Rational Expectations or Distorted Beliefs? Measuring Beliefs From Asset Prices*

Anisha Ghosh† Guillaume Roussellet‡

February 28, 2018

Preliminary and Incomplete - Please do not circulate nor cite

Abstract

This paper proposes an empirical procedure to identify investors' subjective beliefs from observed asset prices, i.e. deviations from rational expectations in the conditional distribution of macroeconomic and financial variables. Our methodology relies on the smoothed empirical likelihood technique, which is non-parametric, allowing us to be agnostic on the nature of behavioral biases (if any). Conditional Euler equation restrictions for a chosen cross-section of assets and a parametric pricing kernel enable us to infer the subjective conditional distributions, given the investors' conditioning set. When using inflation and consumption growth as conditioning variablea, we show that deviations from the objective distribution can be quite large and that belief distortions seem to affect both the conditional mean and higher order moments of consumption growth. We show that the estimated beliefs distortion are remarkably similar for many of the popular consumption-based asset pricing models.

Keywords: Rational Expectations, Behavioral Biases, Pricing Kernel, Conditioning Set, Relative Entropy Minimization.

JEL Classification Codes: C51, E3, E70, G12, G14, G40

*We benefited from helpful comments from George Constantinides and Lars Hansen. All errors and omissions are our own.

†Desautels Faculty of Management, McGill University; anisha.ghosh@mcgill.ca.

‡Desautels Faculty of Management, McGill University; guillaume.roussellet@mcgill.ca.

I Introduction

The rational expectations paradigm (Muth (1961)) postulates that economic agents use available data objectively, or *rationally*, to make inferences about the true underlying model and its parameters. Therefore, under rational expectations, while forecasts about future economic conditions need not always be correct, but the forecast errors are unbiased and uncorrelated with any current available information that was used in the construction of the forecasts. The rational expectations paradigm is a simple yet powerful one, and has, to date, constituted the predominant maintained hypothesis in economics. The efficient markets hypothesis (Fama (1969)), for instance, is based on the underlying premise of rational expectations. Most of the leading asset pricing models proposed in the literature – starting with standard Consumption-CAPM (C-CAPM) of Rubinstein (1976) with time-additive power utility as well as more recent models such as the habit formation model of Campbell and Cochrane (1999) the long run risks model of Bansal and Yaron (2004), and the rare disasters model of Rietz (1988) and Barro (2006) – also assume that agents are fully rational.

The rational expectations hypothesis, while intuitive and appealing, has difficulty explaining a number observed features of the aggregate stock market, the cross section of returns, and individual trading behavior. Moreover, these shortcomings pertain not only to the standard C-CAPM, but also the more recent models intended to overcome the limitations of the C-CAPM. In fact, Ghosh, Julliard, and Taylor (2016), using an information-theoretic approach to recover a multiplicative *missing* component of the stochastic discount factor (SDF) for a broad class of consumption-based asset pricing models, show that the missing component is remarkably similar across seemingly very different models.

The shortcomings of rational models have led to the advent of *behavioral models* where economic agents are assumed to not be fully rational (see Barberis and Thaler (2001) for a survey of behavioral finance). This class of models assumes certain behavioral biases in agents in the processing of available information to form beliefs about the future. More recently, researchers have shown that departures from rational expectations can occur even in the absence of behavioral biases. For instance, if investors have robust control (uncertainty aversion) preferences, then the difficulty in distinguishing between alternative data generating processes using the finite available data leads investors to make consumption-investment decisions from the perspective of the *worst-case* model (see, e.g., Hansen and Sargent (2001)). Unlike rational expectations models that typically rely on implausibly high risk aversion levels to explain stock market data, Barillas, Hansen, and Sargent (2009) argue that robust control models replace the need for implausibly large risk aversion with plausible levels of uncertainty aversion.

While behavioral and robust control models constitute attractive alternatives to the rational expectations framework, limited empirical evidence exists for the presence of particular behavioral biases, or the form of uncertainty aversion. Yet assumptions about the parametric forms of these components are crucial to theoretical models that purport to explain observed aspects of financial market data. In the behavioral finance literature, common assumptions about the form of the behavioral bias include prospect theory (e.g., Kahneman and Tversky (1979)) or ambiguity aversion (e.g., Savage (1964)). In the robust control literature, typically assumed specifications include a quadratic specification of the beliefs distortion, that makes investors act as if the dynamics of macro variables are more persistent than what is estimated from the historical data (e.g., Szoke (2017)).

This paper proposes an approach to measuring the *subjective beliefs* of investors using observed data on asset prices, without relying on any specific functional form assumptions about the nature of beliefs distortion relative to the rational expectations benchmark. For a given choice of the SDF, a set of assets that the SDF is challenged to price, and the specification of the conditioning set, i.e. the vector of state variables, our approach enables the recovery of the entire conditional distribution of possible future realizations of macro and financial variables as perceived by the average investor.

Our approach to extracting beliefs from asset prices relies on the information-theoretic (relative entropy minimization) estimation of conditional moment restriction models proposed in Kitamura, Tripathi, and Ahn (2004). The approach is akin to estimating beliefs so as to maximize the nonparametric log-likelihood of the data, subject to the constraint that the estimated beliefs satisfy the conditional Euler equation restrictions for the chosen set of assets. Therefore, the framework retains the desirable properties of a likelihood based approach, while avoiding parametric assumptions on the form of beliefs.

If beliefs are rational, then the extracted subjective beliefs about the macroeconomic variables of interest, e.g. consumption growth, should coincide with the corresponding objective beliefs estimated using historical macro data alone (without using asset price data). Therefore, comparison of these subjective beliefs with objective beliefs (obtained from commonly used statistical models) can help shed light on the nature of beliefs distortion that is most supported by the data and offer guidelines for the construction of theoretical models that build in such distortions.

Note that the econometric feasibility of the above estimation approach crucially relies on the ability to summarize the investors' conditioning set with a small number of variables. Our choice of the conditioning set draws on the insight in Ghosh and Constantinides (2017), who contribute towards identifying the investors' information set. Specifically, Ghosh and Constantinides (2017) show that just two variables, namely the rate of change in the Consumer Price Index (CPI) and the growth in the average hourly earnings of production in private non farm payrolls, along with consumption growth, go a long way toward proxying for investors' relevant information set. Guided by this finding, we use the above two variables (one at a time as well as jointly), along with consumption growth, as constituting the conditioning set.

Note that our methodology requires the specification of the SDF, the conditioning set, and the cross-section of assets that the SDF is asked to price. We present results for different choices of the above to demonstrate the robustness of our results. First, we show that the estimated subjective beliefs are remarkably similar for several different choices of the pricing kernel. In particular, using the excess return on the market portfolio as the sole test asset and the CPI growth and consumption growth as conditioning variables, we recover the subjective beliefs for the pricing kernels implied by following specifications of the investors' preferences – the standard C-CAPM with power utility, Epstein and Zin (1989) recursive preferences, and the external habit formation preferences of Campbell and Cochrane (1999). The correlations between the subjective beliefs estimated from these different SDFs are higher than XX for all values of the conditioning variables.

Second, the estimated beliefs suggest significant beliefs distortion relative to the objective specifications commonly assumed in the literature. Once again, to demonstrate the robustness of our results, we consider a few different objective specifications. In particular, we consider a standard latent VAR model for the conditioning variables, namely CPI growth and consumption growth. We estimate the model using data on CPI growth and consumption growth alone, without using any asset price data. We also consider a regime-switching model, where the means and volatilities of CPI growth and consumption growth vary across latent regimes, and estimate the model using macro data alone. We show that the dimensions of the deviations of the estimated subjective beliefs from the objective ones are quite similar across the various objective specifications considered. Specifically, in line with the robust control literature, we show that the subjective beliefs appear more pessimistic relative to the corresponding rational ones, for each value of the conditioning variables. However, the estimated beliefs also reveal significant differences relative to commonly assumed specifications of the beliefs distortion in the robust control literature. The important differences are: (i) whereas the literature argues that beliefs distortion make macro variables appear to be more persistent and, therefore, more volatile than what is observed in the data, our estimated beliefs suggest that the distortion affects primarily the skewness making the underlying macro variables appear more negatively skewed than what can be estimated from the data alone, and (ii) the deviation between the subjective and objective beliefs increase substantially in bad states of the world, defined by particular realizations of the conditioning variables.

Our work extends Ghosh, Julliard, and Taylor (2016) who also use an information-theoretic approach to recover a multiplicative 'missing' component of the SDF for a broad class of consumption-based asset pricing models, such as the standard Consumption-CAPM of Rubinstein (1976), habit formation models of Campbell and Cochrane (1999) and Menzly, Santos, and Veronesi (2004), the long run risks model of Bansal and Yaron (2004), and models with complementarities in consumption as in Piazzesi, Schneider, and Tuzel (2007). Note that these models are all based on the maintained hypothesis of rational expectations. Therefore, if the true underlying belief formation process differs from that under rational expectations, then this belief distortion would be captured as a multiplicative missing component of the SDF. Consistent with such an interpretation of the missing component of the SDF, Ghosh, Julliard, and Taylor (2016) find that the missing components of the SDFs extracted from the above seemingly very different asset pricing models are remarkable similar. The present study extends Ghosh, Julliard, and Taylor (2016) in that, whereas Ghosh, Julliard, and Taylor (2016) focus on unconditional Euler equation restrictions, this paper considers conditional Euler equations. While this complicates the analysis in that it requires specification of the conditioning set, it also enables us to estimate the conditional distribution of investors' (subjective) beliefs. In other words, it enables us to estimate the time series of conditional moments of variables of interest, such as returns, consumption growth, and other macroeconomic variables. These forecasts can then be compared to survey-based forecasts of these variables. The forecasts can also be used to shed light on the strengths and weaknesses of beliefs processes commonly assumed in the literature, along with insights for future theoretical modeling.

Our paper contributes to a growing literature that emphasizes the importance of discriminating between investors' subjective beliefs and the objective beliefs obtained from commonly used statistical models, in explaining various aspects of asset market data. Piazzesi, Salomao, and Schneider (2015) show that the subjective bond risk premia are less volatile and less cyclical compared to the premia estimated using standard statistical models. Wang (2017) shows that investors' subjective beliefs has significant explanatory power for a broad cross section of stock portfolios. These studies all use professional survey forecasts data to form estimates of the subjective beliefs. The survey forecasts provide the median (across a group of professional forecasters) forecasts of a variety of future macroeconomic or financial variables. Our approach to recovering investors' subjective beliefs differs markedly from these studies in that, rather than using survey forecasts data, we extract the beliefs from observed asset prices via the conditional consumption Euler equations. Our approach enables the recovery of the *entire conditional distribution of beliefs*, rather than only the distribution of the conditional means of the variables of interest as is possible with the survey based forecasts data. Therefore, our approach can be used to shed light on which characteristics of the distribution are distorted the most under the subjective beliefs, relative to commonly assumed objective specifications.

The remainder of this paper is organized as follows. Section II describes our method of estimating investors' subjective beliefs from observed asset prices. The data used in the empirical analysis are described in Section III. The empirical results are presented in Section V. Section VI concludes with suggestions for future research.

II Non-Parametric Estimation of Beliefs

In this section, we describe the details of our methodology to extract investors' subjective beliefs from observed asset prices. Section II.1 describes the framework wherein there may be a divergence between investors' subjective beliefs and the true underlying objective distribution of the data. The econometric approach to recovery of the investors' subjective beliefs, without making parametric assumptions about the latter, is discussed in Section II.2. An alternative information-theoretic interpretation of the extracted beliefs is provided in Section II.3. Section III.1 describes the choice(s) of the conditioning set, and Section II.4 discusses the various SDFs considered. Throughout, uppercase letters denote random variables, while the corresponding lowercase letters denote particular realizations of these random variables.

II.1 General framework

We assume absence of arbitrage opportunities, such that a strictly positive pricing kernel (hereafter referred to as the SDF), denoted by M_{t+1}, exists. The equilibrium returns $\mathbf{R}^e_{t+1} \in \mathbb{R}^k$ of any set of k traded assets in excess of the risk-free rate satisfy the Euler equation,

$$\mathbb{E}^{\mathbb{P}_t}\left[M_{t+1}\mathbf{R}^e_{t+1}|\underline{\mathcal{F}_t}\right] = \mathbf{0}, \tag{1}$$

where $\underline{\mathcal{F}_t} = \{\mathcal{F}_t, \mathcal{F}_{t-1}, \ldots\}$ denotes the investors' information set at time t, and $\mathbb{E}^{\mathbb{P}_t}\left[\cdot|\underline{\mathcal{F}_t}\right]$ is the expectation operator conditional on $\underline{\mathcal{F}_t}$.[1] If investors are fully rational, then $\mathbb{P}_t$ denotes the objective conditional probability measure given the optimal filtration $\underline{\mathcal{F}_t}$. Therefore, for any random process Y_τ taking values on $\mathrm{supp}(Y_\tau)$, where $\tau > t$,

$$\mathbb{E}\left[Y_\tau|\underline{\mathcal{F}_t}\right] = \int_{\mathrm{supp}(Y_\tau)} Y_\tau \,\mathrm{d}\,\mathbb{P}_t(Y_\tau) \tag{2}$$

Macro models usually identify the SDF as a parametric function of consumption growth between t and $t+1$ denoted by C_{t+1}/C_t, and a set of other possible risk factors that we

[1] Note that both the SDF and the excess returns are expressed in real terms and not in nominal.

denote by Y_{t+1}:

$$M_{t+1} = M\left(\frac{C_{t+1}}{C_t},\, Y_{t+1};\, \theta_0\right)\,, \tag{3}$$

where θ_0 is the vector of parameters driving the SDF.

For Equation (1) to hold with respect to the objective conditional probability measure $\mathbb{P}_t$, we need to assume rational expectations. However, investors may have certain behavioral biases or robust control preferences that distort their beliefs relative to rational expectations and beliefs need not be rational. For the sake of generality, we assume deviations from the rational expectations framework can exist but we do not take a stand on the sources of belief distortions. When for any reason beliefs deviate from rational expectations, Equation (1) needs to be taken with respect to the representative agent's subjective probability measure that we denote by $\widetilde{\mathbb{P}}_t$. We thus have:

$$\mathbb{E}^{\widetilde{\mathbb{P}}_t}\left[M\left(\frac{C_{t+1}}{C_t},\, Y_{t+1};\, \theta_0\right)\mathbf{R}^e_{t+1}\,\middle|\,\underline{\mathcal{F}_t}\right] = \mathbf{0}\,. \tag{4}$$

Assuming the measures $\mathbb{P}_t$ and $\widetilde{\mathbb{P}}_t$ are absolutely continuous, we can write:

$$\frac{\mathrm{d}\,\widetilde{\mathbb{P}}_t}{\mathrm{d}\,\mathbb{P}_t} = Z_t\,, \tag{5}$$

where Z_t is the family of Radon-Nikodym derivatives of $\widetilde{\mathbb{P}}_t$ with respect to $\mathbb{P}_t$. Note that, in the absence of any beliefs distortion relative to the objective measure, we have $\widetilde{\mathbb{P}}_t = \mathbb{P}_t$ at each point in time, and $Z_t = 1$ almost surely. In this paper, our objective is to identify the manifestations of distortions from $\mathbb{P}_t$ to $\widetilde{\mathbb{P}}_t$ in the mostly agnostic possible way. The following sub-sections detail how our econometric methodology allow us to fulfill this goal.

II.2 The smoothed empirical likelihood estimator

Our identification scheme relies on the non-parametric *smoothed empirical likelihood* estimation approach (SEL henceforth) developed by Kitamura, Tripathi, and Ahn (2004). This is akin to the notion of a non-parametric maximum likelihood family of estimators. We detail below the general procedure and how it fits into our framework.

