Identifying Beliefs from Asset Prices

Anisha Ghosh† Guillaume Roussellet‡

April 24, 2019

Abstract

This paper proposes a novel information-theoretic procedure to identify investors’ beliefs about future macroeconomic and financial outcomes from asset prices. Given a cross-section of assets, a pricing kernel, and a conditioning set, our approach recovers the beliefs that are consistent with the observed asset prices, i.e. the conditional distribution of macro and financial variables that satisfy the conditional Euler equations. Our approach is non-parametric, not requiring any functional-form assumptions about the dynamics of the variables or assumptions regarding investor rationality or lack thereof. The recovered beliefs imply strong cyclicality in the conditional mean and skewness of macro variables, such as the aggregate real consumption growth rate, while the conditional volatility is mostly flat over the business cycle. This contrasts with the widely assumed conditionally normal dynamics for these variables in existing literature. A comparison of these price-consistent beliefs with a judicious objective benchmark suggests large beliefs distortions, whereby investors underestimate the expected real growth during severe economic downturns but also simultaneously underestimate, by a much larger amount, the magnitude of the negative skewness in future real growth rates during good and bad states alike. The recovered beliefs about the stock market are positively correlated with survey data on institutional investors’ expectations of future stock market returns.

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 Marine Carrasco, George Constantinides, Marco Del Negro, Domenico Gian-none, Lars Hansen, Scott Joslin, Hanno Lustig, Stefan Nagel, Jean-Paul Renne and seminar participants at the University of Cambridge and McGill University. We thank Qing Xu for excellent research assistance. All errors and omissions are our own.
†Desautels Faculty of Management, McGill University; anisha.ghosh@mcgill.ca, Corresponding author.
‡Desautels Faculty of Management, McGill University; guillaume.roussellet@mcgill.ca.
I  Introduction

Asset prices reflect investors’ beliefs about future economic and financial outcomes. Therefore, understanding how investors form their beliefs and how the beliefs evolve over time with changing economic conditions is crucial in explaining and predicting the behavior of asset prices. Not surprisingly, substantial research effort has been devoted to this topic and this has led to the emergence of two opposing paradigms. On the one hand, the *rational expectations paradigm* (Muth (1961)) postulates that economic agents use available data objectively, or rationally, to form beliefs about the future. On the other hand, in *behavioral models*, agents are assumed to have certain behavioral biases (see, e.g., Barberis and Thaler (2001) for a survey of behavioral finance) that distort their beliefs about the future relative to what would obtain under rationality.

The rational expectations hypothesis, while intuitive and appealing, has difficulty explaining a number observed features of the aggregate stock market, the cross sections of returns of different classes of financial assets, and individual trading behavior. Structural assumptions on investors’ preferences (represented by the pricing kernel) and on the dynamics of the data generating process (representing investors’ beliefs about the future), are the cornerstone of asset pricing models building on rational expectations, but are prone to misspecification error. Moreover, existing literature has shown that, across a wide range of structural assumptions, these models all ‘miss’ a remarkably similar component (see e.g. Ghosh, Julliard, and Taylor (2016)).

Behavioral models present an attractive new approach to help overcome some of the difficulties faced by the rational expectations framework. However, even for this class of models, specific assumptions about the forms of behavioral biases or beliefs distortions relative to the rational benchmark are crucial to explain the key aspects of financial market data.

This paper proposes a non-parametric approach to identify investors’ beliefs from observed asset prices, that bypasses the need for any functional-form assumptions about the dynamics of the data generating process or assumptions regarding investor rationality or lack thereof. Given a pricing kernel, a cross-section of test assets, and a set of conditioning variables, our approach recovers the entire conditional distribution of macroeconomic and financial variables as perceived by the representative investor, i.e. the investor’s beliefs about future macroeconomic and financial outcomes. Our methodology relies on the non-parametric smoothed empirical likelihood (SEL) estimator developed by Kitamura, Tripathi, and Ahn (2004). This approach estimates the conditional density of macroeconomic and financial variables by maximizing the non-parametric (multinomial) log-likelihood of the data, subject to the constraint that the density so estimated satisfies the conditional Euler equation restrictions for the set of test assets. It is in this sense that the estimated beliefs
are consistent with the observed asset prices.

Note that our methodology for the recovery of investors’ beliefs requires three inputs – the pricing kernel, a set of test assets, and a conditioning set underlying the conditional Euler equation restrictions. Our baseline results are based on the most standard choices for these inputs. Specifically, the pricing kernel is derived from the time- and state-separable power utility preferences with a constant coefficient of relative risk aversion (the C-CAPM of Rubinstein (1976)), the excess return on the market portfolio is the sole test asset, and the conditioning set consists of an exponentially-weighted moving average of past consumption growth.

The recovered beliefs show that investors perceive the mean of macro variables such as the aggregate consumption growth rate to be strongly procyclical – the conditional mean varies from a low of 0.8% during the recent financial crisis to a high of 2.4% during the expansionary episode of the mid-sixties. The conditional volatility of consumption growth, on the other hand, is perceived to be largely flat over the business cycle.

An important component of the beliefs process is the time-varying skewness of the consumption growth rate. Our results suggest that the perceived skewness is negative in almost all time periods, but, perhaps more surprisingly, is strongly cyclical becoming less negative during bad times characterized by recession episodes. Thus, our results are indicative of investors fearing severe economic downturns even during good states of the world. For instance, in 1965 : Q4 – a period marked by high real economic growth averaging 4.0% per year for the past three years – the expected consumption growth in the next quarter is high at 2.4% (annualized); yet the probabilities attached to consumption growth in the next quarter being less than −1.3% (two standard errors below the current mean) or −3.2% (three standard errors below the current mean) are non-trivial at 4.6% and 1.4%, respectively. These are, respectively, double and an order of magnitude larger than the corresponding probabilities that would obtain if the conditional density were Gaussian. To the contrary, during the recent financial crisis, e.g., in 2009 : Q2, the expected consumption growth is very low at 0.8%. The probability attached to consumption growth in the next quarter being less than −1.3% is 14.0%, an order of magnitude bigger than this probability during the good state in 1965 : Q4; however, the probability of consumption growth being lower than −3.2% is quite similar in both states −1.6% in 2009 : Q2 versus 1.4% in 1965 : Q4. Thus, the extreme left tail of the distribution does not expand in bad states beyond that in good times, while the mean of the distribution is substantially lower during bad times, causing the skewness to become less negative during these latter times. The above results call into question the widely used assumption of a conditionally Gaussian data generating process in structural asset pricing models.
The recovered beliefs also suggest that expected consumption growth is highly persistent, with a quarterly first-order autocorrelation coefficient of 0.88, compared to realized consumption growth that has an autocorrelation coefficient of only 0.31. The former is also substantially less volatile, with an annualized volatility of 0.1%, compared to an order of magnitude higher volatility of 1.0% for the latter. Finally, we show that the beliefs about next period’s consumption growth have strong forecasting power for the subsequent realization of consumption growth – a regression of the realized consumption growth on its expected value calculated using the recovered beliefs produces a highly statistically significant slope coefficient and an $R^2$ of 9.5% at the quarterly frequency.

We show that our estimated beliefs are robust to choices of pricing kernels, cross-sections of test assets, and conditioning sets. We provide evidence that pricing kernels implied by Epstein and Zin (1989) recursive preferences in the presence of long run risks in consumption growth and the external habit formation preferences of Campbell and Cochrane (1999) yield remarkably similar results. Next, we show that the recovered beliefs look almost indistinguishable if the cross section of assets includes the excess returns on the top and bottom deciles of portfolios formed by sorting stocks on the basis of market capitalization and the book-to-market-equity ratio, in addition to the market portfolio. Finally, the beliefs are robust to a wide range of specifications of the conditioning set – adding inflation, labor market variables, principal components extracted from a broad cross section of over a hundred macro variables, or asset returns to the conditioning set makes little difference to the results.

We also present the implications of investors’ beliefs for the stock market. Specifically, we show that the recovered beliefs imply a strongly time-varying and countercyclical expected equity premium – the premium varies from being lower than a percent at 0.43% in 1965 to more than fifteen times higher at 7.3% in 2009. The conditional volatility is also countercyclical, varying from 16.0% to 24.0%. Finally, the strong time-variation in the mean leads to a highly countercyclical conditional Sharpe ratio of the market portfolio, varying from 0.027 to 0.392. The magnitude and the nature of time-variation in these moments of the market portfolio are in line with commonly held beliefs about the dynamics of these moments in the academic literature, and have represented a challenge for structural models to reproduce.

We compare our recovered beliefs about consumption growth with time series models for consumption growth extensively used in the literature. In particular, we consider two time series specifications: (a) a standard ARMA(1,1) model, and (b) a regime-switching model where the mean of consumption growth varies across latent regimes. We estimate these models using consumption data alone, i.e. without any asset returns data. Our results suggest that the time series models for consumption growth that are commonly assumed in
the literature are different in several ways from investors’ beliefs about consumption growth extracted from observed asset prices. First, the expected consumption growth implied by the latter is much more persistent and less volatile than the former time series models would imply. Second, the investors’ beliefs suggest a fatter left-tail in the distribution of future consumption growth, i.e. the conditional skewness is much more negative during good and bad times, compared to that implied by the time series models. Third, the skewness implied by investors’ beliefs is strongly cyclical, becoming less negative in bad times, a feature that is, once again, missed by these commonly assumed time series models.

Finally, we apply our methodology to shed light on whether the recovered investors’ beliefs are rational or whether they reflect departures from rationality. We show that the SEL approach used to recover the beliefs offers one possible way to address this question. Specifically, we show that the SEL estimator has an alternative information-theoretic interpretation – the recovered conditional distribution (of, e.g., the consumption growth rate) is the one that is minimally distorted relative to the objective distribution, so as to satisfy the pricing restrictions given by the conditional Euler equations for the test assets. In this context, the objective measure is the conditional distribution that would obtain in the absence of the conditional Euler equation restrictions, i.e. without requiring the distribution to be consistent with asset prices. Also, the objective measure is akin to a non-parametric kernel density estimator for the conditional distributions of the macro and financial variables of interest and, therefore, does not rely on functional-form assumptions about the distributions. For the above two reasons, the objective measure may be an attractive candidate for measuring rational beliefs and constitutes our objective benchmark with respect to which investors’ beliefs distortions can be assessed.

Our results suggest that investors’ beliefs about future macroeconomic outcomes, e.g. consumption growth, seem distorted relative to the objective benchmark in two important respects. First, the expected growth rate is lower under the price-consistent beliefs compared to the objective measure, particularly during exceptionally bad states of the world. Second, the conditional skewness is less negative under the price-consistent beliefs compared to the objective measure, during good and bad times alike, with the magnitude of the distortions as well as the percentage distortions far exceeding those in the mean. The conditional volatility, on the other hand, is almost identical between the two measures. To the extent that the objective benchmark can be regarded as a measure of rational beliefs, our results are indicative of investor exuberance during good times – stemming from a perceived truncation of the left tail of the distribution of consumption growth relative to the rational measure,

1The word ‘minimal’ is meant in the information-theoretic sense, and we show that our estimator minimizes the Kullback-Leibler Information Criterion divergence (or relative entropy) between the estimated measure and the objective measure (see also Kitamura and Stutzer (1997), Owen (2001)).
while distortions in the expected mean and volatility of consumption growth are negligible. During bad times, our results are more mixed, suggesting evidence of pessimism via under-estimation of the expected growth rate on the one hand, and optimism on the other hand via less negative skewness in the distribution of consumption growth relative to the rational measure.

Our work builds on a burgeoning literature trying to recover risk premia components with as few assumptions as possible. Ross (2015) argues that, under conditions later clarified by Borovicka, Hansen, and Scheinkman (2016), one can recover simultaneously investors’ preferences and beliefs using a set of Arrow-Debreu securities. In a rational expectations framework, this pursuit of recovery is akin to identifying the pricing kernel using only minimal assumptions. The literature has progressed towards breaking down restrictive assumptions, one after the other, to arrive to an (almost) model-free recovery (see Schneider and Trojani (2017)).

Our approach is similar in spirit, since our non-parametric estimator discretizes the attainable states to recover beliefs. However, the smoothed empirical likelihood does not request a large cross-section of assets to provide the identification result, and the discretized space is assumed a simplification of a true absolutely-continuous conditional density with respect to Lebesgue. Also, it can be used to extract information not only from options, as is common in the literature, but also from any type of financial asset, thereby enabling the use of a much longer data sample.

Closest papers to ours in terms of the objective and methodology are the works of Ghosh, Julliard, and Taylor (2016) and Qiu and Otsu (2017a,b). The former uses an information-theoretic approach to recover a multiplicative missing component of the pricing kernel for a broad class of consumption-based asset pricing models. Whereas Ghosh, Julliard, and Taylor (2016) focus on unconditional Euler equation restrictions (thus, unconditional densities), this paper considers conditional Euler equations. While this complicates the analysis in that it requires the specification of the conditioning set and a different methodology to recover beliefs, it enables us to estimate the conditional distributions of variables of interest as perceived by investors. Qiu and Otsu (2017a,b) develop the theory for the conditional empirical likelihood estimator that we use in this paper in the case of a high dimensional cross-section of assets to identify the pricing kernel.

