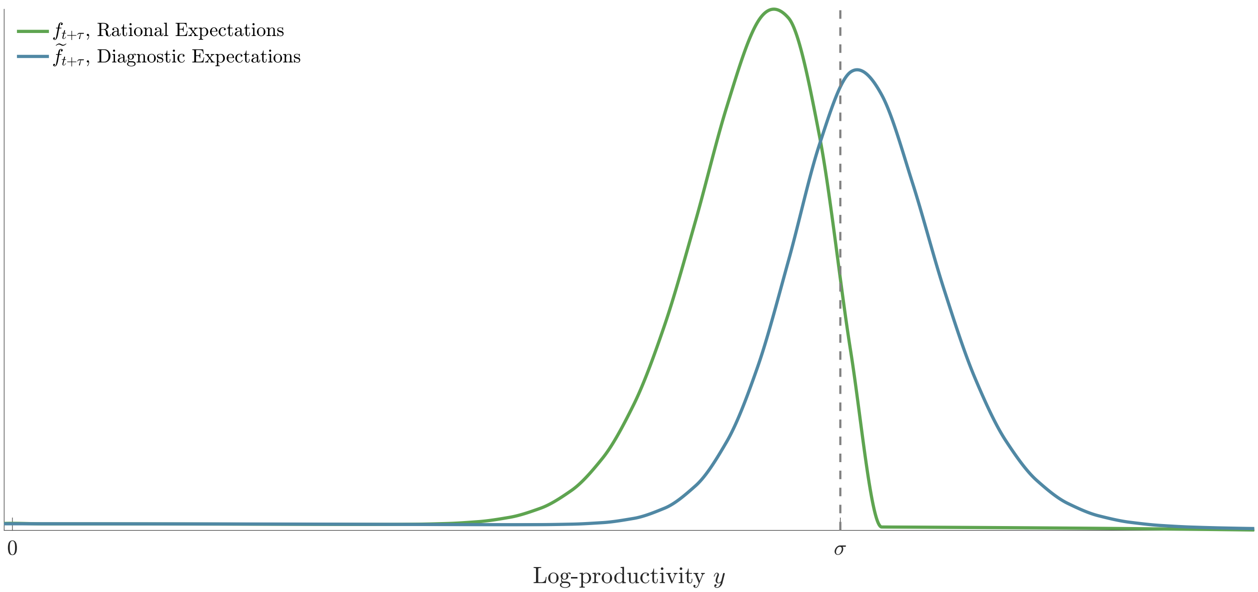
subject to the boundary condition $\widetilde{f}_0(x|x_0) = \delta\{x_0\}$, where $\delta\{\cdot\}$ is the Dirac delta function. We can use equation (7) in two ways. First, for any initial state $x_t$, we can compute the true statistical evolution of $x_{t+\tau}$, denoted by $f(x_{t+\tau}|x_t)$, by solving the equation forward when $\Psi^*(\mathcal{S}) = 0$, which corresponds to the rational expectations case. Second, we can compute household's expected income in period $t+\tau$ under diagnostic expectations $\widetilde{\mathbb{E}}_t(y_{t+\tau})$ given its states at the forecast date $t$ using

$$\widetilde{\mathbb{E}}_t(y_{t+\tau}|x_t) = \int y \widetilde{f}_{t+\tau}(x|x_t) dx. \tag{8}$$

Using equation (7), we can visualise the way diagnostic agents' expectations depart from rational agents'. Figure 1 shows how diagnostic expectations distorts the agent's beliefs following an income shock. We consider a one standard deviation positive shock to log-productivity, starting with the median log-productivity $(y = 0)$ and no sentiment. Instantaneously after the news, the agent forecasts the future evolution of its log-productivity for each future period $t + \tau$. The green density corresponds to the rational (and therefore accurate) forecast for log-productivity two years after the shock.[6] The green line shows how diagnostic expectations distorts beliefs. Like in standard diagnostic expectations, the agent over-weights the probability of states that were made more likely after the shock, and in particular here perceives a probability distribution that is over-optimistic. While the diagnostic expectations framework has so far typically been applied to AR(1) processes for tractability, our rationality wedge approach allows to generalize the intuition of diagnostic expectations to very general stochastic processes such as the jump drift process we're using.

[6] We use two years in this example to align with our data, described in Section 3.

Figure 1: Belief distortion under diagnostic expectations




*Note: This plot shows the subjective expected distribution of log-productivity $y_{t+\tau}$, where $\tau$ corresponds to two years. $\widetilde{f}_{t+\tau}$ is based on diagnostic expectations (blue line) and $f_{t+\tau}$ is based on rational expectations (green line). The beliefs are formed right after a positive shock of magnitude $\sigma$ at $t$, with initial conditions $y = \mathcal{S} = 0$.*

# 3 The effect of household sentiment on consumption behaviour

We can now describe the effect of sentiment on households' consumption-saving behaviour. We first show how sentiment, as a new state variable, distorts the agent's policy functions, and then focus on the dynamics of sentiment to study the consumption response to income shocks.

## 3.1 Sentiment as a state

In this section we characterize how sentiment distorts the consumption-saving behavior of households in partial equilibrium. When agents are naïve, all the results in this section are independent from the law of motion of sentiment itself.

**Distortion on the Euler equation** Proposition 1 provides a tractable formulation of households' Euler equation in the presence of sentiment:

**Proposition 1.** *Consumption obeys the following Euler equation*

$$\mathbb{E}_t \quad \frac{du'(c(x_t))/dt}{u'(c(x_t))} = \left[\rho + \frac{\theta \mathcal{S}_t \times \eta(x_t)}{IES(x_t)}\right] - r \tag{9}$$

*where* $\mathbb{E}_t$ *is the rational expectations operator over a and y,* $IES(x_t) \equiv -u'(c(x_t))/(c(x_t)u''(c(x_t)))$ *and* $\eta(x_t)$ *is the income elasticity of consumption* $\eta(x) \equiv \partial \log c(x)/\partial y$.

*Proof.* Appendix B.1 □

The left hand side of equation (9) is the expected growth rate in marginal utility holding sentiment fixed.[7] The right hand side has the standard $\rho - r$ term, but with the perceived rate of return on savings being distorted by sentiment. When $\theta = 0$ we nest rational expectations for any value of $\mathcal{S}$. Similarly, independently of the value of $\theta$ , when sentiment $\mathcal{S} = 0$, the distortion vanishes and the agent behaves as if they were rational. When sentiment is positive, however, the perceived utility return on wealth is depressed and the saving motive is dampened. As a result, positive sentiment leads to more consumption, and vice-versa. This is intuitive: when agents expect their income to rise from a lifetime prospective, they consume more. This mechanism is captured by the other terms in the wedge. The extent to which sentiment distorts the Euler equation depends on the income elasticity of consumption $\eta$. When sentiment is equal to $\mathcal{S}$, agents wrongly expect their labor income to go up by $\theta \mathcal{S} dt$ percent over $dt$ units of time. This increase in income should lead to an increase in future consumption by $\theta S_t \times \eta(x_{t+dt})dt$ percent in the next period, which depresses the marginal utility of consumption in the future, thus reducing the saving motive today depending on the inter-temporal elasticity of substitution.

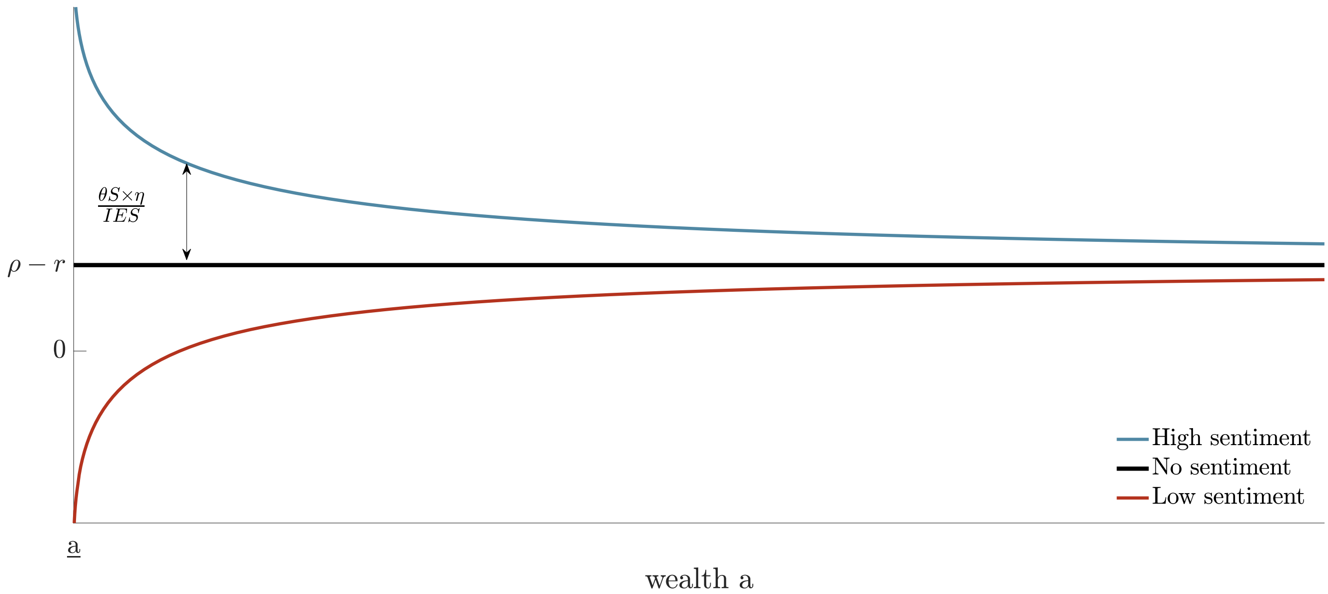
An important implication of (9) is that the right hand side of the Euler equation is state dependent when markets are incomplete, as $\eta$ depends on the household's states. In particular, in our framework, the distortions are decreasing in wealth, since the income elasticity of consumption is higher when households are closer to the borrowing limit. Figure 2 illustrates these distortions for a given level of productivity $y$. The black line corresponds to the standard right hand side of the rational Euler equation, $\rho - r$. When sentiment is positive (negative), the right hand side of the Euler Equation is inflated (depressed), and heterogeneously so depending on wealth as the blue and red lines show in Figure 2. This intuition also leads to the following lemma.

**Lemma 1.** *Suppose that log-productivity is constrained to be in* $[\underline{y}, \bar{y}]$ *and sentiment is constrained to be in* $[\underline{\mathcal{S}}, \bar{\mathcal{S}}]$. *Then when* $r < \rho$ *and with CRRA utility sentiment and income do not matter at the top of the wealth distribution and in particular as* $a \to \infty$

$$s(a, y, \mathcal{S}) = \frac{r - \rho}{\gamma} a \tag{10}$$

[7] $\mathcal{S}$ is held constant in this Euler equation because the naïve agent does not realize that sentiment moves over time. The Euler equation of the sophisticated agent looks identical to (9), with the exception that the expectation operator is now taken over $\mathcal{S}$ as well.

Figure 2: Euler equation distortions along the wealth distribution




*Note: The calibration used is our benchmark calibration. The graph is plotted fixing a value of log-productivity y for illustration.*

*where $\gamma$ is the inverse IES.*

Intuitively, at the top of the wealth distribution, not only is the income elasticity of consumption $\eta$ low, but also labor income plays a minor role in total income. As a result, agents' misperceptions about labor income are just as irrelevant as its fluctuation.[8] This is a special case of Proposition 2 in Achdou et al. (2022). This observation motivates us to focus the rest of the analysis on agents close to the borrowing limit.

