# Identifying Beliefs from Asset Prices[*]

Anisha Ghosh[†]        Guillaume Roussellet[‡]

June 3, 2019

## Abstract

This paper proposes a novel information-theoretic procedure to identify investors' beliefs about future macroeconomic and financial outcomes from observed asset prices. Our approach recovers *price-consistent beliefs*, i.e. the conditional distribution of macro and financial variables that satisfy the conditional Euler equations, given a cross-section of assets, a pricing kernel, and a conditioning set. Our procedure is non-parametric, not requiring any functional-form assumption about the dynamics of the variables, or regarding investor rationality or lack thereof. The price-consistent beliefs show strong cyclicality in the conditional mean and skewness of aggregate consumption growth, while the conditional volatility is mostly flat over the business cycle. This contrasts with the widely assumed conditionally normal dynamics in the existing literature. Looking at stock market returns, we observe that the price-consistent beliefs contain similar information as survey data on institutional investors' expectations. A comparison of these price-consistent beliefs with a non-parametric objective benchmark suggests large beliefs distortions. Investors underestimate the expected consumption growth mostly during recessions, but consistently overestimate the skewness of the consumption growth rate to a much larger extent.

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

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

[†]Desautels Faculty of Management, McGill University; anisha.ghosh@mcgill.ca, *Corresponding author*.

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

# I  Introduction

The conditional Euler equation is the cornerstone equilibrium relationship in asset pricing theory. It stipulates that the price of any traded security equals the expected discounted value of its future cashflows. Asset prices thus reflect investors' beliefs about future economic and financial outcomes. Understanding how these beliefs are formed and how they evolve over time is crucial in explaining and predicting the behavior of asset prices. Existing asset pricing models usually either assume that economic agents have rational expectations and use all available data objectively to forecast the future (Muth (1961)), or that they possess behavioral biases distorting their beliefs relative to the true data generating process of endogenous variables (see Barberis and Thaler (2001) for a survey of behavioral finance).

Whichever assumption is used, these models prove limited and have difficulties explaining the time series behavior of the aggregate stock market returns, of the cross-section of financial asset returns, and individual trading behavior. Indeed, both types of models need structural assumptions about either the true data generating process or the way investors beliefs depart from it. These frameworks are rather driven by computational simplicity than being empirically grounded, and are both prone to potentially large misspecification errors. Misspecification is hard to detect in theory since the true beliefs of investors are unknown. It however translates into the models' inability to fit the time series of asset returns, producing large conditional pricing errors.

This paper proposes a non-parametric approach to identify investors' beliefs from observed asset prices, which bypasses the need for any functional-form assumption about the dynamics of the data generating process or 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. These recovered beliefs are *price-consistent* because they are constructed such that each test asset at each point in time satisfies the conditional Euler equation restrictions.

Our empirical methodology borrows heavily from the non-parametric smoothed empirical likelihood (hereafter referred to as SEL) estimator developed by Kitamura, Tripathi, and Ahn (2004). This method approximates the conditional density of macroeconomic and financial variables at each point in time with a multinomial distribution whose support is given by the available data sample. It then estimates the multinomial probabilities to maximize the likelihood, enforcing the constraint that the test assets are perfectly priced, i.e. the conditional Euler equations are satisfied.

Our methodology requires three inputs: a pricing kernel specification, a set of test assets,

and a relevant information set for the representative investor. 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 excess return on the market portfolio is the sole test asset, and the conditioning set consists of past consumption growth.

The price-consistent beliefs convey four salient features. First, the conditional distribution of the quarterly consumption growth rate is strongly non-Gaussian. This non-normality cannot be explained by time-varying volatility, which is perceived to be largely flat over the business cycle. Instead, consumption growth possesses a very persistent and cyclical component in both its conditional mean and skewness. The former shows pro-cyclical movements 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 skewness, on the other hand, is negative in almost all time periods. Perhaps surprisingly, it is pro-cyclical, becoming less negative during recessionary episodes. These results call into question the widely used assumption of a conditionally Gaussian data generating process in structural asset pricing models.

Second, we show that the price-consistent beliefs are robust to alternative pricing kernels, cross-sections of test assets, and conditioning sets. 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) give remarkably similar results. The results are essentially unchanged upon the inclusion of the excess returns on Small, Big, Growth and Value portfolios. Finally, the estimated 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.

Third, we provide evidence that commonly assumed consumption growth DGPs cannot reproduce the empirical features captured by the price-consistent beliefs. We consider two time-series specifications: *(i)* 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. For both models, the implied expected consumption growth series miss the high persistence implicit in the price-consistent beliefs. The price-consistent beliefs also suggest a fatter left tail in the distribution of future consumption growth, i.e. the conditional skewness is much more negative compared to those implied by the two time series models throughout the sample. Finally, the strong cyclicality of the price-consistent skewness is missed by the commonly assumed time series specifications. Our work thus provides an empirical assessment of what these assumed data generating processes are missing to explain the behavior of asset prices.

Last and importantly, we present evidence suggesting that our recovered beliefs reflect departures from rationality for the representative investor. We compare the price-consistent beliefs with a simple non-parametric estimator of the conditional density of consumption growth, obtained without using any asset pricing restrictions (hereafter referred to as the 'objective' beliefs). The price-consistent beliefs can be interpreted as the ones that are *minimally* distorted relative to the *objective* distribution while satisfying the pricing restrictions.[1] Our estimates show that even relative to this agnostic choice of the objective DGP, price-consistent consumption dynamics are substantially distorted. Our results are indicative of investor exuberance during good times – stemming from an underweighting of the left tail of the distribution of consumption growth relative to the objective measure. During bad times, our results suggest evidence of pessimism on the one hand, via underestimation of the expected growth rate, and optimism on the other hand, via less negative skewness relative to the objective measure.

The recovered price-consistent beliefs also have implications for the expected stock market return. We show that they imply a strongly time-varying and countercyclical expected equity premium and conditional Sharpe ratio. We compare the time series of the price-consistent expected stock market returns with survey data. Specifically, we consider U.S. One-Year Confidence Index from Robert Shiller's investor survey, released by the Investor Behavior Project at Yale University. We find a statistically significant positive relation between the price-consistent beliefs and the Confidence Index. This lends further support to the information-theoretic methodology proposed in this paper to extract investors' beliefs.

Our work builds on a burgeoning literature trying to recover risk premia components with as few assumptions as possible. Ross (2015) shows that, under conditions later clarified by Borovicka, Hansen, and Scheinkman (2016), one can recover simultaneously the investors' preferences and beliefs using a set of Arrow-Debreu securities.[2] 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)).[3] Our approach is similar in spirit, since our non-parametric estimator discretizes

---

[1] The 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 price-consistent measure and the objective measure (see also Kitamura and Stutzer (1997), Owen (2001)).

[2] see also Carr and Yu (2012), Jensen, Lando, and Pedersen (2018), Walden (2017) and Qin and Linetsky (2016, 2017) for related approaches. Earlier papers have also documented the empirical properties of the pricing kernel using parametric assumptions, such as Ait-Sahalia and Lo (1998), Jackwerth (2000), Rosenberg and Engle (2002), Bakshi and Madan (2008), Kistul and Wright (2013) or Song and Xiu (2016).

[3] This 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

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.

Our methodology in this paper relies on the broader class of empirical likelihood methods (see Owen (2001)). These methods are appealing from an empirical point of view since they maximize the Kullback-Leibler information criterion, as a standard MLE would do (see e.g. Kitamura (2007)). Empirical likelihood measures belong to the broader class of Cressie-Read divergence measures between probability distributions (see Cressie and Read (1984)). This family of divergence has gained popularity in financial econometrics for assessing the degree of misspecification of asset pricing models (see Almeida and Garcia (2012)) or generalizing Hansen and Jagganathan (1997) bounds for the pricing kernel (see Almeida and Garcia (2016)). This family of divergence gave rise to various empirical works: Backus, Chernov, and Zin (2014) and Bakshi and Chabi-Yo (2018) look at pricing kernel entropy bounds, and Almeida, Ardison, and Garcia (2018) look at fund performance measures through Cressie-Read divergences. Our work is similar in spirit, in the sense that we are using asset prices to assess the missing component from most standard models in terms of consumption dynamics. We show that this missing component can be essentially attributed to beliefs distortion instead of pricing kernel misspecification. Note that 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, Altissimo and Mele (2009), Gagliardini, Gouriéroux, and Renault (2011), Gospodinov and Otsu (2012), and Crudu and Sandor (2017). Antoine, Bonnal, and Renault (2007) develop the theory for Euclidean Likelihood estimation.[4]

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, recovering 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

---

Orlowski, Sali, and Trojani (2018).

[4]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 Dominguez and Lobato (2004).

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.