---


3This literature is tightly linked to the identification of pricing kernel bounds as provided by financial instruments paying off statistical moments of index return. Notable contributions include Martin (2013, 2017), Kozhan, Neuberger, and Schneider (2013), Schneider (2015, 2018), Schneider and Trojani (2018) and Orlowski, Sali, and Trojani (2018).
Finally, the paper contributes to a growing literature departing from the Muth (1961) rational expectation paradigm to explain various aspects of asset market data. Behavioral finance models represent the bulk of this literature. This class of models assumes that economic agents have certain behavioral biases (see, e.g., Kahneman and Tversky (1979)) that distort their beliefs about the future relative to the rational benchmark. Robust control (or uncertainty averse) preferences represent a promising alternative to the rational expectations framework. In this framework, investors make consumption-investment decisions from the perspective of the worst-case data-generating process (see, e.g., Hansen and Sargent (2001, 2016)). Barillas, Hansen, and Sargent (2009) argue that robust control models replace the need for implausibly large risk aversion in rational models with plausible levels of uncertainty aversion. 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 have significant explanatory power for a broad cross section of stock portfolios. These studies all make specific parametric (conditionally Gaussian) assumptions on the nature of the beliefs distortions and use professional survey forecasts data to estimate the parameters of the subjective beliefs. Our approach differs markedly from these studies in that we abstract from using any parametric assumptions on beliefs distortions and rely only on asset pricing Euler equation restrictions in our identification scheme. Also, we recover the entire conditional distribution of beliefs rather than only the distribution of the conditional means of the variables of interest as authorized by survey-based forecasts data.

The remainder of this paper is organized as follow. Section II describes the conditional empirical likelihood method of estimating investors’ beliefs from observed asset prices. Section III demonstrates, via simulations, the effectiveness of the estimation approach in recovering the conditional distribution of macro variables as well as in identifying beliefs distortions relative to rationality. The data used in the empirical analysis are described in Section IV. The empirical results are presented in Sections V and VI, which report the estimated beliefs and their comparison to time series specifications commonly assumed in the literature, respectively. Section VII sheds light on whether investors’ beliefs are rational. Section VIII concludes with suggestions for future research.

II Non-Parametric Estimation of Beliefs

In this section, we describe the details of our methodology. Section II.1 presents a general framework to illustrate that asset prices reflect the beliefs of investors and, therefore, can be used to recover these beliefs. Our econometric approach, that allows for the identification
of investors’ beliefs, is presented in Section II.2. We provide an alternative information-theoretic interpretation of the extracted beliefs in Section VII.1. Two crucial inputs in our beliefs recovery procedure are the choices of the underlying conditioning set that investors use to form their beliefs and the pricing kernel (hereafter referred to as the SDF) that investors use to discount possible future states of the world. Section II.3 describes our various choices of the conditioning set, and Section II.4 discusses the various SDFs considered. We later show that the recovered beliefs look remarkably similar for a large set of specifications of the conditioning set and the SDF.

Throughout, uppercase letters denote random variables, while the corresponding lowercase letters denote particular realizations of these random variables.

II.1 General framework

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

\[
\mathbb{E}^{P_t} \left[ M_{t+1} R^e_{t+1} \mid \mathcal{F}_t \right] = 0,
\]

(1)

where \( \mathcal{F}_t = \{ \mathcal{F}_t, \mathcal{F}_{t-1}, \ldots \} \) denotes the investors’ information set at time \( t \), and \( \mathbb{E}^{P_t} \left[ \cdot \mid \mathcal{F}_t \right] \) is the expectation operator conditional on \( \mathcal{F}_t \). Therefore, for any random process \( Y_\tau \) taking values on \( \text{supp}(Y_\tau) \), where \( \tau > t \),

\[
\mathbb{E} \left[ Y_\tau \mid \mathcal{F}_t \right] = \int_{\text{supp}(Y_\tau)} Y_\tau \, d\mathbb{P}_t(Y_\tau).
\]

(2)

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

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

(3)

where \( \theta_0 \) is the true value of the vector of parameters driving the SDF.

If investors are fully rational, the probability measure \( \mathbb{P}_t \) in Equation (1) denotes the objective probability measure given the filtration \( \mathcal{F}_t \). Thus, Equation (1) holds with respect to the objective probability measure if and only if investors have rational expectations. However, if investors have any behavioral biases that make their beliefs deviate from rational expectations, the \( \mathbb{P}_t \) in Equation (1) denotes the representative agent’s subjective probability measure. Equation (1), in that case, does not hold under the rational (or objective) measure.

Our objective is to recover \( \mathbb{P}_t \) from observed asset prices. Our econometric procedure,
described in the next sub-section, does not require taking a stance on whether beliefs are rational or whether they are distorted relative to rationality.

II.2 The smoothed empirical likelihood estimator

Our identification approach 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.\(^4\)

To provide some intuition and fix ideas, let us first consider a multinomial model with as many possible states as observation dates, \(T\). In the absence of any constraints, other than requiring the probability \(p_t\) assigned to each state to be positive and the probabilities \((p_t)_{t=1}^T\) to sum to unity, a standard non-parametric maximum likelihood estimator will yield every probability estimate as: \(\hat{p}_t = \frac{1}{T}\), for \(t \in \{1, \ldots, T\}\). Now assume that we perform the same likelihood maximization, but enforcing that the estimated multinomial distribution satisfies the unconditional moment restrictions implied by Equation (1). The method will now distort the probability estimates \(\hat{p}_t\) relative to the \(1/T\) benchmark in order to satisfy the moment restrictions. Specifically, the estimated probabilities will be such that they maximize the log-likelihood of the observed data, subject to the constraint that the unconditional Euler equations are satisfied when evaluated under the estimated probability measure. This is the empirical likelihood (EL) estimator of Owen (2001). The SEL estimator, used in this paper, relies on the same principle, but incorporates additional constraints through conditional Euler equation restrictions. Note that the EL and SEL estimators are non-parametric in the sense that they do not require any parametric functional-form assumptions about the distribution of the data.

In what follows, we assume that the information set of the investors at time \(t\) can be summarized by a finite vector of random variables, that we denote by \(X_t \in \mathbb{R}^n\). Suppose that the historical realizations of consumption growth, other variables in the SDF, excess returns, and the conditioning variables are given by \((g_t, y_t, r_t^e, x_t)_{t=1}^T\). Let \(p_{i,j}\) denote the conditional probability of observing the joint outcome \((g_j, y_j, r_j^e, x_j)\) at time \(t+1\), i.e. the probability of state \(j\) being realized at time \(t + 1\), given that state \(i\) was realized at date \(t\).\(^5\)

---

\(^4\)We refer to the SEL estimator as non-parametric while it arguably belongs to semi-parametric methods. However, while we assume a parametric form for the SDF as given by Equation (3), our goal does not lie in the estimation of the parameters driving the SDF, but rather in the identification of the conditional densities of our endogenous variables. We, therefore, use the terminology *non-parametric* without abuse.

\(^5\)Since \(X_t\) contains lagged endogenous variables such as consumption growth, we consider that our first observation \(g_0\) does not belong to the space of attainable values. This is equivalent to the “loss” of informa-
The SEL estimator for the conditional probabilities \((p_{i,j})\) for \(i, j = \{1, \ldots, T\}\), is such that it belongs to the simplex:

\[
\Delta := \bigcup_{i=1}^T \Delta_i = \bigcup_{i=1}^T \left\{ (p_{i,1}, \ldots, p_{i,T}) : \sum_{j=1}^T p_{i,j} = 1, p_{i,j} \geq 0 \right\}
\]

and that: \(\forall i \in \{1, \ldots, T\}, \forall \theta \in \Theta\),

\[
\left( \hat{p}_{i,.}^{SEL}(\theta) \right) = \arg \max_{(p_{i,.}) \in \Delta_i} \sum_{j=1}^T \omega_{i,j} \log(p_{i,j}) \quad \text{s.t.} \quad \sum_{j=1}^T p_{i,j} \times M(g_j, y_j; \theta) r^*_j = 0, \quad (4)
\]

where \(p_{i,.}\) denotes the \(T\)-dimensional vector of probabilities \((p_{i,1}, \ldots, p_{i,T})\), \(\Theta\) is the set of all admissible parameters \(\theta\), and \(\omega_{i,j}\) are non-negative weights used to smooth the objective function. In the spirit of non-parametric estimators:

\[
\omega_{i,j} = \frac{K\left( \frac{x_i - x_j}{b_T} \right)}{\sum_{t=1}^T K\left( \frac{x_i - x_t}{b_T} \right)}, \quad (5)
\]

where \(K\) is a kernel function belonging to the class of second order product kernels,\(^6\) and the bandwidth \(b_T\) is a smoothing parameter.\(^7\)

The objective function in Equation (4) is simply a ‘smoothed’ log-likelihood, with the constraints enforcing the conditional Euler equation restrictions in Equation (1). The weights \(\omega_{i,j}\) used to smooth the log-likelihood are standard non-parametric kernel weights. The intuition behind the estimator may be understood as follows. Note that we are interested in recovering the conditional distribution of the data. For each value of the current state \(x_i\), the SEL estimator focuses on a fixed neighbourhood around \(x_i\), where the neighbourhood is defined in terms of the distance of other possible values of the state from the current state, i.e. \(|x_i - x_j|\), and not in terms of proximity in time. The estimator then assigns positive weights \(\omega_{i,j}\) only to those states that lie within the fixed neighbourhood of the current state, with the exact values of the weights determined by the kernel function, the distance implied by the treatment of the first observation for building the likelihood in a dynamic time series parametric model. The effect of such an assumption fades out asymptotically.

\(^6\)\(K\) should satisfy Assumption 3.3 in Kitamura, Tripathi, and Ahn (2004), that is restated here for convenience. For \(X = (X^{(1)}, X^{(2)}, \ldots, X^{(n)})\), let \(K = \prod_{i=1}^n k(X^{(i)})\). 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 bounded away from zero on \([-a, a]\).

\(^7\)In theory, \(b_T\) is a null sequence of positive numbers such that \(T b_T \to \infty\). See Assumption 3.7 in Kitamura, Tripathi, and Ahn (2004) for additional restrictions on the choice of \(b_T\).
$|x_i - x_j|$, and the bandwidth parameter $b_T$ (see Equation (5)). The states that lie outside the neighbourhood each receive a weight of zero. Finally, the SEL approach determines the conditional probabilities of each of the states with non-zero weight so as to maximize the smoothed log-likelihood of the data subject to the constraint that the estimated conditional distribution satisfies the conditional Euler equation restrictions (see Equation (4)). The states with zero weight are each assigned a conditional probability of zero.

The solution to Equation (4) is analytical and given by:

$$\forall i, j \in \{1, \ldots, T\},$$

$$\hat{p}_{i,j}^{SEL}(\theta) = \frac{\omega_{i,j}}{1 + M(g_j, y_j; \theta) \cdot \lambda_i(\theta)' r^e_j},$$

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

$$\hat{\lambda}_i(\theta) = \arg \max_{\lambda_i \in \mathbb{R}^k} \sum_{j=1}^{T} \omega_{i,j} \log \left[1 + M(g_j, y_j; \theta) \cdot \lambda_i^T r^e_j\right].$$

Equations (6) and (7) show that the SEL procedure delivers a $(T \times T)$ matrix of probabilities $(\hat{p}_{i,j}^{SEL}(\theta))$ for each value of the parameter vector $\theta$. Each row $i : i = \{1, 2, \ldots, T\}$ contains the probabilities of moving to each of the $T$ possible states $j : \{j = 1, 2, \ldots, T\}$ in the next period, conditional on state $i$ having been realized in the current period. Therefore, the SEL approach recovers the entire conditional distribution of the data. This $(T \times T)$ probability matrix denotes the 'beliefs' of the representative investor that are consistent with observed asset prices, i.e. beliefs that satisfy the conditional Euler equations.

Note that the SEL approach recovers the conditional distribution of the data, without the need for any parametric functional-form assumptions on the form of the distribution. It does so by approximating the conditional distribution, for each possible value of the current state, as a multinomial on the observed data sample. It may seem that this requires the estimation of $T^2$ conditional probabilities, given a sample size of only $T$. However, the number of parameters that the approach needs to estimate in order to generate the $(T \times T)$ conditional probability matrix is only $(T \times k)$, where $k$ denotes the number of conditional Euler equation restrictions, i.e. the number of assets used in the estimation. Specifically, for each date (or, current state), the SEL procedure only requires the estimation of the vector of Lagrange multipliers associated with the conditional Euler equation restrictions. Therefore, for each date, the number of parameters to be estimated is the same as the number of test assets that the SDF is asked to price (see Equations (6) and (7)). In our baseline case, the return on the market is the sole asset used in the estimation. Thus, the overall number of Lagrange multipliers, i.e. the total number of parameters, that need to be estimated is $T$.  

10
In practice, it can happen that the argument of the log function in Equation (7) becomes arbitrarily close to zero or even negative at certain dates. This creates numerical instability in estimation and makes \( \lambda_i \) a corner solution to the optimization problem (7). In order to avoid this case, we use the Owen (2001) normalization described in Appendix A.1. Also, note that the notation \((\hat{p}_{t,j}^{SEL}(\theta))\) emphasizes that the estimated conditional distribution is a function of the chosen value of the parameter vector \( \theta \). The true value \( \theta_0 \) of the parameter vector is, in principle, unknown to the econometrician. Although the SEL method also allows to estimate it, we will assume, for simplicity, that \( \theta_0 \) is known.

II.3 Conditioning Set

An important input to the beliefs extraction procedure described in Section II.2 is the specification of the information set \( F_t \) used by the representative investor to form her beliefs and price assets, i.e. the number and identity of the variables \( X_t \). As is standard in the literature, our non-parametric estimator suffers from the curse of dimensionality and we cannot include as many variables as possible. In order to address this potential shortcoming, we show that the recovered beliefs are remarkably similar for a wide set of choices of the conditioning variables.

We first present results when the conditioning set consists of the history of consumption alone – an assumption commonly made in a large class of macro finance models.