To give intuition and fix ideas, let us first consider a multinomial model with as many possible outcomes as observation dates, that we denote by T. It is easy to show that, in the absence of any constraints other than requiring the probabilities to be positive and sum to unity, a standard maximum likelihood estimator will yield that every probability estimate $\widehat{p}_t = \frac{1}{T}$, for $t \in \{1, \ldots, T\}$. Of course, this is barely a model since pooling together the probabilities in bins will produce a standard histogram representation and the multinomial assumption provides us with nothing more than standard descriptive statistics. Now assume we perform the same likelihood maximization enforcing that the moments restrictions given by Equation (4) hold true given the SDF, $M\left(\frac{C_{t+1}}{C_t}, Y_{t+1}; \theta_0\right)$. The method will now distort the probability estimates $\widehat{p_t}$ to satisfy the conditional moment restrictions.

In the following, we assume that the information set of the investors can be represented by a finite vector of random variables that we denote by $X_t \in \mathbb{R}^n$. The choice of X_t will be crucial for our analysis, but we delay this discussion to the next section and consider for now that it can be anything relevant to explain the joint behavior of macro variables and asset returns. Suppose that the historical realizations of consumption growth, other variables in the SDF, excess returns, and the conditioning variables are given by $\left(g_{t+1} = \frac{c_{t+1}}{c_t},\, y_{t+1},\, \mathbf{r}^e_{t+1}, x_t\right)_{t=0}^{T-1}$. Let $p_{i,j}$ be the probability of observing the joint outcome $(g_j, y_j, \mathbf{r}^e_j)$, given the conditioning variables value x_i. Without any moment restrictions, we have that the only non-zero $p_{i,j}$'s are for $i = t-1$ and $j = t$, such that $p_{t-1,t} = 1$. In general, conditional moment conditions will imply deviations from this simple case. The SEL estimator for the conditional probabilities $(p_{i,j})$ for $i = \{0, \ldots, T-1\}$ and $j = \{1, \ldots, T\}$ are such that they belong to the simplex:

$$\Delta := \cup_{i=0}^{T-1} \Delta_i = \cup_{i=0}^{T-1} \left\{ (p_{i,1}, ..., p_{i,T}) : \sum_{t=1}^{T} p_{i,t} = 1,\, p_{i,t} \geq 0 \right\}$$

and that: $\forall i \in \{0, \ldots, T-1\}, \quad \forall \theta \in \Theta$,

$$\left(\widehat{p}_{i,\cdot}^{SEL}(\theta)\right) = \underset{(p_{i,\cdot}) \in \Delta_i}{\arg\max} \sum_{t=1}^{T} \omega_{i,t} \log(p_{i,t}) \quad \text{s.t.} \quad \sum_{t=0}^{T-1} p_{i,t+1} \times M\left(g_{t+1},\, y_{t+1};\, \theta\right) \mathbf{r}^e_{t+1} = \mathbf{0}, \qquad (6)$$

where $p_{i,\cdot}$ is a shortcut for the $T-$dimensional vector of probabilities $(p_{i,1}, ..., p_{i,T})$, Θ is the set of all admissible parameters θ, and $\omega_{i,t}$ are non-negative weights used to smooth the objective function. In the spirit of non-parametric estimators:

$$\omega_{i,t} = \frac{\mathcal{K}\left(\dfrac{x_i - x_t}{b_T}\right)}{\displaystyle\sum_{j=0}^{T-1} \mathcal{K}\left(\dfrac{x_i - x_j}{b_T}\right)}, \qquad (7)$$

where $\mathcal{K}$ is a kernel function belonging to the class of second order product kernels,[2] and

[2] $\mathcal{K}$ should satisfy Assumption 3.3 in Kitamura, Tripathi, and Ahn (2004), that is restated here for convenience. For $X = (X^{(1)}, X^{(2)}, ..., X^{(n)})$, let $\mathcal{K} = \prod_{i=1}^{n} k(X^{(i)})$. Here $k : \mathbb{R} \to \mathbb{R}$ is a continuously differentiable p.d.f. with support $[-1, 1]$. k is symmetric about the origin, and for some $\alpha \in (0, 1)$ is

the bandwidth b_T is a smoothing parameter.[3] Equation (6) is simply a transformation of a multinomial log-likelihood, using a different set of weights than the observed values, and enforcing the conditional moment conditions for a given specification of the SDF. It is easy to show that the solution to Equation (6) is analytical and given by:

$\forall i \in \{0, \ldots, T-1\}, \quad \forall t \in \{1, \ldots, T\},$

$$\widehat{p}_{i,t}^{SEL}(\theta) = \frac{\omega_{i,t}}{1 + M(g_t,\, y_t;\, \theta) \cdot \widehat{\lambda}_i(\theta)' \, \mathbf{r}_t^e}\,, \tag{8}$$

where $\widehat{\lambda}_i(\theta) \in \mathbb{R}^k : i = \{0, \ldots, T-1\}$ are the Lagrange multipliers associated with the Euler equation constraints, and solve the following unconstrained maximization problem:

$$\widehat{\lambda}_i(\theta) = \underset{\lambda_i \in \mathbb{R}^k}{\text{argmax}} \sum_{t=1}^{T} \omega_{i,t} \log \left[1 + M(g_t, y_t; \theta) \cdot \lambda_i' \, \mathbf{r}_t^e\right]. \tag{9}$$

Equations (8) and (9) show that, although the SEL procedure delivers a $(T \times T)$ matrix of probabilities $(\widehat{p}_{i,t}(\theta))$ for each value of the parameter vector θ, the number of parameters that it needs to estimate in order to generate the probability matrix is only $T \times k$. Indeed, for each date, the SEL procedure requires estimation of the vector of Lagrange multipliers associated with the conditional Euler equation restrictions and the number of parameters equals the number of test assets that the SDF is asked to price. For instance, if the market return is the only asset used in the estimation, the overall number of Lagrange multipliers – and, therefore, the total number of parameters – that need to be estimated is T.[4]

In practice, it can happen that the argument of the log function becomes negative at certain dates. This creates numerical instability in estimation and makes λ_i a corner solution to the optimization problem (9). In order to prevent this case, we use Owen (2001) normalization:

$$\widehat{\lambda}_i^{(o)}(\theta) \quad = \quad \underset{\lambda_i \in \mathbb{R}^k}{\text{argmax}} \sum_{t=1}^{T} \omega_{i,t} \cdot \Psi_\nu \left[1 + M(g_t, y_t; \theta) \cdot \lambda_i' \, \mathbf{r}_t^e\right] \tag{10}$$

$$\text{where} \quad \Psi_\nu(x) = \begin{cases} \log(x) & \text{if} \quad x > \nu \\ \log(\nu) - \frac{3}{2} + 2\frac{x}{\nu} + \frac{1}{2}\left(\frac{x}{\nu}\right)^2 & \text{if} \quad x \leqslant \nu \end{cases} \tag{11}$$

bounded away from zero on $[-a, a]$.

[3] In theory, b_T is a null sequence of positive numbers such that $Tb_T \to \infty$. See Assumption 3.7 in Kitamura, Tripathi, and Ahn (2004) for additional restrictions on the choice of b_T.

[4] This dramatic reduction in the dimensionality of the optimization problem is achieved because the SEL estimator is the solution to a convex optimization problem, and, therefore, the Fenchel duality applies (see, e.g., Borwein and Lewis (1991)).

Equation (11) defines a continuously differentiable function which is easier to manipulate when the argument is close to zero. Owen (2001) recommends using $\nu = 1/T$, which we follow in our empirical approach. Using the previous transformation of the objective function can make the sum of the estimated probabilities with $\widehat{\lambda}_i^{(o)}$ (see Equation (8)) deviate from one. Again, Owen (2001) suggests to normalize the probabilities *ex-post* so that they add up to one:

$$\widehat{p}_{i,t}^{SEL^{(o)}}(\theta) = \frac{\omega_{i,t}}{1 + M(g_t,\, y_t;\, \theta) \cdot \widehat{\lambda}_i^{(o)}(\theta)' \, \mathbf{r}_t^e} \times \left(\sum_{t=1}^{T} \frac{\omega_{i,t}}{1 + M(g_t,\, y_t;\, \theta) \cdot \widehat{\lambda}_i^{(o)}(\theta)' \, \mathbf{r}_t^e} \right)^{-1}, \quad (12)$$

The notation $\left(\widehat{p}_{i,t}^{SEL}(\theta)\right)$ emphasizes that the estimated probabilities are functions of the chosen value of θ. Since the true θ_0 is unknown to the econometrician, the SEL method also allows to estimate it. Let us denote by $\ell_{i,T}(\theta)$ each of the conditional log-likelihood functions to maximize given by Equation (6). The non-parametric log-likelihood function is

$$\ell_T(\theta) = \sum_{i=0}^{T-1} \ell_{i,T}(\theta) \quad (13)$$

Using Equation (8), the smoothed empirical log-likelihood (SEL) at θ is defined as:

$$\ell_T(\theta) \;\; = \;\; \sum_{i=0}^{T-1} \sum_{t=1}^{T} \mathbb{T}_{i,t} \cdot \omega_{i,t} \cdot \log\left(\widehat{p}_{i,t}^{SEL^{(o)}}(\theta) \right),$$

where $\mathbb{T}_{i,t}$ is a sequence of trimming functions, incorporated in the log-likelihood to deal with the well-known *denominator problem* associated with kernel estimators. The maximum SEL estimator of θ is then defined as:

$$\widehat{\theta}^{SEL} = \underset{\theta \in \Theta}{\text{argmax}} \quad \ell_T(\theta)\,. \quad (14)$$

Kitamura, Tripathi, and Ahn (2004) show that the SEL approach delivers an efficient estimator of θ, i.e. the estimator achieves the semi-parametric efficiency bound for conditional moment restriction models.

II.3 An Alternative Interpretation of the SEL estimator

The SEL estimator also has an important information-theoretic interpretation (see, e.g., Kitamura and Stutzer (1997)). To see this, let $\mathcal{P}_t$ be the set of all conditional probability measures defined on $\mathbb{R}^q \times \mathbb{R}^k$, where q denotes the dimension of the variables entering the

SDF and k denotes the dimension of the set of assets used in the estimation. For any set of admissible SDF parameters $\theta \in \Theta$, we define the set of probability measures absolutely continuous with respect to the objective measure $\mathbb{P}_t$ that satisfy the consumption Euler equations:

$$\mathcal{P}_t(\theta) := \left\{ \pi \in \mathcal{P}_t : \mathbb{E}^{\pi} \left[M\left(\frac{C_{t+1}}{C_t}, Y_{t+1}; \theta \right) \mathbf{R}^e_{t+1} \middle| X_t \right] = \mathbf{0} \right\} \tag{15}$$

Therefore, $\cup_{\theta \in \Theta} \mathcal{P}_t(\theta)$ is the set of all the conditional probability measures that are consistent with the model characterized by the moment conditions in Equation (17). The EL estimation can then be shown to select a probability measure π and the estimator of the parameter vector θ so as to:

$$\begin{aligned} \inf_{\theta \in \Theta} \inf_{\pi \in \mathcal{P}_t(\theta)} K(\mathbb{P}_t, \pi) \quad &\equiv \quad \inf_{\theta \in \Theta} \inf_{\pi \in \mathcal{P}_t(\theta)} \int \log \left(\frac{\mathrm{d}\, \mathbb{P}_t}{\mathrm{d}\, \pi} \right) \mathrm{d}\, \mathbb{P}_t \\ \text{s.t.} \quad & \mathbb{E}^{\pi} \left[M\left(\frac{C_{t+1}}{C_t}, Y_{t+1}; \theta \right) \mathbf{R}^e_{t+1} \middle| X_t \right] = \mathbf{0}, \end{aligned} \tag{16}$$

where $K(\mathbb{P}, \pi)$ is the Kullback-Leibler information criterion (KLIC) divergence from π to $\mathbb{P}_t$ (White (1982)). Our representation of beliefs distortion equalizes the π identified from the SEL estimator with the conditional subjective beliefs of investors, $\widetilde{\mathbb{P}}_t$. Note that, $K(\mathbb{P}_t, \widetilde{\mathbb{P}}_t) \geqslant 0$ and it will hold with equality if and only if $\widetilde{\mathbb{P}}_t = \mathbb{P}_t$, that is if investors are rational. On the other hand, if beliefs are not rational, $\mathbb{P}_t$ is not an element of $\mathcal{P}_t(\Theta)$ and for each θ there is a positive minimum KLIC distance attained by the solution of the inner minimization in Equation (16). Thus, the SEL approach searches for an estimate of $\widetilde{\mathbb{P}}_t$ that makes the estimated subjective beliefs as close as possible – in the information-theoretic sense – to the objective one.

II.4 Parametric specification of the SDF

Different modeling assumptions leading to different reformulations of Equation (3), our benchmark case in the following will consider the standard C-CAPM of Breeden (1979), Lucas (1978) and Rubinstein (1976), where the utility function is time and state separable with a constant coefficient of relative risk aversion. For this specification of preferences, the SDF takes the form:

$$M\left(\frac{C_{t+1}}{C_t}, Y_{t+1}; \theta_0 \right) = \left(\frac{C_{t+1}}{C_t} \right)^{-\theta_0} \quad \text{and} \quad \mathbb{E}^{\widetilde{\mathbb{P}}_t} \left[\left(\frac{C_{t+1}}{C_t} \right)^{-\theta_0} \mathbf{R}^e_{t+1} \,|\, X_t \right] = \mathbf{0}, \tag{17}$$

where $Y_{t+1} = \emptyset$ and $\theta_0 \in \mathbb{R}^+$ is the representative agent's coefficient of relative risk aversion (CRRA), and C_{t+1}/C_t denotes the real per capita aggregate consumption growth (gross).[5]

We are well-aware that the above pricing kernel fails empirically to explain *i)* the historically observed levels of returns, giving rise to the *Equity Premium* and *Risk Free Rate* Puzzles (e.g. Mehra and Prescott (1985), Weil (1989)), and *ii)* the cross-sectional dispersion of returns between different classes of financial assets (e.g. Mankiw and Shapiro (1986), Breeden, Gibbons, and Litzenberger (1989), Campbell (1996), Cochrane (1996)). As a result, we also consider two alternative specifications of the SDF that were designed to overcome some of the limitations of the C-CAPM and have substantially superior empirical performance compared to the latter. Our two alternatives are the external habit formation model of Campbell and Cochrane (1999), andthe long-run risk model of Bansal and Yaron (2004) with Epstein and Zin (1991) preferences. Since they are standard in this literature, we refer the reader to Appendix A.1 for more detail.

III Data Description

We present empirical results at the quarterly and annual frequencies. At the annual frequency, because of the different starting periods of available total and nondurable consumption series, we focus on two data samples: a baseline data sample starting at the onset of the Great Depression (1929–2016) and a longer data set (1890–2009) obtained from Campbell (2003) and Robert Shiller's Web site. We use the shorter sample as our baseline because only over this period total consumption can be disaggregated into its nondurable and durable components, but we also report results for the longer sample as a robustness check.

For the 1929–2016 data sample, our proxy for the market return is the Center for Research in Security Prices (CRSP) value-weighted index of all stocks on the NYSE, AMEX, and NASDAQ. The proxy for the risk-free rate is the one-month Treasury-bill rate. Annual returns for the above assets are computed by compounding monthly returns within each year and are converted to real returns using the personal consumption deflator. For consumption, we use per capita real personal consumption expenditures on nondurable goods and services from the National Income and Product Accounts (NIPA).

For the longer data set, the return on the S&P composite index is used as a proxy for the market return. Because of data availability issues, we use the prime commercial paper rate as a proxy for the risk-free rate, therefore partially underestimating the magnitude of the EPP. Consumption refers to the real per capita total personal consumption expenditures. See Campbell (1999) for a detailed data description.

[5] Note that θ_0 is usually assumed unknown to the econometrician but is known to the economic agents.