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. Greenwood and Shleifer (2014) show that investors' expectations about future stock market returns, obtained from survey data, are negatively correlated with estimates of expected returns implied by representative agent rational expectations models. These studies all make specific parametric (conditionally Gaussian) assumptions on the nature of the beliefs distortions or use professional survey forecasts data to perform inference about 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 SEL 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 price-consistent 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 the general framework to illustrate that asset prices reflect the beliefs of investors and can be used to recover them . Our econometric approach is presented in Section II.2. Two crucial inputs in our 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 they use to discount possible future states of the world. Section IV.1 describes our various choices of the conditioning set, and Section IV.2 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 positive SDF, denoted by $M_{t+1}$ exists. The equilibrium returns $\mathbf{R}_{t+1}^e \in \mathbb{R}^k$ of any set of $k$ traded assets in excess of the risk-free rate satisfy the conditional Euler equation,

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

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

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

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

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

where $\theta_0$ is the true value of the vector of parameters driving the SDF. If investors are rational, the probability measure $\mathcal{P}$ in Equation (1) is equivalent to the objective probability measure $\mathcal{P}_0$, given the filtration $\underline{\mathcal{F}_t}$. Thus, if investors have rational expectations, Equation (1) rewrites:

$$\mathbb{E}^{\mathcal{P}_0} \left[ M_{t+1}\mathbf{R}_{t+1}^e | \underline{\mathcal{F}_t} \right] = \mathbf{0_k}, \tag{4}$$

However, if investors have any behavioral biases that make their beliefs deviate from rational expectations, $\mathcal{P}$ is equivalent to the investors' subjective probability measure, denoted by $\mathcal{P}^*$.

Our objective is to recover $\mathcal{P}$ from observed asset prices at each point in time, in order to uncover the conditional distributions formed by investors given $\underline{\mathcal{F}_t}$. Our econometric procedure, described below, does not require taking a stance on whether beliefs are rational or whether they are distorted relative to rationality. We will then use our procedure to assess whether investors form rational beliefs, i.e. whether $\mathcal{P} = \mathcal{P}_0$.

## 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.[5]

To provide some intuition and fix ideas, let us first consider the problem of recovering unconditional beliefs. This refers to the unconditional distribution of variables such as consumption growth. Let us assume that the consumption growth $G_t$ is best described by a multinomial model, with as many possible states as observation dates, denoted by $T$. Provided that the probability $p_t \equiv \mathcal{P}(G_t = g_t)$ assigned to each state is non-negative and that the probabilities sum to unity, a standard non-parametric maximum likelihood estimator will yield every probability estimate as: $\widehat{p}_t = \frac{1}{T}$, for $t \in \{1, \ldots, T\}$. This is a standard histogram. Now assume that we perform the same likelihood maximization, but enforcing that the estimated multinomial distribution satisfies the unconditional Euler Equations:

$$\mathbb{E}^{\mathcal{P}} \left[ M\left(G_t, Y_t; \theta_0\right) \mathbf{R}_t^e \right] = \mathbf{0_k} \iff \sum_{t=1}^{T} p_t \times M\left(g_t, y_t; \theta_0\right) \mathbf{r_t^e} = \mathbf{0_k}. \tag{5}$$

The method will distort the probability estimates $\widehat{p}_t$ relative to the $1/T$ benchmark in order to satisfy the $k$ moment restrictions. Specifically, the estimated probabilities will be such that they maximize the log-likelihood of the observed data, and that the $k$ assets are priced perfectly. This is the empirical likelihood (EL) estimator of Owen (2001). The SEL

---

estimator, used in this paper relies on the same principle, but incorporates the additional constraints implied by conditional Euler equation restrictions. Note that both the EL and SEL estimators are non-parametric in the sense that they do not require any parametric functional-form assumptions about the data generating process.

To incorporate the conditional information, 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 $X_t \in \mathbb{R}^n$. At each point in time, the representative investor observes the realizations of consumption growth $g_t$, other variables in the SDF $y_t$, excess returns $\mathbf{r}_t^e$, and the conditioning variables $x_t$. Let $p_{i,j}$ denote the conditional probability of observing the joint outcome $(g_j, y_j, \mathbf{r}_j^e, x_j)$ at time $t+1$, i.e. the probability of state $j$ being realized at time $t+1$, given that $(g_i, y_i, \mathbf{r}_i^e, x_i)$ was realized at date $t$.[6] The $T \times T$ matrix with entries $p_{i,j}$ is thus the transition probability matrix of the observed states.

The SEL estimator for the conditional probabilities $(p_{i,j})$ for $i, j = \{1, \ldots, T\}$, is such that it belongs to the simplex:

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

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

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

where $p_{i,\cdot}$ denotes the $T-$dimensional vector of probabilities $(p_{i,1}, ..., p_{i,T})$, $\Theta$ 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:

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

where $\mathcal{K}$ is a kernel function belonging to the class of second order product kernels, and the bandwidth $b_T$ is a smoothing parameter.[7] The objective function in Equation (6) is simply

---

[6] 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 information 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.

[7] $\mathcal{K}$ should satisfy Assumption 3.3 in Kitamura, Tripathi, and Ahn (2004), that is restated here for

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 solution to Equation (6) is analytical and given by: $\forall i, j \in \{1, \dots, T\}$,

$$\widehat{p}_{i,j}^{SEL}(\theta) = \frac{\omega_{i,j}}{1 + M(g_j, \, y_j; \, \theta) \cdot \widehat{\lambda}_i(\theta)' \, \mathbf{r}_j^e} \, , \tag{8}$$

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

$$\widehat{\lambda}_i(\theta) = \arg \max_{\lambda_i \in \mathbb{R}^k} \sum_{j=1}^{T} \omega_{i,j} \log \left[ 1 + M(g_j, y_j; \theta) \cdot \lambda_i' \, \mathbf{r}_j^e \right]. \tag{9}$$

Equations (8) and (9) show that the SEL procedure delivers the entire $(T \times T)$ matrix of probabilities $\left( \widehat{p}_{i,j}^{SEL}(\theta) \right)$ for any given value of the parameter vector $\theta$. Each row $i = \{1, 2, ..., T\}$ contains the probabilities of moving to each of the $T$ possible states $\{j = 1, 2, ..., T\}$ 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 represents the 'beliefs' of the representative investor that are consistent with observed asset prices, i.e. the data-generating process perfectly prices the test assets. Hereafter, we refer to this matrix as the *price-consistent beliefs*.

The intuition behind the SEL estimator can be understood as follows. $\omega_{i,j}$ will overweight states where the realized conditioning variables are *close* to each other, i.e. states for which $|x_i - x_j|$ is small, and not in terms of proximity in time. The bandwidth parameter $b_T$ then controls this notion of closeness, as standard in non-parametric methods. Absent the asset pricing constraints in Equation 6, the best estimates for the probabilities are given by $p_{i,j} = \omega_{i,j}$. Imposition of the asset pricing constraints distorts the probabilities $p_{i,j}$ relative to $\omega_{i,j}$.

Note that the SEL approach does not rely on any parametric assumption about the shape of the distribution or the dynamics. Instead, it approximates the conditional distribution, for each possible value of the current state, as a multinomial p.d.f. on the observed data sample. It may seem that this requires the estimation of $T \times T$ conditional probabilities, given a sample size of only $T$. However, the approach only needs to estimate a $(T \times k)$ matrix, where $k$ denotes the number of assets used in the estimation. Specifically, for each

---

convenience. For $X = (X^{(1)}, X^{(2)}, ..., X^{(n)})$, let $\mathcal{K} = \prod_{i=1}^{n} k(X^{(i)})$. Here $k : \mathbb{R} \to \mathbb{R}$ is a continuously differentiable p.d.f. with support $[-1, 1]$. $k$ is symmetric about the origin, and for some $\alpha \in (0, 1)$ is bounded away from zero on $[-a, a]$. 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 $b_T$.

date, the SEL procedure only requires the estimation of the $k$-vector of Lagrange multipliers associated with the conditional Euler equation restrictions $\lambda_i(\theta)$. 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 (8) and (9)). 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 $T$. [8]

In practice, it can happen that the argument of the log function in Equation (9) 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 (9). In order to avoid this case, we use the normalization introduced by Owen (2001) and described in Appendix A.1. Also, note that the notation $\left(\widehat{p}_{i,j}^{SEL}(\theta)\right)$ 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.

# 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. Our first example focuses on a hypothetical economy where investors have distorted beliefs the underlying *i.i.d.* consumption growth process. Our second example considers a long-run risk economy where both the mean and volatility of the consumption growth rate are persistent, and investors know perfectly the data generating process.