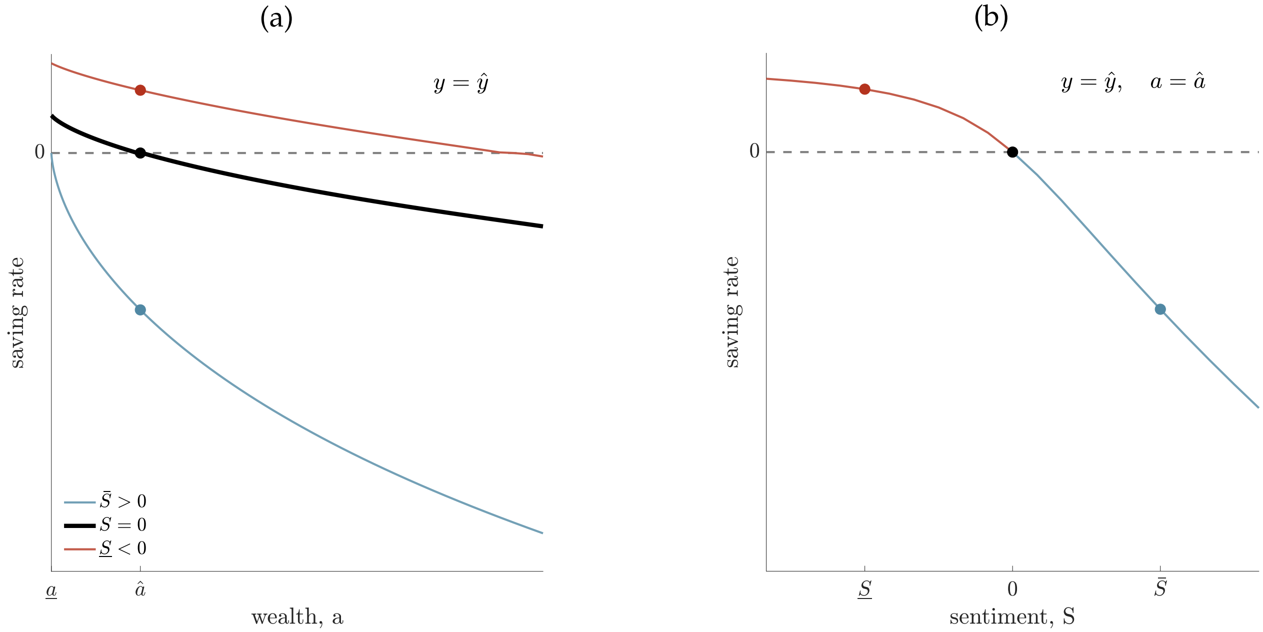
**Effect on policy functions** To go beyond analytical results we solve the model numerically and derive the policy functions.[9] Figure 3 plots the saving rate policy functions, first over assets $a$ in panel 3a and then over our additional state variable sentiment $\mathcal{S}$ in panel 3b. We define the saving rate as the flow of savings $\dot{a}_t$ divided by the total flow of income $ra_t + we^{y_t}$. A negative saving rate implies the agent is dissaving. Each panel plots the policy functions for a specific value of log-productivity, denoted by $\hat{y}$.

In panel 3a, the black curve depicts the policy function for the case with zero senti-

[8]Note that introducing a risky asset (as in, for example, Benhabib et al. (2015)) would break this result. In particular, in the case in which idiosyncratic risk comes from capital income, the distortion on the Euler Equation would be equal to $\frac{\theta \mathcal{S}_t \times a \times MPC(x_t)}{IES(x_t)}$ where $MPC$ represents the derivative of the consumption function with respect to wealth. This extensions is outside of the scope of this paper, but we see this as a promising direction for future work.

[9]Section 4 describes our numerical approach and calibration in more detail and Section 6 derives the quantitative implications.

Figure 3: Saving rate policy function



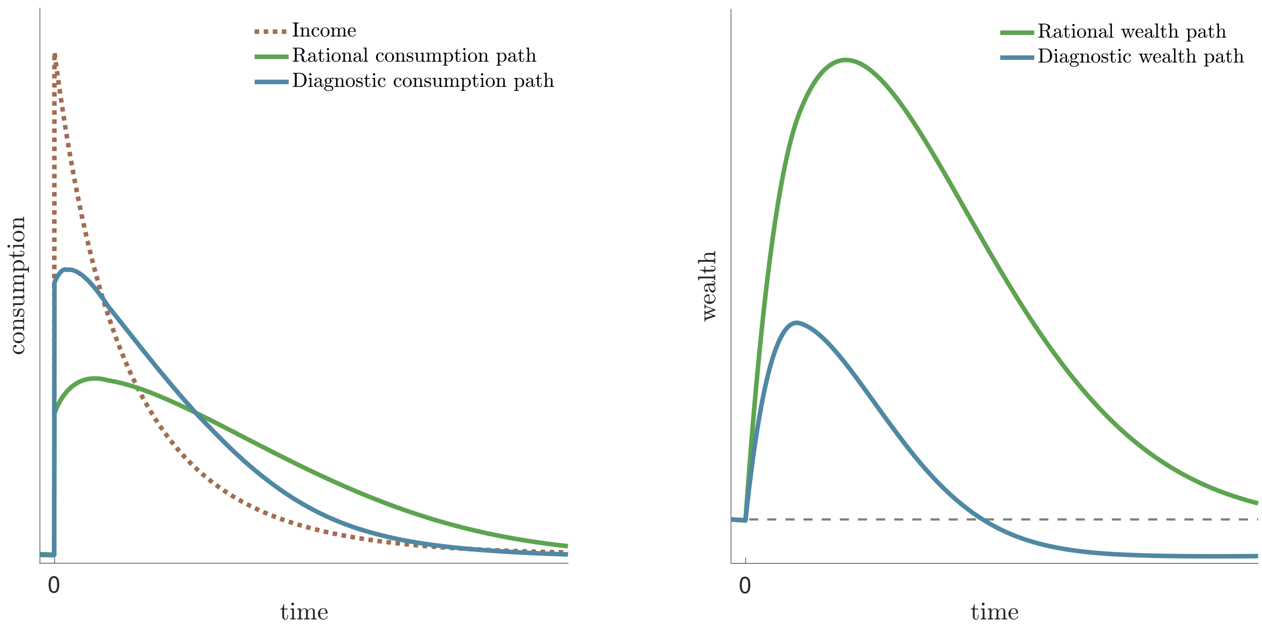

*Note: The dots represent the saving rate for a given level of wealth $\hat{a}$ and income $\hat{y}$ and different levels of sentiment $\mathcal{S} \in \{\underline{\mathcal{S}}, 0, \bar{\mathcal{S}}\}$*

ment, which, as discussed before, coincides with the policy function under rational expectations, since we are focusing on the naïve case. The blue line illustrates the effects of positive sentiment on the agent's decisions. Positive sentiment leads households to consume more than they would under full rationality, because they are now over-optimistic about their future income prospects. In the particular case we are consuming, positive sentiment pushes the agent to a hand-to-mouth state with zero savings, when they are at the borrowing limit. The red line, instead, depicts the case of negative sentiment. Quite intuitively, the effect on the agents' behavior are opposite to the previous case. Because households are now overly pessimistic about the future evolution of their income, they save more today. Panel 3a plots the savings rate for different levels sentiment, keeping fixed the asset state at $\hat{a}$, which we report in panel 3a for reference. We can see how there are strong *non-linearities* in the role that sentiment plays for households' saving decisions. In particular, while negative level of sentiment only induce mild distortions on agents' behavior, positive sentiment has large effects on the household's choice. These non-linearities will play an important role in the analysis in Section 7.

## 3.2 Consumption with sentiment dynamics

After having analyzed the role of sentiment on households' behavior in the state space, we now turn our attention to the sequence space. In particular, because sentiment reacts instantaneously to income shocks, it turns out to have interesting implications for agents' consumption-saving *dynamics*. Figure 4 illustrates this point, by showing the effects of a positive income shocks on the path of consumption and savings when agents have diagnostic expectations. In particular, we consider a positive shock to log-productivity –here equal to one standard deviation of the typical income shock $\sigma$– which takes place at time 0. To ease the exposition, we shut down all shocks to the income path after time 0.[10]

Figure 4: Consumption and wealth dynamics




*Note: We start at $y_0 = 0$ and $S_0 = 0$ and wealth equal to 23% of the average wealth in the stationary distribution (this is so that the consumption path before the shock is flat, for exposition purposes). We shock log-productivity at time 0. We shut down all the other shocks over time to isolate the effects of the shock. We use our benchmark calibration described in Table 1.*

The income path is represented by the dotted brown line in Figure 4. The green line then depicts the paths for consumption and savings that would be chosen under rational expectations. It can be seen how consumption jumps upon impact, then still increases slowly for some periods as the agent builds assets, before decreasing continuously as income keeps mean reverting. Savings display a similar path, as the household engages

[10] Note that the shock we are considering is different from standard "MIT shocks". In fact, in our case agents do not have perfect foresight about the future path of their idiosyncratic income and hence still face risk. This is the reason why in Figure 4 consumption doesn't immediately jump to its maximum before mean-reverting.

in intertemporal smoothing, but with a smoother pattern. When agents are diagnostic, the positive jump in sentiment induced by the shock leads the agent to overreact on impact, thus increasing consumption by more and savings by less than they would under rational expectations. The reason for this over-reaction is that the diagnostic household perceives their income to be higher in the future than what it actually will be. For this reason, they accumulate less assets to be used for future consumption smoothing, and thus consume more today. Over time, however, sentiment reverts back to zero and the consumption path of the diagnostic agent progressively reverts to the rational one. In fact, the diagnostic agent's consumption level eventually goes below that of the rational agent. This is because during the "enthusiasm phase", the diagnostic household over-consumes, thus depletes their assets. As sentiment fades, the household is left with less wealth than it would have in the rational counterfactual, and hence ends up consuming less. This illustrates how in our framework short run over-consumption translates into under-consumption in the long run, as inter-temporal mistakes propagate through time via wealth adjustments.

# 4 Evidence of sentiment in survey data

In this section, we provide evidence that households have diagnostic expectations when forecasting their own future income. We show this by relying on data from the Survey of Household Income and Wealth. We then rely on this empirical evidence to calibrate the psychology parameters of our model.

## 4.1 Data

Our data come from the Survey of Household Income and Wealth (SHIW), which is run biannually by the Bank of Italy.[11] The SHIW is a representative survey of Italian households, featuring a rotating panel component. It includes detailed and disaggregated data on households' income, assets, and liabilities, and also provides information on consumption and saving behavior, as well as demographic characteristics. Crucially for our analysis, the 2012 and 2014 waves of the survey also asked respondents to report their income expectations for the following year.[12] Throughout the rest of our empirical analysis, the focus is going to be on these two waves. In addition, we also obtain data on realized income for the neighboring 2010 and 2016 waves. Overall, a total of 4,140 and 8,156 households

[11]This survey has been used extensively in the literature, see for example Jappelli and Pistaferri (2014, 2020), Auclert (2019).

[12]For the exact wording of the questions, see Appendix D.1.

reported their income expectations in the 2012 and 2014 waves respectively. However, once we restrict the sample to those respondents who also appear in the following wave of the survey –in order to be able to compare expected and realized income– the effective size of our sample shrinks to 1,288 households in 2012 and 2,038 in 2014.

Respondents' expectations for future individual income were also elicited in the 1989 and 1991 waves of the survey. In Appendix D.2 we show that all our results hold when we consider this different time period.

**Income Expectations** Because of the biannual nature of the survey, there is always a one year gap between our data on income expectations and realizations. For example, in the 2014 wave of the survey respondents were asked to report their expected income for the year 2015, but data on realized income are available only for the next survey wave, conducted in 2016.[13] To address this issue, we construct our measure of time $t$'s expectations for income in $t+2$ –which we denote by $y^e_{i,t+2}$– by extrapolating expectations for income growth between $t$ and $t+1$. In particular, households are asked to report their expected income growth from year $t$ to $t+1$, which we define as $g^e_{i,t+1|t} \equiv \frac{y^e_{i,t+1}}{y_{i,t}}$, where $y_{i,t}$ denotes household $i$'s realized income in year $t$. We then construct income expectations for year $t+2$ as:[14] $y^e_{i,t+2} = y_{i,t} \times \left(g^e_{i,t+1|t}\right)^2$.

**Realized Income** For each household $i$ and year $t$ in our sample we also collect data on total net realized income, which is defined as the sum of income from labor, pension and transfers, self-employment, and capital net of income taxes. We denote this variable by $y_{i,t}$. In the rest of our analysis, we drop retired households, which we define as those households having zero labor and self-employment income, but positive income from pension and transfers. However, all of our results still hold when we consider the full sample, or when we drop self-employed households.

Because the 1989 and 1991 waves of the survey asked respondents about their expected future *individual* labor income, $y_{i,t}$ denotes individual labor income when we analyze these years in Appendix D.2.

**Forecast Errors** Armed with data on both realized and expected income, we are now ready to construct our main variable of interest. In particular, we define household's

[13]Note that this feature is also shared by Rozsypal and Schlafmann (2023). In their setting, respondents are asked to report expected income for the next year, but data on realized income are available only for 6 months ahead income.

[14]Note that our results are virtually unchanged if we just assume that $y^e_{i,t+2} = y_{i,t} \times \left(g^e_{i,t+1|t}\right) \equiv y^e_{i,t+1}$.