Similarly, at the quarterly frequency, we use the longest available data sample covering the period 1947:Q1–2016:Q4. The measure of consumption and the proxies for the market return and risk free rate are identical to those for the baseline annual sample.

We make the standard "end-of-period" timing assumption that consumption during year (quarter) t takes place at the end of the year (quarter).

As discussed in Section V, we estimate investors' (subjective) beliefs for a few different choices of the conditioning set. The conditioning variables used include the growth rate in the CPI-U and the growth in the average hourly earnings of production on private nonfarm payrolls. The availability of this data dictates the starting periods of our estimation using each of these conditioning variables. Historical data on the CPI-U and average hourly earnings, available from 1890 and 1940 onwards, respectively, are obtained from the Bureau of Economic Analysis.

III.1 Conditioning Set

The econometric feasibility of the beliefs extraction procedure described in Section II.2 relies crucially on being able to characterize the conditioning set, X_t, underlying the consumption Euler equations with a small number of variables. Our choice of the conditioning set draws on the insight in Ghosh and Constantinides (2017), who contribute towards identifying the investors' information set.

Ghosh and Constantinides (2017) present evidence that the market-wide price-dividend ratio is strongly correlated with two groups of macro variables – inflation and labor market – both in level as well as in first difference. Moreover, this high correlation is observed not just in the US, but also in all the other G7 countries. They show that a standard learning model, where investors' information sets are expanded to allow learning about the latent economic regime (and, therefore, about the average future growth rates of the economy) not only from the consumption history alone as is common in the existing literature, but also from inflation and labor market variables, can go a long way toward explaining many observed features of the macroeconomy and financial markets. In particular, their results suggest that just two macroeconomic variables – the rate of change in the CPI and the growth in average hourly earnings of production on private non farm payrolls – along with consumption growth go a long way towards proxying for investors' relevant information sets.

Drawing on the findings in Ghosh and Constantinides (2017), we present results for a few different choices of the conditioning set, X_t. Specifically, we consider the following choices for X_t: (a) the rate of growth in the Consumer Price Index for all Urban Consumers (CPI-U) and consumption growth, (b) the growth in the average hourly earnings of production on private nonfarm payrolls in the manufacturing sector and consumption growth, and (c)

CPI-U growth, the growth in the average hourly earnings, and consumption growth.

IV Estimated conditional distributions with moment restrictions

We present results when investors' have power utility preferences with a constant coefficient of relative risk aversion, the conditioning set includes lagged consumption growth and inflation, and the excess return on the market portfolio is the sole test asset. All our results are computed with Epanechnikov kernel function, and the bandwidth parameter $b_{v,T} = 2\hat{\sigma_v}$ where $\hat{\sigma_v}$ is the empirical standard deviation of the conditioning variable v.[6] In our benchmark case, we use Owen's renormalization as presented in Section II.2 and we set the SDF risk aversion coefficient θ_0 equal to 10 (the upper limit of what is considered to be an acceptable range).

In this Section, we do not assume a particular true data generating process for the different variables in our model. As such, we are only able to identify deviations of the conditional distributions estimated through SEL with respect to the marginal distribution. Even if we do not take a stance in this Section on whether the estimated probabilities are consistent with rational expectations, one can view this first set of results as guidance for modeling assumptions in consumption-based asset pricing models.

IV.1 Estimated conditional probabilities

For each possible value of the vector of conditioning variables, our approach delivers the conditional probabilities attached to possible states of the world in the next period. To facilitate interpretation and characterization of the results, we present these probabilities for a few different values of the conditioning state vector. In particular, we divide the realizations of consumption growth into three groups – low consumption growth (LG, $g_t < 0\%$), medium consumption growth (MG, $g_t \in [0\%, 1.5\%]$), and high consumption growth (HG, $g_t > 1.5\%$). Using a similar procedure, we divide the realizations of inflation into low inflation (LI, $\pi_t < 1\%$), medium inflation (MI, $\pi_t \in [1\%, 4\%]$), and high inflation (HI, $\pi_t > 4\%$). We then sort the joint realizations of consumption growth and inflation into nine groups formed from the intersection of the two sets. For instance, if the consumption growth at time t falls in the LG group and the inflation at time t falls in the LI group, then the joint realization of

[6] The results are robust to alternative choices of the kernel function and smoothing parameters within three standard deviations of the respective conditioning variables. These results are omitted for brevity and are available from the authors upon request.

consumption growth and inflation at time t is assigned to the group (LG, LI). Our sorting procedure implies that each group typically has one or more observation.

[Insert Figure 1 about here.]

Figure 1 plots the conditional probabilities for the above nine consumption growth-inflation groups and compares them to what would be obtained without any moment restriction, that is $1/T$ (red solid line). In essence, this tells us for each values of the conditioning variables and when the method needs to significantly depart from the i.i.d. assumption to be consistent with the conditional moment condition as given by the Euler equation (17). We can thus identify the periods where the conditional distributions differ from the marginal. At least three key features emerge from the picture.

First, most of the spikes of the estimated conditional probabilities are observed during recessions, whatever the value of the conditioning set. This confirms that the method is clearly able to identify crises as unconventional states of the world, that is states where the conditional distributions depart greatly from the stationary state. Second, for medium growth or medium inflation (panels 2 or panel b's), the conditional probabilities are overweight with respect to $1/T$ during normal times. This shows that the estimation method is able to capture the persistence observed in stable macroeconomic periods. Last, the extreme states of the world show probabilities spikes during different crises periods. The most informative are the (LO, LI) and (LO, HI) conditioning states (panels a.1 and c.1 respectively). For the low growth - low inflation state, we observe conditional probabilities jumping up to [10%-18%] during the great depression, the second world war and right after the 2008 financial crisis. In comparison, the low growth - high inflation state has peaks of 8% to 18% during the aftermath of the second world war, and the two oil crashes. This is particularly consistent with the nature of the shocks leading to the depressions, with deflationary pressures for the former and stagflation for the latter. Thus, by only incorporating one asset pricing moment condition, the estimation method is already able to identify economically significant conditional macroeconomic densities. Interestingly, for the 2008 financial crisis, the conditional probabilities are all higher than $1/T$ regardless of the value of the conditioning variables. This can be a reflection of the systemic nature of the crisis, where the causes are to be found outside of the set of conditioning variables we are including in the estimation. Another possible explanation is that agents tend to be pessimistic and overestimate the probability of the last crisis because of distorted beliefs. We explore the possibility of the latter in the next Section.

IV.2 Conditional distributions of consumption growth and inflation

To complement the previous approach, we explore deviations from marginal to conditional distributions estimated with the SEL procedure by representing histograms for each conditioning set. Because of the generality of the estimation method, we are able to derive the conditional distribution of both consumption growth and inflation. The results are presented on Figures 2 and 3 respectively.

[Insert Figures 2 and 3 about here.]

Looking at the distribution of consumption growth first, we observe a distinct pattern of over-weighting medium and moderately low growth states whatever the values of past consumption and past inflation. For negative past growth and medium past inflation for instance (panel b.1 of figure 2) we see that the conditional probabilities for falling in any bin of consumption growth below 2% is either virtually equal or higher than the marginal ones. Although for medium and high growth (panels 2. and 3.) the estimated conditional probabilities resemble qualitatively the marginal distribution of consumption, the highest deviations are observed for low growth states. For negative past growth and low past inflation, the probabilities of observing negative growth can be nearly 10 times higher than the marginal counterparts (see panel a.1). In the same fashion, for negative past growth and high past inflation, probabilities of observing medium growth next period are at least twice as big as their marginal counterparts. These findings emphasize that conditional distributions do not only vary in mean and variance but their entire shape is affected by the values of the conditioning macro variables.

For conditional distributions of inflation, the same sort of pattern can be detected. For nearly all conditioning variables values, the conditional probabilities of observing medium inflation are higher than their marginal counterpart. However, deviations happening for extreme states of the world are even more blatant than for consumption growth. When past inflation is high (panel c.) the conditional probabilities of observing high inflation are largely overweighted, especially when past consumption growth is negative (see panel c.1 of Figure 3). When both past consumption growth and inflation are low, the conditional probabilities of observing low inflation can be as much as 10 times higher than their marginal counterparts.

These results show that conditional probabilities of macoreconomic variables shape in non-trivial ways with respect to their past values. These results question the traditional assumptions of DGP of any traditional rational expectation consumption-based asset pricing model. Before turning to a comparison with potential DGPs, we provide more precise information about these conditional distributions by looking at their conditional moments.

IV.3 Conditional moments

Since the SEL estimates provide the conditional probabilities of the different observables, it is easy to compute conditional moments given the different values of the conditioning set. Let us consider (net) consumption growth $\log(g_{t+1})$ for instance. The conditional mean as of date t is given by:

$$\widehat{\mathbb{E}}^{\widetilde{\mathbb{P}}_t}[\log(g_{t+1})|X_t] = \sum_{i=1}^{T} \widehat{p}_{i,t}^{SEL^{(o)}}(\theta) \cdot \log(g_i)\,. \tag{18}$$

Note that this conditional expectation can be computed for any date t, but also for any nonlinear transformation of the observable variables, and for the moments of any order. We perform this computation for both consumption growth and inflation for the first 4 moments, and group the resulting moments in the same conditioning sets as in the previous Sections. Results are presented in Tables 1 and 2 for consumption growth and inflation respectively.

[Insert Tables 1 and 2 about here.]

For consumption growth, the conditional mean increases with the past consumption growth emphasizing some persistence in the consumption process. the impact of past inflation is however either hump-shaped or decreasing with inflation values. When past consumption is between 0% and 1.5% (MG), the highest conditional mean values are obtained for low to medium inflation (LI and MI) and correspond to the marginal consumption growth mean, and high inflation has a negative impact on future consumption on average. For the conditional mean of inflation, the same sort of persistence and the conditional mean of inflation goes up with values of past inflation (see top-left panel of Table 2). The conditional mean of inflation tends to increase with values of past consumption growth except for high past inflation (HI) where it decreases. Again, past consumption growth has a nonlinear effect on the conditional mean of inflation, and can have a steepening effect on the conditional means with respect to past inflation: the high minus low inflation (HI-LI) when consumption growth is low (LG) is at 7.1pp for a 1.3pp when consumption growth is high (HG). These non-linearities represent a first source of discrepancy with respect to commonly assumed DGPs for these two variables (VARMA-type unobservable component models for instance).

Besides conditional means, recent consumption-based asset pricing models focus greatly on second-order conditional moments of endogenous variables (long-run risk models for instance). However, values of consumption growth conditional standard deviation are extremely similar at about 1.5pp for medium and high past growth (MG and HG), whatever the past value of inflation (see top-right panel of Table 1). When past growth is negative and inflation is low however (LG and LI), the conditional standard deviation increases to 3.5pp emphasizing how bad this state of the world is. For inflation, conditional standard deviations tend to vary more with respect to the conditioning values: higher past consumption growth tends to decrease inflation conditional standard deviations, while higher past inflation tends to make conditional standard deviations increase. Again, most of the action comes from the low growth state (LG), where conditional standard deviations can go as high as 4.4% (LG, LI state) against a 2.4% in a medium state (MG, MI). This shows that non-linearities are also present for second order conditional moments, and those can actually be very close to time-invariant in some periods.

Last, we can see that the discrepancies between the conditioning bins are the biggest for third and fourth conditional moments of both consumption growth and inflation. Looking at their marginal moments, they are both negatively skewed (-1.48 and -0.35 rep.) and both have fat tails (5.76 and 2.93 excess kurtosis resp.). However, the picture changes once we look at the conditional moments. For both variables the negative skewness is observed only when past consumption growth was negative (LG). For virtually all the remaining states, the conditional distributions are positively skewed. For most cases, conditional skewness increases with past values of consumption growth, and decreases with past values of inflation. The deviations can be substantial from one conditioning value to the other: for medium past inflation, the difference in conditional skewness between high and low past consumption growth (HG-LG) is of 2.73, and these two conditioning sets have nearly opposite skew of -1.50 and 1.23. For conditional kurtosis, again, the low past growth low past inflation state (LG, LI) seems to be much more different than the other states, implying negative excess kurtosis of both future growth and inflation. For consumption growth, kurtosis grows with past growth whereas inflation kurtosis tends to decrease with past inflation.

All in all, these results suggest three key messages. First, the linearity assumption used in most asset pricing models DGPs is not adequate to represent the conditional means and variances of endogenous variables. Second, third and fourth order conditional moments are where most of the deviations from time-invariance can be expected. Third, these third and fourth order conditional moments change the identification of *good* and *bad* states of the world for the representative investor. For instance, higher growth is usually considered a *good* state of the world, but is usually accompanied with high growth kurtosis and high positive inflation skew which are less desirable features.

V Identifying subjective beliefs

In the previous Section, we identified conditional distributions of macroeconomic variables from asset pricing restrictions using the SEL method. So far, we have remained silent on whether these probabilities defined objective distributions or represent subjective beliefs since we did not specify a true DGP. In this Section, we consider standard modeling assumptions from the most popular consumption-based asset pricing models and assume they represent the true DGPs. We identify how our SEL estimates deviate from these models' and explore where the belief distortions are the highest.

V.1 Models for the objective distribution

We follow the long run risk literature, that hypothesizes the presence of a small predictable component in expected consumption growth. The level of persistence of consumption growth is usually calibrated to a high value, with the argument that asset prices support the presence of a highly persistent component although this is difficult to identify using consumption data alone.[7] More recently, researchers have argued that even if the true underlying consumption dynamics are as postulated by the long run risks model, the statistical difficulty in distinguishing between alternative models with a finite available data history leads economic agents with robust control preferences to act from the perspective of the *worst case* model. Thus, agents act as if macro variables such as consumption growth and inflation are more persistent than what is observed in the data (see, e.g., Szoke (2017) and Bidder and Dew-Becker (2016).

Our first specification of the objective data generating process follows Szoke (2017) and assumes the following underlying model for the joint dynamics of consumption growth and inflation:

$$\begin{aligned} \begin{pmatrix} \Delta c_{t+1} \\ \pi_{t+1} \end{pmatrix} &= \beta_0 + X_t + \alpha \cdot \epsilon_{t+1} \\ X_{t+1} &= \kappa \cdot X_t + \sigma \cdot \epsilon_{t+1} \end{aligned}$$

where ϵ_{t+1} $N(0, I)$ is i.i.d., and X_t is latent to the econometrician. It is easy to show that this model is exactly a VARMA(1,1) where:

$$\begin{pmatrix} \Delta c_{t+1} \\ \pi_{t+1} \end{pmatrix} = (I - \kappa)\beta_0 + \kappa \begin{pmatrix} \Delta c_t \\ \pi_t \end{pmatrix} + (\sigma + \alpha)\epsilon_{t+1} - \kappa\alpha\epsilon_t\,, \tag{19}$$

and the model is easily estimated through (quasi) maximum likelihood. Results from the model are presented in Table 3. We also report R-squared giving the in-sample fit of the

[7] When estimated using consumption data alone, i.e. without using any asset price data in the estimation, the estimate of the persistence of expected consumption growth is much smaller than what is needed for the model to explain the historically observed level of the equity premium (Constantinides and Ghosh (2011)).

conditional mean. Looking at the estimates, the VARMA model does a reasonable job explaining both consumption growth and inflation, explaining around one and two thirds of the variance of the system. Szoke (2017) shows that if investors act from the point of view of the worst-case scenario, the parameter estimates are distorted and involves a pessimistic adjustment to the first and second moments of these variables. Under the subjective belief, expected inflation is more persistent and more strongly correlated with consumption growth than what the baseline model suggests.