## III.1    Distorted Beliefs When Consumption Growth is *i.i.d.*

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):

$$M_{t+1} = \delta \cdot G_{t+1}^{-\theta_0} \tag{10}$$

---

[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)).

where $\delta$ denotes the subjective discount factor and $\theta_0$ the relative risk aversion coefficient. Suppose that consumption growth is *i.i.d.* log-normal:

$$\log (G_{t+1}) \overset{\mathcal{P}_0}{\sim} \mathcal{N} \left( \mu, \sigma^2 \right) . \tag{11}$$

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 $\mu^* = (1 - \chi)\mu$, where $\chi \in (0, 1)$ is the severity of pessimism, and there are no distortions in the beliefs about the volatility :

$$\log (G_{t+1}) \overset{\mathcal{P}^*}{\sim} \mathcal{N} \left( \mu^*, \sigma^2 \right) . \tag{12}$$

In the above scenario, equilibrium asset prices reflect the subjective beliefs of investors. Solving the equilibrium, we have that both the riskfree rate $R_{f,t} = R_f$ and the price-dividend ratio $P_t/D_t = Z$ are constant and equal to:

$$R_f(\chi) = \frac{1}{\delta} \cdot \exp \left( \theta_0 (1 - \chi)\mu - \frac{\theta_0^2 \sigma^2}{2} \right) , \tag{13}$$

and,

$$Z(\chi) = \frac{1}{\frac{1}{\delta} \cdot \exp \left\{ (1 - \theta_0)(1 - \chi)\mu + \frac{(1-\theta_0)^2 \sigma^2}{2} \right\} - 1} . \tag{14}$$

The equilibrium market return is obtained as:

$$R_{m,t+1}(\chi) = \frac{P_{t+1}/C_{t+1} + 1}{P_t/C_t} \cdot G_{t+1} = \frac{Z(\chi) + 1}{Z(\chi)} \cdot G_{t+1} ,$$

where $Z(\chi)$ is defined in Equation (14). In this economy the Euler equation holds under the subjective probability measure $\mathcal{P}^*$. We have:

$$\mathbb{E}^{\mathcal{P}^*} \left[ G_t^{-\theta_0} \left( R_{m,t}(\chi) - R_f(\chi) \right) \right] = 0 .$$

We can use the EL approach to estimate the unconditional distorted distribution of consumption growth. Our probabilities lies on the simplex:

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

11

and they solve the following problem:

$$\left(\widehat{p}^{EL}(\theta_0)\right) = \underset{(p)\in\Delta}{\arg\max} \sum_{t=1}^{T} \log(p_t) \quad \text{s.t.} \quad \sum_{t=1}^{T} p_t \cdot \frac{r_{m,t}(\chi) - r_f(\chi)}{g_t^{\theta_0}} = 0. \tag{15}$$

To perform our simulation, we calibrate $\mu$ and $\sigma^2$ to the sample mean and variance, respectively, of (log) consumption growth in our data (quarterly U.S. postwar data, see Section IV). 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 ($T = 267$).

Using the estimated probabilities $\widehat{p}^{EL}$, 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 $\chi = \{0.05, 0.1, 0.15\}$, 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 underestimate the mean of consumption growth by 5% (quarterly mean $\mu^* = 0.46\%$ compared with $\mu = 0.48\%$). Our identification is obtained through the fact that 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 distorted mean under subjective beliefs. Moreover, the 90% confidence interval for the mean is very tight and does not include the true historical mean $\mu$ of consumption growth. The estimated volatility of consumption growth has a mean of 0.51% across the 500 simulations – almost identical to the historical value. Our method is thus effective in backing out the subjective beliefs of investors from asset pricing data, whether they are distorted or not. Finally, the coefficient of skewness across the simulations is 0.003, very close to the true value of 0 both statistically and economically. For more severe distortions (see rows 2 and 3), very similar results are obtained. 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.

12

Table 1 – Estimated Subjective Beliefs (in percent)

| | Mean | Volatility | Skewness |
|---|---|---|---|
| | true values | | |
| $\mu^* = \mu$ | 0.48 | 0.5 | 0 |
| $\mu^* = 0.95\mu$ | 0.46 | 0.5 | 0 |
| $\mu^* = 0.90\mu$ | 0.43 | 0.5 | 0 |
| $\mu^* = 0.85\mu$ | 0.41 | 0.5 | 0 |
| **Panel A: T=267** | | | |
| $\mu^* = 0.95\mu$ | .46 [.45,.46] | .51 [.48,.55] | .003 [−.24,.24] |
| $\mu^* = 0.90\mu$ | .43 [.43,.44] | .51 [.48,.55] | .002 [−.25,.20] |
| $\mu^* = 0.85\mu$ | .41 [.41,.41] | .52 [.48,.55] | −.014 [−.27,.26] |
| **Panel B: T=500** | | | |
| $\mu^* = 0.95\mu$ | .46 [.45,.46] | .51 [.48,.54] | −.010 [−.19,.18] |
| $\mu^* = 0.90\mu$ | .43 [.43,.44] | .51 [.49,.54] | .004 [−.22,.19] |
| $\mu^* = 0.85\mu$ | .41 [.41,.41] | .52 [.49,.55] | −.006 [−.21,.19] |
| **Panel C: T=1000** | | | |
| $\mu^* = 0.95\mu$ | .46 [.46,.46] | .51 [.49,.53] | −.008 [−.12,.13] |
| $\mu^* = 0.90\mu$ | .43 [.43,.44] | .51 [.50,.54] | −.007 [−.13,.13] |
| $\mu^* = 0.85\mu$ | .41 [.41,.41] | .52 [.50,.54] | −.015 [−.16,.12] |

## III.2   Recovering the True DGP in a Long-Run Risk Economy

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, under the objective probability measure $\mathcal{P}_0$, aggregate consumption and dividend growth have a small persistent predictable component and stochastic volatility that captures time-

varying economic uncertainty:

$$
\begin{aligned}
\log\left(G_{t+1}\right) &= \mu_c + X_t + \Sigma_t \epsilon_{c,t+1}, \\
\log\left(G_{d,t+1}\right) &= \mu_d + \phi X_t + \phi_d \Sigma_t \epsilon_{d,t+1}, \\
X_{t+1} &= \rho X_t + \phi_e \Sigma_t \epsilon_{x,t+1}, \\
\Sigma_{t+1}^2 &= (1-\nu)\sigma^2 + \nu\Sigma_t^2 + \sigma_w \epsilon_{w,t+1},
\end{aligned}
\tag{16}
$$

where $G_{d,t+1} = D_{t+1}/D_t$ is the dividend growth 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:

$$
\mathbb{E}^{\mathcal{P}_0}\left[G_{t+1}^{-\frac{\alpha}{\psi}} \cdot R_{c,t+1}^{\alpha-1} \cdot \mathbf{R}_{t+1}^e \,\big|\, X_t, \Sigma_t^2\right] = \mathbf{0},
\tag{17}
$$

where $R_{c,t+1}$ is the unobservable return on total wealth, $\alpha = \frac{1-\theta_0}{1-\frac{1}{\psi}}$, $\theta_0$ is the CRRA, and $\psi$ is the elasticity of intertemporal substitution.

We solve for the equilibrium as in the original article and set the model parameters equal to the authors' calibrated values (see Bansal and Yaron (2004)). We simulate a time series of the same length as the historical data ($T = 267$) for the two state variables $X_t$ and $\Sigma_t$, consumption growth $G_t$, dividend growth $G_{d,t}$ and the market return and risk free rate. We then use the SEL approach to recover $\mathcal{P} = \mathcal{P}_0$ using the excess returns on the market portfolio as the sole test asset and $(X_t, \Sigma_t^2)$ as conditioning variables. The challenge for the SEL method is to identify that $X_t$ and $\Sigma_t$ exactly represent the conditional mean and volatility of the consumption growth process, without the information that Equation (16) was used to simulate the data. In turn, the SEL tries to recover the conditional distribution of consumption growth from the simulated time series, and the Euler Equation (17).
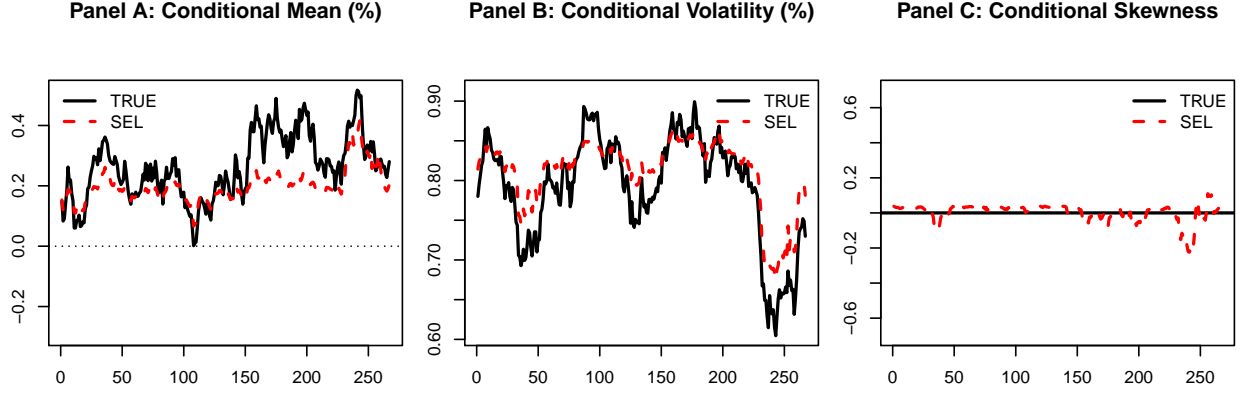
Our implementation follows Equations (8)-(9). Using the recovered probabilities, we compute the time series of the conditional mean, volatility, and skewness of consumption growth. We then repeat this process 500 times. Results are presented on Figure 1.