Next we add additional macro variables to the conditioning set. Our choices for these additional variables draws on the insight in Ghosh and Constantinides (2017), who contribute towards identifying the investors’ information set. The authors present evidence that the market-wide price-dividend ratio is strongly correlated with inflation and labor market variables that also have forecasting power for consumption, dividend, and GDP growth. In particular, their results suggest that just two macroeconomic variables – the rate of change in the CPI (inflation) 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. Moreover, these features are observed not just in the U.S., but also in most of the G7 countries. Drawing on the above findings, we also present results for the following choices of the conditioning set, \( X_t \):

(i) past consumption growth and inflation,
(ii) past consumption growth and the growth in the average hourly earnings of production on private non farm payrolls, and
(iii) past consumption growth, inflation,

\(^8\)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)).
and the growth in the average hourly earnings of production on private non farm payrolls.

Finally, to address the concern that additional variables, not included in the above choices, may be in the information sets of investors, we extract principal components (PCs) from a broad cross-section of over a hundred macro variables. The variables cover six broad categories of macroeconomic data: output, labor market, housing sector, orders and inventories, money and credit, and price levels. The first two PCs explain about 60% of the variation in these variables. We use these two PCs, along with consumption growth, as additional specifications of the conditioning set.

Key to our analysis is also our ability to capture a possible non-Markovian structure of the information set. For example, consider a specification where only the first lag of the consumption growth is included in the conditioning set. If the true beliefs dynamics involves not only \( C_t/C_{t-1} \) but the entire past of the process, then we have to incorporate more lags in the conditioning set. We deal with this issue using exponentially-weighted moving averages of our conditioning variables.\(^9\)

\[ \text{II.4 Parametric specifications of the SDF} \]

Different assumptions about investors’ preferences lead to different specifications of the SDF in Equation (3). In our benchmark case, we 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 (CRRA). For this specification of preferences, the SDF takes the form:

\[
M\left(\frac{C_{t+1}}{C_t}, Y_{t+1}; \theta_0\right) = \delta \left(\frac{C_{t+1}}{C_t}\right)^{-\theta_0} \quad \text{and} \quad \mathbb{E}^P_t \left[ \left(\frac{C_{t+1}}{C_t}\right)^{-\theta_0} R_{t+1}^e | X_t \right] = 0, \quad (8)
\]

where \( Y_{t+1} = \emptyset \) and \( \theta_0 \in \mathbb{R}^+ \) is the representative agent’s CRRA.

It is well known 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 (see e.g. Mankiw and Shapiro (1986), Breeden, Gibbons, and Litzenberger (1989), Campbell (1996), Cochrane (1996)).

As a result, to demonstrate the robustness of our results, 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. These

\(^9\)Our exponentially weighted variables are computed as \( X_t^{(EW)} = \alpha X_t + (1-\alpha)X_t^{(EW)} \), thus incorporating efficiently the entire past of our conditioning process. In practice, we set \( \alpha = 0.28 \), whereby the past 13 quarters receive 99% of the weight. Our results are not sensitive to the value of \( \alpha \).
include the external habit formation model (see, e.g., Campbell and Cochrane (1999)) and Epstein and Zin (1991) recursive preferences in the presence of long run risks in consumption growth (see, e.g., Bansal and Yaron (2004)). Since these models are standard in the literature, we refer the reader to Appendix A.2 for more details on the associated functional forms of the SDF.

III Performance of SEL Estimator: Two Example Economies

In this section, we demonstrate, via two hypothetical simulated economies, the performance of the SEL estimation approach described in II.

In our first simulation exercise, we show that the SEL estimator is remarkably successful in recovering the subjective beliefs of investors’ when the latter diverges from the true underlying distribution of the data. Specifically, we consider an endowment economy where a representative agent has power utility preferences with a constant coefficient of relative risk aversion (CRRA). Suppose that consumption growth is i.i.d. log-normal:

\[ \log \left( \frac{C_{t+1}}{C_t} \right) \sim N(\mu, \sigma^2) , \]

where \( \mu \) is the mean of the consumption growth. We assume that the representative investor is pessimistic and acts as if the average consumption growth was lower than \( \mu \). Specifically, she acts as if consumption growth has a mean of \((1 - \lambda)\mu\), where \( \lambda \in (0, 1) \) is the severity of pessimism:

\[ \log \left( \frac{C_{t+1}}{C_t} \right) \sim N(\tilde{\mu}, \sigma^2) , \]

where \( \tilde{\mu} = (1 - \lambda)\mu \). We assume that there are no distortions in the beliefs about the volatility and the higher moments of consumption growth.

In the above scenario, equilibrium asset prices reflect this subjective belief of investors. In particular, the equilibrium price-dividend ratio is \( \frac{P_t}{D_t} = z \), a constant, where

\[ z = \frac{\exp \left[ \log(\delta) + (1 - \theta_0)\tilde{\mu} + \frac{(1 - \theta_0)^2\sigma^2}{2} \right]}{1 - \exp \left[ \log(\delta) + (1 - \theta_0)\tilde{\mu} + \frac{(1 - \theta_0)^2\sigma^2}{2} \right]} , \]

and the equilibrium risk free rate is:

\[ R_{f,t} = \frac{1}{\exp \left( \log(\delta) + (1 - \theta_0)\tilde{\mu} + \frac{(1 - \theta_0)^2\sigma^2}{2} \right)} . \]
In the above equations, $\delta$ denotes the subjective discount factor and $\theta_0$ the risk aversion coefficient.

To perform our simulation exercise, we calibrate $\mu$ and $\sigma^2$ to the sample mean and variance, respectively, of (log) consumption growth in our data. The preference parameters are calibrated at $\delta = 0.99$ and $\theta_0 = 10$. We simulate a time series of consumption growth of the same length as the historical data ($T = 267$ with quarterly data, see Section IV). Using the simulated consumption growth, we obtain the market return as:

$$R_{m,t+1} = \frac{P_{t+1}}{C_{t+1}} \cdot \frac{C_t}{C_{t+1}} = \frac{z + 1}{z} \cdot \frac{C_{t+1}}{C_t},$$

where $z$ is defined in Equation (11). The time series of the risk free rate is simply a constant given by Equation (12).

We then use the EL approach to estimate the unconditional distribution of consumption growth. We have:

$$\Delta = \left\{(p_1, ..., p_T) : \sum_{t=1}^T p_t = 1, p_t \geq 0 \right\}$$

and:

$$\left(\hat{p}^{EL}(\theta_0)\right) = \arg \max_{(p) \in \Delta} \sum_{t=1}^T \log(p_t) \quad \text{s.t.} \quad \sum_{t=1}^T p_t \frac{C_t}{C_{t-1}} (r_{m,t} - r_{f,t}) = 0, \quad (13)$$

Using the estimated probabilities $\hat{p}^{EL}$, we compute the mean, volatility, and the skewness of consumption growth. Note that these are the moments of consumption growth that are consistent with the asset prices, i.e. the moments as perceived by the representative investor. We repeat the above estimation for 500 simulated samples. We report the means and 95% confidence intervals of the moments of consumption growth across these simulations. To demonstrate the power of the estimation approach, we present results for different magnitudes of the beliefs distortion, i.e. for $\lambda = \{0.05, 0.1, 0.15\}$, and for different simulated sample sizes, i.e. $T_{sim} = \{267, 500, 1,000\}$.

The results are reported in Table 1. Panel A presents results for simulated samples of the same length as the historical time series. Consider first Row 1, where investors are assumed to underestimate the mean of consumption growth by 5%, i.e. the quarterly mean of 0.46% under subjective beliefs is 5% below the historical mean of 0.48%. The equilibrium market return and risk free rate reflect these subjective beliefs of investors. Row 1 shows that the EL method is successful at capturing these subjective beliefs of investors. Specifically, the SEL-implied mean of consumption growth has a mean of 0.46% across the 500 simulations, coinciding with the true value of the mean under subjective beliefs. Moreover,
the 90% confidence interval for the mean is very tight and does not include the historical mean of consumption growth. The SEL implied volatility of consumption growth has a mean of 0.51% across the 500 simulations – identical to the historical value. Note that, in our experiment, there are no beliefs distortions in the volatility and the SEL method successfully identifies the volatility observed in the historical data. Finally, the average of the coefficient of skewness across the simulations is 0.003, very close to the true value of 0, and the 90% confidence interval excludes values greater than 0.24 in absolute terms. Rows 2 and 3 show that very similar results are obtained for more severe beliefs distortions in the mean of consumption growth – the SEL method correctly identifies the subjective mean, and the estimated volatility and skewness are very close to their historical values with tight confidence bands. Finally, Panels B and C show the effect of increasing sample size on the
performance of the SEL estimator – the performance at samples sizes of 500 and 1,000 are quite similar to those observed for available sample sizes in the historical data.

In our second example, we show that the SEL estimator is successful in recovering the true conditional distribution of macro variables in the absence of any beliefs distortions. Specifically, we consider the long run risks model of Bansal and Yaron (2004). In this model, aggregate consumption and dividend growth have a small persistent predictable component and stochastic volatility that captures time-varying economic uncertainty:

\begin{align*}
\Delta \log (C_{t+1}) & = \mu_c + x_t + \sigma_t \epsilon_{c,t+1}, \\
\Delta d_{t+1} & = \mu_d + \phi x_t + \phi_d \sigma_t \epsilon_{d,t+1}, \\
x_{t+1} & = \rho x_t + \phi_x \sigma_t \epsilon_{x,t+1}, \\
\sigma^2_{t+1} & = (1 - \nu)\sigma^2 + \nu \sigma^2_t + \sigma_w \epsilon_{w,t+1},
\end{align*}

where \( d_{t+1} \) is the dividend process, and the shocks are all standard normal and mutually independent. The representative agent in this economy has Kreps-Porteus recursive preferences. Thus, in equilibrium, the following conditional Euler equations are satisfied:

\[ \mathbb{E}^{P_t} \left[ \left( \frac{C_{t+1}}{C_t} \right)^{-\alpha/\psi} R_{c,t+1}^{\alpha-1} R^e_{t+1} | x_t, \sigma^2_t \right] = 0, \]

where \( R_{c,t+1} \) is the unobservable return on total wealth, \( \alpha = \frac{1-\theta_0}{1-\psi} \), \( \theta_0 \) is the CRRA, and \( \psi \) is the elasticity of intertemporal substitution. Note that in the above equation, \( P_t \) denotes the objective distribution of the data summarized in Equations (14)-(17).

We solve for the equilibrium as in the original article and set the model parameters equal to the authors’ calibrated values. We simulate a time series of the two state variables, consumption growth, dividend growth, market return and risk free rate of the same length as the historical data. We then use the SEL approach to recover \( P_t \) using the excess returns on the market portfolio as the sole test asset and \((x_t, \sigma^2_t)\) as conditioning variables. Using the recovered conditional distribution, we compute the time series of the conditional mean, volatility, and skewness of consumption growth. We then repeat this process 500 times.

Figure 1-I, Panel A plots, for a randomly chosen sample from among the 500 simulated samples, the true time series of the conditional mean of consumption growth, \( \mu_c + x_t \), (red line) along with the time series of the mean implied by the estimated SEL probabilities (black line). The correlation between the two time series is 95.5%. Panel B plots the true, \( \sigma_t \), (red line) and SEL-implied (black line) time series of the conditional volatility of consumption growth. The correlation between these two time series is 90.1%. Panel C plots the true (red line) and SEL-implied (black line) time series of the conditional skewness of consumption growth.
Notes: The left panel of this figure plots the subjective conditional moments as estimated by the SEL on one simulated trajectory of the long-run risk model. The right panel presents the distribution of the errors on the estimated subjective conditional moments. For each simulation, we obtain time series of subjective conditional moments that we compare to the true simulated moments. For each trajectory we form the time series of errors and report both mean (black solid line) and 95% confidence bands (red dashed lines). Our calibration uses parameters from Bansal and Yaron (2004).

growth. In the model, consumption growth is conditionally Gaussian and, therefore, has a coefficient of skewness equal to 0 at all times. For the skewness implied by the estimated SEL probabilities, the estimated skewness is less than 0.15 in magnitude in 94.4% of the time periods.

To summarize the performance of the SEL estimator across the simulated samples, in each sample, we obtain the time series of the deviations of the conditional mean, volatility, and skewness of consumption growth from their respective true values. Figure 1-II plots the median and 90% confidence interval of the deviations for the conditional mean (Panel A), volatility (Panel B), and skewness (Panel C) across the 500 samples. The figure shows the median deviations are zero for each time period for all three conditional moments and the 90% confidence intervals of the deviations are quite tight.

IV Data Description

We present empirical results at the quarterly frequency over the sample period 1947:Q1–2013:Q4. For consumption, we use per capita real personal consumption expenditures on nondurable goods and services from the National Income and Product Accounts (NIPA).
We make the standard “end-of-period” timing assumption that consumption during quarter $t$ takes place at the end of the quarter.

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 real risk free rate is obtained as follows: the quarterly nominal yield on three-month Treasury bills is deflated using the realized growth in the Consumer Price Index (CPI) to obtain the ex-post real three-month Treasury-bill rate. The ex-ante quarterly risk free rate is then obtained as the fitted value from the regression of the ex-post three-month Treasury-bill rate on the three-month nominal yield and the realized growth in the CPI over the previous year.

In addition to using the excess returns on the market portfolio as the sole asset in the extraction of the subjective beliefs of investors, we also present results when the set of assets include portfolios of small market capitalization, large market capitalization, growth and value stocks. Monthly returns on these portfolios are obtained from Kenneth French’s data library, and correspond to the the smallest and largest deciles of portfolios formed by sorting the universe of U.S. stocks on the basis of size and book-to-market-equity.

Quarterly returns for the above assets are computed by compounding monthly returns within each quarter and are converted to real returns using the CPI.