Our second specification of the objective data generating process is the widely used regime-switching model, where the means of consumption growth and inflation differ across latent regimes:

$$\begin{pmatrix} \Delta c_{t+1} \\ \pi_{t+1} \end{pmatrix} = \begin{pmatrix} \mu_{c,s_{t+1}} \\ \mu_{\pi,s_{t+1}} \end{pmatrix} + \begin{pmatrix} \sigma_c & \rho \\ \rho & \sigma_\pi \end{pmatrix} \begin{pmatrix} \epsilon_{c,t+1} \\ \epsilon_{\pi,t+1} \end{pmatrix}$$

s_t is the scalar state variable that denotes the latent economic regime. Regime switching models of the type in Equation (20) have been extensively used in the macroeconomics and asset pricing literature (see, e.g., David and Veronesi (2013), Johannes, Lochstoer, and Mou (2016), Ghosh and Constantinides (2017)) and offer a flexible approach to modeling the underlying dynamics of macro and financial variables. Moreover, researchers have argued that this modeling choice offers a good empirical fit for the observed dynamics of the particular macro variables such as consumption growth and inflation that we include in our conditioning set. Therefore, it constitutes a good choice for the objective distribution of the data.

We present results when the number of regimes is set to five. We estimate the model with maximum likelihood, using historical data on consumption growth and inflation. This provides the objective beliefs, i.e. the conditional distribution of future consumption growth and inflation estimated from macro data alone without using any asset price data.

V.2 The distribution of subjective beliefs

For comparison with the results obtained through SEL, we need to form conditional distributions using our possible objective models. A slight difficulty lies in the fact that the conditioning sets used for the SEL method and the objective models are not the same. Indeed, it is well-known that a VARMA model is non-Markovian and that one should either include the entire history of the past realized values of the dependent variables to obtain conditional distributions. To overcome this issue, we simulate the models using the estimated parameters for 1,000,000 periods and group the simulated values with respect to the conditioning past inflation and consumption growth bins. For simplicity, we use a multivariate Gaussian distribution although sampling from the historical distribution leads to no economically significant difference. Using this particular procedure aligns the conditioning sets and allows use to compare similar quantities as produced by different estimation methods. Results for consumption growth and inflation are presented respectively on Figures 4 and 5.[8]

[Insert Figures 4 and 5 about here.]

For most cases, the differences between SEL estimates and the VARMA-based conditional distributions are more pronounced than for the marginal histograms. Because of its autoregressive structure, the VARMA is able to produce a distribution that moves with the values of the conditioning set. For the conditional distributions of consumption growth, the inflation conditioning plays little role in the VARMA while past growth has a positive impact on future consumption growth emphasizing its positive persistence, and its conditional distributions shift to the right from panels 1 to 3 (see Figure 4). However, the SEL method tends to put higher probabilities to the right side of the distribution than the VARMA for low and medium past growth, and lower probabilities to the right side of the distribution than the VARMA for high past growth. This means that the representative investor tends to be optimistic in bad and normal times, but rather pessimistic in good times. Second, contrary to the literature assumptions, the consumption growth persistence is perceived lower by the representative investor than estimated by the VARMA. Indeed, the probabilities of falling close to 2% (the marginal mean of consumption growth) is nearly always higher for the SEL than for the VARMA (see Figure 4).

Since inflation is more persistent that consumption growth, the VARMA-based conditional distributions tend to move more with respect to past inflation values (see Figure 5). When past inflation is low (LI), the SEL method underweights probabilities in the left side of the distribution which indicates that investors believe that inflation will come back to normal rapidly. The underweighting can represent nearly half of the entire distribution probability mass for (LI, MG) (see panel a2 of Figure 5). For medium past inflation (MI), we still observe a huge peak in medium inflation values given by the SEL, but the distortions now depends on the past growth value: for low and medium past growth, investors tend to overweight the right side of the distribution while they underweight the right side for high past growth (MI,

[8]One key feature to note is that although the VARMA model is conditionally i.i.d. Gaussian when the conditioning set includes both the previous period's endogenous variables and the shocks, it is not the case when the conditioning set contains only the previous period's endogenous variables and not the shocks. The same reasoning applies for the regime switching model.

HG). This feature is even more blatant for high past inflation (HI) where investors believe harder in long-lasting hyperinflation when the economy is in a low growth environment, but they believe in convergence to medium inflation values for medium and high growth. Overall, this also implies that investors tend to attribute less persistence to inflation than the one implied by historical data and are confident in the convergence to medium inflation except in extreme states of the world, such as deflationary and hyperinflationary recessions.

To better observe the sources of distortion in the distributions, we plot the change of measure from objective to subjective beliefs by dividing the values obtained with the SEL by the values obtained for the VARMA (Values of 0 are replaced by missing values). The results are presented on Figures 6 and 7.

[Insert Figures 6 and 7 about here.]

The obtained deviations from rational expectations are more diverse than those postulated in empirical asset pricing models. For instance, the U-shaped change of measure as assumed by Szoke (2017) is valid only for the past low growth and inflation state and past low growth and high inflation state for future growth and inflation respectively. In most cases, the change of measure is hump-shaped and has a peak for high values. This means beliefs are more highly distorted for medium values than for extreme values. A few exceptions appear: a decreasing change of measure is observed for consumption in the (MI, LG) state and for inflation in the (LI, LG) state. This would correspond to a standard pessimistic view when the states of the world are bad. The same pessimistic views are observed for inflation distributions in high past inflation states. Panels c.2 and c.3 of Figure 7 show an increasing pattern implying an overweighting of high inflation probabilities.

In sum, extreme cases seem to push investors to be pessimistic while more normal times push them towards optimism. We thus observe a persistence in beliefs distortions.

V.3 Distorted conditional moments

As in the previous Section, we further explore the distortions from objective to subjective beliefs by computing the conditional moments implied by the VARMA model. The results of these computations are presented in Tables 4 and 5.

[Insert Tables 4 and 5 about here.]

Having a look to conditional means of both consumption growth and inflation confirm the identified deviations from the distributions in the previous Section. The most interesting supplementary information is now given by the higher order conditional moments. For consumption growth variance, we see that investors usually believe in lower conditional standard deviations than what the VARMA implies. For skewness and kurtosis however, it is quite the opposite: for nearly all values of the conditioning variables, investors beliefs attribute higher skewness but higher excess kurtosis than the VARMA. Looking at Table 5, the same sort of conclusions can roughly be drawn except for the LG state: when past consumption growth is medium or high, investors believe in lower inflation standard deviation, higher skewness and higher kurtosis. However, for low past growth, the results on standard deviations and skewness are completely reversed and investors believe in high inflation variance and low negative skew. Beliefs distortions therefore affect third and fourth order moments to a significant order.

VI Conclusion and Extensions

The rational expectations hypothesis endows investors with the ability to rationally process available information to form beliefs about the true underlying data generating process and, therefore, about the future. Most prominent asset pricing models are based on the premise of rational expectations. While the assumption of rational expectations is intuitive and appealing and has, to date, constituted the dominant maintained hypothesis in financial economics, the above models have had considerable difficulty in explaining several observed aspects of asset markets. The shortcomings of rational models have led to the advent of behavioral models that assume that investors' have certain behavioral biases that influence their decision-making in the face of uncertainty. These biases distort the investors' beliefs relative to those that would be formed under rational expectations. Behavioral models typically assume certain specific forms of biases, about which there is limited empirical evidence. Moreover, specific biases often help explain certain aspects of the data better relative to rational models, while performing worse with other aspects of the data.

If behavioral biases are pervasive, then they should be reflected in observed asset prices. Relying on this insight, this paper proposes an information-theoretic (relative entropy minimization) approach to extracting investors' beliefs from asset prices, without making any specific functional form assumptions about the belief formation process. If investors have rational expectations, the extracted beliefs about future macroeconomic outcomes should coincide with objective beliefs estimated from historical macro data alone (without reference to asset market data). On the other hand, if investors are not rational and suffer from certain behavioral biases, then the beliefs extracted from asset prices would appear distorted relative to the rational benchmark.

Our methodology relies on the smooth empirical likelihood estimation approach proposed in Kitamura, Tripathi, and Ahn (2004). The procedure delivers a nonparametric maximum likelihood estimate of the investors' subjective beliefs and, therefore, the most likely beliefs given the data. Implementation of this approach requires the specification of the underlying stochastic discount factor (SDF), the investors' conditioning set underlying the conditional Euler equation restrictions, and the set of assets that the SDF is asked to price. We present our results for several choices of the above to demonstrate robustness.

We show that our estimated subjective beliefs differ substantially from both, widely used objective specifications of the data generating process as well as commonly assumed forms of beliefs distortion. Specifically, the conditional distributions of macro variables such as future consumption growth and inflation exhibit strongly countercyclical negative skewness, while the variances of the conditional distributions remain relatively flat. This strongly contrasts with the conditional log-normal specifications widely assumed in the literature. Also, whereas the literature argues that beliefs distortion make macro variables appear to be more persistent and, therefore, more volatile (under the commonly assumed autoregressive dynamics) than what is observed in the data, our estimated beliefs suggest that the distortion affects primarily the skewness making the underlying macro variables appear more negatively skewed than what can be estimated from the data alone. Also, the deviation between the subjective and objective beliefs increase substantially in bad states of the world, defined by particular realizations of the conditioning variables.

The analysis in the present paper assumed the existence of a representative investor. Important extensions of the work include allowing for heterogeneous investors with potentially asymmetric information and, therefore, heterogeneous beliefs. These are outside the scope of the current paper but present interesting directions for future research.

References

BANSAL, R., AND A. YARON (2004): "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 59(4), 1481–1509.

BARBERIS, N., AND R. THALER (2001): "A Survey of Behavioral Finance," in *Handbook of the Economics of Finance (forthcoming)*, ed. by G. Constantinides, M. Harris, and R. Stulz. Elsevier Science B.V., Amsterdam.

BARILLAS, F., L. P. HANSEN, AND T. J. SARGENT (2009): "Doubts or Variability?," *Journal of Economic Theory*, 144(6), 2388–2418.

BARRO, R. J. (2006): "Rare Disasters and Asset Markets in the Twentieth Century," *Quarterly Journal of Economics*, 121(3), 823–866.

BIDDER, R., AND I. DEW-BECKER (2016): "Long-Run Risk is the Worst Case Scenario," *American Economic Review*, 106(9), 2494–2527.

BORWEIN, J. M., AND A. LEWIS (1991): "Duality Relationships for Entropy-type Minimization Problems," *SIAM Journal of Control and Optimization*, 29, 325–338.

BREEDEN, D. T. (1979): "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities," *Journal of Financial Economics*, 7, 265–96.

BREEDEN, D. T., M. R. GIBBONS, AND R. H. LITZENBERGER (1989): "Empirical Test of the Consumption-Oriented CAPM," *The Journal of Finance*, 44(2), 231–262.

CAMPBELL, J. Y. (1996): "Understanding Risk and Return," *Journal of Political Economy*, 104(2), 298–345.

——— (1999): "Asset Prices, Consumption, and The Business Cycle," in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and M. Woodford, vol. 1C, chap. 19, pp. 1231–1303. Elsevier Press.

——— (2003): "Consumption-Based Asset Pricing," in *Handbook of the Economics of Finance*, ed. by G. Constantinides, M. Harris, and R. Stulz, chap. 13. North-Holland, Amsterdam.

CAMPBELL, J. Y., AND J. H. COCHRANE (1999): "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107(2), 205–51.

COCHRANE, J. H. (1996): "A Cross-Sectional Test of an Investment-Based Asset Pricing Model," *Journal of Political Economy*, 104(3), 572–621.

CONSTANTINIDES, G. M., AND A. GHOSH (2011): "Asset Pricing Tests with Long Run Risks in Consumption Growth," *The Review of Asset Pricing Studies*, 1, 96–136.

DAVID, A., AND P. VERONESI (2013): "What Ties Return Volatilities to Fundamentals and Price Valuations?," *Journal of Political Economy*, (121), 682–746.

EPSTEIN, L. G., AND S. E. ZIN (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–968.

——— (1991): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis," *Journal of Political Economy*, 99(2), 263–286.

Fama, E. F. (1969): "Efficient Capital Markets: A Review of Theory and Empirical Work," *Journal of Finance*, 25, 383–417.

Ghosh, A., and G. Constantinides (2017): "What Information Drives Asset Prices?," Manuscript, McGill University and University of Chicago.

Ghosh, A., C. Julliard, and A. P. Taylor (2016): "What is the Consumption-CAPM Missing? An Information-Theoretic Framework for the Analysis of Asset Pricing Models," forthcoming in *The Review of Financial Studies*.

Hansen, L. P., and T. J. Sargent (2001): "Robust Control and Model Uncertainty," *AER*, 91(2), 60–66.

——— (2016): "Set of Models and Prices of Uncertainty," NBER Working Papers 22000, NBER Inc.

Johannes, M., L. A. Lochstoer, and Y. Mou (2016): "Learning About Consumption Dynamics," *Journal of Finance*, 71, 551–600.

Kahneman, D., and A. Tversky (1979): "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47, 263–291.

Kitamura, Y., and M. Stutzer (1997): "An Information-Theoretic Alternative To Generalized Method Of Moments Estimation," *Econometrica*, 65(4), 861–874.

Kitamura, Y., G. Tripathi, and H. Ahn (2004): "Empirical Likelihood-Based Inference in Conditional Moment Restriction Models," *Econometrica*, 72(6), 1667–1714.

Kreps, D. M., and E. L. Porteus (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 46(1), 185–200.

Lucas, Jr., R. E. (1978): "Asset Prices in an Exchange Economy," *Econometrica*, 46, 1426–14.

Mankiw, N. G., and M. D. Shapiro (1986): "Risk and Return: Consumption Beta Versus Market Beta," *Review of Economics and Statistics*, 68, 452–59.

Mehra, R., and E. C. Prescott (1985): "The Equity Premium: A Puzzle," *Journal of Monetary Economics*, 15(2), 145–61.

Menzly, L., T. Santos, and P. Veronesi (2004): "Understanding Predictability," *Journal of Political Economy*, 112(1), 1–47.

Muth, J. F. (1961): "Rational Expectations and the Theory of Price Movements," *Econometrica*, 29(3), 315–335.

Owen, A. B. (2001): *Empirical Likelihood*. Chapman and Hall.

Piazzesi, M., J. Salomao, and M. Schneider (2015): "Trend and Cycle in Bond Premia," mimeo.

PIAZZESI, M., M. SCHNEIDER, AND S. TUZEL (2007): "Housing, consumption and asset pricing," *Journal of Financial Economics*, 83(3), 531–569.

RIETZ, T. A. (1988): "The Equity Risk Premium – a Solution," *Journal of Monetary Economics*, 22(1), 117–131.

RUBINSTEIN, M. (1976): "The Valuation of Uncertain Income Streams and the Pricing of Options," *Bell Journal of Economics*, 7, 407–25.

SAVAGE, L. (1964): *The Foundations of Statistics*. Wiley, New York.

SZOKE, B. (2017): "Estimating Robustness," *Working paper*, NYU.

WANG, X. (2017): "Investors' Subjective Beliefs and the Cross-Section of Stock Returns," Working Paper.

WEIL, P. (1989): "The Equity Premium Puzzle and the Risk-Free Rate Puzzle," *Journal of Monetary Economics*, 24, 401–421.

WHITE, H. (1982): "Maximum Likelihood of Misspecified Models," *Econometrica*, 50, 1–25.