On Figure 1-I, we plot our results for a randomly chosen sample from among the 500 simulated samples. The three panels represent the simulated time series of the conditional mean ($\mu_c + x_t$), volatility ($\sigma_t$) and skewness (0), respectively in red. The corresponding SEL estimated time series are in black. For these three conditional moments, the correlation between our estimates and the true series are respectively 95.5%, 90.1% and 94.4%. Our method thus shows good performance for recovering the conditional moments of the consumption growth data.
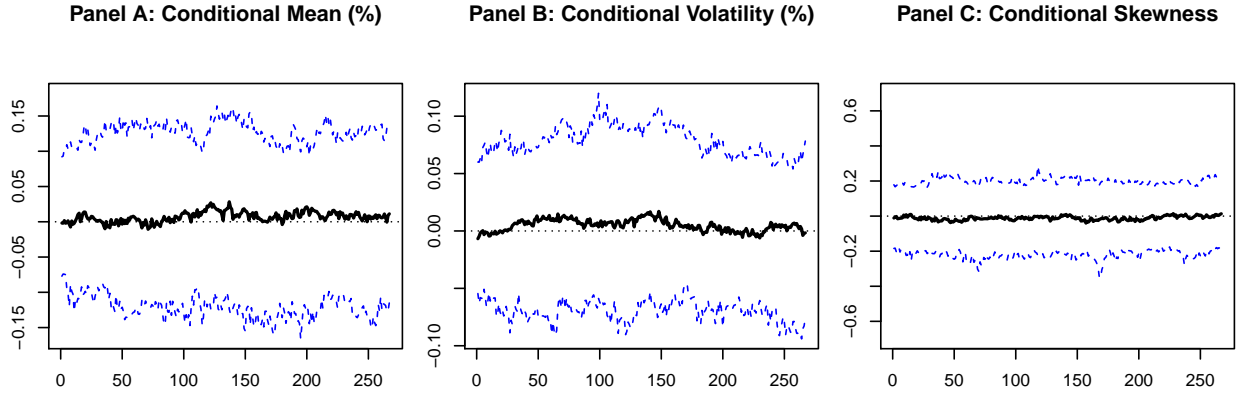
the above results might of course be only by chance or cherry-picked across simulations.

14

## Figure 1 – Simulation of long-run risk model

### (I) Time series of conditional beliefs

**Panel A: Conditional Mean (%)**     **Panel B: Conditional Volatility (%)**     **Panel C: Conditional Skewness**



### (II) Confidence bands

**Panel A: Conditional Mean (%)**     **Panel B: Conditional Volatility (%)**     **Panel C: Conditional Skewness**



*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).

To summarize the performance of the SEL estimator across the simulated samples, our measure looks at the residuals between the simulated conditional moments of consumption growth and the SEL-implied ones. 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.

Our results suggest that the SEL can recover subjective consumption growth dynamics from asset prices with sufficient precision. Our two examples illustrate that whether or not there are significant distortions in the investors beliefs $\mathcal{P}^*$ with respect to the true dynamics $\mathcal{P}_0$, our methodology is equally effective to recover $\mathcal{P}$.

# IV    Implementation Details and Data Description

We describe our choices of inputs for the implementation of the SEL method. In Section IV.1, we present our various specifications of the conditioning set. In Section IV.2, we present the range of possible SDFs that we use to characterize preferences. Finally, in Section IV.3, we detail the data we select for the estimation.

## IV.1    Conditioning Set

An important input to the beliefs extraction procedure described in Section II.2 is the specification of the information set $\underline{\mathcal{F}_t}$ used by the representative investor to price assets, i.e. which variables $X_t$ are relevant to her. A conservative approach would be to include as many variables as possible in the conditioning set. However, as for all non-parametric estimators, our approach suffers from the curse of dimensionality. We thus consider several alternative specifications as conditioning set.

In our baseline specification, 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. 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. We thus perform the estimation for the following choices of the conditioning sets, $X_t$: $(i)$ 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, 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 inven-

tories, 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 the first lag of consumption growth but the entire past of the process (like a moving-average dynamics for instance), 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, denoted by $X_t^{(EW)}$.[9]

## IV.2 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:

$$M\left(G_{t+1}, Y_{t+1}; \theta_0\right) = \delta \cdot G_{t+1}^{-\theta_0}, \tag{18}$$

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)).

We thus 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 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.[10]

---

[9]Our exponentially weighted variables are computed as $X_t^{(EW)} = \alpha X_t + (1-\alpha)X_{t-1}^{(EW)}$, 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 fairly insensitive to the value of $\alpha$.

[10]Other specifications of the SDF have been introduced more recently to better capture asset pricing

## IV.3 Data Sources

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 non-durable 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 IV.1, 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.

---

dynamics, such as models featuring ambiguity aversion (see e.g. Cuesdeanu and Jackwerth (2018) for a review of recent pricing kernel specifications). We leave these for future research.

# V  Empirical Results: Characterizing Beliefs

We estimate the beliefs and present the results for our baseline specification. In this specification, the representative investor has power utility preferences with a constant relative risk aversion parameter $\theta_0 = 10$ (the upper bound of what is generally considered to be an acceptable range), 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.[11] For parameters linked to the non-parametric essence of our estimator, all our results are computed with the Epanechnikov kernel function and with the bandwidth parameters $b_{v,T} = 3\hat{\sigma}_v$, where $\hat{\sigma}_v$ is the empirical standard deviation of the conditioning variable $v$.[12] We estimate the probabilities $\{p_{i,j}\}_{i,j=1,\dots,T}$ by implementing the method described in Section II.2.

Our results are detailed hereafter. After briefly reviewing the pricing errors produced by the SEL in Section V.1, we present the SEL-implied conditional densities of consumption growth in two chosen (good and bad) states of the world (Section V.2). In Section V.3, 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.4, 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.5 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 modeling assumptions typically required in macro-finance models to match the observed dynamics of asset prices.

## V.1  Conditional Pricing Errors

Before turning to the estimated probabilities, we highlight the merits of the SEL method by comparing its implied conditional pricing errors with those obtained with competing methods. The conditional pricing error $\eta_t^{SEL}$ for the market excess return at date $t$ is given by:

$$\hat{\eta}_t^{SEL} = \hat{\mathbb{E}}^{\mathcal{P}}\left[G_{t+1}^{-\theta_0} \cdot \mathbf{R}_{m,t+1}^e | G_t^{(EW)}\right] = \sum_{j=1}^{T} \hat{p}_{t,j}^{SEL} \cdot g_j^{-\theta_0} \cdot \mathbf{r}_{m,j}^e. \tag{19}$$

---

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

[12]The 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.

In comparison, we also obtain the pricing error at each date when the conditional expectation underlying the expression for the pricing error is evaluated using the non-parametric local linear regression (LLR) method. The LLR method estimates the conditional mean function at the current state $g_t^{(EW)}$ by fitting a linear regression locally, with weighted least squares in a fixed neighbourhood of $g_t^{(EW)}$. As with the SEL estimator, the neighborhood is defined in terms of the distance of other possible values of the state from the current state, i.e. $\left|g_i^{(EW)} - g_j^{(EW)}\right|$, and not in terms of proximity in time. The weights are determined by the kernel function, the distance $\left|g_i^{(EW)} - g_j^{(EW)}\right|$, 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 $g_t^{(EW)}$ 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 SEL and the LLR methods (red and black lines, respectively). Consistent with theory, SEL produces zero pricing errors in the time series. In comparison, the conditional pricing errors obtained with LLR 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 conclude that models are unable to simultaneously match the cross section and time series of asset returns and 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 and Singleton (2011), it does not require the underlying SDF to be affine in the parameters.

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.

## V.2   Price-Consistent 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.

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:$Q$3, given the information available in the preceding quarter. The annualized mean and volatility of this distribution are 0.8% and 1.1%, respectively. Superimposed in the same graph is a Gaussian

Figure 2 – Time Series of Conditional Pricing Errors
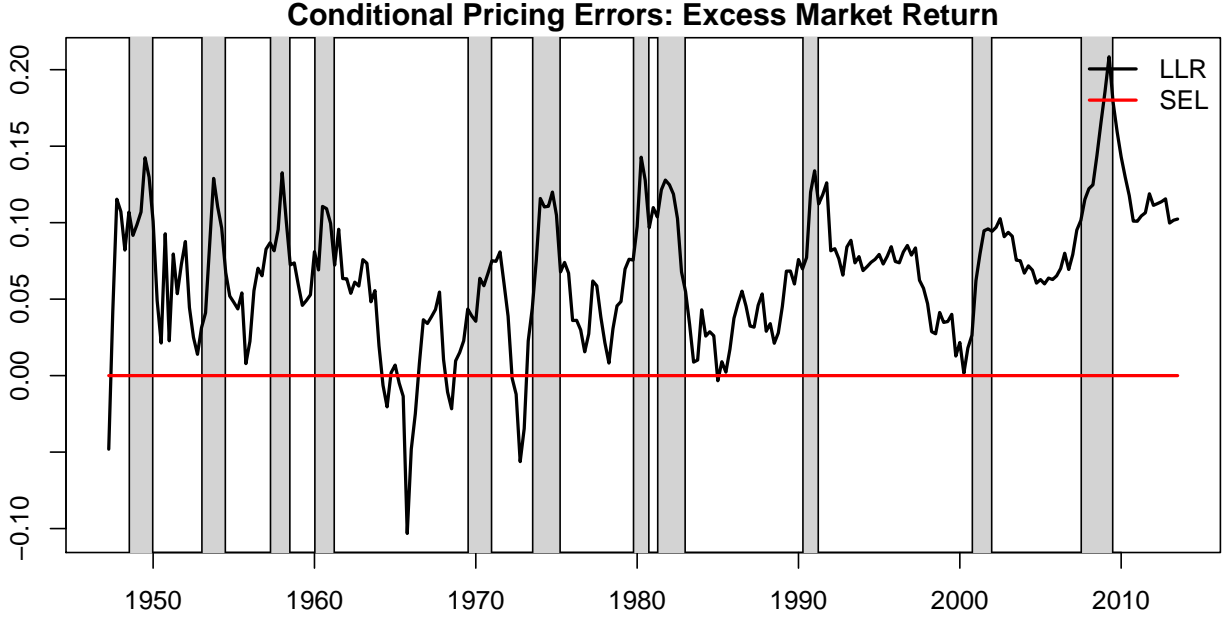
**Conditional Pricing Errors: Excess Market Return**



*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 (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. 'LLR' stands for the non-parametric local linear regression method.

density (red line) with the same mean and variance as the SEL density. The skewness of the SEL density is 0.17, mildly positive, despite the bi-modal nature of the density, with the second mode occurring to the left of the main mode.