$i$ forecast error in period $t$ as the percentage difference between realized and expected income in year $t+2$, that is:

$$FE_{i,t} = \frac{y_{i,t+2}}{y^e_{i,t+2}} - 1 \tag{11}$$

Thus, according to our definition, households who make positive forecast errors turn out to be overly pessimistic when predicting their future income, and vice-versa. Note that because of the way we define it, our forecast error variable is bounded below by $-1$, but is unbounded above. Throughout the rest of our analysis, we thus trim our forecast error variable at the 98$^{\text{th}}$ percentile.[15]

**Other Data** Finally, we also collect data on households' area of residence, number of components, and net wealth –defined as the total value of real estate, businesses, valuables, and financial wealth owned, net of debt and mortgages– as well as age, sex, educational attainment, occupation, and sector of employment of the main respondent within the household. Throughout our analysis, we always report results weighted by survey weights.

## 4.2 Evidence of Over-Extrapolation in Households' Income Expectations

We now document three motivating facts on households' income expectations formation process. Our focus is on the forecast error households make when predicting their future income, as defined in (11).[16]

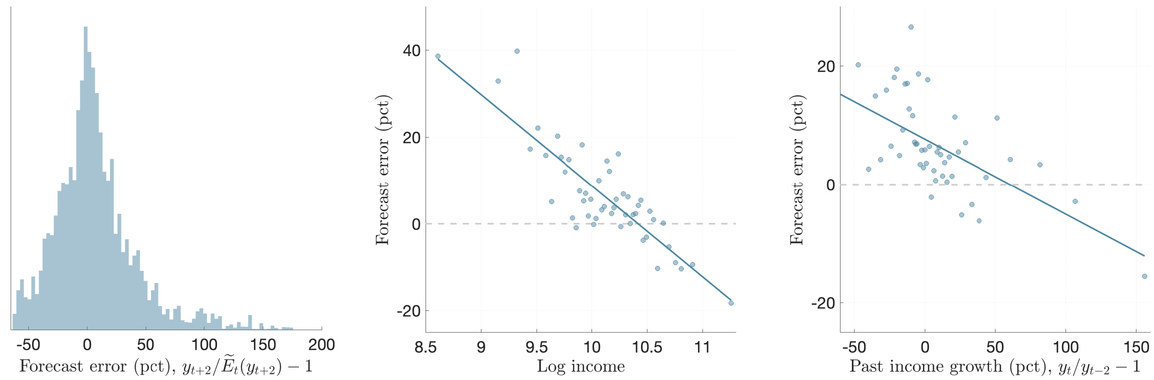
The first panel of Figure 5 shows that there is large dispersion in the errors households make when forecasting future income. Note that this fact alone is not in direct contradiction with rational expectations. In fact, even in the rational expectations benchmark, the cross-section of forecast errors follows a distribution with some non-zero dispersion. In particular, under rational expectations, the distribution of forecast errors would simply mimic that of idiosyncratic income shocks.

Second, households' forecast errors are negatively correlated with households' income. This is shown in the middle panel of Figure 5, where we plot a binned scatter plot of the logarithm of household income against their forecast error. We residualize both axis by time fixed effects, number of members of the household, area of residence, as well as

[15] Note that the particular cut-off for trimming data is immaterial for our results, as we show in Appendix D.2.

[16] In Appendix D.2 we show that all our facts also hold when we consider the 1989 and 1991 waves of the survey, in which questions about expected future income were also asked.

Figure 5: Three motivating facts on households' income expectations




*Note: Forecast errors are computed according to (11). The second panel displays a binned scatter plot of $FE_t$ against $\log(y_t)$, where we define current income $y_t$ as total household income net of taxes and capital income. Results are unchanged if we consider total household income including capital income. The third panel shows a binned scatter plot of $FE_t$ against the past income change. We group observations in 75 bins. In both panels we residualize the x and y axis by time fixed effects, number of members of the household, area of residence, as well as age category, sex, and educational attainment of the respondent. In the third panel we also control for household's income quintile. We trim forecast errors at the $98^{th}$ percentile.*

age category, sex, and educational attainment of the respondent.[17] We find that high income households tend to be overly optimistic when forecasting future income, while low income ones tend to be excessively pessimistic. This fact is at odds with the predictions of the rational expectations hypothesis. In fact, under rational expectations forecast errors should be unpredictable. This result is in line with previous evidence for the US context based on the Michigan Survey of Consumers (Rozsypal and Schlafmann, 2023).

Third, and finally, we document that households' errors in predicting future income not only correlate with income *levels*, but also with past income *changes*. More precisely, even after controlling for the level of income, households that experienced an income increase in the past tend to be excessively optimistic about the path of their future income, and vice-versa. We show this in the third panel of Figure 5, by means of a binned scatter plot of income growth from year $t-2$ to $t$ against the forecast error $FE_{i,t}$ defined in (11). We include the same controls as for the second panel, but in this case we also control for the household's income quintile. To the best of our knowledge, we are the first to document correlation between past income changes and future forecast errors. Once more, this fact cannot be rationalized by rational expectations.

Taken together, we believe these three facts provide motivating evidence that households' expectations formation process for future idiosyncratic income cannot be completely approximated by rational expectations. In particular, our second and third facts

[17] In Appendix D.2 we show that this pattern is robust to the particular set of controls included, and in particular also holds unconditionally, as well as after controlling for household wealth.

suggest that households tend to over-extrapolate past income changes when forming expectations about the future, a pattern which has already been showed for aggregate variables, see in particular Bordalo et al. (2020). Moreover, insofar as households' income process is at least partly idiosyncratic across households, so that different households experience different histories of past income shocks, this is going to lead to a distribution of households' optimism/pessimism about future income, exactly as documented in our first fact. We now use these last two facts to put some discipline on our theory.

# 5 Calibration of behavioral parameters

We discipline the model using the Italian Survey of Household Income and Wealth (SHIW) described in Section 4. In our baseline, we use the 2012 and 2014 waves, and replicate the exercise for 1989 and 1991 as a robustness check.

**Target moments** We target two moments that identify the two "psychology" parameters $(\theta, \kappa)$:

**Slope of FE on income:** We match the coefficient from regressing the two-year forecast error on current log income in the model and the data. Specifically, we estimate the OLS slope $\beta_y$ from

$$FE_{i,t} = \alpha_y + \beta_y \log y_{i,t} + u_{i,t},$$

where in the data both variables are residualized with respect to time fixed effects, household size, area of residence, age category, sex, and educational attainment of the respondent.

**Autocorrelation of FE across waves:** We also match the coefficient from regressing the forecast error on its two-year lag in the model and the data:

$$FE_{i,t} = \alpha + \rho_{fe} FE_{i,t-2} + u_{i,t},$$

estimated on the 2012–2014 and 1989–1991 panels, which allow us to compute forecast errors. As before, in the data we residualize forecast errors using the same set of control variables.

Intuitively, $\theta$ governs the extent to which sentiment $S_{i,t}$ biases beliefs (the level of distortion), while $\kappa$ determines the *persistence* of sentiment and thus of forecast errors.

Given parameter values, we simulate a long time series of income, sentiment, and forecast errors for households. For each simulated household at time $t$, the model delivers a perceived distribution for future income, $\tilde{f}(y_{t+2} \mid s_t, y_t)$ and a true realization $y_{t+2} \sim$

$f(\cdot \mid s_t, y_t)$. This allows us to compute the expected income under the distorted perceived process and the (naive) agent's expected income. We then use the minimum distance method to calibrate $\kappa$ and $\theta$ by matching the target moments described above.

## 5.1 Calibration of structural parameters

We solve the model numerically using the finite-differences scheme proposed in Achdou et al. (2022).[18] Table 1 provides the parameter values we use for the calibration. We calibrate the model to a quarterly frequency.

Table 1: Model Calibration

| **Parameter** | **Description** | **Value** | **Justification** |
|---|---|---|---|
| *Preferences* | | | |
| $\rho^{DE}$ | Discount rate (p.a.) | 3.9% | Match wealth to income ratio |
| $\rho^{RE}$ | Discount rate (p.a.) | 4% | Match wealth to income ratio |
| $\gamma$ | Inverse IES | 1 | Standard |
| *Diagnostic Expectations Parameters* | | | |
| $\theta$ | Diagnosticity | 4% | Calibrated |
| $\kappa$ | Decay of New Information | 0.25% | Calibrated |
| *Income process* | | | |
| $\lambda$ | Arrival rate of income shocks | 3.46% | Kaplan et al. (2020) |
| $\mu$ | Mean reversion rate of income | 3.48% | Kaplan et al. (2020) |
| $\sigma$ | Standard deviation of shocks | 0.736 | Match standard deviation of log-income |
| *Other structural parameters* | | | |
| $r$ | Interest rate | 1% | Kaplan and Violante (2022) |
| $w$ | Wage | 0.82 | Normalize mean income to 1 |
| $\underline{a}$ | Hard borrowing limit | 0 | Kaplan and Violante (2022) |

*Note: All parameters are expressed at quarterly frequency unless indicated otherwise.*

**Income** We follow the procedure in Kaplan et al. (2018) and calibrate the parameters of the income process by matching three empirical moments: (i) the variance of log earnings, (ii) the variance of two-year log changes in earnings, and (iii) the kurtosis of two-year log changes in earnings. To construct the empirical counterparts, we use residualized labor income of household heads in the SHIW sample between ages 25 and 55. Residualization is obtained by regressing log labor income on three age-category dummies and a dummy

[18] See Appendix C for a detailed description of our numerical algorithm.

for the sector of employment. The parameters $\sigma$, $\lambda$, and $\mu$ are then calibrated using the minimum distance method.

**Wealth** Following Kaplan and Violante (2022), we calibrate $\rho$ such that the average wealth to average income ratio in our model is equal to the one observed in the data, which is 9.69 in our case. We calibrate $\rho$ separately for the rational and diagnostic model, to ensure that the wealth to income ratio is the same in both cases. However, calibrating both models with a common $\rho$ does not change our results.

**Other structural parameters** As is standard in the literature, we set the intertemporal elasticity of substitution to one. Finally, we follow Kaplan and Violante (2022) we set the partial equilibrium interest rate to 1% per annum and the borrowing limit to 0.

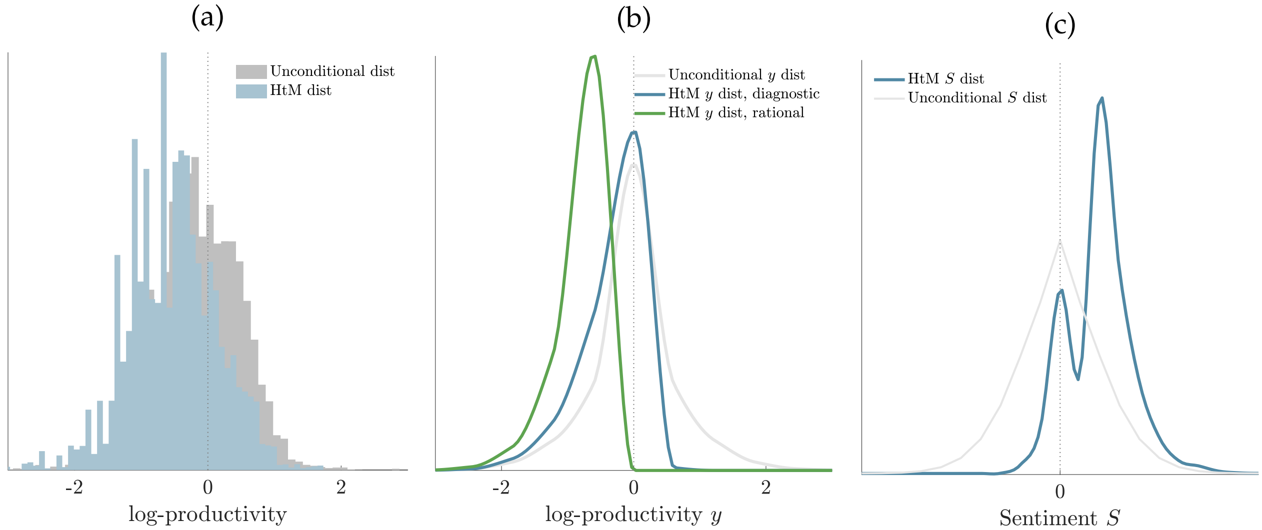
# 6 Steady state household behaviour

In this section, we present three key predictions for steady-state household behavior under our diagnostic expectations framework. First, we show that sentiment amplifies the consumption response to income shocks, thus providing a potential rationalization for the "excess sensitivity" observed in the data. Second, sentiment generates latent heterogeneity in the consumption response to income and wealth shocks, which has been recently emphasized in the empirical literature. Third, our diagnostic expectation framework predicts a larger persistence of the hand-to-mouth state than the rational expectations benchmark. We also show that the stationary distribution of our model features a substantial fraction of hand-to-mouth households, which more closely matches the distribution observed in the data, without the need to rely on the presence of an illiquid asset. Finally, we show that our behavioral friction has non-trivial welfare cost. Moreover, these welfare costs are heterogeneous in the cross-section of households, and are substantially larger for poorer agents.