As discussed in Section II.3, we recover investors’ beliefs for several different choices of the conditioning set. The conditioning variables used include consumption growth, the growth rate in the CPI, the growth in the average hourly earnings of production on private non farm payrolls, and principal components extracted from a broad cross section of 106 macroeconomic variables (that includes the CPI and earnings variable). We obtain panel data on the 106 macroeconomic variables from Sydney Ludvigson’s web site, based on the Global Insights Basic Economics Database and The Conference Board’s Indicators Database. The variables cover six broad categories of macroeconomic data: output, labor market, housing sector, orders and inventories, money and credit, and price levels. We refer the reader to Ludvigson’s website for a detailed description of these variables.

V Empirical Results: Characterizing Beliefs

We first estimate the beliefs using the SEL approach for our baseline specification – namely, where the representative investor has power utility preferences with a constant CRRA, the excess return on the market portfolio is the sole test asset, and the conditioning set consists of an exponentially-weighted moving average of lagged consumption growth. We set the risk aversion coefficient $\theta_0$ equal to 10 (the upper bound of what is generally considered to be
an acceptable range). Note that the SEL estimation approach, like all other nonparametric procedures, requires specification of the kernel function and the associated bandwidth parameter. All our results are computed with the Epanechnikov kernel function and with the bandwidth parameter $b_{v,T} = 3\hat{\sigma}_v$, where $\hat{\sigma}_v$ is the empirical standard deviation of the conditioning variable $v$.

Note that the SEL approach delivers the conditional distribution of future outcomes, for each possible value of the conditioning set, i.e. for each date. Therefore, we can obtain the conditional pricing error for the underlying asset (the excess market return in this case) at each date $t$ as $\sum_{j=1}^{T} \hat{p}_{t,j}^{SEL} \times g^{-\theta}_j r_{m,j}$. Figure 2 plots the time series of the pricing errors (black line). The figure shows that the pricing errors are identically equal to zero at each time period, demonstrating the strength of the SEL method.

To further highlight the merits of the SEL method, we also obtain the pricing error at each date when the conditional expectation underlying the expression for the pricing error, $\mathbb{E} \left[ \frac{C_{t+1}}{C_t} \right] \mathbf{R}^{e}_{m,t+1} | X_t \right]$, is evaluated using the widely used nonparametric local linear regression (LLR) method. The LLR method estimates the conditional mean function at the current state $x_t$ by fitting a linear regression locally, with weighted least squares in a fixed neighbourhood of $x_t$. As with the SEL estimator, the neighbourhood is defined in terms of the distance of other possible values of the state from the current state, i.e. $|x_i - x_j|$, and not in terms of proximity in time. The weights are determined by the kernel function, the distance $|x_i - x_j|$, and the bandwidth parameter $b_T$. These are chosen to be identical to those used in the SEL approach to facilitate comparison. The fitted value from the regression at $x_t$ provides an estimate of the conditional pricing error for the excess return on the market portfolio at date $t$.

Figure 2 plots the time series of the pricing errors obtained using the LLR method (red line). The conditional pricing errors, unlike those obtained using the SEL approach, are large and volatile, varying from $-10\%$ and $21\%$ (annualized). These results are consistent with the findings in Nagel and Singleton (2011) who show that asset pricing models, even the ones that produce small average or unconditional pricing errors, typically produce large and volatile conditional pricing errors. They, therefore, conclude that models are unable to simultaneously match the cross section and time series of asset returns. They then go on to propose an econometric procedure for the estimation of the parameters of conditional asset pricing models, aimed as reducing the conditional pricing errors. The SEL approach, on the other hand, successfully sets the conditional pricing errors to zero. Moreover, unlike Nagel

---

10The results are robust to values of $\theta_0$ between 1 and 10.

11The results are robust to alternative choices of the kernel function and smoothing parameters within four standard deviations of the respective conditioning variables. These results are omitted for brevity and are available from the authors upon request.
and Singleton (2011), it does not require the underlying SDF to be affine in the parameters.

Figure 2 – Time Series of Conditional Pricing Errors

Notes: The figure plots the time series of the conditional pricing errors for the excess return on the market portfolio. The conditional expectation underlying the calculation of the pricing error is evaluated using the SEL probabilities. The pricing kernel is that implied by power utility preferences with a CRRA, the excess return on the market is used as the sole asset, and the conditioning set consists of an exponentially-weighted moving average of past consumption growth. The sample is quarterly covering the period 1947:Q2-2013:Q4.

Having shown the success of the SEL approach in pricing assets, we next turn to a characterization of the recovered beliefs. Note that these beliefs are consistent with observed asset prices, i.e. they satisfy the conditional Euler equation restrictions. We begin by presenting the SEL-implied conditional densities of consumption growth in two chosen (good and bad) states of the world. Section V.1. In Section V.2, we present the time series of the first three moments from the conditional distributions of consumption growth. This helps shed light on the dynamic evolution of the beliefs of investors, i.e. the beliefs formation process. In Section V.3, we present the beliefs for alternative choices of the three key inputs required in our approach, namely the SDF, the cross section of asset returns, and the conditioning set. Finally, Section V.4 presents the implications of the recovered beliefs for the expected equity premium, and the conditional volatility and Sharpe ratio of the market return.

Note that, so far, we do not take a stance on whether or not the recovered beliefs are rational. One can view our results as providing guidance on modelling assumptions typically
required in macro-finance models to match the observed dynamics of asset prices.

V.1 Investors’ Beliefs About Consumption Growth

For each possible realization of the conditioning set, i.e. the current state, the SEL approach delivers the conditional probabilities attached to the different 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 dates.

Figure 3 – Estimated conditional beliefs densities, $\theta_0 = 10$

Notes: The figure plots the conditional densities of consumption growth in 2009:Q3 (Panel A) and 1966:Q1 (Panel B). The conditional densities are obtained using the estimated SEL distributions for the realizations of the conditioning state vector in 2009:Q2 and 1965:Q4, respectively. The pricing kernel corresponds to the time and state separable power utility model with a constant CRRA, the excess return on the market portfolio is the sole test asset, and the conditioning set includes an exponentially-weighted moving average of past consumption growth. The sample is quarterly covering the period 1947:Q2-2013:Q4.

Consider, first, the period of the recent financial crisis. Figure 3, Panel A presents the SEL-implied conditional density of consumption growth (black line) in 2009:Q3, given the information available in the preceding quarter 2009:Q2. The annualized mean and volatility of this distribution are 0.8% and 1.1%, respectively. Superimposed in the same graph is a Gaussian density (red line) with the same mean and variance as the SEL density. The
coefficient of skewness of the SEL density is 0.17, close to the corresponding value of 0 for the normal density. The skewness of the SEL density is close to 0 (in fact, mildly positive), despite the bi-modal nature of the density, with the second (lower mode) occurring to the left of the main mode. This is driven by investors assigning larger probabilities to events higher than two and three standard deviations from the current mean compared to events lower than two and three standard deviations from the current mean. Specifically, in 2009:Q2, the average investor in the stock market assigns a probability of 14.0% to the event that consumption growth in the subsequent quarter will be lower than one standard deviation from its conditional mean, compared to a very similar probability of 14.4% for the event that consumption growth will be higher than one standard deviation from its mean; assigns a probability of 1.6% to the event that consumption growth in the subsequent quarter will be lower than two standard deviations from its conditional mean, compared to a higher probability of 4.3% for the event that consumption growth will be higher than two standard deviations from its mean; and assigns a probability of 0.0% to the event that consumption growth in the subsequent quarter will be lower than three standard deviations from its conditional mean, compared to a probability of 0.5% for the event that consumption growth will be higher than three standard deviations from its mean.

Consider, next, Figure 3, Panel B that presents the conditional density of consumption growth in 1966:Q1, given the information available in the preceding quarter 1965:Q4. This period, unlike the financial crisis, was characterized by high economic growth with real per capita consumption growth averaging 1.0% per quarter or 4.0% annualized over the past three years. The annualized mean of the SEL distribution in 1966:Q1 is 2.4% – a 200% increase compared to the mean of only 0.8% in 2009:Q3. On the other hand, the annualized volatilities of the distributions at these two dates are almost identical (0.94% in 1966:Q1 versus 1.1% in 2009:Q3). Interestingly, Panel B reveals that the conditional distribution of consumption growth is highly negatively skewed in this good state, with a coefficient of skewness of −0.65. Specifically, the average investor assigns a probability of 4.6% to the event that consumption growth in the subsequent quarter will be lower than two standard deviations from its conditional mean, compared to a corresponding probability of only 2.3% for an investor whose beliefs are conditionally Gaussian; and assigns a probability of 1.4% to the event that consumption growth in the subsequent quarter will be lower than three standard deviations from its conditional mean, compared to an order of magnitude smaller probability of 0.1% for the same event for an investor whose beliefs are conditionally Gaussian. On the other hand, the investor assigns lower probabilities of 1.0% and 0.0%, respectively, to the events that consumption growth in the subsequent quarter will be higher than two and three standard deviations from its conditional mean compared to an investor.
Figure 4 – Estimated conditional beliefs densities, $\theta_0 = 10$

Notes: The figure plots the conditional densities of consumption growth in 2009:Q3 and 1966:Q1. The conditional densities are obtained using the estimated SEL distributions for the realizations of the conditioning state vector in 2009:Q2 and 1965:Q4, respectively. The pricing kernel corresponds to the time and state separable power utility model with a constant CRRA, the excess return on the market portfolio is the sole test asset, and the conditioning set includes an exponentially-weighted moving average of past consumption growth. The sample is quarterly covering the period 1947:Q2-2013:Q4.

To further facilitate comparison of the SEL densities between good and bad states, Figure 4 superimposes the estimated densities in periods 2009:Q3 and 1966:Q1 in the same plot.

To summarize, three main conclusions emerge from this section. First, the conditional distribution of consumption growth, as perceived by the average investor, has a markedly lower mean during bad times compared to good times, i.e. the mean of the distribution shifts to the left during bad states of the world. Second, the conditional volatility of consumption growth is perceived to be fairly flat across the business cycle. Third, the perceived distribution of consumption growth is more negatively skewed during good times. In other words, even during particularly good times, when expected consumption growth is very high, investors still attach substantial probabilities to severe economic downturns. As an illustration, in 1965 : Q4, the expected consumption growth is 2.4% (annualized); yet the probabilities
attached to consumption growth in the next quarter being less than \(-1.3\%\) or \(-3.2\%\) are non-trivial at 4.6\% and 1.4\%, respectively.

The above results call into question the widely used assumption of a conditionally Gaussian data generating process (DGP) in asset pricing models. Before turning to a more formal comparison with widely assumed DGPs, we provide further characterization of the estimated conditional distributions by looking at their time series evolution.

V.2 Evolution and Dynamics of Beliefs

The SEL approach provides estimates of the conditional probabilities of the different possible future states for each possible value of the current state. Therefore, it is easy to compute the conditional moments of the macro variables for different values of the conditioning state. For instance, let us consider the log consumption growth, \(\log(g_t)\). The mean of \(\log(g_{t+1})\), conditional on the information available on date \(t\), is given by:

\[
\hat{\mathbb{E}}_t[\log(g_{t+1})|X_t] = \sum_{j=1}^{T} \hat{p}_{t,j}^{SEL}(\theta) \cdot \log(g_j).
\] (19)

Note that this conditional expectation can be computed for each date \(t\), i.e. for each realized value of the conditioning set, to obtain a time series of the conditional mean of consumption growth as perceived by the average investor. The conditional expectation of any nonlinear transformation of the consumption growth rate can be similarly computed. We perform this computation to obtain the time series of the conditional mean, volatility, and skewness of consumption growth.

The results are presented in Figure 5. Panel A presents the time series of the conditional mean of consumption growth. Several patterns are evident from the panel. First, the conditional mean is strongly procyclical. The conditional mean is at its peak at or shortly before the onset of a recessionary episode (denoted by shaded grey areas), declines steadily through the recession, reaches its trough around the end of the recession before rising back up. The correlation between the conditional mean and a dummy variable that takes the value one in a given quarter if that quarter is classified as an NBER-designated recession period and zero otherwise is \(-48.1\%\). Procyclicality in the conditional mean of consumption growth is not a surprising result and has been extensively documented in the literature. Second, beliefs about expected consumption growth are more persistent and less volatile than realized consumption growth. Specifically, the conditional mean has an annualized volatility of 0.1\%, that is an order of magnitude smaller than the 1.0\% volatility of realized consumption growth. And the conditional mean has a much higher persistence with a first-order autocor-
Figure 5 – Time series of Conditional Moments, $\theta_0 = 10$

**Panel A: Conditional Mean**

**Panel B: Conditional Volatility**

**Panel C: Conditional Skewness**

Notes: The figure plots the time series of the conditional mean (Panel A), conditional volatility (Panel B), and conditional coefficient of skewness (Panel C) of consumption growth. Shaded areas denote NBER designated recession periods. The conditional moments are obtained using the estimated SEL distributions. The pricing kernel corresponds to the time and state separable power utility model with a constant CRRA, the excess return on the market portfolio is the sole test asset, and the conditioning set includes exponentially-weighted moving averages of lagged consumption growth and inflation. The sample is quarterly covering the period 1947:Q2-2013:Q4.

relation coefficient of 0.88 than that of realized consumption growth that has a first-order autocorrelation coefficient of only 0.31.

Panel B presents the time series of the conditional volatility of consumption growth. The figure shows that the conditional volatility is fairly flat over the time period 1947-2013. Countercyclicality in the conditional volatility of consumption growth is a more debated feature of the data for which limited direct empirical evidence exists. Our results suggest
that the time-variation in the volatility is miniscule compared to the variation in the mean. Over the sample period, the conditional mean varies from 0.8% is 2009:Q3 to 2.4% in 1966:Q1 – a 200% increase. On the other hand, the conditional volatility varies from 0.9% to 1.1% – an only 22% increase.