This positive skewness is driven by the differential probabilities assigned to events sufficiently far in the tails. Specifically, for events located more than one standard deviation away from the mean, the representative investor assigns a probability of 14.0% to the left tail, compared to a very similar probability of 14.4% to the right tail. In comparison for two and three standard deviations away from the mean, she assigns probabilities of 1.6% and 0.0% to the left tail and higher probabilities of 4.3% and 0.5% to the right tail, respectively.

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. 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

**Figure 3 – Estimated conditional beliefs densities, $\theta_0 = 10$ (%)**

**Panel A: 2009:Q3**
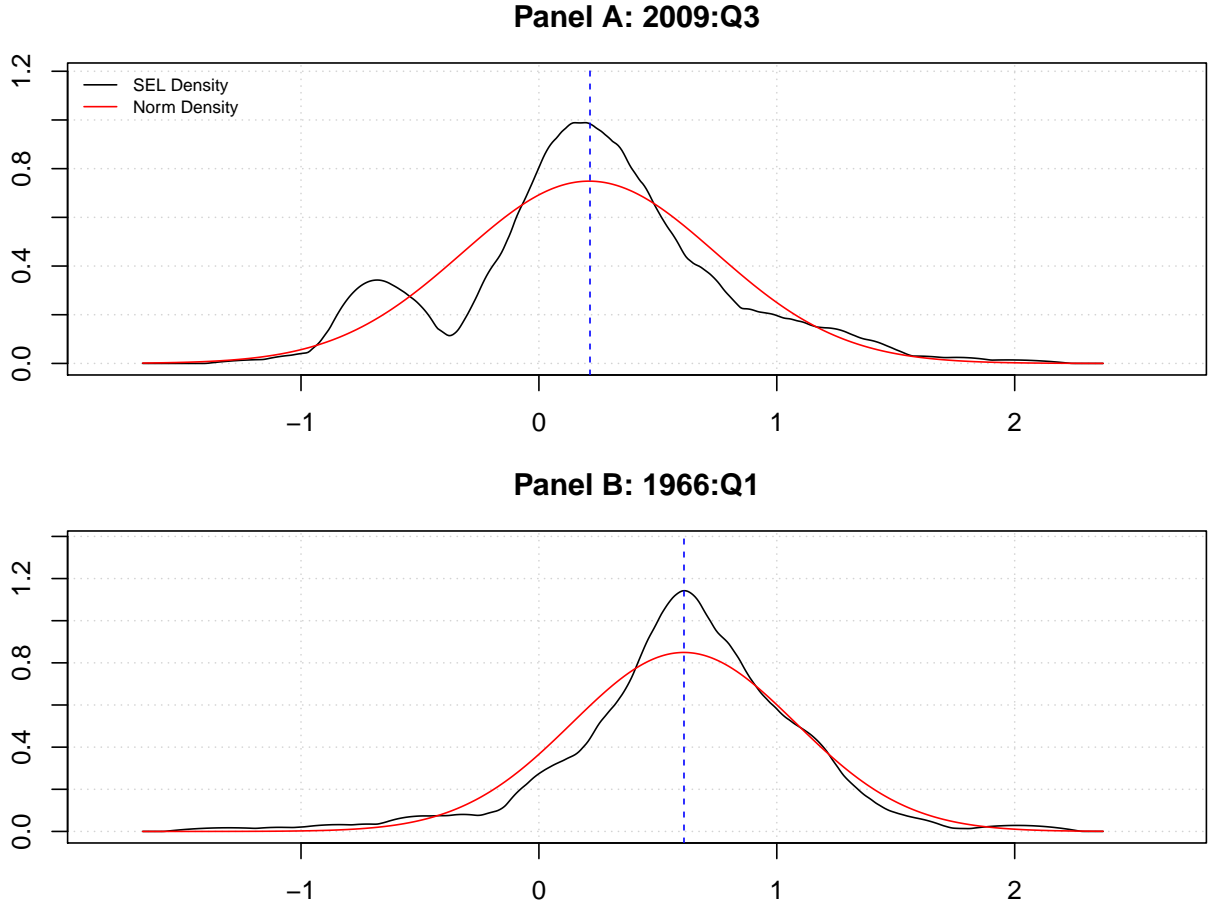


**Panel B: 1966:Q1**



*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.

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). Panel B reveals that the conditional distribution of consumption growth is highly negatively skewed in this good state, with a coefficient of skewness of $-0.65$.

This highly negative skewness emphasizes the strong non-Gaussianity of the data. This is particularly apparent in the left tail of the consumption growth distribution. Specifically, for events happening two and three standard deviations below the mean, the representative investor assigns a probability of 4.6% and 1.4%, respectively, compared to a probability of 2.3% and 0.1%, respectively, for an investor with conditionally Gaussian beliefs.

To summarize, three main conclusions emerge from this section. First, the conditional distribution of consumption growth as perceived by the representative 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 economically significant 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 moments and look at their time series evolution.

## V.3 Price-Consistent Moments of Consumption Growth

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:

$$\widehat{\mathbb{E}}^{\mathcal{P}}\left[\log(G_{t+1})|G_t^{(EW)}\right] = \sum_{j=1}^{T} \widehat{p}_{t,j}^{SEL} \cdot \log(g_j). \tag{20}$$

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 representative investor. The conditional expectation of any nonlinear transformation of the consumption growth rate can be similarly computed. For any moment of order $k$, we have:

$$\widehat{\mathbb{E}}^{\mathcal{P}}\left[\log(G_{t+1})^k|G_t^{(EW)}\right] = \sum_{j=1}^{T} \widehat{p}_{t,j}^{SEL} \cdot \log(g_j)^k. \tag{21}$$

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 4. Panel A presents the time series of the conditional mean of consumption growth. Several patterns are evident from the panel. First, the condi-

23

Figure 4 – Time series of Conditional Moments, $\theta_0 = 10$
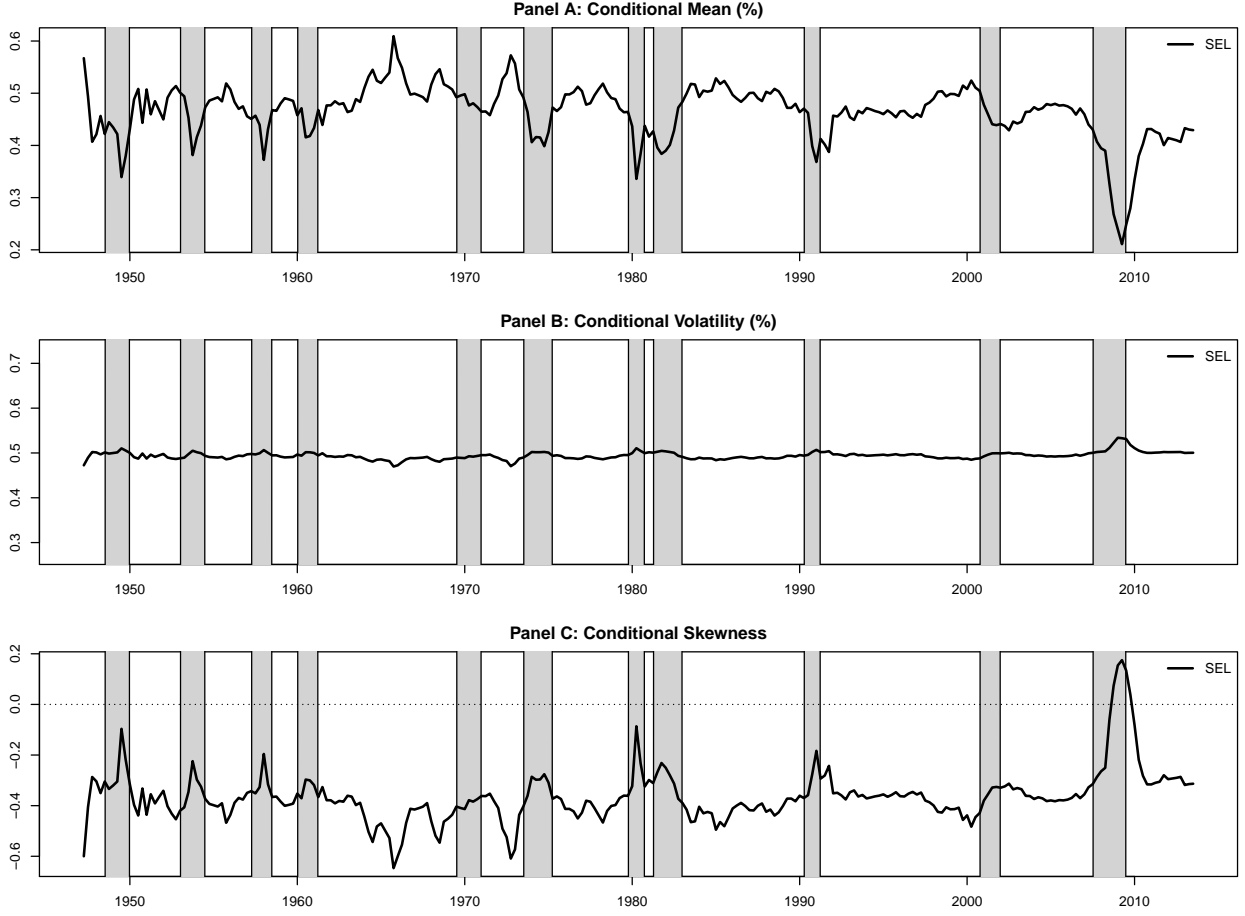
**Panel A: Conditional Mean (%)**