A Appendix

A.1 Different Stochastic Discount Factors

Note that, while the C-CAPM implies that the SDF depends on contemporaneous consumption growth alone, more recent asset pricing models identify the SDF as a function of consumption growth *and* certain additional variables. The external habit formation model of Campbell and Cochrane (1999), for example, models the SDF as a function of contemporaneous consumption growth and the growth in the so-called surplus consumption ratio, that reflects the deviation of current consumption from a slow-moving habit level. In models with Epstein and Zin (1989) recursive preferences, such as the long run risks model of Bansal and Yaron (2004), the SDF is a function of consumption growth and the return on the total wealth portfolio. Therefore, in order to accommodate a variety of SDFs in our framework, we allow the SDF to be a parametric function of consumption growth $\left(\frac{C_{t+1}}{C_t}\right)$, certain additional variables summarized by the vector Y_{t+1}, and a set of parameters θ_0: $M_{t+1} = M(\frac{C_{t+1}}{C_t}, Y_{t+1}; \theta_0)$. The Euler equation rewrites:

$$\mathbb{E}^{\mathbb{P}_t}\left[M(\frac{C_{t+1}}{C_t}, Y_{t+1}; \theta_0)\mathbf{R}^e_{t+1} \,\middle|\, \underline{\mathcal{F}_t}\right] = \mathbf{0}, \tag{20}$$

We present our results for several different choices of the SDF, to demonstrate robustness.

We also consider a couple of alternative specifications of the SDF that were designed to overcome some of the limitations of the C-CAPM and that have substantially superior empirical performance compared to the latter. The first corresponds to the external habit formation paradigm (see, e.g., Campbell and Cochrane (1999)), where identical agents maximize power utility defined over the difference between consumption and a slow-moving habit or time-varying subsistence level. The SDF is given by

$$M^m_t = \delta(C_t/C_{t-1})^{-\gamma}\left(S_t/S_{t-1}\right)^{-\gamma}, \tag{21}$$

where δ is the subjective time discount factor, γ is the curvature parameter that provides a lower bound on the time varying coefficient of relative risk aversion, $S_t = \frac{C_t - X_t}{C_t}$ denotes the surplus consumption ratio, and X_t is the habit level.

Note that the SDF depends on the surplus consumption ratio, S, that is not directly observed. We extract the time series of the surplus consumption ratio from observed consumption data as follows.

In the Campbell and Cochrane (1999) model, the aggregate consumption growth is assumed to follow an $i.i.d.$ process:

$$\Delta \log(C_t) = g + \upsilon_t, \quad \upsilon_t \sim i.i.d.N\left(0, \sigma^2\right).$$

The log surplus consumption ratio evolves as a heteroskedastic $AR(1)$ process:

$$\log(S_t) = (1-\phi)\,\overline{\log(S)} + \phi \log(S_{t-1}) + \lambda\left(\log(S_{t-1})\right)\upsilon_t, \tag{22}$$

where $\overline{\log(S)}$ is the steady state log surplus consumption ratio and

$$\lambda\left(\log(S_t)\right) = \begin{cases} \frac{1}{\overline{S}}\sqrt{1-2\left(\log(S_t) - \overline{\log(S)}\right)} - 1, & \text{if } \log(S_t) \le s_{max} \\ 0, & \text{if } \log(S_t) > s_{max} \end{cases},$$

$$s_{max} = \overline{\log(S)} + \frac{1}{2}\left(1 - \overline{S}^2\right), \qquad \overline{S} = \sigma\sqrt{\frac{\gamma}{1-\phi}}.$$

For each value of γ, we use the calibrated values of the model's preference parameters (δ, ϕ), the sample mean (g) and volatility (σ) of the consumption growth process, and the innovations in real consumption growth, $\widehat{\upsilon}_t = \Delta c_t - g$, to extract the implied time series of the surplus consumption ratio using equation (22).[9]

Our final specification of the SDF is that implied by the long run risks model of Bansal and Yaron (2004). This model assumes that the representative consumer has the version of Kreps and Porteus (1978) preferences adopted by Epstein and Zin (1989) and Weil (1989), for which the SDF is given by

$$M^m_{t+1} = \delta^\theta \left(\frac{C_{t+1}}{C_t}\right)^{-\frac{\theta}{\rho}} R^{\theta-1}_{c,t+1},$$

[9]Note that the above approach to obtaining historical realizations of the surplus consumption ratio relies on the specific functional form assumptions about the dynamics of consumption growth and the surplus consumption ratio made in Campbell and Cochrane (1999), which are arguably restrictive. Therefore, we also adopt a second, more agnostic approach to measuring the surplus consumption ratio. We assume that the log habit level is a weighted average of past log consumption levels:

$$\log(X_t) = \sum_{j=0}^{\infty} a_j \log(C_{t-1-j}). \tag{23}$$

The above dynamics of log habit would emerge as a log-linear approximation around the non-stochastic steady-state for a variety of parametric assumptions about the consumption growth and the surplus consumption ratio. We assume an exponential specification for the coefficients, $a_j = \rho^j$ and lag length of five years (or sixty quarters) to impute the historical time series of the external habit level. The estimated subjective beliefs with this more model-free specification of the habit level are qualitatively similar to those obtained with the parametric specifications in Campbell and Cochrane (1999) and are omitted for brevity.

where $R_{c,t+1}$ is the unobservable gross return on an asset that delivers aggregate consumption as its dividend each period, δ is the subjective time discount factor, ρ is the elasticity of intertemporal substitution, $\theta := \frac{1-\gamma}{1-1/\rho}$, and γ is the relative risk aversion coefficient.

The aggregate consumption and dividend growth rates, Δc_{t+1} and Δd_{t+1}, respectively, are modeled as containing a small persistent expected growth rate component, x_t, that follows an AR(1) process with stochastic volatility, and fluctuating variance, σ_t^2, that evolves according to a homoscedastic linear mean reverting process.

Constantinides and Ghosh (2011) show that, for the log-linearized model, the log of the SDF is given by

$$\ln M_{t+1}^m = c_1 + c_2 \Delta c_{t+1} + c_3 x_{t+1} + c_4 \sigma_{t+1}^2 + c_5 x_t + c_6 \sigma_t^2 \tag{24}$$

where the parameters $(c_1, c_2, c_3, c_4, c_5, c_6)$ are known functions of the underlying time series and preference parameters of the model.

Note that the conditional mean of consumption growth, x_t, and its stochastic volatility, σ_t, are not directly observable. Using the calibrated parameter values from Bansal and Yaron (2004), we extract the state variables, x_t and σ_t^2, from observed consumption data, using a Bayesian smoother. The details of the method are described in Section C.1.

B Tables

Table 1 – Consumption growth conditional moments

	Mean (0.01923)				*Standard Deviation (0.02051)*			
	LG	MG	HG	HG−LG	LG	MG	HG	HG−LG
LI	0.00829	0.01933	0.02215	0.01386	0.03544	0.01539	0.01537	-0.02006
MI	0.01441	0.01928	0.02184	0.00743	0.01712	0.01531	0.01557	-0.00155
HI	0.01155	0.01682	0.01965	0.0081	0.0138	0.01511	0.01579	0.00199
HI−LI	0.00326	-0.00251	-0.0025		-0.02164	-0.00028	0.00041	
	Skewness (−1.47902)				*Excess Kurtosis (5.75789)*			
	LG	MG	HG	HG−LG	LG	MG	HG	HG−LG
LI	-0.5583	0.0935	0.1591	0.7173	-0.3229	0.5168	0.7828	1.1057
MI	-0.6689	0.0712	0.0698	0.7388	0.3645	0.5452	0.8175	0.453
HI	0.3201	0.0712	-0.0239	-0.3439	0.3046	-0.0173	0.4766	0.1721
HI−LI	0.8784	-0.0223	-0.1829		0.6274	-0.534	-0.3062	

Notes: This table contains the conditional moments of consumption growth as estimated by the SEL method. (LI, MI, HI) correspond to low (<1%), medium (∈ [1%, 4%]) and high inflation (>4%). (LG, MG, HG) correspond to low (<0%), medium (∈ [0%, 1.5%]) and high consumption growth (>1.5%). The numbers in italic are the marginal moments computed from the plug-in estimator $T^{-1}\sum_{i=1}^{T}\left(\frac{x_i-\bar{x}}{\hat{\sigma}(x)}\right)^k$.

Table 2 – Inflation conditional moments

	Mean (0.03175)				*Standard Deviation (0.03766)*			
	LG	MG	HG	HG−LG	LG	MG	HG	HG−LG
LI	-0.00893	0.02806	0.03222	0.04115	0.04414	0.0201	0.02278	-0.02136
MI	0.02222	0.03205	0.03605	0.01383	0.03529	0.02353	0.02635	-0.00893
HI	0.06226	0.04383	0.04527	-0.01699	0.04051	0.03101	0.03218	-0.00833
HI−LI	0.07119	0.01577	0.01305		-0.00363	0.01091	0.0094	
	Skewness (-0.34945)				*Excess Kurtosis (2.93326)*			
	LG	MG	HG	HG−LG	LG	MG	HG	HG−LG
LI	-0.8547	1.021	1.1963	2.051	-0.5186	3.0475	2.299	2.8176
MI	-1.4954	1.1917	1.2361	2.7314	4.6633	2.4246	1.8603	-2.803
HI	-0.9776	0.7879	0.895	1.8726	1.5751	0.2091	0.2913	-1.2838
HI−LI	-0.1229	-0.2331	-0.3012		2.0937	-2.8384	-2.0077	

Notes: This table contains the conditional moments of inflation as estimated by the SEL method. (LI, MI, HI) correspond to low (<1%), medium (∈ [1%, 4%]) and high inflation (>4%). (LG, MG, HG) correspond to low (<0%), medium (∈ [0%, 1.5%]) and high consumption growth (>1.5%). The numbers in italic are the marginal moments computed from the plug-in estimator $T^{-1}\sum_{i=1}^{T}\left(\frac{x_i-\bar{x}}{\hat{\sigma}(x)}\right)^k$.

Table 3 – VARMA coefficient estimates

	$(I-\kappa)\beta_0$	κ		$-\kappa\alpha$		R^2
$\log g_{t+1}$	0.01200	0.281	0.0663	0.2509	-0.1732	0.338
t-stat	2.6246	1.0922	0.7316	0.936	-1.1646	
π_{t+1}	0.0092	0.5186	0.4346	-0.1152	0.6559	0.684
t-stat	1.2800	1.5961	4.0677	-0.4349	4.7039	

Notes: This table contains the QML estimates of a bivariate VARMA(1,1) model estimated on real consumption growth $\log g_{t+1}$ and inflation π_{t+1} on annual data from 1930 to 2015. $(I-\kappa)\beta_0$ is the intercept, κ is the autoregressive matrix, and $\kappa\alpha$ is the moving average matrix. The R^2 column measures the fit of the conditional mean with respect to the realized values.

Table 4 – Consumption growth VARMA conditional moments

	Mean				*Standard Deviation*			
	LG	MG	HG	HG−LG	LG	MG	HG	HG−LG
LI	*0.00561*	*0.01443*	0.02481	0.01919	*0.01846*	0.018	0.01858	0.00011
MI	*0.00564*	*0.01457*	0.02561	0.01997	0.01827	0.01779	0.01879	0.00052
HI	*0.00588*	*0.01444*	0.02657	0.02069	0.01834	0.01794	0.01878	0.00044
HI−LI	0.00027	10^{-5}	0.00177		-0.00012	$-6\cdot 10^{-5}$	$2\cdot 10^{-4}$	
	Skewness				*Excess Kurtosis*			
	LG	MG	HG	HG−LG	LG	MG	HG	HG−LG
LI	0.0025	*-0.0673*	*0.0202*	0.0177	*0.0662*	*0.0406*	*-0.0082*	-0.0744
MI	0.005	*0.0133*	*0.0185*	0.0135	*-0.082*	*-0.0392*	*-0.0451*	0.0369
HI	*-0.0475*	*-0.0268*	*0.0246*	0.0721	*-0.0228*	*-0.0673*	*-0.0789*	-0.0561
HI−LI	-0.05	0.0406	0.0044		-0.089	-0.1079	-0.0707	

Notes: This table contains the conditional moments of consumption growth as estimated by a simple VARMA on consumption growth and inflation. (LI, MI, HI) correspond to low (<1%), medium (∈ [1%, 4%]) and high inflation (>4%). (LG, MG, HG) correspond to low (<0%), medium (∈ [0%, 1.5%]) and high consumption growth (>1.5%). Numbers in italic are lower than those observed in Table 1.

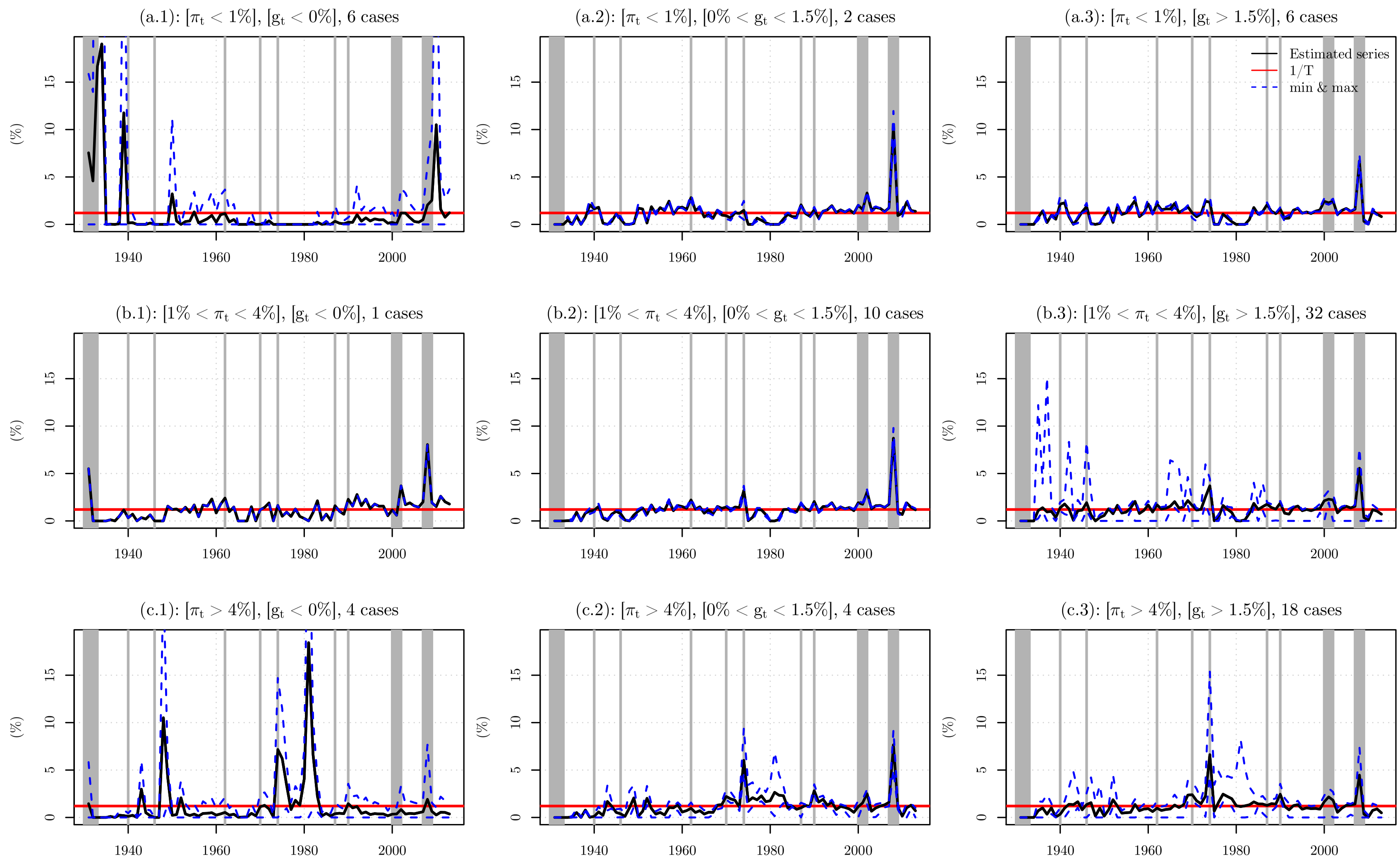
Table 5 – Inflation VARMA conditional moments

	Mean				*Standard Deviation*			
	Lg	Mg	Hg	Hg−Lg	Lg	Mg	Hg	Hg−Lg
Li	-0.0126	*-0.00176*	*0.01008*	0.02267	*0.02786*	0.02682	0.02743	-0.00044
Mi	*0.01446*	*0.02292*	*0.03422*	0.01975	*0.02489*	0.02479	*0.02529*	$4{\cdot}10^{-4}$
Hi	*0.03918*	0.04887	0.06389	0.02471	*0.02729*	*0.0274*	*0.02892*	0.00164
Hi−Li	0.05177	0.05064	0.05381		-0.00058	0.00058	0.00149	
	Skewness				*Excess Kurtosis*			
	Lg	Mg	Hg	Hg−Lg	Lg	Mg	Hg	Hg−Lg
Li	-0.12	*-0.0967*	*-0.0835*	0.0365	0.0372	*-0.0944*	*0.0711*	0.034
Mi	-0.011	*-0.0288*	*0.0307*	0.0417	*0.0087*	*-0.0034*	*-0.0035*	-0.0123
Hi	0.1214	*0.1246*	*0.154*	0.0325	*0.0049*	*0.062*	*0.0984*	0.0935
Hi−Li	0.2414	0.2214	0.2375		-0.0322	0.1564	0.0273	

Notes: This table contains the conditional moments of inflation as estimated by a simple VARMA on consumption growth and inflation. (Li, Mi, Hi) correspond to low (<1%), medium (∈ [1%, 4%]) and high inflation (>4%). (Lg, Mg, Hg) correspond to low (<0%), medium (∈ [0%, 1.5%]) and high consumption growth (>1.5%). Numbers in italic are lower than those observed in Table 2.