Panel C presents the time series of the conditional skewness of consumption growth. This panel reveals that investors perceive the skewness of consumption growth to be negative in almost all states and, perhaps more importantly, strongly time varying. The skewness varies from −0.65 in 1966:Q1 to 0.17 in 2009:Q3, with an average of −0.36. Moreover, the time-variation is cyclical – the correlation between the skewness and the recession dummy is 44.0%. Thus, the skewness is more negative during good states of the world, even though the expected consumption growth is markedly higher during these periods compared to bad states. This suggests that, even during particularly good times, when expected consumption growth is high, investors still attach substantial probabilities to severe economic downturns.

Finally, we show that investors’ beliefs about consumption growth has strong forecasting power for consumption growth. A regression of the realized consumption growth on the SEL-implied expected consumption growth produces a highly statistically significant coefficient (with a t-statistic of 5.27) and an $R^2$ of 9.5%. Figure 6 presents the time series of realized consumption growth (black line) and the fitted value from a forecasting regression of the realized consumption growth on its SEL-implied conditional mean (red line).

Overall, our results suggest that cyclical variation in the first and third moments of macro variables such as consumption growth are important components of the dynamics of these variables as perceived by the average investor, i.e. an important component of the beliefs formation process. Time-variation in the second moment (volatility), on the other hand, seems to be economically small. Variation in the skewness is typically missing from modelling assumptions about the underlying data generating process or beliefs formation process often made in the literature.

V.3 Alternative Pricing Kernels, Cross Sections of Assets, and Instruments

Note that the SEL approach delivers estimates of investors’ beliefs, given three key inputs – namely, the SDF that describes investors’ preferences, the cross section of assets that the SDF is required to price, and the investors’ conditioning set. To demonstrate robustness, we present results for different choices of these three inputs.

First, we consider alternative choices of the SDF. Indeed, the shortcomings of the time and state separable power utility model with a constant CRRA have been extensively documented in the literature. This raises the question as to whether the ’beliefs’ that we recover using
Figure 6 – Forecasting Consumption Growth

Notes: The figure plots the time series of the realized consumption growth (black line) and the fitted value from a forecasting regression of the realized consumption growth on its SEL-implied conditional mean (red line). Shaded areas denote NBER designated recession periods. The conditional mean is obtained using the estimated SEL distributions. The pricing kernel corresponds to the time and state separable power utility model with a constant CRRA, the excess return on the market portfolio is the sole test asset, and the conditioning set includes exponentially-weighted moving averages of lagged consumption growth and inflation. The sample is quarterly covering the period 1947:Q2-2013:Q4.

this specification of the pricing kernel indeed capture the beliefs of investors, or do they identify a component of the true underlying SDF unrelated to beliefs that is missing from this simple power utility kernel. We consider two alternative specifications of the SDF that have been proposed to overcome some of the limitations of the power utility model and that have proven quite successful at explaining a variety of observed features of financial market data. Specifically, we consider the SDFs implied by the long run risks model of Bansal and Yaron (2004) with Epstein and Zin (1989) recursive preferences (hereafter referred to as BY) and the external habit formation preferences of Campbell and Cochrane (1999) (hereafter referred to as CC).

For each of these pricing kernels, we use the SEL approach to recover the conditional distribution of future consumption growth, for each possible value of the conditioning set. Using the estimated conditional distributions, we obtain the time series of the conditional
Figure 7 – Time series of Conditional Moments, Alternative SDFs

Notes: The figure plots the time series of the conditional mean (Panel A), conditional volatility (Panel B), and conditional coefficient of skewness (Panel C) of consumption growth. Shaded areas denote NBER designated recession periods. The conditional moments are obtained using the estimated SEL distributions. The pricing kernel corresponds to the standard CCAPM (CCAPM), external habit model (CC), and the long run risks model with recursive preferences (BY). The excess return on the market portfolio is the sole test asset and the conditioning set includes an exponentially-weighted moving averages of lagged consumption growth. The sample is quarterly covering the period 1947:Q2-2013:Q4.

moments of consumption growth. Figure 7 plots the time series of the conditional mean (Panel A), the conditional volatility (Panel B), and the conditional coefficient of skewness (Panel C) of consumption growth, obtained using the \( CC \) kernel (red-dashed line), and the \( BY \) kernel (blue-dotted line). To facilitate comparison, we also plot the time series of these moments recovered using the \( CCAPM \) kernel (black line).

The main conclusion that emerges from Figure 7 is that the beliefs recovered from these three models, with very different specifications of preferences, are remarkably similar. Consider Panel A for example. The time series of the conditional mean of consumption growth as perceived by the average investor is virtually indistinguishable, regardless of the specification of the SDF. Similar conclusions are obtained for the time series of the conditional
volatility and skewness in Panels B and C, respectively. Table II reports the correlations between the time series of the conditional mean (Panel A), conditional volatility (Panel B), and conditional skewness (Panel C) of consumption growth for the three different SDFs considered. The correlations are all very high, varying from 92.2% and 99.9%. Overall, the results suggest that the characteristics of beliefs are quite robust to the choice of investor preferences.

Table II: Correlation Between Beliefs: Alternative SDFs

<table>
<thead>
<tr>
<th></th>
<th>CCAPM</th>
<th>CC</th>
<th>BY</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A. Correlation Between Means</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCAPM</td>
<td>1.0</td>
<td>0.999</td>
<td>0.998</td>
</tr>
<tr>
<td>CC</td>
<td></td>
<td>1.0</td>
<td>0.993</td>
</tr>
<tr>
<td>BY</td>
<td></td>
<td></td>
<td>1.0</td>
</tr>
<tr>
<td><strong>Panel B. Correlation Between Volatilities</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCAPM</td>
<td>1.0</td>
<td>0.961</td>
<td>0.929</td>
</tr>
<tr>
<td>CC</td>
<td></td>
<td>1.0</td>
<td>0.922</td>
</tr>
<tr>
<td>BY</td>
<td></td>
<td></td>
<td>1.0</td>
</tr>
<tr>
<td><strong>Panel C. Correlation Between Skewness</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCAPM</td>
<td>1.0</td>
<td>0.960</td>
<td>0.965</td>
</tr>
<tr>
<td>CC</td>
<td></td>
<td>1.0</td>
<td>0.957</td>
</tr>
<tr>
<td>BY</td>
<td></td>
<td></td>
<td>1.0</td>
</tr>
</tbody>
</table>

The table reports the correlations between the time series of the conditional mean (Panel A), conditional volatility (Panel B), and conditional skewness (Panel C) of consumption growth for the three different SDFs considered.

The next key input in the SEL approach is the choice of the cross section of assets. The results presented so far were for the scenario where the excess return on the market portfolio is the sole test asset. We examine the robustness of the results by expanding the cross section of assets to include the excess returns on the 'Small' and 'Big' portfolios (the bottom and top deciles of portfolios formed by sorting stocks on the basis of market capitalization) and 'Growth' and 'Value' portfolios (the bottom and top deciles of portfolios formed by sorting stocks on the basis of the book-to-market-equity ratio), in addition to the market portfolio. The extracted beliefs, presented in Figure 8, are almost identical to those obtained when the excess return on the market portfolio is the sole asset used in the estimation. The correlations between the conditional means, volatilities, and skewness for the two choices of test assets are 99.7%, 98.8%, and 97.8%, respectively.

Next, we show that the extracted beliefs are also fairly robust to the specification of the conditioning set – adding to the exponentially weighted moving average of consumption
Figure 8 – Time series of Conditional Moments, Alternative Test Assets

**Panel A: Conditional Mean**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.002</td>
<td>0.004</td>
<td>0.006</td>
<td>Mkt</td>
<td>Mkt,S,L,G,V</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Panel B: Conditional Volatility**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Volatility</td>
<td>0.003</td>
<td>0.005</td>
<td>0.007</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Panel C: Conditional Skewness**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Skewness</td>
<td>-0.6</td>
<td>-0.2</td>
<td>0.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** The figure plots the time series of the conditional mean (Panel A), conditional volatility (Panel B), and conditional coefficient of skewness (Panel C) of consumption growth. Shaded areas denote NBER designated recession periods. The conditional moments are obtained using the estimated SEL distributions. The pricing kernel corresponds to the standard CCAPM (CCAPM). The test assets consist of the excess return on the market portfolio (black line) and excess returns on the market, Small, Big, Growth, and Value portfolios (red-dashed line). The conditioning set includes an exponentially-weighted moving averages of lagged consumption growth. The sample is quarterly covering the period 1947:Q2-2013:Q4.

growth moving averages of additional conditioning variables such as the rate of inflation, labor market variables such as the growth in average hourly earnings of production in private non farm payrolls, principal components extracted from a broad cross section of over a hundred macro variables, or the excess stock market return, makes little difference to the results. Our choices for the conditioning set (beyond past consumption growth) include variables that have strong correlation with the market-wide price-dividend ratio and that also forecast future consumption and dividend growth rates, thereby suggesting that investors’ use such information to form their beliefs.

Figure 9 presents the time series of the first three moments of consumption growth recovered from six different choices of the conditioning set. The choices include
include an exponentially-weighted moving averages of lagged consumption growth (light blue line), consumption growth and inflation (black line), consumption growth, inflation, and market return (red line), consumption growth, inflation, and growth in the average hourly earnings of production in private non farm payrolls (green line), consumption growth, inflation, growth in the average hourly earnings of production in private non farm payrolls, and market return (dark blue line), and consumption growth and a principal component extracted from a broad cross section of over a hundred macro variables (pink line). Table III presents the correlations between the extracted conditional moments of consumption growth for these different choices of the conditioning set.

The table and figure show that our results are quite robust to the choice of the conditioning set. Specifically, the conditional mean is strongly procyclical, regardless of the choice of the conditioning set. Including lagged asset returns in the conditioning set reduces further the estimated conditional mean of consumption growth during bad times, compared to specifications where only lagged macro variables are in the conditioning set. The pairwise correlations between the time series of conditional means recovered from the different choices of the conditioning set vary from 0.75 to 0.98, with an average of 0.87.

The pattern in the conditional volatility of consumption growth is also similar (albeit less so than the conditional mean) across the various specifications of the conditioning set. The average pairwise correlation between the time series of conditional volatilities recovered from the different choices of the conditioning set is 0.63. For some choices of the conditioning set, a countercyclical pattern in the volatility is somewhat more pronounced than when the conditioning set consists of the consumption history alone, e.g when the conditioning set consists of the history of consumption growth and a principal component extracted from a broad cross section of macro variables. However, even in the latter case, the range of variation in the conditional volatility of consumption growth is quite small – the annualized volatility varies from 0.72% to 1.1%, compared to varying from 0.94% to 1.1% when consumption growth alone is in the conditioning set.

And the conditional skewness is largely negative and varies cyclically for most choices of the conditioning set. The average pairwise correlation between the time series recovered from the different choices of the conditioning set is 0.58.

Finally, our results are robust to the data frequency. Repeating the SEL approach to estimating the beliefs on annual data over the entire available sample period 1890 – 2009 or over the sample 1929 – 2013 when disaggregated data on nondurables and services consumption is available gives very similar results.\textsuperscript{12}

\textsuperscript{12}The results are omitted for brevity and are available from the authors upon request.
Figure 9 – Time series of Conditional Moments, Alternative Instruments

Panel A: Conditional Mean

Panel B: Conditional Volatility

Panel C: Conditional Skewness

Notes: The figure plots the time series of the conditional mean (Panel A), conditional volatility (Panel B), and conditional coefficient of skewness (Panel C) of consumption growth. Shaded areas denote NBER designated recession periods. The conditional moments are obtained using the estimated SEL distributions. The pricing kernel corresponds to the standard CCAPM (CCAPM) and the test asset consist of the excess return on the market portfolio. The conditioning set includes an exponentially-weighted moving averages of lagged consumption growth (light blue line), consumption growth and inflation (black line), consumption growth, inflation, and market return (red line), consumption growth, inflation, and growth in the average hourly earnings of production in private non farm payrolls (green line), consumption growth, inflation, growth in the average hourly earnings of production in private non farm payrolls, and market return (dark blue line), and consumption growth and a principal component extracted from a broad cross section of over a hundred macro variables (pink line). The sample is quarterly covering the period 1947:Q2-2013:Q4.

V.4 Beliefs About the Stock Market

So far, we have focused on beliefs about consumption growth. The recovered beliefs, however, also have implications for asset returns. In this sub-section, we present the implications of the recovered beliefs for the aggregate stock market return.