**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.

tional 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 during a recession 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%, i.e. ten times smaller than the 1.0% volatility of realized consumption growth, and its first-order autocorrelation coefficient 0.88 compared to only 0.31 for realized consumption growth.

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. Specifically, the conditional volatility varies from 0.9% to 1.1%.

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 and 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, investors' beliefs about consumption growth forecasts future 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 5 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 consumption growth is an important component of its dynamics *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.4 Alternative Pricing Kernels, Cross Sections of Assets, and Instruments

All the preceding results presented beforehand depend heavily on our choice of three key inputs: the SDF that describes investors' preferences, the cross section of test assets that the SDF is required to price, and the investors' conditioning set. To demonstrate robustness,

Figure 5 – Forecasting Consumption Growth (%)
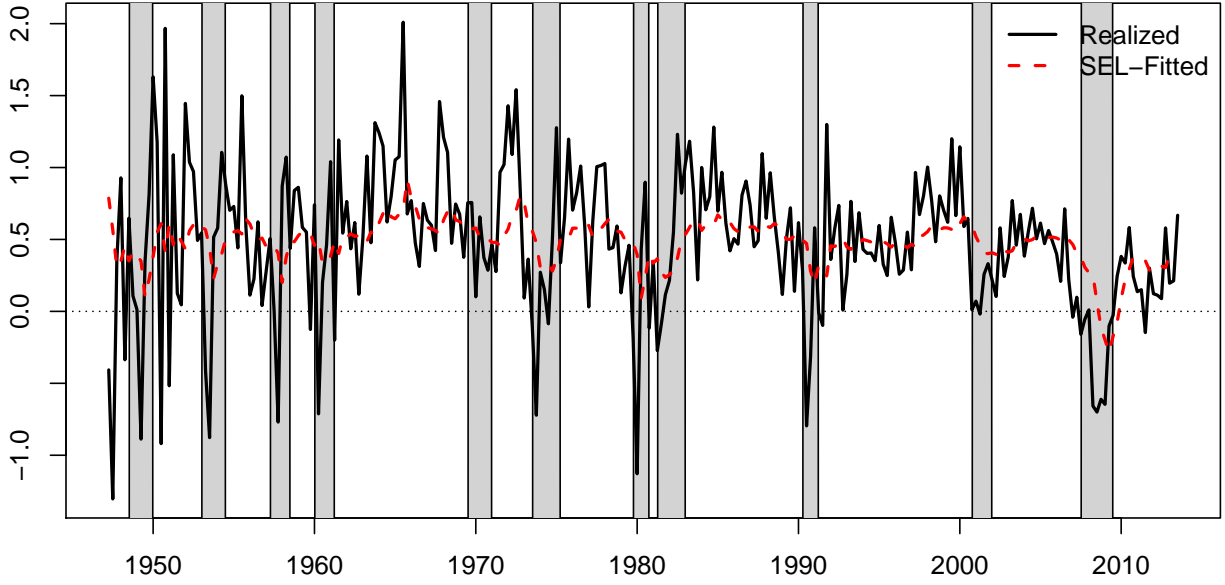


*Notes:* The figure plots the time series of the realized consumption growth (black line) and the fitted value from a forecasting regression of the realized consumption growth on its SEL-implied conditional mean (red line). Shaded areas denote NBER designated recession periods. The conditional mean is 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.

we present results for the different choices of these three inputs detailed in Sections IV.1-IV.2.

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 probabilities that we recover using 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, namely the SDFs implied by the long-run risk 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 both of these pricing kernels, we use the SEL approach to recover the conditional distribution of future consumption growth and form the time series of consumption growth conditional moments, as in Section V.3. Figure 6 plots the time series of the conditional mean (Panel A), the conditional volatility (Panel B), and the conditional skewness (Panel C)

Figure 6 – Time series of Conditional Moments, Alternative SDFs
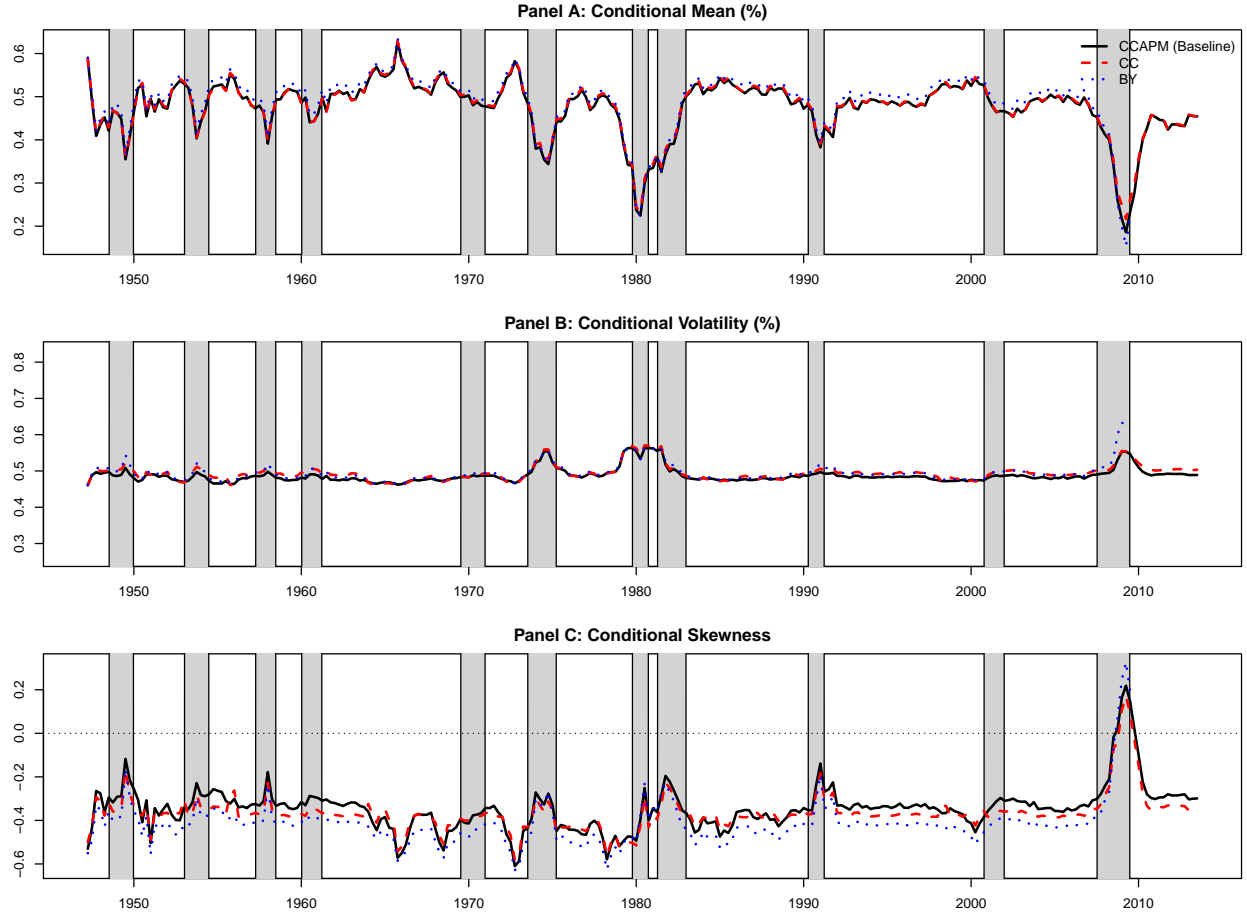
*Notes:* The figure plots the time series of the conditional mean (Panel A), conditional volatility (Panel B), and conditional coefficient of skewness (Panel C) of consumption growth. Shaded areas denote NBER designated recession periods. The conditional moments are obtained using the estimated SEL distributions. 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.