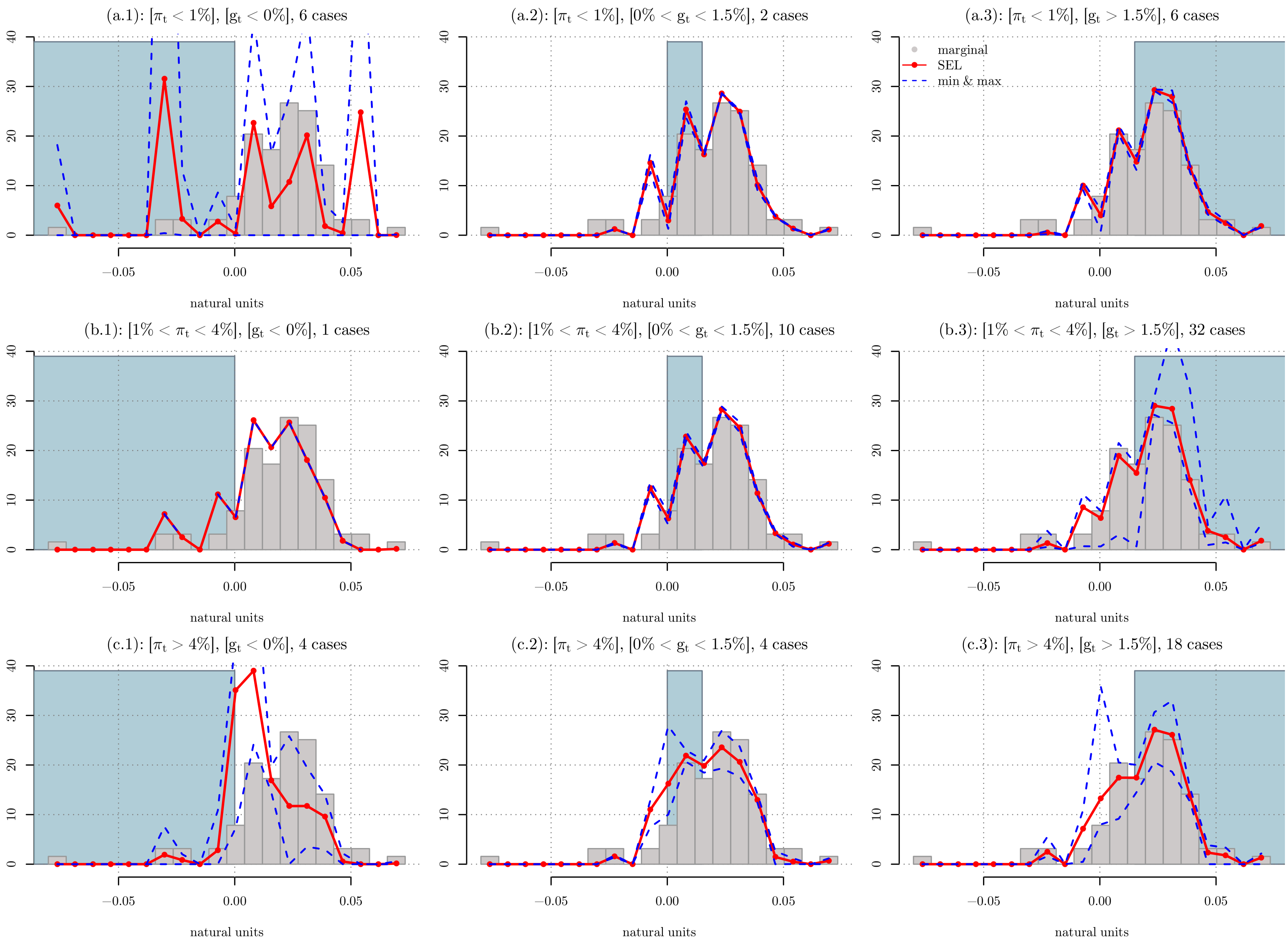
C Figures

Figure 1 – Time series of conditional beliefs, $\gamma = 10$



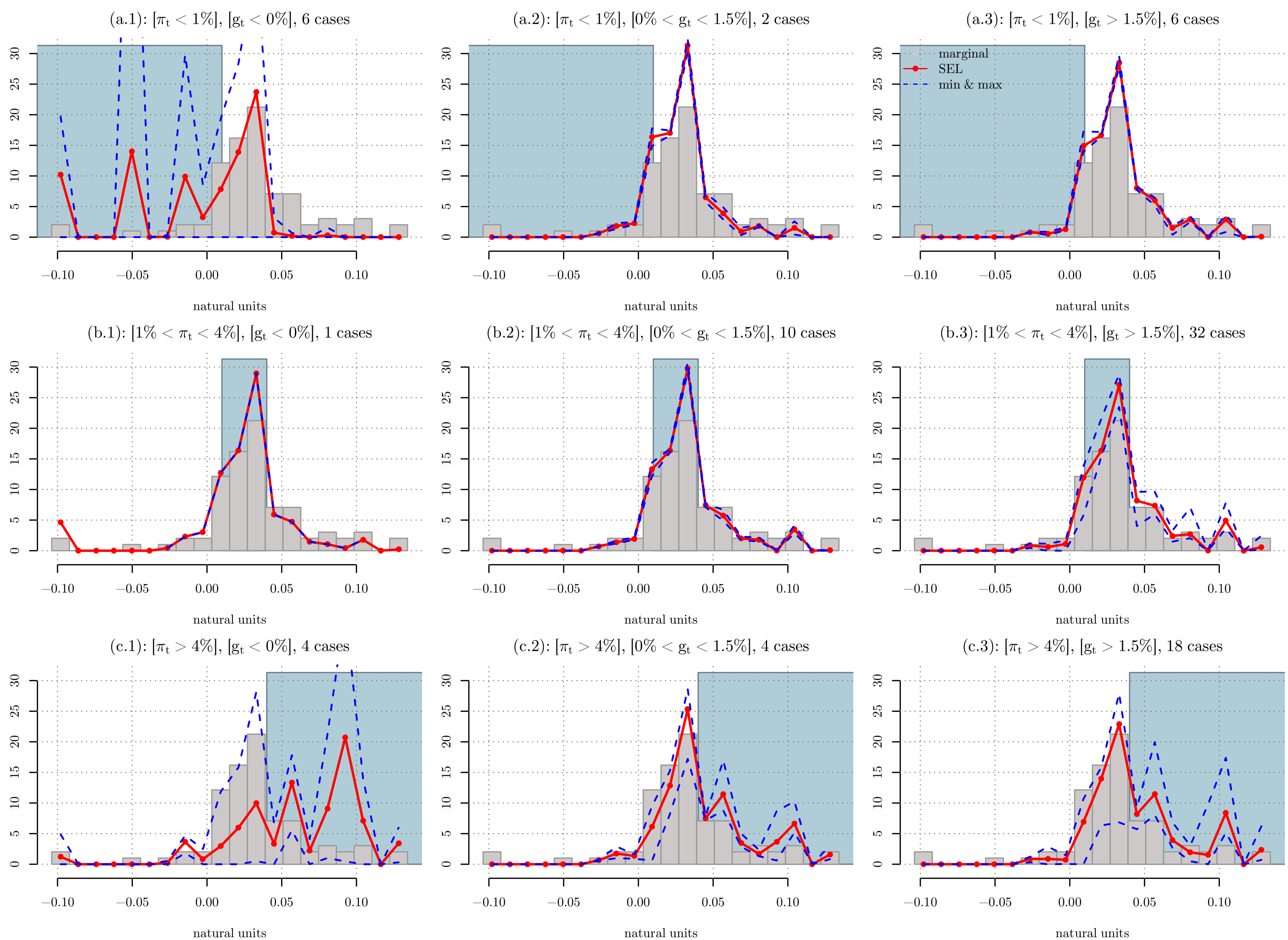
Notes: The figure plots the conditional distribution of future consumption growth for several sets of the conditioning variables, namely the past rate of inflation and past consumption growth. Red line is the $1/T$ result that would be obtained without the conditional moment conditions. Dotted lines are the min and max series in each set of the conditioning variables. The time series of probabilities are obtained with the non-parametric method described in the main text. Smoothing parameter b_n is set to be equal to twice the standard deviation of each conditioning variable. Investors are assumed to have power utility preferences, and the excess return on the market portfolio is the single asset used in the estimation.

Figure 2 – Conditional histogram of consumption growth, $\gamma = 10$



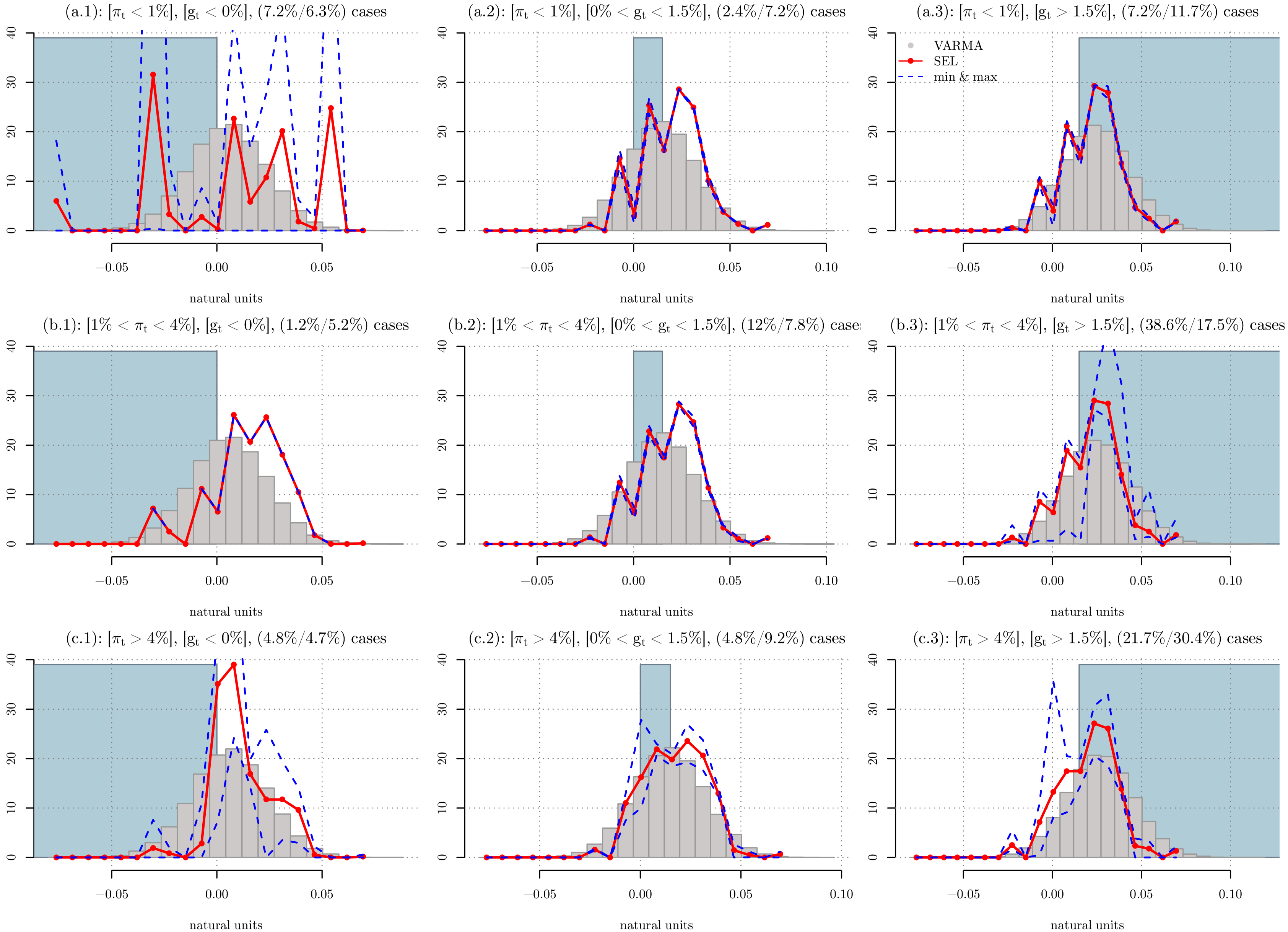
The figure plots the conditional histogram of future consumption growth for several sets of the conditioning variables, namely the past rate of inflation and past consumption growth. Dotted lines are the min and max series in each set of the conditioning variables. Blue bars express the difference between the estimated probabilities(red) and the historical probabilities (grey histogram). The conditional probabilities are obtained with the non-parametric method described in the main text. Smoothing parameter b_n is set to be equal to twice the standard deviation of each conditioning variable. Investors are assumed to have power utility preferences, and the excess return on the market portfolio is the single asset used in the estimation.