## 6.1 Anatomy of diagnostic hand-to-mouth households

In this section we show that our economy features a larger mass of hand-to-mouth (HtM) households than in the rational benchmark, thus providing a much better fit of the data. Moreover, diagnosticity makes the HtM state "stickier" in the sense that, compared to the rational benchmark, once diagnostic agents enter the HtM state they have a lower probability of escaping it at any point in time.

Figure 6: Hand-to-mouth composition




*Note: In panel 6a, we compute log-productivity as the data analogue of our model: we divide income by average income and take log. We follow section 6.1 to define the hand-to-mouth households in our model and data (we follow the literature in using liquid wealth for the data definition).*

**The hopeful Hand-to-Mouths** We follow Kaplan et al. (2014) and adopt a standard definition of hand-to-mouth households:

**Definition 1** (Hand-to-Mouth (HtM))**.** *The region of the state space such that the household is HtM is denoted by $\boldsymbol{\mathcal{H}}$ and is the set of states such that household's wealth is less than half their monthly income:* $\boldsymbol{\mathcal{H}} = \{(a, y, \mathcal{S}) | a \leq we^y/6\}$

The stationary distribution of our diagnostic economy features a share of 20.1% of agents in the HtM state as defined in Section 6.1. Despite the fact that we do not target this moment when calibrating our model, this figure is very close to the empirical estimates for the Italian economy. In fact, Kaplan et al. (2014) find that the share of HtM households in the Italian economy is around 20%, using data from the Household Finance and Consumption Survey (HFCS) for the period 2008-2010. When we estimate the share of HtM agents in our data, we find it to be approximately 21%.[19] Note that our rational expectations benchmark model is not able to match the share of HtM households observed in the data. In particular, only 10.2% of agents are HtM in the stationary distribution of this model. In fact, the literature usually resorts to ad-hoc modeling tools in order to match the share of HtM households observed in the data, such as the introduction of illiquid assets (Kaplan and Violante, 2014) or preference heterogeneity (Aguiar et al., Forthcoming).

[19] We define HtM households as those households whose liquid assets are below half of their non-capital monthly income.

It turns out that our model is also able to produce a better empirical fit when it comes to the *composition* –as opposed to the *mass*– of HtM households. Notably, compared to the rational benchmark, HtM households in the diagnostic model exhibit substantially higher average income. This can be seen from Figure 6b, which plots the conditional log-productivity distribution for hand-to-mouth households, comparing the rational case (green) and diagnostic case (blue).[20] The reason for this discrepancy is depicted in Figure 6c: a significant portion of HtM households in the diagnostic model exhibit positive sentiment. These households, which we call the"hopeful" hand-to-mouth, are middle-income households that have experienced extended spells of positive sentiment. As a result, they depleted their wealth and have become financially constrained, without necessarily being in a low income state. Thus, the presence of sentiment gives rise to a mass of middle-income, optimisitc, hand-to-mouth households. This is qualitatively in line with the empirical evidence in our data. In fact, Figure 6a shows the empirical analog of Figure 6b in the SHIW data. We find that around 20% of italian households that are classified as HtM have an income level which is above the cross-sectional average. Once more, our diagnostic model is decently able to match this untargeted moment, with around 29% of HtM agents featuring above average income. In the rational benchmark, on the other hand, virtually all HtM households have below average income.

**Sticky hand-to-mouth** Before discussing the dynamics of the HtM state, it is useful to define the Marginal Propensity to Save (MPS) out of log-productivity shocks $\Delta$.

**Definition 2.** *The Marginal Propensity to Save out of productivity shocks $\Delta$, for a household with state vector $x \equiv (a, y, \mathcal{S})$ over a period $\tau$ is given by*

$$\mathfrak{s}_\tau(\Delta; x) = \frac{Sav_\tau(a, y + \Delta, \mathcal{S} + \Delta) - Sav_\tau(a, y, \mathcal{S})}{e^\Delta - 1} \tag{12}$$

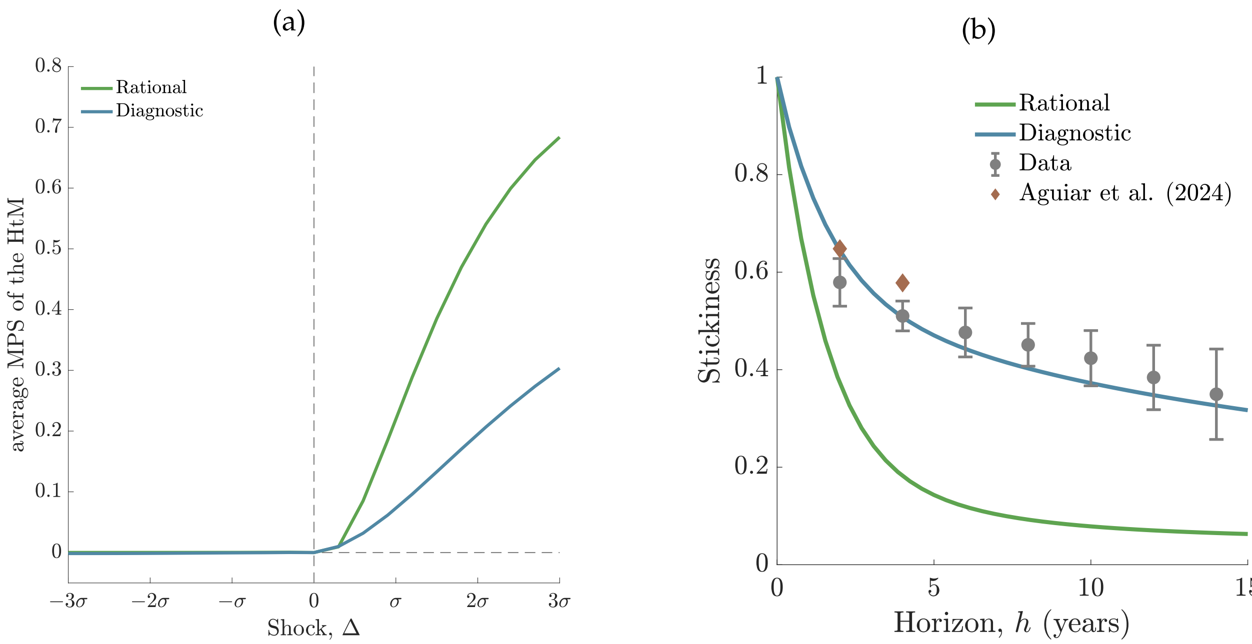
$$\text{where} \quad Sav_\tau(x) = \mathbb{E}_0 \left[ \int_0^\tau s^{DE}(x_t) dt \mid x_0 = x \right] \tag{13}$$

*where $s^{DE}(\cdot)$ is the diagnostic expectations savings policy function and $\mathbb{E}_0$ is the rational expectations operator with respect to all the state variables.*

Consistently with the non-linearities generated by sentiment and emphasized before, our model also predicts asymmetric effects of negative and positive log-productivity shocks on the savings rate of HtM agents. This is displayed in Figure 7a, which plots the MPS

[20] For reference, we also report the unconditional distribution in grey. Naturally, this distribution is unaffected by the level of rationality.

Figure 7: Stickiness




out of income shocks of different sign and size for HtM agents under the rational and diagnostic benchmark. Because we are focusing on HtM households, negative shocks have no effect on the savings rate both under rational and diagnostic expectations, since the agent is constrained both before and after the shock. Positive shocks, on the other hand, may push the agent out of the HtM region, thus inducing them to save out of the shock. When agents are diagnostic, however, the MPC out of sentiment puts downward pressure on this saving motive. In fact, sentiment jumps up in response to the shock and, as a result, agents don't save as much as in the rational benchmark, thus remaining closer to the borrowing limit. The lower MPS out of positive shocks induced by sentiment implies that under diagnostic expectations it is more difficult for agents to escape the HtM state. To quantify this effect, we now define a concept of "stickiness" to evaluate the extent to which agents are "trapped" in the HtM state.

**Definition 3** (Stickiness)**.** *We define the horizon-h stickiness of $\boldsymbol{\mathcal{H}}$, $\mathfrak{S}_h(\boldsymbol{\mathcal{H}}|x)$ as the average probability that a HtM households is still HtM h years in the future:*

$$\mathfrak{S}_h(\boldsymbol{\mathcal{H}}) = \mathbb{E}_x\left[\mathbb{P}(x_h \in \boldsymbol{\mathcal{H}}|x_0 = x)|x \in \boldsymbol{\mathcal{H}}\right]$$

*where the expectation $\mathbb{E}_x$ is taken over the stationary distribution.*

Stickiness thus captures how easy it is to "escape" the hand-to-mouth condition. To characterize $\mathfrak{S}_h(\boldsymbol{\mathcal{H}})$ we once again rely on the true evolution of the statistical distribution of the states $x_t$:

$$\mathfrak{S}_h(\boldsymbol{\mathcal{H}}) = \int_{\boldsymbol{\mathcal{H}}} dF_h(dx) \tag{14}$$

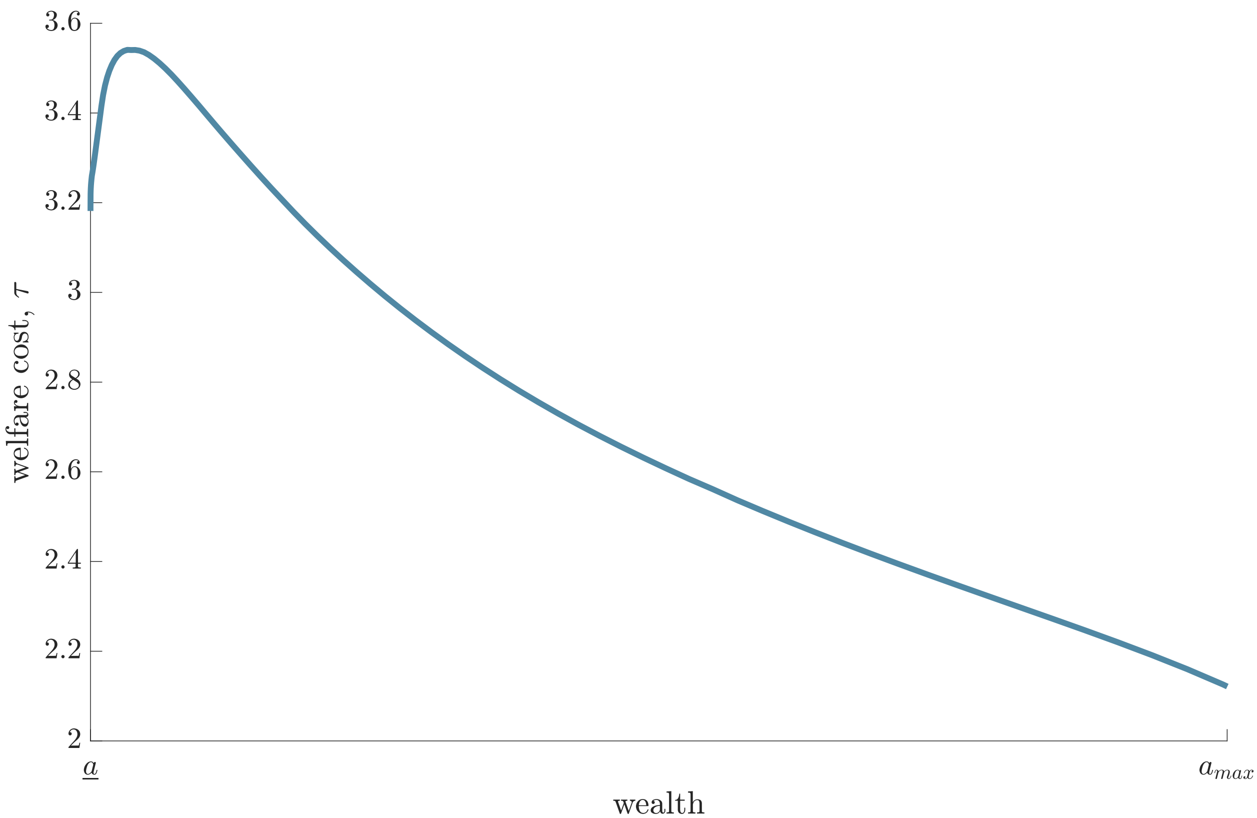
where $dF_h(dx)$ is obtained using equation (7) with boundary condition $F_0(dx) = G(dx)$, where $G(dx)$ is the stationary distribution.[21] Clearly, stickiness for horizon $h = 0$ is equal to 1, i.e., $\mathfrak{S}_0(\boldsymbol{\mathcal{H}}) = 1$. Moreover, $\mathfrak{S}_h(\boldsymbol{\mathcal{H}})$ converges to the stationary mass of HtM households as the horizon $h$ goes to infinity. Figure 7b depicts the stickiness of the HtM state for both the rational and the diagnostic model. In the rational model, stickiness decays at a much higher rate than in the diagnostic one, taking about 40 years to reach its stationary value. In the diagnostic model instead, the average probability that an agent is still HtM in 40 years is still nearly twice its stationary level. Because of the mistakes induced in consumption-saving decisions, diagnostic expectations thus makes it more difficult for agents to escape the HtM state and in this sense it generates a poverty trap. To check that we are indeed capturing a poverty trap and not generally slower transition dynamics generated by diagnostic expectations, we run a placebo test in Appendix B.3 . In particular, Figure B1 analyzes the stickiness of the top 0.1% of the wealth distribution state in the rational and diagnostic benchmarks. The difference in stickiness between the two models appear to be substantially smaller than in the HtM case, thus suggesting that diagnostic expectations does not produce slower dynamics in general, but has particularly stronger effects on the stickiness of the HtM state.