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 6 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 are virtually indistinguishable from each other. Similar conclusions are obtained for the time series of the conditional volatility and skewness in

Panels B and C, respectively. Table 2 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 2 – Correlation Between Beliefs: Alternative SDFs

|  | (A): Mean | | | (B): Volatility | | | (C): Skewness | | |
|---|---|---|---|---|---|---|---|---|---|
|  | CCAPM | CC | BY | CCAPM | CC | BY | CCAPM | CC | BY |
| CCAPM | 1 | 0.999 | 0.998 | 1 | 0.961 | 0.929 | 1 | 0.960 | 0.965 |
| CC | – | 1 | 0.993 | – | 1 | 0.922 | – | 1 | 0.957 |
| BY | – | – | 1 | – | – | 1 | – | – | 1 |

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 7, 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 robust to the specification of the conditioning set. Figure 8 presents the time series of the first three moments of consumption growth recovered from six different choices of the conditioning set. The choices 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 3 presents the

## Figure 7 – Time series of Conditional Moments, Alternative Test Assets
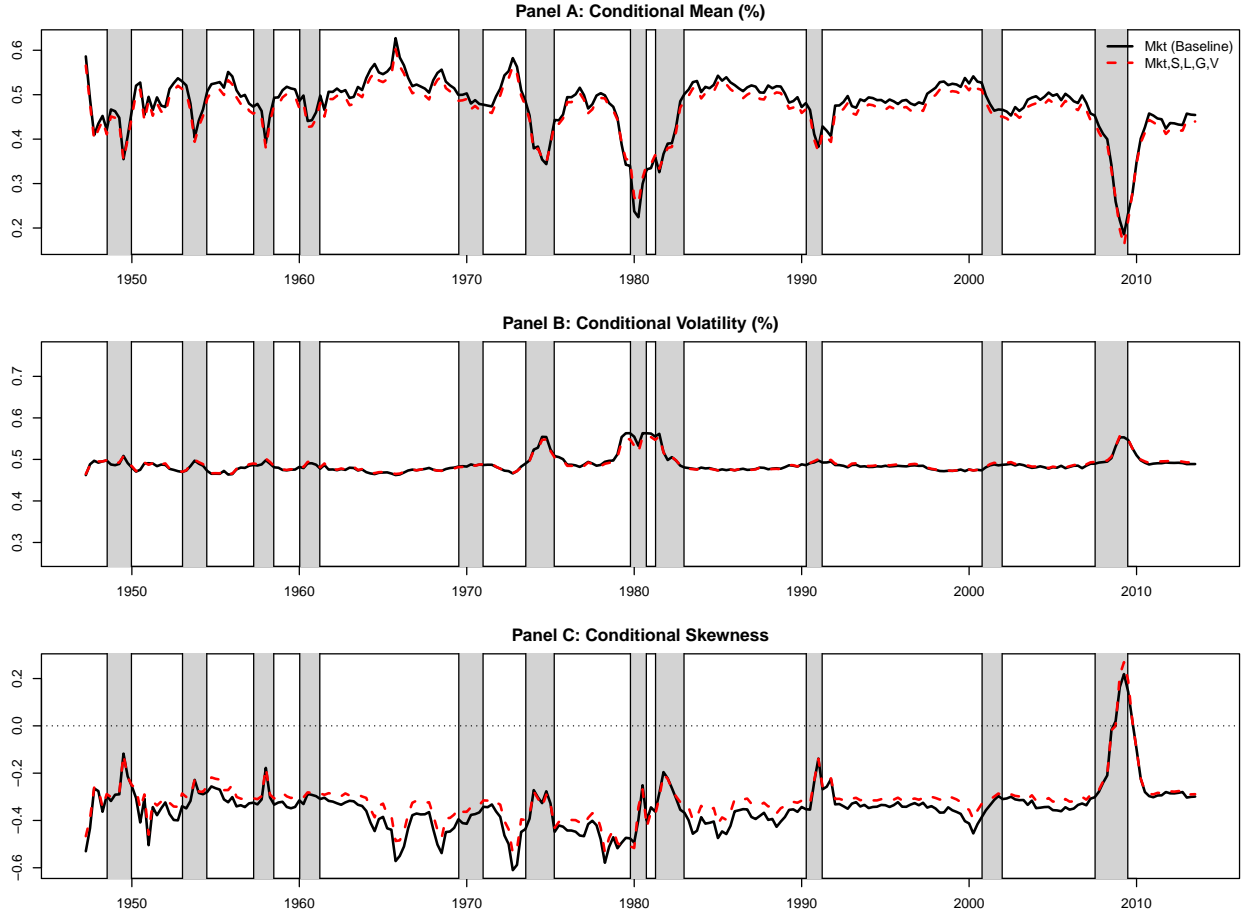


**Panel A: Conditional Mean (%)**

**Panel B: Conditional Volatility (%)**
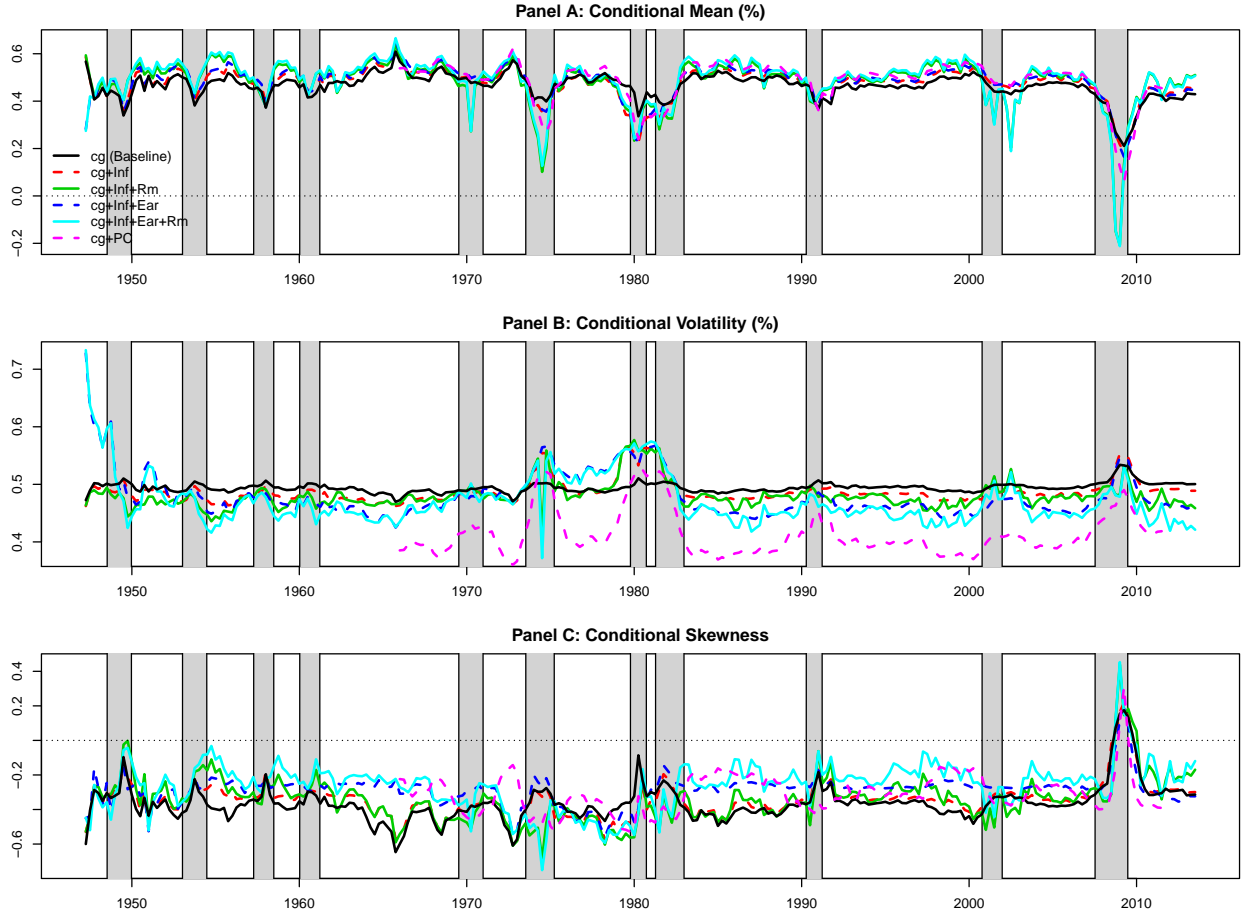
**Panel C: Conditional Skewness**

*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). 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.

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

Figure 8 – 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.

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

Table 3 – Correlation Between Recovered Beliefs: Alternative Conditioning Sets

|  | $cg, Inf$ | $cg, Inf, rm$ | $cg, Inf, Ear$ | $cg, Inf, Ear, Rm$ | $cg$ | $cg, PC$ |
|---|---|---|---|---|---|---|
| Panel A. Correlation Between Means | | | | | | |
| $cg, Inf$ | 1.0 | 0.826 | 0.953 | 0.809 | 0.906 | 0.965 |
| $cg, Inf, Rm$ | – | 1.0 | 0.784 | 0.976 | 0.755 | 0.844 |
| $cg, Inf, Ear$ | – | – | 1.0 | 0.823 | 0.877 | 0.967 |
| $cg, Inf, Ear, Rm$ | – | – | – | 1.0 | 0.750 | 0.850 |
| $cg$ | – | – | – | – | 1.0 | 0.938 |
| $cg, PC$ | – | – | – | – | – | 1.0 |
| 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 |
| $g, Inf, Ear, Rm$ | – | – | – | 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

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 economically 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.