Figure 3 – Conditional histogram of inflation, $\gamma = 10$



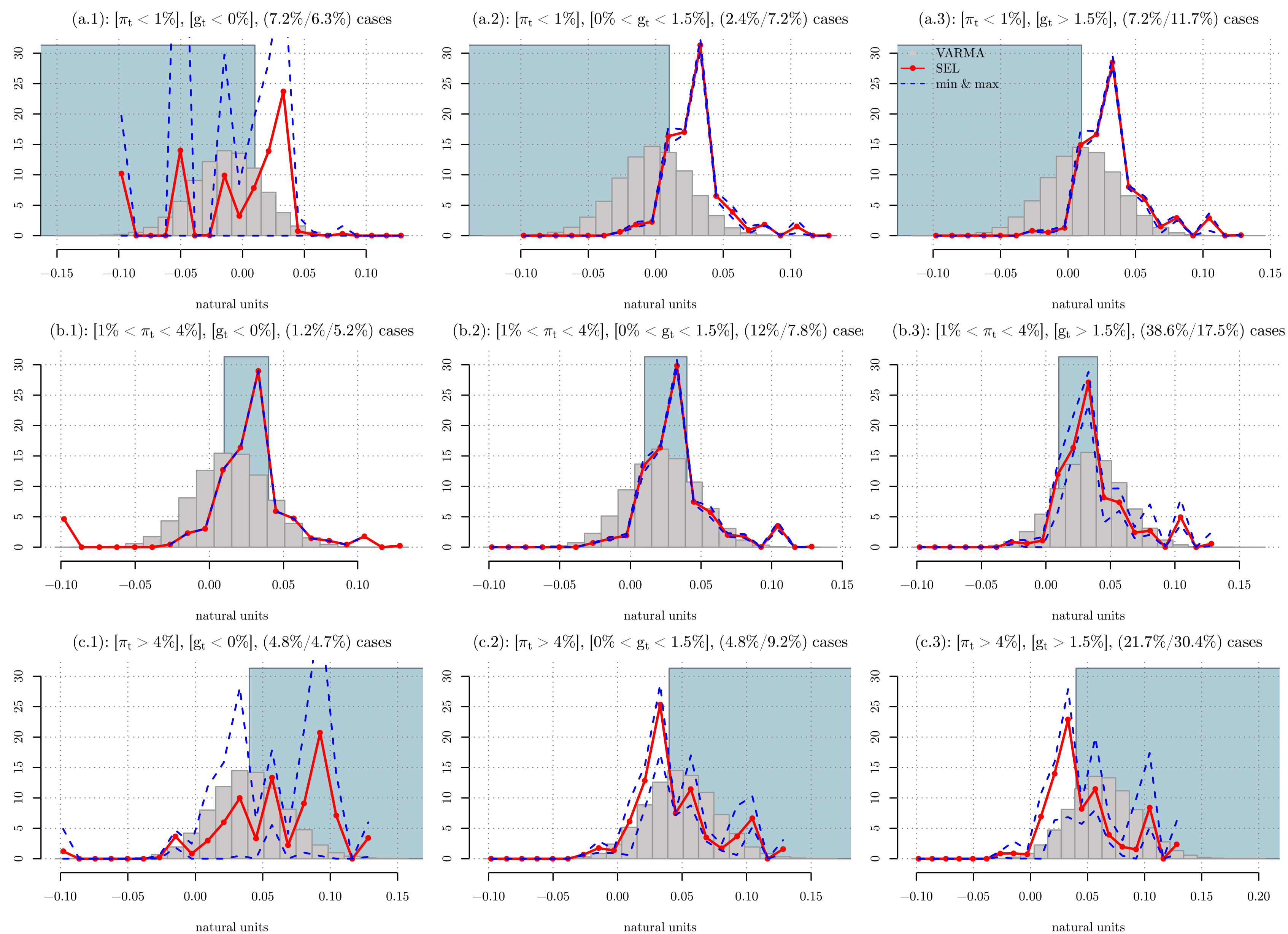
The figure plots the conditional histogram of future inflation for several sets of the conditioning variables, namely the past rate of inflation and past consumption growth. Dotted lines are the min and max series in each set of the conditioning variables. Blue bars express the difference between the estimated probabilities(red) and the historical probabilities (grey histogram). The conditional probabilities are obtained with the non-parametric method described in the main text. Smoothing parameter b_n is set to be equal to twice the standard deviation of each conditioning variable. Investors are assumed to have power utility preferences, and the excess return on the market portfolio is the single asset used in the estimation.

Figure 4 – Conditional histogram of consumption growth v.s. VARMA estimates, $\gamma = 10$



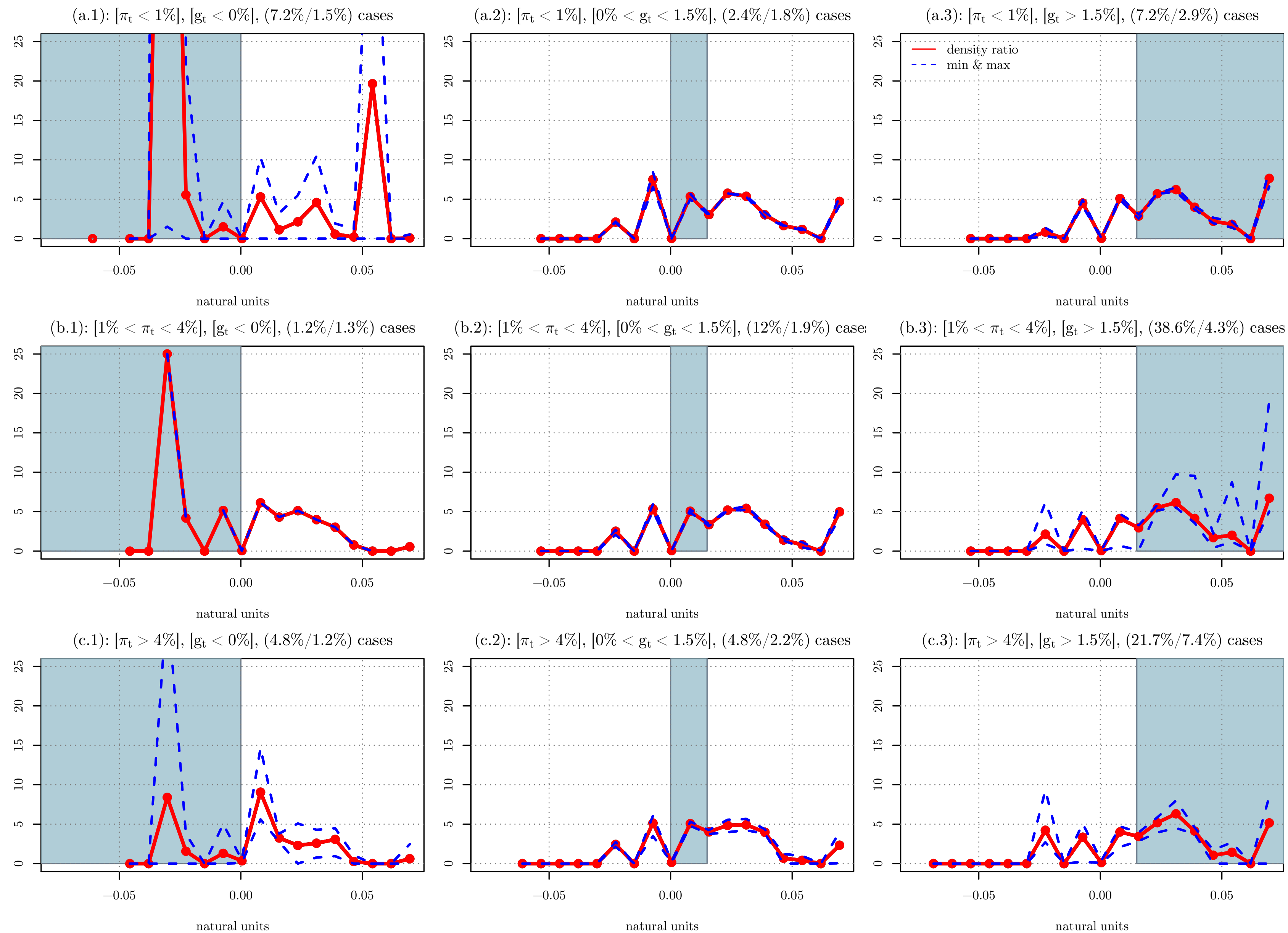
The figure plots the conditional histogram of future consumption growth for several sets of the conditioning variables, namely the past rate of inflation and past consumption growth. Dotted lines are the min and max series in each set of the conditioning variables. Blue bars express the difference between the estimated probabilities(red) and the historical probabilities (grey histogram). The conditional probabilities are obtained with the non-parametric method described in the main text. Smoothing parameter b_n is set to be equal to twice the standard deviation of each conditioning variable. Investors are assumed to have power utility preferences, and the excess return on the market portfolio is the single asset used in the estimation.

Figure 5 – Conditional histogram of inflation v.s. VARMA estimates, $\gamma = 10$



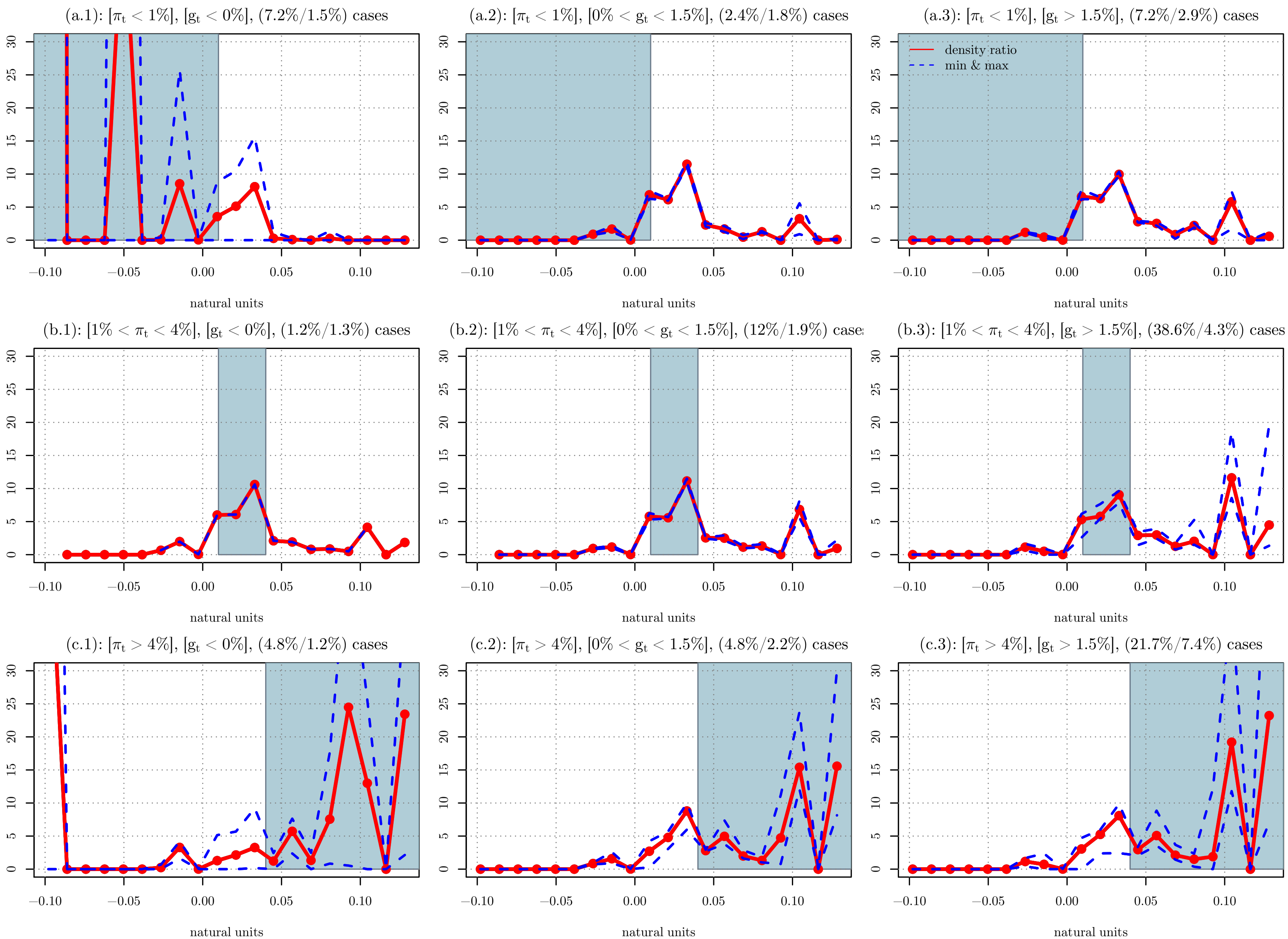
The figure plots the conditional histogram of future inflation for several sets of the conditioning variables, namely the past rate of inflation and past consumption growth. Dotted lines are the min and max series in each set of the conditioning variables. Blue bars express the difference between the estimated probabilities(red) and the historical probabilities (grey histogram). The conditional probabilities are obtained with the non-parametric method described in the main text. Smoothing parameter b_n is set to be equal to twice the standard deviation of each conditioning variable. Investors are assumed to have power utility preferences, and the excess return on the market portfolio is the single asset used in the estimation.

Figure 6 – consumption growth SEL v.s. VARMA density ratio, $\gamma = 10$



The figure plots the conditional histogram of future consumption growth for several sets of the conditioning variables, namely the past rate of inflation and past consumption growth. Dotted lines are the min and max series in each set of the conditioning variables. Blue bars express the difference between the estimated probabilities(red) and the historical probabilities (grey histogram). The conditional probabilities are obtained with the non-parametric method described in the main text. Smoothing parameter b_n is set to be equal to twice the standard deviation of each conditioning variable. Investors are assumed to have power utility preferences, and the excess return on the market portfolio is the single asset used in the estimation.

Figure 7 – Inflation SEL v.s. VARMA density ratio, $\gamma = 10$



The figure plots the conditional histogram of future inflation for several sets of the conditioning variables, namely the past rate of inflation and past consumption growth. Dotted lines are the min and max series in each set of the conditioning variables. Blue bars express the difference between the estimated probabilities(red) and the historical probabilities (grey histogram). The conditional probabilities are obtained with the non-parametric method described in the main text. Smoothing parameter b_n is set to be equal to twice the standard deviation of each conditioning variable. Investors are assumed to have power utility preferences, and the excess return on the market portfolio is the single asset used in the estimation.

C.1 Additional Figures

In order to extract the state variables, x_t and σ_t, from consumption data, we assume the same time series specification for the aggregate consumption growth process as in Bansal and Yaron (2004), with the only exception that we introduce a square-root process for the variance (as in Hansen, Heaton, Lee, and Roussanov, HB of Econometrics, 2007):

$$\Delta c_{t+1} = \mu + x_t + \sigma_t \eta_{t+1} \tag{25}$$

$$x_{t+1} = \rho x_t + \phi_e \sigma_t e_{t+1} \tag{26}$$

$$\sigma_{t+1}^2 = \sigma^2(1-\nu_1) + \nu_1 \sigma_t^2 + \sigma_w \sigma_t w_{t+1}. \tag{27}$$

Note that the model is calibrated at the monthly frequency with the *monthly* parameter values being: $\mu = .0015$, $\rho = .979$, $\phi_e = .044$, $\sigma = .0078$, $\nu_1 = .987$, $\sigma_w = .00029487$. We need to extract the *quarterly* state variables, $x_{t,q}$ and $\sigma_{t,q}^2$. As a first step, we simulate a long sample (five million observations) from the above system, treating the given parameter values as the truth and retaining the simulated state variables. As a second step, we aggregate the simulated data into quarterly non-overlapping observations:

$$\Delta c_{t,q} = \Delta c_t + \Delta c_{t-1} + \Delta c_{t-2}, \text{ for } t = 3, 6, 9, \ldots$$

$$x_{t,q} = x_t + x_{t-1} + x_{t-2}$$

$$\sigma_{t,q}^2 = \sigma_t^2 + \sigma_{t-1}^2 + \sigma_{t-2}^2$$

As a third step, we estimate the model parameters in equations (25)-(27) using these quarterly observations and treating the state variables as observed. This step produces the following quarterly estimates of the parameters:

$$\rho_q = \rho_m^3 = .9383137$$

$$v_{1,q} = v_{1,m}^3 = .9615048$$

$$\mu_q = 3 \times \mu_m = .0045$$

$$\sigma_q^2 = Mean\left(\sigma_{t,q}^2\right) = .0001822490$$

$$\phi_{e,q} = \sqrt{\frac{Var\left(x_{t+1,q} - \rho_q x_{t,q}\right)}{\sigma_q^2}} = .1084845$$

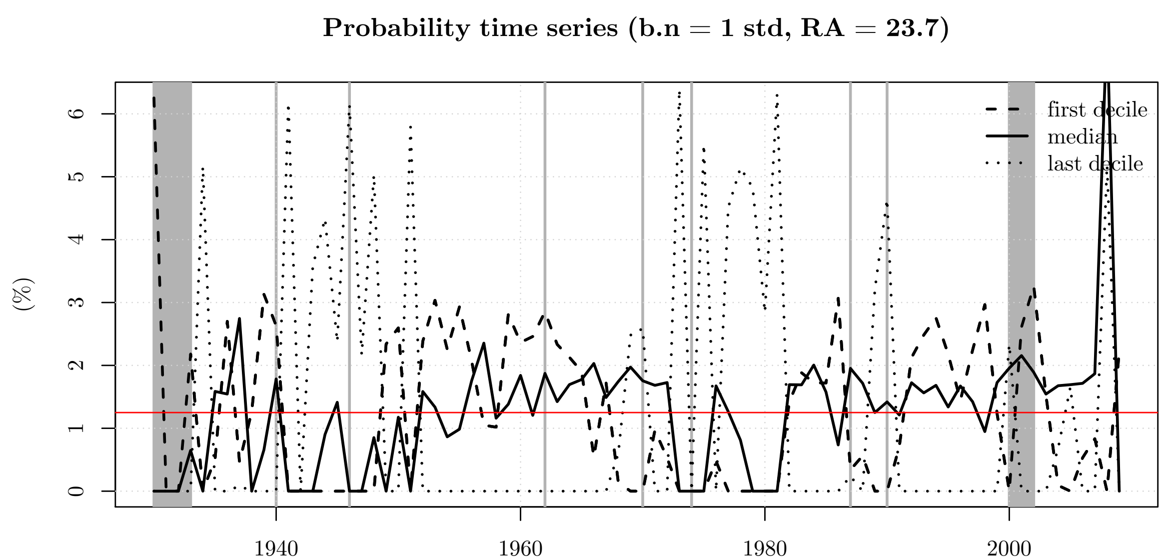
$$\sigma_{w,q} = \sqrt{\frac{Var\left(\sigma_{t+1,q}^2 - \sigma_q^2\left(1 - v_{1,q}\right) - v_{1,q}\sigma_{t,q}^2\right)}{\sigma_q^2}} = 0.0007328592,$$

where the variables with subscript m are the monthly calibrated values, and the means and variances are the ones obtained in the simulated sample. As a fourth step, we run a Bayesian smoother through the historical quarterly consumption growth treating the quarterly parameters as being known with certainty. The smoother produces estimates of the quarterly state variables $\hat{x}_{t,q}$ and $\hat{\sigma}^2_{t,q}$.