**Empirical estimate** We estimate the empirical counterpart of the stickiness of the HtM state using the SHIW panel data, following the approach of Aguiar et al. (Forthcoming). A household is classified as hand-to-mouth if its liquid net wealth is below two months of labor income. Exploiting the long panel dimension of SHIW, we track these households for up to 14 years in biennial intervals, computing at each step the fraction that remain in the HtM state. Our estimates for the first four years closely align with the evidence in Aguiar et al. (Forthcoming) based on the PSID: roughly 60% of households remain HtM after two years and about 55% after four years. The longer horizon afforded by SHIW further reveals the persistence of the HtM state in the long run: even after 14 years, nearly 40% of households that were initially HtM continue to be so.

As shown in Figure 7b, the rational expectations model fails to capture the persistence of the hand-to-mouth (HtM) state, whereas our calibrated diagnostic model replicate it to

[21] Since there is a Dirac mass point at the borrowing limit we define $\mathfrak{S}$ in terms of an integral over a measure rather than over a density.

Figure 8: Welfare cost along the wealth distribution




*Note: The range for wealth used in this graph goes from0to the maximum net worth observed in our SHIW data, which is around 15 million euros. The welfare cost is expressed in percentage terms.*

a close extent. The mechanism is straightforward. Under rational expectations, households correctly foresee future income fluctuations: after a positive income shock, they restrain consumption and accumulate sufficient savings buffers, which then insulate them against future negative shocks. As a result, the HtM state does not display meaningful persistence. By contrast, diagnostic expectations distort households' beliefs. Following a positive shock, households become overly optimistic about their future income, consume more aggressively, and save less. This behavioral response leaves them more likely to arrive and stay at the HtM state, as seen in Figure 7b.

## 6.2 Welfare evaluation

**Welfare metric** Diagnostic expectations generate non-trivial welfare losses. To get a better understanding on how quantitatively important these losses may be, we develop a welfare evaluation in the spirit of Lucas (1987). First we evaluate welfare function from a paternalistic view point based on the true law of motion of the states. In particular we define $W^{DE}(a, y, \mathcal{S})$ to be the welfare of an agent using the diagnostic expectations policy function $c^{DE}(a, y, \mathcal{S})$ and starting with initial conditions $a_0 = a, y_0 = y, \mathcal{S}_0 = \mathcal{S}$:

$$W^{DE}(a_0, y_0, S_0) = \mathbb{E}_0 \int_0^{\infty} e^{-\rho t} u(c^{DE}(a_t, y_t, S_t)) dt$$

where $\mathbb{E}_0$ is the rational expectations operator capturing the true evolution of the states. Now suppose the diagnostic agent with no sentiment has access to a technology making their beliefs rational, with a cost expressed as a flow consumption tax $\tau$. This consumption tax will depend on the household's initial conditions and will serve as our welfare metric. Formally:

$$W^{DE}(a_0, y_0, 0) = \mathbb{E}_0 \int_0^{\infty} e^{-\rho t} u \left[ (1 - \tau(a_0, y_0)) c^{RE}(a_t, y_t) \right] dt$$

Where $c^{RE}$ is the rational expectations policy functions. Assuming log utility we immediately get the welfare cost schedule as a function of the initial states $a_0 = a$ and $y_0 = y$:

$$\tau(a, y) = 1 - \frac{e^{\rho W^{DE}(a,y,0)}}{e^{\rho W^{RE}(a,y)}} \tag{15}$$

**Distribution of welfare cost** Given our calibration, we find the welfare cost of diagnosticity,averaged across the wealth and income distribution, to be 3.3 percent. Compared to the rule of thumb in Lucas (1987) that a cost of 0.5 percent of lifetime consumption is "large", our estimate indeed reveals a substantial welfare toll of diagnostic expectations. In particular, the interaction of income volatility and non-rational expectations leads households to commit intertemporal errors that prove to be very expensive in our model. Furthermore, these welfare costs are not uniformly distributed across the wealth distribution. As depicted in Figure 8, cost tends to be decreasing in wealth. This is connected to our prior discussion that diagnostic expectations particularly distort decisions of low wealth households. Interestingly, though,the cost turns out to be increasing in wealth in the vicinity of the borrowing constraint. This is because when agents are exactly at the borrowing limit, the scope for making mistakes is largely reduced, as intertemporal decisions are constrained and beliefs do not matter as much. Finally, as wealth goes to infinity the welfare cost converges to zero.[22]

[22] This result is reminiscent of the result in Allais et al. (2020), who find that the welfare costs of volatile inflation vanish for agents at the top of the wealth distribution.

# 7 Conclusion

In this paper we introduced diagnostic expectations in an otherwise standard incomplete market model. We proposed that over-extrapolation of recent income news can maintain agents into a state of financial constraint, as positive income shocks drive overoptimism and overconsumption. To develop this idea, we adapted the theory of diagnostic expectations to idiosyncratic shocks in incomplete markets, introducing a 'rationality wedge' to handle deviations from rational expectations with heterogeneous agents. This enables generalizing diagnostic expectations beyond traditional AR(1) processes and introduce this behavioral friction to more standard quantitative models of income fluctuation. We suggested that households' perceptions of future income are distorted by 'sentiment', a new state variable in an incomplete market framework, together with wealth and productivity. We validated empirically the way sentiment can distort households' expectations using survey data on households' expectations from the Survey of Household Income and Wealth. Households having experienced growth in their income tend to forecast future income above their actual realized value, a bias linked to our sentiment variable. Lastly, we found that diagnostic expectations increase the persistence of the hand-to-mouth state. When agents are financially constrained, diagnostic expectations make it harder to escape financial constraints as agents have a higher marginal propensity to consume out of positive income shocks. In incomplete market models, diagnostic expectations can represent a significant cost, averaging 3.3 percent of lifetime consumption.

# References


ACHDOU, Y., J. HAN, J.-M. LASRY, P.-L. LIONS, AND B. MOLL (2022): "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach," *The Review of Economic Studies*, 89, 45–86.

AGUIAR, M., M. BILS, AND C. BOAR (Forthcoming): "Who Are the Hand-to-Mouth?" *The Review of Economic Studies*.

AIYAGARI, R. (1994): "Uninsured Idiosyncratic Risk and Aggregate Saving," *Quarterly Journal of Economics*, 109(3), 659–684.

ALLAIS, O., Y. ALGAN, E. CHALLE, AND X. RAGOT (2020): "The Welfare Cost of Inflation Risk under Imperfect Insurance," *Annals of Economics and Statistics*, 1–20.

ARELLANO, M., R. BLUNDELL, S. BONHOMME, AND J. LIGHT (2023): "Heterogeneity of Consumption Responses to Income Shocks in the Presence of Nonlinear Persistence," TSE Working Papers 23-1435, Toulouse School of Economics (TSE).

AUCLERT, A. (2019): "Monetary Policy and the Redistribution Channel," *American Economic Review*, 109(6).

BAKER, S. R. AND C. YANNELIS (2017): "Income changes and consumption: Evidence from the 2013 federal government shutdown," *Review of Economic Dynamics*, 23, 99–124.

BANERJEE, A. AND S. MULLAINATHAN (2010): "The Shape of Temptation: Implications for the Economic Lives of the Poor," Working Papers id:2484, eSocialSciences.

BENHABIB, J., A. BISIN, AND S. ZHU (2015): "The wealth distribution in Bewley economies with capital income risk," *Journal of Economic Theory*, 159, 489–515.

BERNHEIM, B. D., D. RAY, AND ŞEVIN YELTEKIN (2015): "Poverty and Self-Control," *Econometrica*, 83, 1877–1911.

BEWLEY, T. (1977): "The permanent income hypothesis: A theoretical formulation," *Journal of Economic Theory*, 16, 252 – 292.

BIANCHI, F., C. ILUT, AND H. SAIJO (2024): "Diagnostic business cycles," *Review of Economic Studies*, 91, 129–162.

BLUNDELL, R., L. PISTAFERRI, AND I. PRESTON (2008): "Consumption Inequality and Partial Insurance," *American Economic Review*, 98, 1887–1921.

BORDALO, P., N. GENNAIOLI, Y. MA, AND A. SHLEIFER (2020): "Overreaction in Macroeconomic Expectations," *American Economic Review*, 110, 2748–82.

BORDALO, P., N. GENNAIOLI, AND A. SHLEIFER (2018): "Diagnostic Expectations and Credit Cycles," *The Journal of Finance*, 73, 199–227.

BROER, T., A. N. KOHLHAS, K. MITMAN, AND K. SCHLAFMANN (2021): "Information and Wealth Heterogeneity in the Macroeconomy," *CEPR Discussion Paper No. 15934*.

——— (2022): "On the possibility of Krusell-Smith Equilibria," *Journal of Economic Dynamics and Control*, 141, 104391.

DALTON, P. S., S. GHOSAL, AND A. MANI (2016): "Poverty and Aspirations Failure," *Economic Journal*, 126, 165–188.

GENNAIOLI, N., Y. MA, AND A. SHLEIFER (2016): "Expectations and Investment," *NBER Macroeconomics Annual*, 30, 379–431.

HALL, R. E. AND F. S. MISHKIN (1982): "The Sensitivity of Consumption to Transitory Income: Estimates from Panel Data on Households," *Econometrica*, 50, 461–481.

HARRIS, C. AND D. LAIBSON (2013): "Instantaneous Gratification," *The Quarterly Journal of Economics*, 128, 205–248.

HUGGETT, M. (1990): "The Risk-Free Rate in Heterogeneous Agent Economies," *Manuscript, University of Minnesota*.

IMROHOROGLU, A. (1989): "Costs of Business Cycles with Indivisibilities and Liquidity Constraints," *Journal of Political Economy*, 1364–83.

JAPPELLI, T. AND L. PISTAFERRI (2014): "Fiscal Policy and MPC Heterogeneity," *American Economic Journal: Macroeconomics*, 6, 107–36.

——— (2020): "Reported MPC and Unobserved Heterogeneity," *American Economic Journal: Economic Policy*, 12, 275–97.

KAPLAN, G., B. MOLL, AND G. VIOLANTE (2018): "Monetary Policy According to HANK," *American Economic Review*, 108(3).

——— (2020): "The Great Lockdown and the Big Stimulus: Tracing the Pandemic Possibility Frontier for the U.S," Tech. rep.

KAPLAN, G. AND G. L. VIOLANTE (2014): "A Model of the Consumption Response to Fiscal Stimulus Payments," *Econometrica*, 82, 1199–1239.

——— (2022): "The Marginal Propensity to Consume in Heterogeneous Agent Models," *Annual Review of Economics*, 14, 747–775.

KAPLAN, G., G. L. VIOLANTE, AND J. WEIDNER (2014): "The Wealthy Hand-to-Mouth," *Brookings Papers on Economic Activity*, 45, 77–153.

LAIBSON, D. AND P. MAXTED (2023): "The Beta-Delta-DELTA Sweet Spot," NBER Working Papers

30822, National Bureau of Economic Research, Inc.