We start by presenting the beliefs about the expected equity premium. Just like with consumption growth, we can use the SEL probabilities attached to the different possible
Table III: Correlation Between Recovered Beliefs: Alternative Conditioning Sets

<table>
<thead>
<tr>
<th></th>
<th>cg, Inf</th>
<th>cg, Inf, rm</th>
<th>cg, Inf, Ear</th>
<th>cg, Inf, Ear, Rm</th>
<th>cg</th>
<th>cg, PC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A. Correlation Between Means</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cg, Inf</td>
<td>1.0</td>
<td>0.826</td>
<td>0.953</td>
<td>0.809</td>
<td>0.906</td>
<td>0.965</td>
</tr>
<tr>
<td>cg, Inf, Rm</td>
<td></td>
<td>1.0</td>
<td>0.784</td>
<td>0.976</td>
<td>0.755</td>
<td>0.844</td>
</tr>
<tr>
<td>cg, Inf, Ear</td>
<td></td>
<td></td>
<td>1.0</td>
<td>0.823</td>
<td>0.877</td>
<td>0.967</td>
</tr>
<tr>
<td>cg, Inf, Ear, Rm</td>
<td></td>
<td></td>
<td></td>
<td>1.0</td>
<td>0.750</td>
<td>0.850</td>
</tr>
<tr>
<td>cg</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0</td>
<td>0.938</td>
</tr>
<tr>
<td>cg, PC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0</td>
</tr>
</tbody>
</table>

| Panel B. Correlation Between Volatilities |         |            |             |                  |     |       |
| cg, Inf          | 1.0     | 0.763      | 0.649       | 0.526            | 0.653 | 0.918 |
| cg, Inf, Rm      |         | 1.0        | 0.530       | 0.692            | 0.325 | 0.749 |
| cg, Inf, Ear     |         |            | 1.0         | 0.908            | 0.329 | 0.841 |
| cg, Inf, Ear, Rm |         |            |             | 1.0              | 0.169 | 0.703 |
| cg               |         |            |             |                  | 1.0  | 0.680 |
| cg, PC           |         |            |             |                  |      | 1.0   |

| Panel C. Correlation Between Skewness |         |            |             |                  |     |       |
| cg, Inf          | 1.0     | 0.836      | 0.794       | 0.605            | 0.903 | 0.230 |
| cg, Inf, Rm      |         | 1.0        | 0.638       | 0.851            | 0.710 | 0.394 |
| cg, Inf, Ear     |         |            | 1.0         | 0.718            | 0.592 | 0.443 |
| cg, Inf, Ear, Rm |         |            |             | 1.0              | 0.392 | 0.574 |
| cg               |         |            |             |                  | 1.0  | 0.059 |
| cg, PC           |         |            |             |                  |      | 1.0   |

The table reports the correlations between the time series of the conditional mean (Panel A), conditional volatility (Panel B), and conditional skewness (Panel C) of consumption growth for the six different choices of the conditioning considered.

states of the world in the subsequent period, for each possible value of the current state (current period), to determine the expected equity premium in that state. The results are presented in Figure 10, Panel A. The figure shows that the (annualized) expected equity premium is quite volatile, varying from 0.43% in 1965 : Q4 to 7.3% in 2009 : Q3. The strongly countercyclical nature of the expected equity premium is also evident from the figure. The correlation between the expected premium and a recession dummy is 42.5%. A regression of the realized equity premium on the expected premium produces a statistically significant slope coefficient (with a t-statistic of 2.56) and an $R^2$ of 2.5%.

Figure 10, Panel B presents the time series of the conditional volatility of the market return, obtained using the estimated SEL probabilities. The conditional volatility varies from 15.7% to 23.7% and, like the conditional mean, it is strongly countercyclical with the correlation with the recession dummy being equal to 45.1%.
Figure 10 – Time series of Conditional Moments of Market Return

**Panel A: Expected Equity Premium**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.01</td>
<td>0.04</td>
<td>0.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Panel B: Conditional Volatility of Market Return**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.16</td>
<td>0.20</td>
<td>0.24</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Panel C: Conditional Sharpe Ratio of Market Return**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.05</td>
<td>0.20</td>
<td>0.35</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** The figure plots the time series of the conditional mean (Panel A), conditional volatility (Panel B), and conditional Sharpe ratio (Panel C) of the market return. Shaded areas denote NBER designated recession periods. The conditional moments are obtained using the estimated SEL distributions. The pricing kernel corresponds to the Campbell-Cochrane external habit model (CC) and the test asset consist of the excess return on the market portfolio. The conditioning set includes an exponentially-weighted moving averages of lagged consumption growth. The sample is quarterly covering the period 1947:Q2-2013:Q4.

Finally, Figure 10, Panel C presents the time series of the conditional Sharpe ratio of the market return. The conditional Sharpe ratio is highly volatile, varying from 0.027 in 1965:Q4 to 0.37% in 2009:Q1 and strongly countercyclical with a correlation of 39.2% with the recession dummy.

The magnitude and the nature of time-variation in these moments of the market portfolio are in line with commonly held beliefs in the academic literature about the dynamics of these moments, and have represented a challenge for structural models to reproduce.

We compare the time series of the expected stock market returns implied by the price-consistent beliefs with survey data on investors’ expectations about future stock market returns. Specifically, we consider Robert Shiller’s investor survey, released by the Investor
Behavior Project at Yale University. The survey, conducted at six-month intervals prior to July 2001 and monthly thereafter, asks a sample of institutional investors and a sample of wealthy individual investors how much of a change in percentage terms do they expect in the Dow Jones Industrial Index in the coming year. The U.S. Institutional One-Year Confidence Index is the percentage of institutional investors giving a number strictly greater than zero, i.e. the percent of the sample expecting an increase in the Dow in the coming year. Similarly, the Individual One-Year Confidence Index is the percentage of individual investors giving a number strictly greater than zero (see Robert Shiller’s website for further details on the survey). The Institutional Index is available from October 1989 onwards, while the Individual Index is available over a shorter sample period beginning in April 1999.

Figure 11 – Expected Stock Market Return

Notes: The figure plots the time series of the U.S. Institutional One-Year Confidence Index (black line), along with the fitted time series from a regression of the Index on the price-consistent beliefs (red line). The price-consistent beliefs are obtained using the estimated SEL distributions. The pricing kernel corresponds to the Campbell-Cochrane external habit model (CC) and the test asset consist of the excess return on the market portfolio. The conditioning set includes an exponentially-weighted moving averages of lagged consumption growth. The sample is quarterly covering the period 1989:Q3-2013:Q3.

A regression of the U.S. Institutional One-Year Confidence Index on the price-consistent beliefs over 1989-2013 gives a positive and strongly statistically significant slope coefficient (with a t-statistic of 2.76) with $R^2$ of 9.6%. Figure 11 plots the time series of the U.S.
Institutional One-Year Confidence Index (black line), along with the fitted time series from a regression of the Index on the price-consistent beliefs (red line).

However, a regression of the U.S. Individual One-Year Confidence Index on the price-consistent beliefs over 1999-2013 gives a strongly statistically significant negative slope coefficient (with a t-statistic of $-2.79$) with $R^2$ of 12.8%. A closer look reveals that this is caused by the divergence between the two series during the recent financial crisis. Specifically, a regression of the Individual Index on the price-consistent beliefs over the period 1999 : Q3-2007 : Q3 that ends right before the onset of the financial crisis, produces a positive and highly significant slope coefficient (with a t-statistic of 4.10) and an impressive $R^2$ of 36.7%. However, the same regression performed over the sample period since the onset of the crisis (2007 : Q4-2013 : Q3) produces a statistically insignificant slope coefficient with a t-statistic of only 0.25 and an $R^2$ of only 0.27%. In fact, even the U.S. Institutional One-Year Confidence Index and the U.S. Individual One-Year Confidence Index diverge during and in the immediate aftermath of the crisis – the correlation between the two indices is 74.2% over the period 1999 : Q3-2007 : Q3, but becomes significantly negative at $-32.9\%$ over the period 2007 : Q4-2011 : Q3.

Overall, the evidence suggests that the price-consistent beliefs about the stock market, extracted using the SEL method, line up with institutional investors’ beliefs about the stock market captured in survey data. This lends further support to the information-theoretic methodology proposed in this paper to extract investors’ beliefs.

## VI Comparisons to Commonly Assumed Data Generating Processes

In the previous section, we identified, using the SEL method, the conditional distribution of macroeconomic variables like consumption growth that satisfy the conditional Euler equation restrictions for a chosen set of assets. The recovered distributions represent the beliefs of the average investor in the stock market. In this section, we compare these beliefs with a few standard time series models of the dynamics of macro variables extensively used in the literature. We identify how our SEL estimates deviate from those implied by the time series models and explore the dimensions along which the deviations are the largest. We would like to emphasize that the time series models that we compare our SEL-implied beliefs to by no means represent an exhaustive set of models considered in the literature. They are, however, among the most widely assumed dynamics.
VI.1 Commonly Assumed DGPs

Our first choice of the data generating process with which to compare the recovered beliefs is a standard ARMA (1,1) model for consumption growth:

$$\Delta c_{t+1} = (1 - \psi_1)g + \psi_1 \Delta c_t + \theta_1 \nu_{1,t} + \nu_{1,t+1},$$

where

$$\nu_{1,t+1} = i.i.d. N(0, \sigma^2)$$

The above specification is perhaps the one that is used most extensively in the macroeconomics and finance literatures. Wachter (2006) assumes this specification in an external habit model to explain the observed real and nominal term structures of interest rates. The ARMA (1,1) specification for realized consumption growth also naturally obtains in the long run risks literature (see, e.g., Bansal and Yaron (2004)), that hypothesizes the presence of a small predictable component in expected consumption growth. More recently, an ARMA (1,1) specification for consumption growth has been shown to emerge in a model with robust control preferences, where the statistical difficulty in distinguishing between alternative ARMA (1,1) models (e.g., with differing levels of persistence) with a finite available data history leads economic agents with such preferences to act from the perspective of the worst case model, i.e., an ARMA (1,1) specification with higher persistence of consumption growth than what is estimated using historical data on consumption growth alone (see, e.g., Szoke (2017) and Bidder and Dew-Becker (2016)).

We estimate the model through (quasi) maximum likelihood, using historical data on consumption growth alone. Figure 12 plots the historical time series of consumption growth (black line) along with the ARMA (1,1) model-implied conditional mean (red line). The figure show that the specification offers a good fit for the observed dynamics of consumption growth. A regression of the realized consumption growth on its model-implied conditional mean produces an impressive $R^2$ of 78.1%. The good fit offered by the ARMA (1,1) specification explains its wide popularity in the literature and motivates this model as one of our choices with which to compare the SEL-estimated beliefs about consumption growth. Note that the ARMA (1,1) specification implies that the conditional distribution of consumption growth is homoscedastic and Gaussian.

Our second choice of the data generating process is the widely used regime-switching model, where the mean of consumption growth differs across latent regimes:

$$\Delta c_{t+1} = \mu_{c,s_{t+1}} + \sigma c \epsilon_{c,t+1}$$

(20)
Notes: The figure plots the historical time series (black line) along with its model-implied conditional mean (red line), of consumption growth over the period 1947:Q1-2013:Q4. The model parameters are estimated via quasi maximum likelihood.

where $\epsilon_{c,t} = i.i.d.N(0, 1)$, and $s_t$ is the scalar state variable that denotes the latent economic regime. Regime switching models have been extensively used in the macroeconomics and asset pricing literatures (see, e.g., David and Veronesi (2013), Ghosh and Constantinides (2017)) and offer a flexible approach to modeling the underlying dynamics of macro and financial variables. Moreover, unlike the ARMA (1,1) specification above that implies that the conditional distribution of consumption growth is lognormal, the regime-switching model generates fat tails in the conditional distribution because of the regimes being latent.

We set the number of regimes to equal four, because the fit of the model is the best for this choice of the number of regimes. We estimate the model through standard maximum likelihood filtering techniques. Figure 13 plots the historical time series of consumption growth (black line) along with its model-implied conditional mean (red line). The figure shows that the specification offers a good fit for the observed dynamics of consumption growth. A regression of the realized consumption growth on its model-implied conditional mean produces an $R^2$ of 10.0%.
VI.2 Recovered Beliefs versus Commonly Assumed Dynamics

In this sub-section, we compare the beliefs about consumption growth recovered using the SEL approach to the two commonly assumed time series models for consumption growth described in Section VI.1.

Consider first the ARMA (1,1) model for consumption growth. This specification is perhaps the most widely used in the macro and finance literatures and, as shown in Section VI.1, offers a good fit to the observed dynamics of consumption growth. Figure 14 plots the time series of the conditional mean (Panel A), volatility (Panel B), and skewness (Panel C) of consumption growth implied by the SEL beliefs (black line) and the ARMA(1,1) specification (red line).

Panel A shows that the ARMA(1,1) model for the consumption growth dynamics, with the parameters estimated using the consumption history alone (i.e., without using any asset price data), implies much more volatile and less persistent beliefs about future consumption growth compared to the beliefs extracted using the SEL approach that are consistent with observed asset prices. Specifically, the conditional mean of consumption growth implied by the ARMA(1,1) model has an annualized volatility of 0.38%, more than three times the
volatility of 0.10% of the conditional mean implied by the SEL beliefs. The former, on the other hand, has a lower persistence, measured by a first-order autocorrelation coefficient of 0.75 compared to 0.88 for the latter. This is consistent with the findings in the literature that if an ARMA model is hypothesized as the true underlying data generating process, then a higher persistence parameter for the ARMA process is typically needed for the model to have success at explaining asset prices compared to the persistence that would be estimated using historical consumption data alone.

Figure 14 – Comparison of ARMA(1,1) and SEL

Notes: The figure plots the time series of the conditional mean (Panel A), conditional volatility (Panel B), and conditional coefficient of skewness (Panel C) of consumption growth obtained using the estimated SEL distributions (black line) and from the ARMA (1,1) specification (red line). Shaded areas denote NBER designated recession periods. The conditional moments are obtained using the estimated SEL distributions. The pricing kernel corresponds to the time and state separable power utility model with a constant CRRA. The test asset consists of the excess return on the market portfolio. The conditioning set includes an exponentially-weighted moving average of lagged consumption growth. The sample is quarterly covering the period 1947:Q2-2013:Q4.

Panel B suggests that the conditional homoscedasticity assumption of the ARMA(1,1) model is close to that implied by the recovered beliefs. Finally, Panel C suggests that the conditional Gaussianity assumption of the ARMA(1,1) model is starkly contrary to that
implied by the recovered beliefs. Specifically, as shown in Section V, investors perceive the conditional skewness of consumption growth to be strongly cyclical—a feature missing from the conditionally Gaussian ARMA(1,1) model.