The same steps can be applied to obtain the parameter estimates and, therefore, the time series of the state variables at the annual frequency. In this case, we have: $\rho_a = .7751617$; $v_{1,a} = .8546845$; $\mu_a = .018$; $\sigma^2_a = .0007299038$; $\phi_{e,a} = .3853643$; $\sigma_{w,a} = .00270020$.

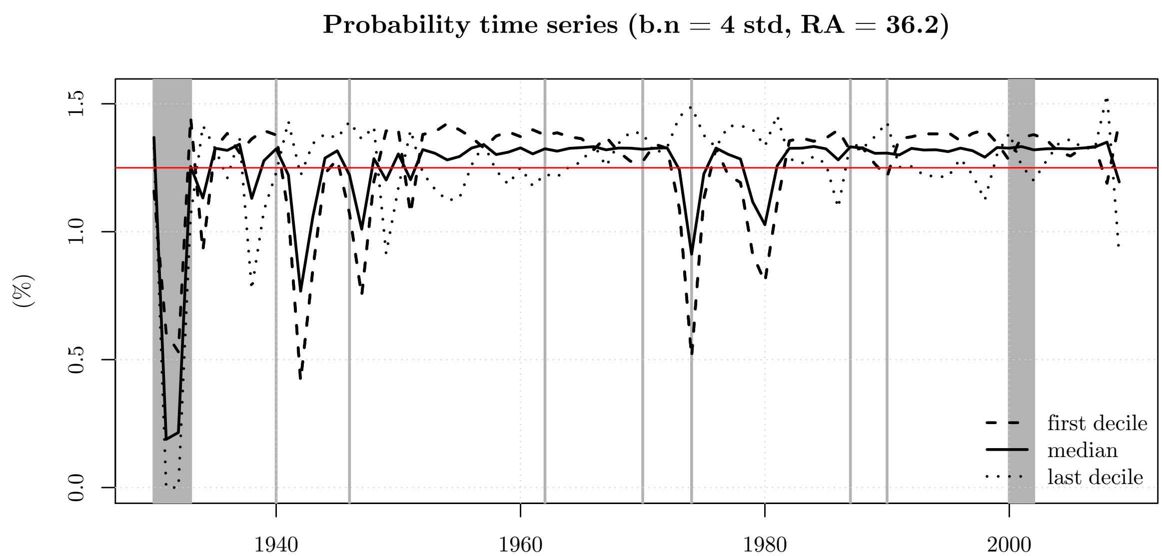
C.2 Additional Figures

Figure 8 – Time series of conditional beliefs, estimated γ



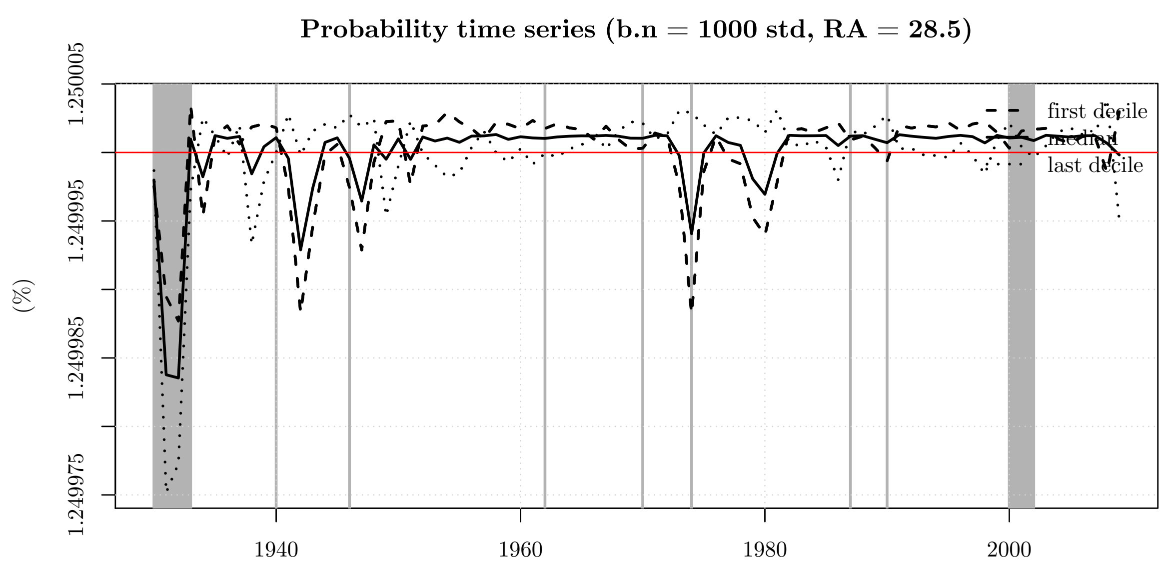
Notes: The figure plots the conditional distribution of future consumption growth for several values of the conditioning variable, namely the rate of inflation. The dashed, solid and dotted line present the time series for inflation equal to its in-sample first decile, median, and 9th decile respectively. The subjective beliefs, extracted using the relative entropy minimization approach, are used to obtain the conditional distribution. Investors are assumed to have power utility preferences, inflation is the sole conditioning variable, and the excess return on the market portfolio is the single asset used in the estimation.

Figure 9 – Time series of conditional beliefs, estimated γ



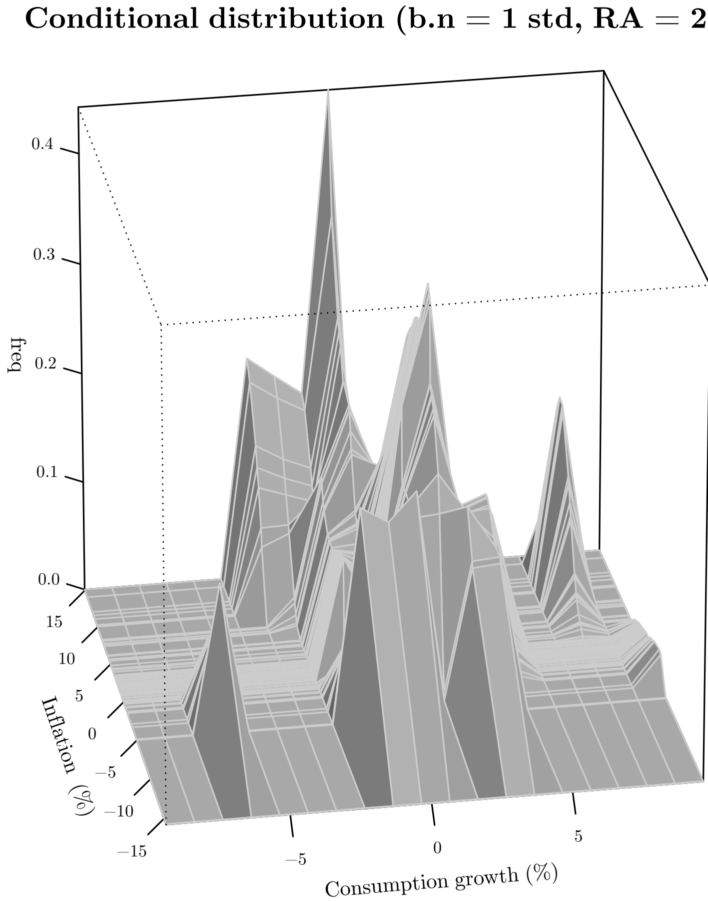
Notes: The figure plots the conditional distribution of future consumption growth for several values of the conditioning variable, namely the rate of inflation. The dashed, solid and dotted line present the time series for inflation equal to its in-sample first decile, median, and 9th decile respectively. The subjective beliefs, extracted using the relative entropy minimization approach, are used to obtain the conditional distribution. Investors are assumed to have power utility preferences, inflation is the sole conditioning variable, and the excess return on the market portfolio is the single asset used in the estimation.

Figure 10 – Time series of conditional beliefs, estimated γ



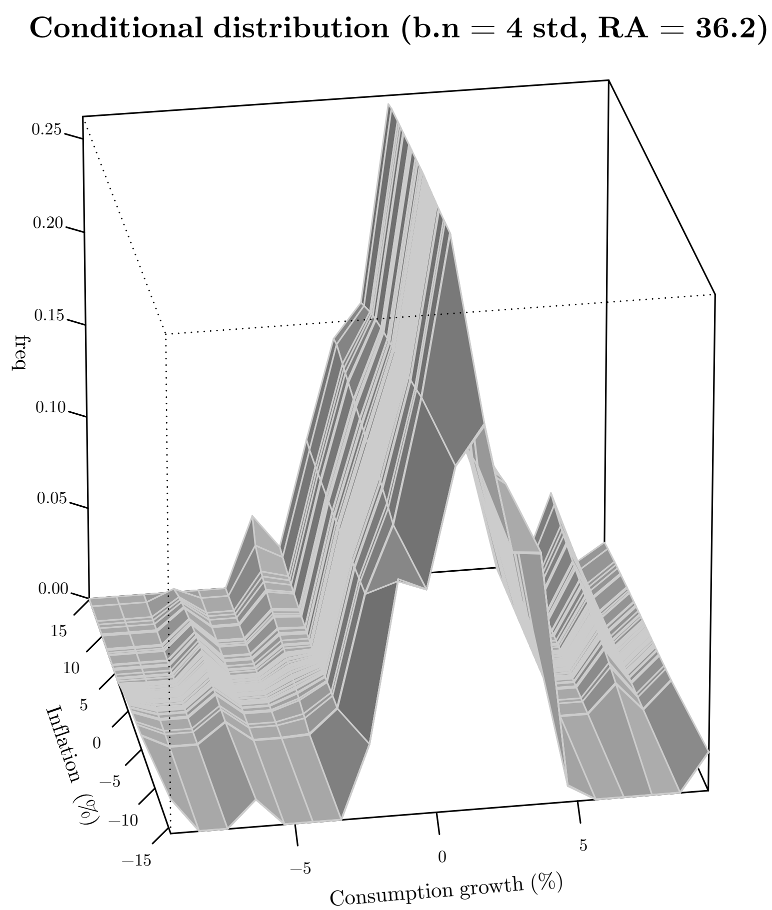
Notes: The figure plots the conditional distribution of future consumption growth for several values of the conditioning variable, namely the rate of inflation. The dashed, solid and dotted line present the time series for inflation equal to its in-sample first decile, median, and 9th decile respectively. The subjective beliefs, extracted using the relative entropy minimization approach, are used to obtain the conditional distribution. Investors are assumed to have power utility preferences, inflation is the sole conditioning variable, and the excess return on the market portfolio is the single asset used in the estimation.

Figure 11 – Conditional histogram of consumption growth, estimated γ



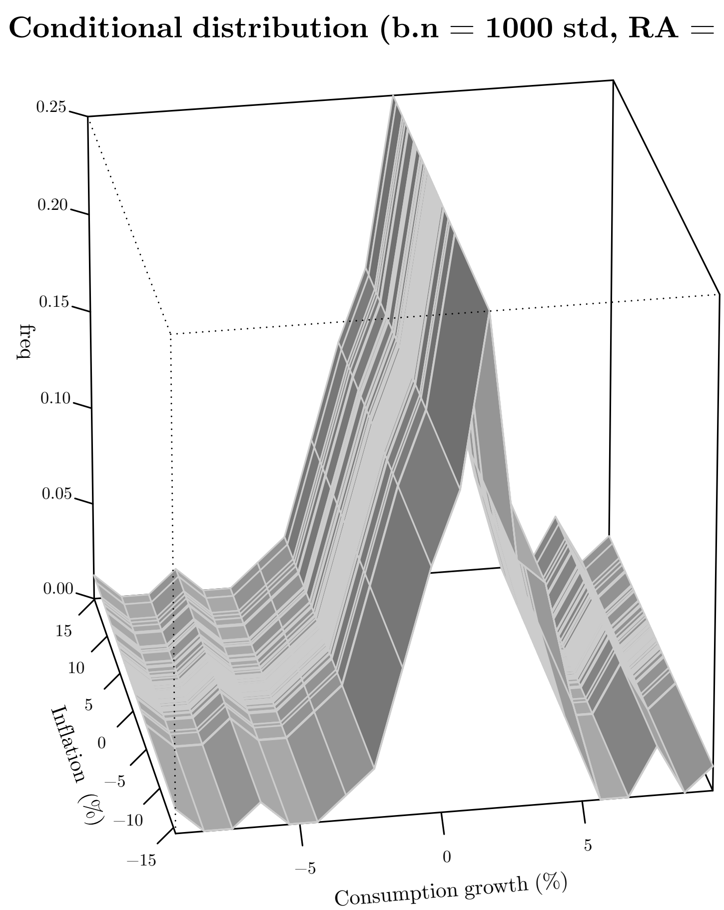
The figure plots the conditional histogram of future consumption growth for each value of the conditioning variable, namely the rate of inflation. The subjective beliefs, extracted using the relative entropy minimization approach, are used to obtain the conditional distribution. Investors are assumed to have power utility preferences, inflation is the sole conditioning variable, and the excess return on the market portfolio is the single asset used in the estimation.

Figure 12 – Conditional histogram of consumption growth, estimated γ



The figure plots the conditional histogram of future consumption growth for each value of the conditioning variable, namely the rate of inflation. The subjective beliefs, extracted using the relative entropy minimization approach, are used to obtain the conditional distribution. Investors are assumed to have power utility preferences, inflation is the sole conditioning variable, and the excess return on the market portfolio is the single asset used in the estimation.

Figure 13 – Conditional histogram of consumption growth, estimated γ



The figure plots the conditional histogram of future consumption growth for each value of the conditioning variable, namely the rate of inflation. The subjective beliefs, extracted using the relative entropy minimization approach, are used to obtain the conditional distribution. Investors are assumed to have power utility preferences, inflation is the sole conditioning variable, and the excess return on the market portfolio is the single asset used in the estimation.