LAIBSON, D., P. MAXTED, AND B. MOLL (2021): "Present Bias Amplifies the Household Balance-Sheet Channels of Macroeconomic Policy," *NBER Working Paper No. 29094*.

LEWIS, D. J., D. MELCANGI, AND L. PILOSSOPH (2019): "Latent Heterogeneity in the Marginal Propensity to Consume," Staff Reports 902, Federal Reserve Bank of New York.

LUCAS, R. (1987): "Models of Business Cycles," *Oxford: Basil Blackwell*.

L'HUILLIER, J.-P., S. R. SINGH, AND D. YOO (2023): "Incorporating Diagnostic Expectations into the New Keynesian Framework," Working Paper Series 2023-19, Federal Reserve Bank of San Francisco.

MA, Y., T. ROPELE, D. SRAER, AND D. THESMAR (2020): "A Quantitative Analysis of Distortions in Managerial Forecasts," *NBER Working Paper No. 26830*.

MAXTED, P. (2023a): "A Macro-Finance Model with Sentiment," *The Review of Economic Studies*, 91, 438–475.

——— (2023b): "Present Bias Unconstrained: Consumption, Welfare, and the Present-Bias Dilemma ," Tech. rep.

PAPPA, E., M. O. RAVN, AND V. STERK (2023): "Chapter 19 - Expectations and incomplete markets," in *Handbook of Economic Expectations*, ed. by R. Bachmann, G. Topa, and W. van der Klaauw, Academic Press, 569–611.

ROZSYPAL, F. AND K. SCHLAFMANN (2023): "Overpersistence Bias in Individual Income Expectations and Its Aggregate Implications," *American Economic Journal: Macroeconomics*, 15, 331–71.

SERGEYEV, D., C. LIAN, AND Y. GORODNICHENKO (2023): "The Economics of Financial Stress ," .

THAKRAL, N. AND L. T. TÔ (2021): "Daily Labor Supply and Adaptive Reference Points," *American Economic Review*, 111, 2417–2443.

# Appendix

## A Extra results

**Greater sensitivity of consumption to income shocks** A rich strand of literature studies how income shocks impact consumption-savings decisions.[23] A pattern which has been well documented empirically is the excess sensitivity of consumption to income shocks (Baker and Yannelis, 2017). As was already clear from Figure 4, our diagnostic expectations framework naturally predicts that the consumption response to income shocks is magnified on impact, thus generating a large income sensitivity of consumption. This is because diagnostic agents not only consume out of the income shock, but also out of the shock to sentiment. In other words, in our framework an income shock has an extra effect on consumption through its effect on agents' expectations, via sentiment. This naturally leads to the following notion of consumption response out of income shocks:

**Definition 4.** *The instantaneous percentage consumption response to log-productivity shocks* $\Delta$ *is given by*

$$\mathfrak{y}(\Delta; x) = \frac{c(a, y+\Delta, \mathcal{S}+\Delta) - c(a, y, \mathcal{S})}{we^{y+\Delta} - we^{y}} \frac{we^{y}}{c(a, y, \mathcal{S})} \tag{A.1}$$

*where* $c(\cdot)$ *represents the diagnostic expectations consumption policy function.*

Equation (A.1) above effectively defines an elasticity, as $\mathfrak{y}(\Delta; x)$ represents the instanteneous percentage change in consumption in response to a log productivity shock $\Delta$ inducing a $e^{\Delta} - 1$ percentage change in labor income at the time of the shock.[24] In particular, when $\Delta$ is small, we can decompose (A.1) in terms of elasticities as a simple directional derivative:

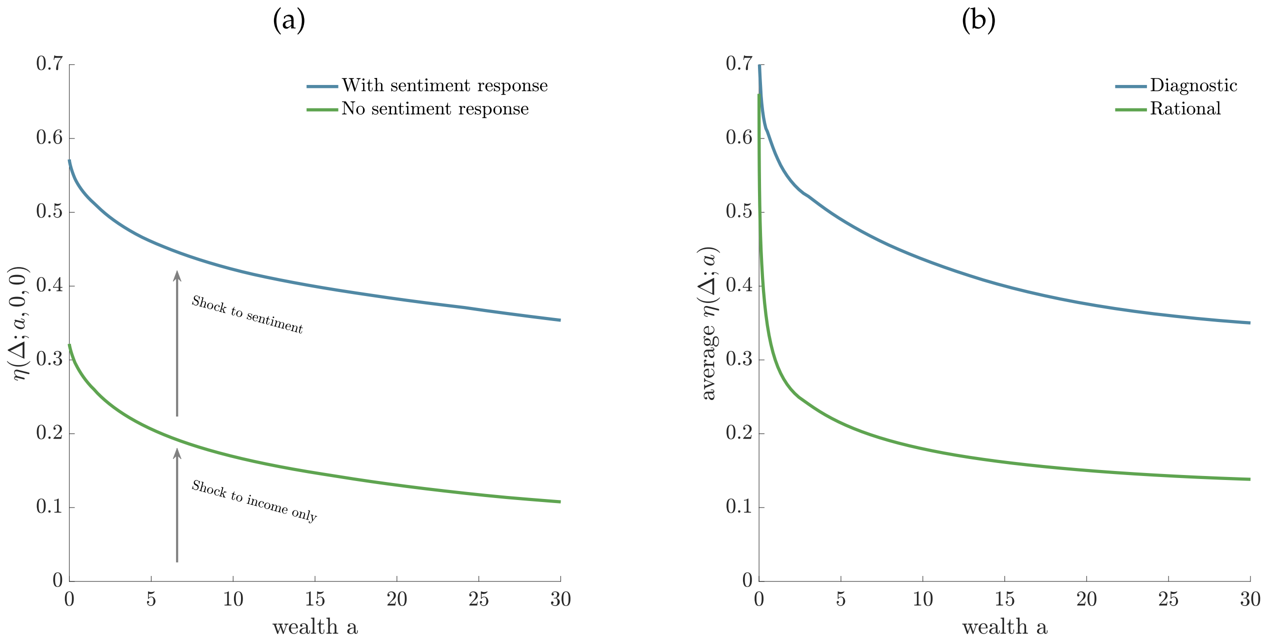
$$\lim_{\Delta \to 0} \mathfrak{y}(\Delta; x)\Delta = \underbrace{\frac{\partial \log c(a, y, \mathcal{S})}{\partial y}}_{\substack{\text{income elasticity} \\ \text{of consumption}}} \frac{\Delta}{\sqrt{2}} + \underbrace{\frac{\partial \log c(a, y, \mathcal{S})}{\partial \mathcal{S}}}_{\substack{\text{sentiment semi-elasticity} \\ \text{of consumption}}} \frac{\Delta}{\sqrt{2}} \tag{A.2}$$

Equation (A.2) reveals that the consumption response to productivity shocks can be decomposed into two effects, respectively denoting the household's reaction to changes in income and sentiment. This is visualized in Figure A.1a, where we plot the instantaneous

[23] See Hall and Mishkin (1982) and Blundell et al. (2008), among many others.

[24] This is coming from the way we define labor income. We have $\frac{we^{y+\Delta} - we^{y}}{we^{y}} = e^{\Delta} - 1$

Figure A.1: Consumption response to productivity shocks




*Note: Panel A.1a depicts $\eta$ for a household with median income and no sentiment initially, and panel A.1b plots the average of $\eta$ over income and sentiment.*

consumption response $\eta(\Delta; x)$ to a shock $\Delta$ equal to one standard deviation of the typical income shock over the wealth distribution, with initial conditions $y = 0$ and $\mathcal{S} = 0$. The green line depicts the consumption response under rational expectations, which only comprises the response to the income shock, i.e., the first term in (A.2). The blue line, instead, shows the consumption response under diagnostic expectations. It can be seen that, for all wealth levels, the sensitivity of consumption is larger in the diagnostic case. This is exactly because, when agents have diagnosticity, they also react to their change in sentiment after they are hit with an income shock, as captured by the second term in (A.2). This result also holds after aggregating over income and sentiment states, as showed in Figure A.1b, which plots the consumption response across all levels of wealth, averaged over income and sentiment.[25]

**Latent heterogeneity** A recent empirical literature has emphasized the fact that, even after controlling for a large array of observables, there remains a lot of unexplained heterogeneity in households' Marginal Propensities to Consume (MPCs) (Lewis et al., 2019) and consumption response to income shocks (Arellano et al., 2023). This inability to pre-

[25] Clearly, the averaging over sentiment levels only matters for the blue line, depicting the diagnostic expectations case.

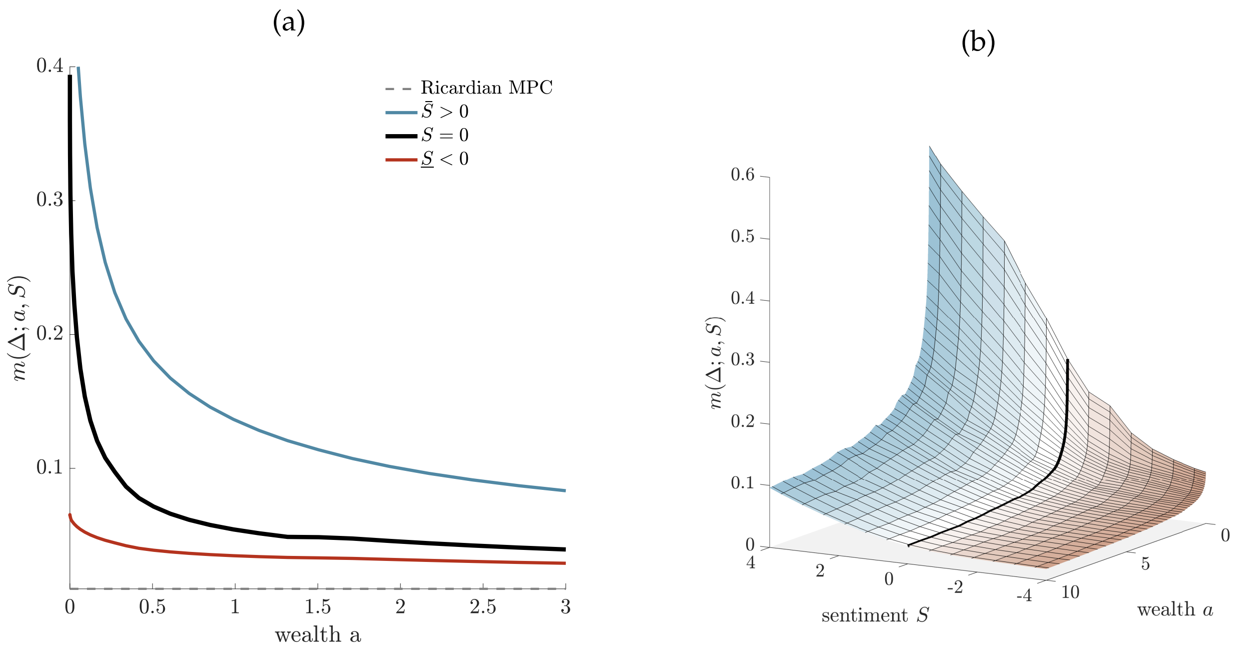
dict agents' response to income shocks and wealth windfalls has been referred to as "latent heterogeneity". Our framework is able to rationalize this heterogeneity. In fact, sentiment in our economy is both unevenly distributed across households and affects agents' response to income and wealth changes. In so far as sentiment is not observable by the econometrician, it thus represents a good candidate to explain the latent heterogeneity observed in the data. In what follows, we focus on the dispersion generated by sentiment in households' MPCs. However, similar arguments apply for the income elasticity of consumption. In Figure 3 we showed how sentiment directly affects households' policy functions for consumption and savings, in a non-linear fashion. It is then easy to see how in our setting sentiment also modifies the traditional notion of MPCs out of unexpected wealth windfalls, even though such windfalls don't have any direct effect on sentiment. To see this, we now extend Achdou et al. (2022) definition of Marginal Propensity to Consume to our diagnostic expectations setting:

**Definition 5.** *The Marginal Propensity to Consume out of a windfall $\Delta$, for a household with state vector $x \equiv (a, y, \mathcal{S})$ over a period $\tau$ is given by*

$$\mathfrak{m}_\tau(\Delta; x) = \frac{C_\tau(a+\Delta, y, \mathcal{S}) - C_\tau(a, y, \mathcal{S})}{\Delta} \quad \textit{where} \quad C_\tau(x) = \mathbb{E}_0\left[\int_0^\tau c^{DE}(x_t)dt \mid x_0 = x\right] \tag{A.3}$$

Following standard practice in the literature, we focus on quarterly MPCs out of a one-time unanticipated windfall of 1000€. Figure fig. A.2a plots the MPC over wealth and sentiment, averaged across log-productivity y. The figure illustrates how MPCs are increasing in sentimentacross the whole wealth distribution, thus implying that sentiment generates variation in MPCs for a given level of wealth. Figure fig. A.2b further visualizes this, by plotting MPCs over the wealth distribution for a fixed level of log-productivity and three different levels of sentiment. This picture clearly shows how even knowing households' income and wealth is not enough to predict MPCs. In fact, sentiment still generates large dispersion in agents' MPC after controlling for their income and wealth. Our setting can thus provide a behavioral explanation for the latent heterogeneity in MPCs (as well as in the income elasticity of consumption) observed in the data. Moreover, a further advantage of our framework is that it provides a characterization of sentiment that can be mapped to the data. In fact, from the definition of sentiment (2) it immediately follows that we can measure $\mathcal{S}_t$ as a weighted average of past income shocks. Panel datasets featuring a large enough time dimension, such as the Panel Study of Income Dynamics (PSID), then look promising to obtain a measure of $\mathcal{S}_t$ that varies both over individuals and time.