We next consider the regime-switching model. Note that, unlike the ARMA(1,1) model, the regime-switching specification allows for conditional heteroscedasticity as well as non-Gaussianity in the conditional distribution (emanating from the latent regimes). Figure 15 plots the time series of the conditional mean (Panel A), volatility (Panel B), and skewness (Panel C) of consumption growth implied by the SEL beliefs (black line) and the regime-switching model (red line).

**Figure 15 – Comparison of Regime-Switching model and SEL**

*Notes:* The figure plots the time series of the conditional mean (Panel A), conditional volatility (Panel B), and conditional coefficient of skewness (Panel C) of consumption growth obtained using the estimated SEL distributions (black line) and from the regime-switching model (red line). Shaded areas denote NBER designated recession periods. The conditional moments are obtained using the estimated SEL distributions. The pricing kernel corresponds to the time and state separable power utility model with a constant CRRA. The test asset consists of the excess return on the market portfolio. The conditioning set includes an exponentially-weighted moving average of lagged consumption growth. The sample is quarterly covering the period 1947:Q1-2013:Q4.

As with the ARMA(1,1) model, the regime-switching model implies higher volatility and
lower persistence of expected consumption growth compared to those obtained with the SEL-
recovered beliefs. In particular, the expected consumption growth implied by the regime-
switching model has an annualized volatility of 0.35%, very similar to the 0.38% implied
by the ARMA(1,1) model and more than three times higher than the volatility of 0.10%
implied by the SEL beliefs. On the other hand, the first-order autocorrelation coefficient of
the expected consumption growth implied by the regime-switching model is 0.83 compared
to 0.88 for the SEL beliefs.

The regime-switching model generates mild countercyclicaliy in the conditional volatility
of consumption growth – the annualized volatility varies from 0.88% to 1.2% over the sample
period (compared to varying from 0.9% to 1.1% for the SEL beliefs) with a correlation
coefficient of 40.2% with a recession dummy. As with the conditional mean, the regime-
switching model implies a more volatile and less persistent conditional volatility process
compared to the SEL beliefs recovered from asset prices.

Finally, the regime-switching model perhaps differs the most from the SEL beliefs in
terms of its implications for the skewness of consumption growth. The differences are many-
fold. First, the average magnitude of the skewness is lower in the former compared to the
latter: the average skewness is −0.10 for the regime-switching model compared to −0.36
for the SEL beliefs. Thus, beliefs consistent with asset prices are more negatively skewed.
Second, the negative skewness is much more persistent in the SEL beliefs compared to that in
the regime-switching model – the first-order autocorrelation coefficient is 0.88 for the former
compared to only 0.50 with the latter. Third, for the SEL beliefs, the skewness becomes
more negative during good times (with investors being concerned about severe economic
downturns even during good times) whereas the opposite is true for the regime-switching
model. The correlation between a recession dummy and the skewness computed using the
SEL beliefs is 44.0%, whereas for the skewness implied by the regime-switching model, the
correlation is −25.9%.

Overall, our results suggest that dynamic time series models for consumption growth
that are commonly assumed in the literature are different in several ways from investors’
beliefs about consumption growth extracted from observed asset prices. First, the expected
consumption growth implied by the latter is much more persistent than the former time series
models would imply. Second, the investors’ beliefs suggest a more fat left-tailed distribution
for future consumption growth, i.e. the conditional skewness is much more negative, during
both good and bad times, compared to that implied by the commonly assumed models.
Third, the skewness implied by investors’ beliefs is strongly cyclical, a feature that is, once
again, missed by commonly assumed models. These results offer modelling guidelines for
macro finance models in order to improve their ability to match the observed dynamics of
VII Are Investors’ Rational?

In Section V, we recovered investors’ beliefs about future macroeconomic outcomes that are consistent with observed asset prices, i.e. the beliefs satisfy the conditional Euler equation restrictions for a chosen set of assets. In Section VI, we identified the major dimensions along which the recovered beliefs differ from commonly assumed time series models for these macro variables. As a next step, the question naturally arises regarding whether investors’ beliefs are rational or whether they are distorted relative to rationality. The rational expectations hypothesis and behavioral finance constitute the two central paradigms in financial economics and remain perhaps one of the most actively debated topic in the discipline. Therefore, a formal data-driven approach to identifying deviations (if any) from rationality represents an important advance in this debate.

A full investigation of this topic is beyond the scope of this paper. However, in this section, we present some preliminary evidence suggesting economically significant distortions in investors’ beliefs relative to a judiciously chosen benchmark. Note that, In Section VI, we showed the discrepancies between the recovered investors’ beliefs about consumption growth and commonly assumed time series models for consumption growth. However, the discrepancy with respect to a particular time series model cannot directly be interpreted as measuring deviations from rationality, without the added assumption of that time series model being the true (or rational) model. Since a multitude of time series models have been assumed in the literature, this immediately raises the question as to which among these models should be regarded as the rational benchmark relative to which beliefs distortions can be measured.

Fortunately, the SEL approach used in this paper offers one possible way to address the above issue. Section VII.1 presents an alternative, information-theoretic interpretation of the SEL estimator. This property makes the SEL objective function measure deviations from, as we will see, a judiciously chosen benchmark. The benchmark is non-parametric, not requiring any parametric functional-form assumptions on the form of the conditional distribution of the variables of interest. It, therefore, offers a more robust approach to constructing a benchmark relative to which beliefs distortions can be measured. It may, therefore, also be argued that this benchmark is an attractive candidate for a rational benchmark and that any deviations from it can be interpreted as deviations from rationality. In Section VII.2, we present and characterize the distortions of the recovered investors’ beliefs from the non-parametric benchmark.
VII.1 An Alternative Interpretation of the SEL estimator

In Section II.2, we described how the SEL estimator is akin to a non-parametric maximum likelihood estimator. In this section, we show that the SEL estimator also has an important information-theoretic interpretation (see, e.g., Kitamura and Stutzer (1997)).

To see this, let \( P_t \) be the set of all conditional probability measures defined on \( \mathbb{R}^{q+k} \), where \( q \) denotes the dimension of the (or, the number of) variables \((C_{t+1}/C_t, Y_{t+1}, X_{t+1})\) entering the SDF and the conditioning set, and \( k \) denotes the dimension of the cross section of assets used in the estimation. For any set of admissible SDF parameters \( \theta \in \Theta \), we define the set of conditional probability measures, absolutely continuous with respect to the (true) underlying objective measure \( P_{\text{obj}}^t \), that satisfy the conditional Euler equations:

\[
P_t(\theta) := \left\{ \tilde{P}_t \in P_t : \mathbb{E}^{\tilde{P}_t} \left[ M \left( \frac{C_{t+1}}{C_t}, Y_{t+1}; \theta \right) R_{t+1}^e | X_t \right] = 0 \right\}, \forall t \in \{1, \ldots, T\}.
\]

(21)

Therefore, \( P_t(\theta) \) is the set of all the conditional probability measures that are consistent with the asset pricing model characterized by the conditional Euler equation restrictions.

The SEL estimation can then be shown to select a probability measure \( \hat{P}_t(\theta) \), for each \( t \), such that:

\[
\hat{P}_t(\theta) = \inf_{\tilde{P}_t \in P_t} \text{KLIC}(P_{\text{obj}}^t, \tilde{P}_t) \equiv \inf_{\tilde{P}_t \in P_t} \int \log \left( \frac{d P_{\text{obj}}^t}{d \tilde{P}_t} \right) d P_{\text{obj}}^t
\]

s.t. \( \mathbb{E}^{\tilde{P}_t} \left[ M \left( \frac{C_{t+1}}{C_t}, Y_{t+1}; \theta \right) R_{t+1}^e | X_t \right] = 0 \),

(22)

where KLIC\((P_{\text{obj}}^t, \tilde{P}_t)\) is the Kullback-Leibler Information Criterion (KLIC) divergence (or, relative entropy) between the two measures \( P_{\text{obj}}^t \) and \( \tilde{P}_t \) (see White (1982)).

Note that the KLIC divergence is non-negative and is exactly equal to zero if and only if \( \hat{P}_t(\theta) = P_{\text{obj}}^t \) almost surely, that is, if the investors' beliefs consistent with asset prices (beliefs that satisfy the conditional Euler restrictions) coincide with the objective beliefs. Thus, the SEL approach searches for an estimate of \( \hat{P}_t \) that makes the estimated beliefs as close as possible – in the information-theoretic sense – to the objective one \( P_{\text{obj}}^t \), while also requiring that the estimated beliefs satisfy the pricing restrictions given by the conditional Euler equations.\(^\text{13}\)

\(^\text{13}\)Most papers using empirical likelihood methods are usually interested in efficient parameter estimation, contrary to our main goal here. Significant contributions in this direction include, among others, Kitamura (2007), Altissimo and Mele (2009), Gagliardini, Gouriéroux, and Renault (2011), Gospodinov and Otsu (2012), and Crudu and Sandor (2017). Alternatives have also been proposed for identifying parameters based on conditional moment conditions by, for instance, Imbens, Donald, and Newey (2003), Carrasco and Florens (2000) or Domínguez and Lobato (2004). Instead of using KLIC, we could have used any of the broader class
An inspection of Equations (4) and (22) that define the SEL estimator as maximizing the smoothed non-parametric log-likelihood of the data and as minimizing the distance between the estimated measure and the underlying objective measure, respectively, show that the two have identical solutions when the objective measure is taken to be $P_{obj}^t = \omega_{t,j}$, $j = 1, 2, ..., T$, for each $t = 1, 2, ..., T$. This is a natural choice for the objective measure because it maximizes the log-likelihood of the data in Equation (4), but without imposing the constraint that the estimated probabilities satisfy the conditional Euler equation restrictions. In other words, the maximum likelihood estimate of the conditional probability measure in the absence of any asset pricing restrictions simply equals the kernel density weights used to smooth the likelihood function. Kernel density estimators are widely used to approximate conditional distributions of variables of interest and have the attractive feature of not requiring any functional-form assumptions on the form of the distributions. The objective measure, being a kernel density estimator, inherits this attractive property. Imposition of the pricing restrictions distorts the probabilities relative to the objective measure and may be viewed as distortions in investors’ beliefs relative to the objective benchmark that are necessary to satisfy the pricing restrictions. The SEL procedure searches for a probability measure, $\hat{P}_t(\theta)$, that satisfies the Euler restrictions, while deviating as little as possible from the objective measure.

Since, by construction, the measures $P_{obj}^t$ and $\hat{P}_t(\theta)$ are absolutely continuous, we can write:

$$\frac{d \hat{P}_t(\theta)}{d P_{obj}^t} = Z_t,$$

where $Z_t$ is the Radon-Nikodym derivative of $\hat{P}_t(\theta)$ with respect to $P_{obj}^t$. In the absence of any beliefs distortions relative to the objective benchmark, we have $\hat{P}_t(\theta) = P_{obj}^t$ at each point in time, and $Z_t = 1$ almost surely.

The SEL procedure, thus, enables the characterization of the deviations (if any) from $P_{obj}^t$ to $\hat{P}_t(\theta)$. Moreover, it does so without the need for any parametric distributional assumptions on either the objective measure, $P_{obj}^t$, or the nature of the beliefs distortions relative to the objective measure, $Z_t$. Thus, the methodology can be used to shed light on whether investor’s beliefs deviate or not from the objective measure, and, in the latter case, the precise nature of

the deviations. If the objective beliefs can be argued to be rational beliefs, then, in the case where the beliefs recovered from prices using the SEL approach coincide with the objective beliefs, it can be argued that investors are rational. On the other hand, if the optimal KLIC distance of Equation (22) is positive, then the beliefs can be argued to exhibit deviations from rationality and the precise nature of the deviations can be characterized by examining the behavior of the Radon-Nikodym derivative in Equation (23).

VII.2 Empirical Results

We start by comparing the pricing performance of the objective measure $P_{t,j}^{\text{obj}} = \omega_{t,j}, j = 1, 2, ..., T$ with that of the measure estimated using the SEL approach, $\hat{P}_{t,j}, j = 1, 2, ..., T$, that satisfies the conditional Euler equation constraints (hereafter referred to as the price-consistent measure). Figure 16 plots the time series of the conditional pricing errors for the excess stock market return, under each of the two measures. The objective measure produces large and highly volatile conditional pricing errors, varying from $-2.9\%$ to $12.1\%$. The pricing errors are larger during bad states of the world, characterized as NBER recession periods – the correlation between the pricing errors and a NBER-recession dummy variable is $47.5\%$. The SEL-implied probabilities $\hat{\pi}_{i,j}, i, j = 1, 2, ..., T$, on the other hand, produce conditional pricing errors that are identically equal to zero in all periods.

Figure 16 suggests that the distortions of the price-consistent beliefs from the objective beliefs are more severe during recessionary episodes. In other words, investors’ beliefs need to be distorted more relative to the objective benchmark during bad times in order for the former to be consistent with observed asset prices. In order to formalize and quantify the time-variation in the beliefs distortions, Figure 17 plots the time series of the KLIC divergence between the probability measures $P_t^{\text{obj}}$ and $\hat{P}_t(\theta)$, as given by the objective function in Equation (22). The strongly countercyclical nature of the divergence is evident – the KLIC varies from 0.005 to 0.040, with a correlation of 48.0% with recessions.

In order to identify the precise nature of the beliefs distortions, Figure 18 plots the time series of the mean, volatility, and skewness of consumption growth under the objective measure and under the price-consistent measure, as well as the percentage changes in these moments from the former measure to the latter.