Figure A.2: Marginal Propensity to Consume




*Note: The MPCs in panel Figure A.2a are plotted for one productivity state y. The MPCs in panel Figure A.2b are averaged over all productivity states using the stationary distributions.*

# B Proof of propositions

## B.1 Proof of proposition 1

We prove the equation for a general log-productivity process. We'll derive the proposition for a general a generator $\mathcal{B}$ for the joint process of $y$ and $\mathcal{S}$.[26] The HJB of the sophisticated agent is given by

$$\rho V = \max_{c} \quad u(c) + V_a(ra + we^y - c) + \mathcal{B}V + \mathcal{S}V_y \tag{B.1}$$

From this we can get the first order condition and the envelop condition:

$$u'(c) = V_a \implies u''(c)c_a = V_{aa}, u''(c)c_y = V_{ay} \tag{B.2}$$

$$\rho V_a = V_{aa}s + rV_a + \mathcal{B}V_a + \mathcal{S}V_{ya} \tag{B.3}$$

[26] We'll impose the regularity condition that $\partial_a BV = B\partial_a V$, which is satisfied for our baseline income process.

From which we get

$$\rho u'(c) = u''(c)c_a s + ru'(c) + \mathcal{B}u'(c) + \mathcal{S}u''(c)c_y \tag{B.4}$$

$$\iff u''(c)c_a s + \mathcal{B}u'(c) = (\rho - r)u'(c) - \mathcal{S}u''(c)c_y \tag{B.5}$$

$$\iff \mathbb{E}(du'(c)/dt) = (\rho - r)u'(c) - u''(c)c_y\mathcal{S} \tag{B.6}$$

$$\iff \frac{\mathbb{E}_t[du'(c_t)/dt]}{u'(c_t)} = \rho - r - \frac{u''(c_t)c_t}{u'(c_t)}\frac{\partial \log c_t}{\partial y}\mathcal{S}_t \tag{B.7}$$

Using the definition of intertemporal elasticity of substitution and income elasticity of consumption we get the desired result.

For the naïve case, the HJB is given by:

$$\rho V = \max_c \quad u(c) + V_a(ra + we^y - c) + \mathcal{I}V + \mathcal{S}V_y \tag{B.8}$$

From this we can get the first order condition and the envelop condition:

$$u'(c) = V_a \implies u''(c)c_a = V_{aa}, u''(c)c_y = V_{ay} \tag{B.9}$$

$$\rho V_a = V_{aa}s + rV_a + \mathcal{I}V_a + \mathcal{S}V_{ya} \tag{B.10}$$

From which we get

$$\rho u'(c) = u''(c)c_a s + ru'(c) + \mathcal{I}u'(c) + \mathcal{S}u''(c)c_y \tag{B.11}$$

$$\iff u''(c)c_a s + \mathcal{I}u'(c) = (\rho - r)u'(c) - \mathcal{S}u''(c)c_y \tag{B.12}$$

$$\iff \mathbb{E}(du'(c)/dt) = (\rho - r)u'(c) - u''(c)c_y\mathcal{S} \tag{B.13}$$

$$\iff \frac{\mathbb{E}_t[du'(c_t)/dt]}{u'(c_t)} = \rho - r - \frac{u''(c_t)c_t}{u'(c_t)}\frac{\partial \log c_t}{\partial y}\mathcal{S}_t \tag{B.14}$$

Using the definition of intertemporal elasticity of substitution and income elasticity of consumption we get the desired result.

## B.2 Feynmann-Kac equation

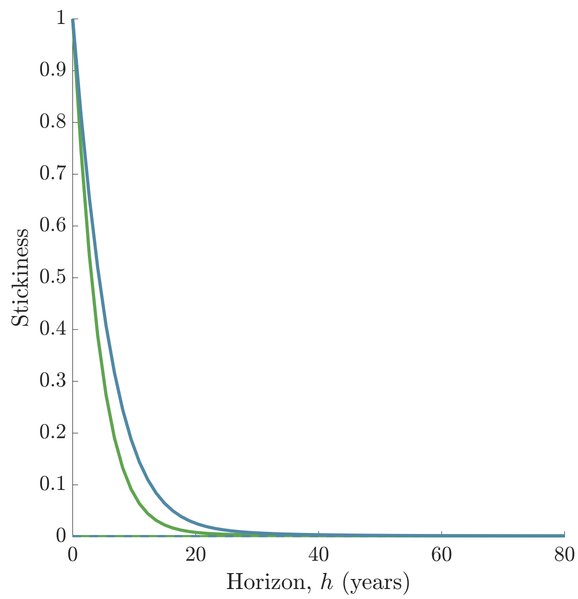
The conditional expectation $C_\tau(x)$ (A.3) can be computed as $C_\tau(x) = \Gamma(x, 0)$ where $\Gamma(x, t)$ satisfies the PDE

$$0 = c^{DE}(x) + \mathcal{G}^*\Gamma(x, t), \quad \text{where} \quad \mathcal{G} \equiv (ra + we^y - c^{DE}(x))\partial_a + \mathcal{B} + \partial_t \tag{B.15}$$

with terminal condition $\Gamma(x, \tau) = 0$ for all $x$.

### B.3 Additional figures

Figure B1: Stickiness of the ultra-rich




*Note: we define the ultra rich as the top* 0.1% *of the wealth distribution in each model.*

# C Numerical appendix

We can define the income process in the more standard way: at some poisson rate $\lambda$, the agent draws a new log-productivity drawn from a normal centered around zero and with variance $\sigma^2$. In this case, upon the new draw, sentiment will move by a jump equal to the difference between the new draw and the previous log-prodictivity. Hence at some poisson rate sentiment is drawn from a normal with mean $-y$ and variance $\sigma^2$. First, we discretize the exogeneous generator in the HJB equation. In our case, this generator is given by

$$\mathcal{B}W(a,y,S) = -\beta y \partial_y V(a,y,S) - \eta S \partial_S V(a,y,S) + \lambda \int \left(V(a,x,S+x-y) - V(a,y,S)\right)\phi(x)dx$$

where $\phi(x)$ is a normal pdf with mean 0 and variance $\sigma^2$.

**Discretize** We want to get the discretized operator $\mathcal{B}$, which captures the true transition for the joint process of income and sentiment.

$$\boldsymbol{B}v_{ijk} = (v_{ij+1k} - v_{ijk})\left(\frac{-\beta y_j}{\Delta y}\right)^{+} + (v_{ijk} - v_{ij-1k})\left(\frac{-\beta y_j}{\Delta y}\right)^{-}$$

$$+ (v_{ijk+1} - v_{ijk}) \left( \frac{-\eta \mathcal{S}_k}{\Delta s} \right)^+ + (v_{ijk} - v_{ijk-1}) \left( \frac{-\eta \mathcal{S}_k}{\Delta s} \right)^-$$
$$+ \lambda \sum_{j'=1}^{J} \left( v_{ij'(k+j'-j)} - v_{ijk} \right) \phi(y_{j'}) \Delta y$$

The crucial step is to rearrange:

$$\begin{aligned}
\rho v_{ijk} = \quad & v_{ijk} Y_{ijk} & Y_{ijk} &\equiv \sum_{m=1}^{4} X^m_{ijk} + \sum_{n=1}^{J} Z^n_{ijk} \\
& + v_{ij+1k} X^1_{ijk}, & X^1_{ijk} &\equiv \left( \frac{-\beta y_j}{\Delta y} \right) \\
& + v_{ij-1k} X^2_{ijk}, & X^2_{ijk} &\equiv \left( \frac{-\beta y_j}{\Delta y} \right)^- \\
& + v_{ijk+1} X^3_{ijk}, & X^3_{ijk} &\equiv \left( \frac{-\eta \mathcal{S}_k}{\Delta s} \right)^+ \\
& + v_{ijk-1} X^4_{ijk}, & X^4_{ijk} &\equiv \left( \frac{-\eta \mathcal{S}_k}{\Delta s} \right)^- \\
& + v_{i1(k+1-j)} Z_{ijk}, & Z_{ijk} &\equiv \lambda \phi(y_1) \Delta y \\
& \vdots & & \\
& + v_{iJ(k+J-j)} Z_{ijk}, & Z_{ijk} &\equiv \lambda \phi(y_J) \Delta y
\end{aligned}$$

We call the $X_{ijk}$ the coefficient meshes. We assume an equal grid for $\mathcal{S}$ and $y$ and hence we take $\min\{\max\{(k + j' - j), 1\}, J\}$.

# D Empirical Appendix

## D.1 Survey Questions

**2012 Wave** The following question was asked in the 2012 wave of the survey:

*Twelve months from now, your household's income will be (please distribute 100 points):*

Respondents were then asked to assign probabilities to 5 different scenarios: (i) higher than today (by 10% or more), (ii) somewhat higher than today (2 to 10%), (iii) basically the same (no more than 2% increase or decrease), (iv) somewhat lower (2 to 10%), (v) much lower than today (by 10% or more).

**2014 Wave** In 2014 the wording of the question changed as follows:

*Consider your household's overall income in 2015. Compared with 2014, how much higher/lower do you think it will be in percentage terms?*

In this case, respondents were asked to report their point forecast.

**1989 Wave** The following question was asked in the 1989 wave of the survey:

*Consider the evolution of your total labor or pension income from now to May 1991. Please distribute 100 points among the following scenarios:*

Respondents were then asked to assign probabilities to 12 different scenarios ranging from negative growth (for which they were also asked to report a point estimate of the percentage decrease in income) to above 25% income growth.

**1991 Wave** The following question was asked in the 1991 wave of the survey:

*Consider your total labor or pension income one year from now. Please distribute 100 points among the following scenarios:*

Respondents were then asked to assign probabilities to 12 different scenarios ranging from negative growth (for which they were also asked to report a point estimate of the percentage decrease in income) to above 25% income growth.

## D.2 Robustness

**'89-'91 Waves** Appendix D.2 below reproduces the last two panels of Figure A.2 using data from the 1989 and 1991 waves of the SHIW. Because of the format of the question, both the y and the x axis of Appendix D.2 are expressed in terms of individual level, rather than household level, income.

**Control for Wealth** Appendix D.2 below reproduces the last two panels of Figure A.2 after residualizing both the x and y axis for the logarithm of household net wealht.

**Unconditional Correlations: No Controls** Appendix D.2 below reproduces the last two panels of Figure A.2, without including any control before plotting the binned scatter plot.

Figure D.1: 1989 and 1991 Survey Waves

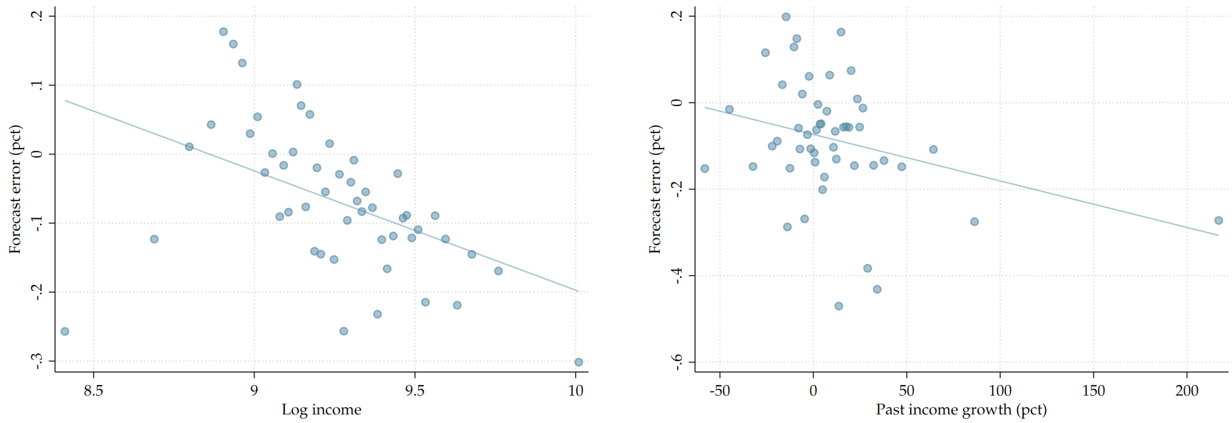
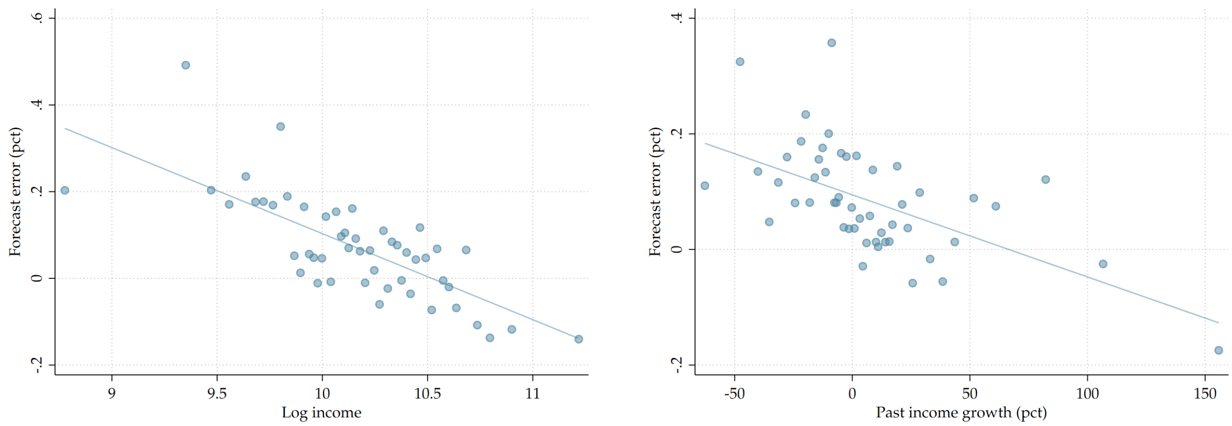



Figure D.2: Controlling for Wealth




**Trim Data at 99th Percentile** Appendix D.2 below reproduces the last two panels of Figure A.2 but trimming the forecast error and income change data at the 99th –rather than 98th– percentile.[27]

**Trim Data at 95th Percentile** Appendix D.2 below reproduces the last two panels of Figure A.2 but trimming the forecast error and income change data at the 95th –rather than 98th– percentile.[28]

[27] Note that for the income change variable, we perform two-sided trimming, by trimming observations based on their absolute value.

[28] Note that for the income change variable, we perform two-sided trimming, by trimming observations based on their absolute value.

Figure D.3: Unconditional Correlations

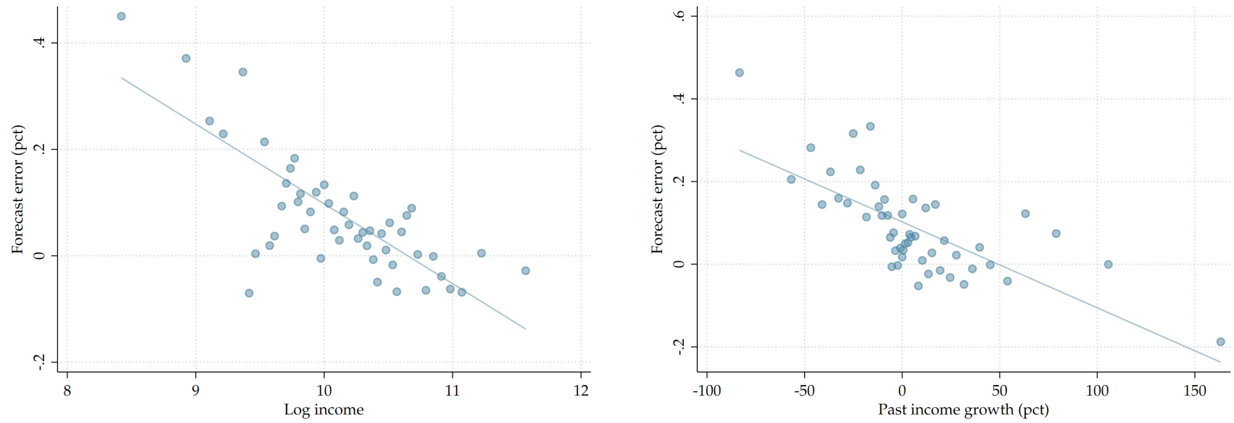
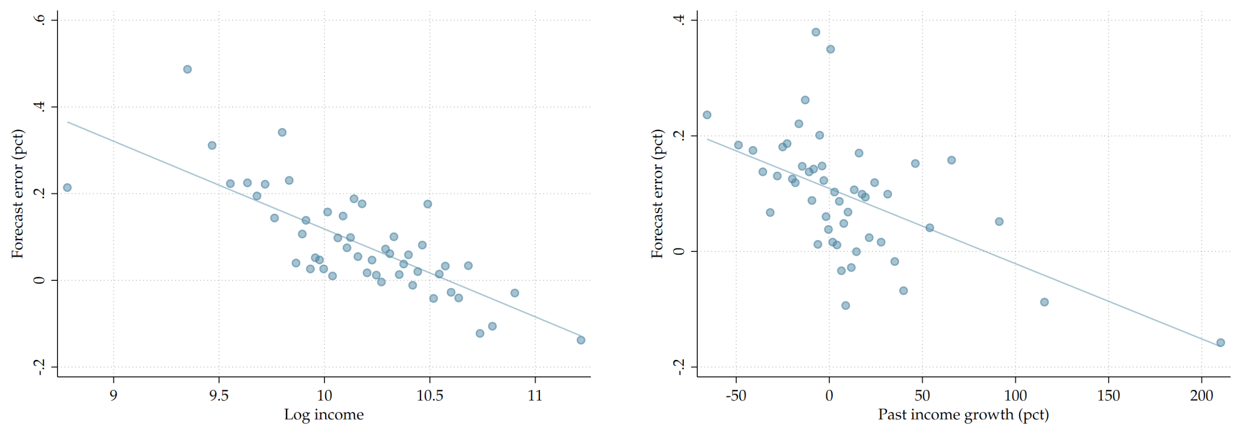



Figure D.4: Trimming Forecast Errors at the 99th Perecntile



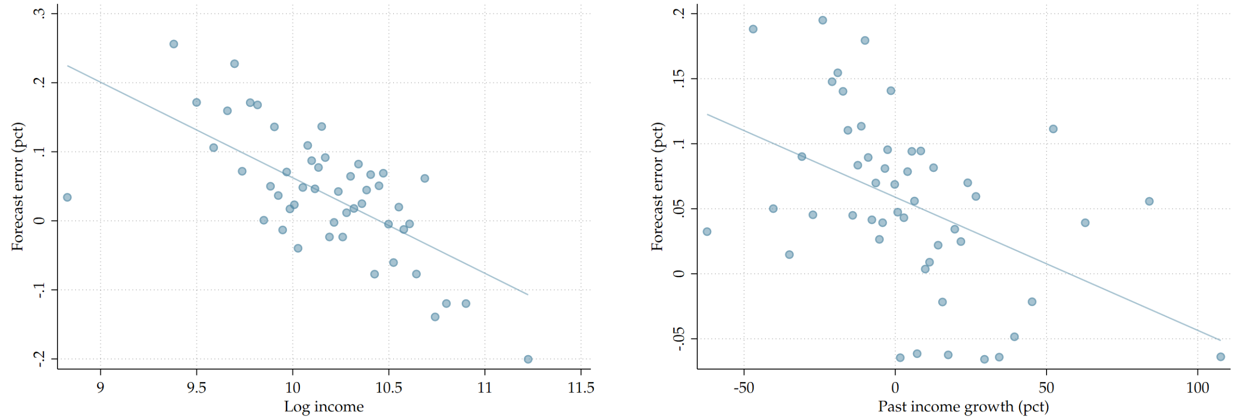

## D.3 Fast belief propagation via ADI splitting

**Problem.** A direct implicit step for belief propagation on the joint $(y, s)$ grid requires solving a very large sparse linear system for each time step and block of right-hand sides. This is the dominant runtime bottleneck in conventional implementations.

**Key observation.** In our environment the forward operator approximately separates as a *Kronecker sum*:

$$\mathcal{L} \;\approx\; \mathcal{L}_y \otimes I_s \;+\; I_y \otimes \mathcal{L}_s \quad (+\,\text{jump/convolution term}),$$

Figure D.5: Trimming Forecast Errors at the 95$^{\text{th}}$ Perecntile




where $\mathcal{L}_y$ and $\mathcal{L}_s$ are one-dimensional Fokker–Planck operators for the income and sentiment coordinates. This structure admits an *Alternating-Direction Implicit* (ADI) time step.

**Scheme (Peaceman–Rachford / CN–ADI).** For time step $\Delta t$ and mass $G^n \in \mathbb{R}^{J\times K\times N}$ on the $(y, s)$ grid:

$$\begin{aligned} \text{(i) Half-step in } y: \quad \left(I_y - \tfrac{\Delta t}{2}\mathcal{L}_y\right) X \;&=\; G^n \left(I_s + \tfrac{\Delta t}{2}\mathcal{L}_s\right)^\top, \\ \text{(ii) Half-step in } s: \quad \left(I_s - \tfrac{\Delta t}{2}\mathcal{L}_s\right)^\top \left(G^{n+1}\right)^\top \;&=\; \left(I_y + \tfrac{\Delta t}{2}\mathcal{L}_y\right) X. \end{aligned}$$

Each half-step reduces to independent banded 1-D solves (tri-/penta-diagonal depending on the stencil) that we *factorize once* and reuse for all steps and RHS columns. When rare jumps are small, we treat them via a diffusion approximation; if needed, a separable discrete convolution can be inserted between half-steps to capture a richer jump kernel.

**Discretization and boundaries.** We use mass-conservative finite differences with zero-flux (Neumann) boundaries along $y$ and $s$ to ensure near-unit mass per page. Optional page-wise renormalization maintains numerical mass exactly.

**Complexity and scaling.** One ADI step costs

$$O\big(JK\,(J+K)\big)$$

versus a single sparse backsolve on the full $JK \times JK$ system. Memory falls from storing a large 2-D LU to two small 1-D factorizations. We stack right-hand sides into panels (e.g., 128–256 columns) to exploit BLAS-3 while avoiding cache pressure. In our implementation this reduces wall-time for the belief-propagation block by an order of magnitude relative to a single large implicit solve.

**Position relative to the literature.** ADI is classical in numerical PDEs. Its application to forward Kolmogorov operators in heterogeneous-agent calibration is, to our knowledge, uncommon. We emphasize it here because it substantially increases the grid sizes and number of right-hand sides that are tractable in our setting, thereby enabling our estimation that targets the two-year FE autocorrelation.

**Algorithm (pseudo-code).** Let $Ay = (I_y - \frac{\Delta t}{2}\mathcal{L}_y)$, $As = (I_s - \frac{\Delta t}{2}\mathcal{L}_s)$ and $By = (I_y + \frac{\Delta t}{2}\mathcal{L}_y)$, $Bs = (I_s + \frac{\Delta t}{2}\mathcal{L}_s)$, with LU factorizations precomputed for $Ay$ and $As^\top$.

[leftmargin=1.1cm]For each panel of pages $G^n_{\text{panel}}$:

1. (a) Compute $R_1 = G^n_{\text{panel}}\, Bs^\top$; solve $Ay\, X = R_1$.
   (b) Compute $R_2 = By\, X$; solve $As^\top\, (G^{n+1}_{\text{panel}})^\top = R_2^\top$.
   (c) (Optional) Renormalize each page of $G^{n+1}_{\text{panel}}$ to unit mass.
2. Advance $n \leftarrow n+1$ and repeat.

**Moment computation.** The model-implied moments $\{\beta_y, \beta_{\Delta y}, \rho^{(2y)}_{\text{FE}}\}$ are recomputed at each parameter trial from the simulated pseudo-panel using the same trimming, weighting, and calendar alignment as in the data.