Consider first Panel A that presents the time series of the conditional mean of consumption growth under the two measures. The figure shows that the price-consistent probability weights imply a lower expected consumption growth compared to that implied by the objective measure, particularly so in exceptionally bad states of the world. For instance, at the height of the recent financial crisis in 2009:Q2, the expected consumption growth under the objective measure was 1.0% annualized whereas that under the price-consistent mea-
Notes: The figure plots the time series of the conditional pricing errors for the excess return on the market portfolio. The conditional expectation underlying the calculation of the pricing error is evaluated using the SEL probabilities (black line) or using the kernel weights (red line). The pricing kernel is that implied by power utility preferences with a CRRA, the excess return on the market is used as the sole asset, and the conditioning set consists of an exponentially-weighted moving average of past consumption growth. The sample is quarterly covering the period 1947:Q2-2013:Q4.

Sure was almost 20% lower at 0.8%. Over the 4 quarters in 2009, the average quarterly consumption growth under the objective measure was 1.2% annualized – 15% higher than the corresponding average growth rate of 0.97% under the price-consistent measure. Much smaller deviations are observed during good times. For instance, during the expansionary episode of the sixties (1961:Q2–1969:Q3), the average expected consumption growth under the objective measure and price-consistent measure were almost identical at 2.1% and 2.0%, respectively. Similarly, the average expected consumption growth under the objective and price-consistent measures were also almost identical at 2.0% and 1.9%, respectively, during the expansionary episode of the nineties (1991:Q2–2000:Q4). Formally, Panel B, that presents the percentage change in the mean growth rate when moving from the objective to the price-consistent measure, shows that the magnitude of the distortion is greater during recessionary episodes. The correlation between an NBER-recession dummy and the absolute difference between the expected consumption growth under the two measures is 48.0%, and that between the recession dummy and the percentage difference between the growth rates
Notes: The figure plots the time series of the KLIC divergence between the SEL-implied probability measure and the objective measure given by the kernel weights. The former is obtained using as inputs the SDF implied by power utility preferences with a CRRA, the excess return on the market is used as the sole asset, and the conditioning set consists of an exponentially-weighted moving average of past consumption growth. The sample is quarterly covering the period 1947:Q1-2013:Q4.

under the two measures is 44.9%. This is indicative of investor pessimism during particularly bad states of the world.

Consider next Panel C, that plots the time series of the conditional volatility of consumption growth under the two measures. The figure clearly shows that the distortions in the conditional volatility are much smaller than those in the mean. This is further highlighted in Panel D that presents the percentage changes in the volatility between the two measures – the distortions are close to zero at all time periods with a maximum of only 1.4%.

Finally, in Panels E and F, we plot the time series of the conditional skewness of consumption growth under the two measures and the percentage change between them, respectively. The figures show large discrepancies in the skewness between the two measures. In fact, the magnitudes of the discrepancy in the skewness when moving from the objective to the price-consistent measure are far greater, in both absolute terms as well as in percentage terms, compared to the distortions in the volatility and the mean. Specifically, we find that the skewness of consumption growth is less negative under the price-consistent measure relative to the objective measure and that this feature holds during both good and bad times. For in-
Notes: The figure plots the time series of the KLIC divergence between the SEL-implied probability measure and the objective measure given by the kernel weights. The former is obtained using as inputs the SDF implied by power utility preferences with a CRRA, the excess return on the market is used as the sole asset, and the conditioning set consists of an exponentially-weighted moving average of past consumption growth. The sample is quarterly covering the period 1947:Q1-2013:Q4.

stance, during the expansionary episode of the sixties, the average skewness of consumption growth under the objective measure was $-0.5$, while that under the price-consistent measure is $16.5\%$ larger at $-0.4$. Similarly, the average skewness of consumption growth under the objective and price-consistent measures were $-0.47$ and $-0.37$, respectively, (corresponding to a $20.8\%$ difference) during the expansionary episode of the nineties (1991:Q2–2000:Q4). During the recent financial crisis spanning 2008:Q1–2009:Q4, the skewness is again more negative under the objective measure compared to the price-consistent measure ($-0.14$ versus $-0.04$, corresponding to a $70.8\%$ difference). This is indicative of investor optimism or
exuberance, that is observed in good as well as bad states of the world.

Interestingly, we do not find much evidence of increased persistence under the price-consistent beliefs relative to the objective benchmark – the persistence of the conditional mean is 0.88 under both the objective and price-consistent measures; the persistence of the conditional volatility is 0.87 under the former compared to 0.88 under the latter; and the persistence of the conditional skewness is 0.88 under both measures.

To summarize, investors’ beliefs about possible macroeconomic outcomes seem distorted relative to the objective benchmark in two important respects. First, the expected growth rate is lower under the price-consistent beliefs compared to the objective measure, during particularly bad states of the world. During particularly good times, on the other hand, the magnitudes of the distortions are typically less than 5.0%. Second, the conditional skewness is less negative under the price-consistent beliefs compared to the objective measure, both during good as well as bad times. Moreover, the magnitude of the distortions in the skewness far exceed those in the mean. The conditional volatility, on the other hand, is almost identical between the two measures. To the extent that the objective measure may be regarded as a measure of rational beliefs, the deviations of the price-consistent beliefs from the objective ones may be viewed as distortions relative to rationality. Under this interpretation, our results are indicative of investor exuberance during good times – stemming from a perceived truncation of the left tail of the distribution of consumption growth relative to the rational measure, while distortions in the expected mean and volatility of consumption growth are negligible. During bad times, our results are more mixed, suggesting evidence of pessimism via underestimation of the expected growth rate on the one hand, and optimism on the other hand via less negative skewness in the distribution of consumption growth relative to the rational measure.

VIII Conclusion and Extensions

Current asset prices reflect investors’ beliefs about future economic and financial outcomes. Relying on this insight, we propose an information-theoretic methodology to recover investors’ beliefs from observed asset prices. Our approach is non-parametric, not requiring any functional-form assumptions about the beliefs as reflected in the dynamics of the variables of interest, or assumptions regarding investor rationality or lack thereof.

Our methodology relies on the smoothed empirical likelihood (SEL) estimator developed by Kitamura, Tripathi, and Ahn (2004), that estimates the conditional density of macroeconomic and financial variables by maximizing the non-parametric (multinomial) log-likelihood of the data, subject to the constraint that the density so estimated satisfies the conditional
Euler equation restrictions for the set of test assets. The inputs required for the approach include a pricing kernel that represents the investors’ preference over risky outcomes, a cross section of assets that the kernel is required to price, and a conditioning set underlying the conditional Euler equations that investors’ use to form their beliefs.

The recovered beliefs suggest that the expected growth rate of macro variables such as the aggregate consumption growth rate is strongly procyclical, while the conditional volatility is mostly flat over the business cycle. The beliefs also exhibit strong non-Gaussian features, with the conditional skewness being almost always negative and becoming more negative during good times. The latter feature of beliefs is suggestive of investors fearing severe economic downturns even during particularly good states of the world characterized by high expected real growth rates. Bad states, on the other hand, correspond to low expected growth rates but the extreme left and right tails do not expand further relative to good states, causing the magnitude of the negative skewness to reduce during bad times. We show that these findings are robust to alternative choices of the pricing kernel, the set of test assets, and the conditioning set.

Finally, we apply our methodology to shed light on whether or not the recovered investors’ beliefs are rational. We show that the SEL estimator has an alternative information-theoretic interpretation – the recovered conditional distribution (that represents the investors’ beliefs) is the one that is minimally distorted relative to the objective measure, so as to satisfy the pricing restrictions given by the conditional Euler equations for the test assets. The objective measure, in this case, corresponds to a non-parametric kernel density estimator, not requiring any functional-form assumptions about the form of the distribution of the variables of interest. It, therefore, may constitute an attractive choice for the rational measure. Our results suggest that investors’ beliefs about possible macroeconomic outcomes seem distorted relative to the objective benchmark in two important respects. First, the conditional mean is lower under the former compared to the latter, particularly during bad times. This is indicative of pessimistic behavior, with the severity of the pessimism increasing during bad times. Second, the conditional skewness is less negative under the former compared to the objective measure. This is a consequence of the mean of the distribution shifting to the left under the investors’ price-consistent measure, while the extreme left and right tails coinciding with the objective measure.

The current paper focuses on characterizing investors’ beliefs about the aggregate consumption growth rate and the aggregate stock market portfolio. However, the methodology is considerably general and may be used to identify beliefs about the joint dynamics of various macro variables (e.g., co-movement of real growth and inflation) or about different asset classes. These are left to future research.
Also, the present paper assumed the existence of a representative investor. Important extensions of the work include allowing for heterogeneous investors with heterogeneous beliefs. These are outside the scope of the current paper but present interesting directions for future research.
References


A Appendix

A.1 Owen normalization

In order to avoid numerical issues associated with the estimation of the Lagrange multipliers, Owen (2001) proposes the following transformation.

\[
\hat{\lambda}_i^o(\theta) = \arg\max_{\lambda_i \in \mathbb{R}^k} \sum_{j=1}^T \omega_{i,j} \cdot \Psi_\nu \left[ 1 + M(g_j, y_j; \theta, \lambda'_i r^e_j) \right]
\]

where

\[
\Psi_\nu(x) = \begin{cases} 
\log(x) & \text{if } x > \nu \\
\log(\nu) - \frac{3}{2} + 2 \frac{x}{\nu} - \frac{1}{2} \left( \frac{x}{\nu} \right)^2 & \text{if } x \leq \nu
\end{cases}
\]

Equation (25) 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 above transformation of the objective function can make the sum of the estimated probabilities with \( \hat{\lambda}_i^o \) (see Equation (6)) deviate from unity. Again, Owen (2001) suggests to normalize the probabilities ex-post so that they add up to one:

\[
\hat{p}_{i,j}^{SEL}(\theta) = \frac{\omega_{i,j}}{1 + M(g_j, y_j; \theta, \lambda'_i r^e_j)} \times \left( \sum_{j=1}^T \frac{\omega_{i,j}}{1 + M(g_j, y_j; \theta, \lambda'_i r^e_j)} \right)^{-1},
\]

A.2 Different Stochastic Discount Factors

In this Appendix, we present the alternative choices for the SDF that we consider when recovering investors’ beliefs using the SEL approach.

The first choice corresponds to the external habit formation preferences (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

\[
M_t = \delta(C_t/C_{t-1})^{-\gamma} (S_t/S_{t-1})^{-\gamma},
\]

where \( \delta \) is the subjective time discount factor, \( \gamma \) is a utility curvature parameter that provides a lower bound on the time varying CRRA, \( S_t = \frac{C_t - X_t}{C_t} \) denotes the surplus consumption ratio, and \( X_t \) is the habit level.

Note that the SDF depends on the surplus consumption ratio, \( S_t \), 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:
\[
\Delta \log(C_t) = g + \nu_t, \quad \nu_t \sim i.i.d. N(0, \sigma^2).
\]
The log surplus consumption ratio evolves as a heteroskedastic \( AR(1) \) process:
\[
\log(S_t) = (1 - \phi) \log(S) + \phi \log(S_{t-1}) + \lambda (\log(S_{t-1})) \nu_t, \quad (28)
\]
where \( \log(S) \) is the steady state log surplus consumption ratio and
\[
\lambda (\log(S_t)) = \begin{cases} 
\frac{1}{S} \sqrt{1 - 2 (\log(S_t) - \log(S))} - 1, & \text{if } \log(S_t) \leq s_{\text{max}} \\
0, & \text{if } \log(S_t) > s_{\text{max}}
\end{cases}
\]
\[
s_{\text{max}} = \log(S) + \frac{1}{2} \left( 1 - S^2 \right), \quad S = \sigma \sqrt{\frac{\gamma}{1 - \phi}}.
\]

We use the calibrated values of the model’s preference parameters \( (\delta, \phi, \gamma) \), the sample mean \( (g) \) and volatility \( (\sigma) \) of the consumption growth process, and the innovations in real consumption growth, \( \hat{\nu}_t = \Delta c_t - g \), to extract the implied time series of the surplus consumption ratio using Equation (28). This renders the SDF fully observable and, therefore, our SEL approach can be applied to recover the investor’s beliefs.

Our second 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 recursive preferences (see, e.g., Epstein and Zin (1989) and Weil (1989)), for which the SDF is given by
\[
M_{t+1} = \delta^\theta \left( \frac{C_{t+1}}{C_t} \right)^{-\frac{\theta}{\rho}} R_{c,t+1}^{\theta-1},
\]
where \( R_{c,t+1} \) is the unobservable gross return on an asset that delivers aggregate consumption as its dividend each period, \( \delta \) is the subjective time discount factor, \( \rho \) is the elasticity of intertemporal substitution, \( \theta := \frac{1-\gamma}{1-1/\rho} \), and \( \gamma \) is the relative risk aversion coefficient.

The aggregate consumption and dividend growth rates, \( \Delta c_{t+1} \) and \( \Delta d_{t+1} \), respectively, are modeled as containing a small persistent expected growth rate component, \( x_t \), that follows a heteroskedastic AR(1) process, and fluctuating variance, \( \sigma_t^2 \), that evolves according to a homoscedastic AR(1) process.

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

\[
\ln M_{t+1} = c_1 + c_2 \Delta c_{t+1} + c_3 x_{t+1} + c_4 \sigma_{t+1}^2 + c_5 x_t + c_6 \sigma_t^2
\]  

(29)

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, \(\sigma_t\), are not directly observable. Using the calibrated parameter values from Bansal and Yaron (2004), we extract the state variables, \(x_t\) and \(\sigma_t^2\), from observed consumption data, using a Bayesian smoother. As with the external habit model, the SDF, therefore, becomes fully observable rendering it amenable to SEL estimation of the investor’s beliefs.