Last, 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. Note that the weakest correlations are observed for the {$cg, PC$} conditioning set, which is only observed over a shorter sample

period, i.e. from 1966:Q1.

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.[13]

## V.5    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 first ask what the recovered beliefs imply about the expected equity premium. Just like with consumption growth, we can use the SEL probabilities attached to the different possible states of the world in the subsequent period, for each possible value of the current state, to determine the expected equity premium in that state:

$$\widehat{\mathbb{E}}^{\mathcal{P}}\left[R_{m,t+1}^{e} \,|\, G_{t}^{(EW)}\right] = \sum_{j=1}^{T} \widehat{p}_{t,j}^{SEL} \cdot r_{m,j}^{e} \,. \tag{22}$$
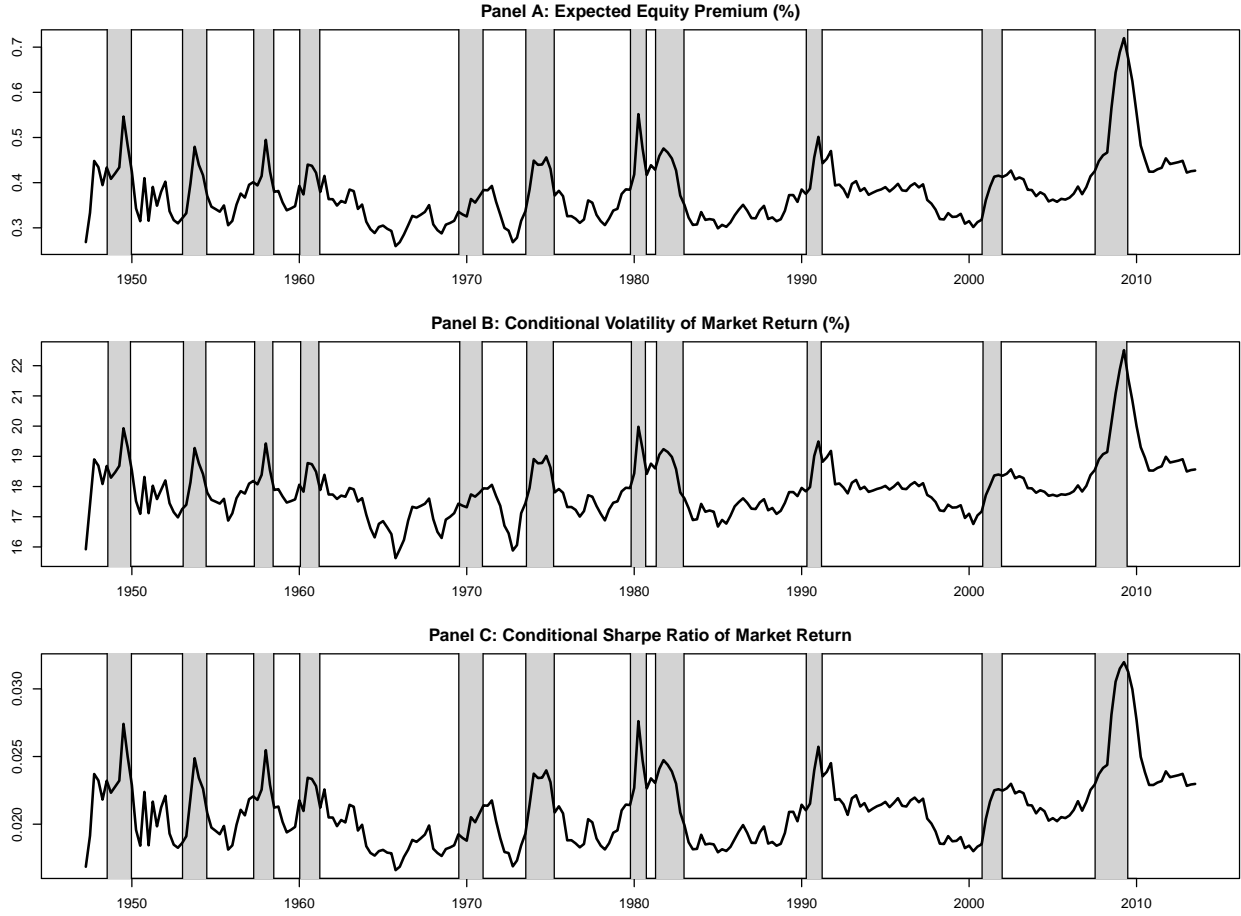
The results are presented in Figure 9, Panel A. The figure shows that the (annualized) expected equity premium is quite volatile, varying from 0.43% in 1965:Q4 to 7.3% in 2009:Q3. The strongly countercyclical nature of the expected equity premium is also evident from the figure. The correlation between 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%.

Panel B and C of Figure 9 present the time series of the conditional volatility and Sharpe ratio of the market return, respectively. The conditional volatility varies from 15.7% to 23.7% and, like the conditional mean, it is strongly countercyclical and the correlation with the recession dummy is equal to 45.1%. 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.

We compare the time series of the price-consistent expected stock market returns with survey data. Specifically, we consider Robert Shiller's investor survey, released by the Investor Behavior Project. The survey, conducted at six-month intervals prior to July 2001 and monthly thereafter, asks a sample of institutional investors how much of a relative change they expect in the Dow Jones Industrial Index in the coming year. The U.S. Institutional

---

[13]The results are omitted for brevity and are available from the authors upon request.

Figure 9 – Time series of Conditional Moments of Market Return

**Panel A: Expected Equity Premium (%)**



**Panel B: Conditional Volatility of Market Return (%)**



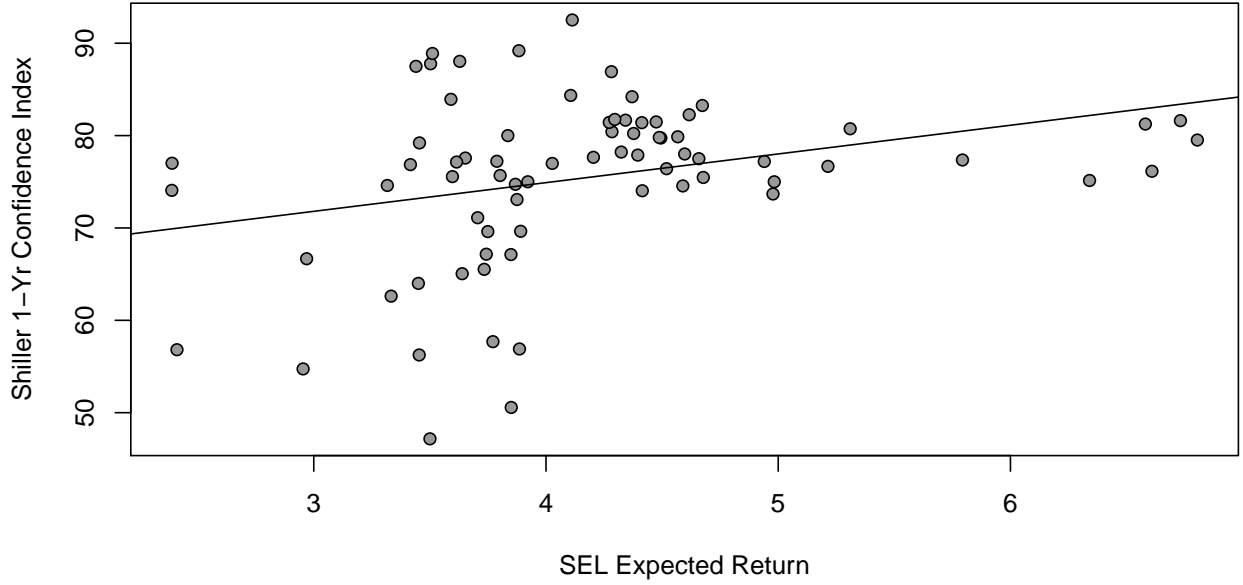**Panel C: Conditional Sharpe Ratio of Market Return**



*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.

One-Year Confidence Index is the percentage of institutional investors expecting an increase in the DJIA in the coming year.[14] The Institutional Index is available from October 1989 onwards.

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 10 presents the scatterplot of the confidence index on the price-consistent expected market returns as an illustration.

---

[14]See Robert Shiller's website for further details on the survey.

Figure 10 – Comparison of Expected Market Returns with Survey Forecasts

Overall, the evidence suggests that the price-consistent beliefs about the stock market, extracted using the SEL method, convey partly similar information as institutional investors' beliefs about the stock market captured in survey data.

# VI  Are Price-Consistent Beliefs Comparable to Commonly Assumed Data Generating Processes?

In the previous section, we identified the conditional distribution of consumption growth that satisfies 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 two standard time series models for the dynamics of macro variables extensively used in the literature. We thus arbitrarily assume that these latter DGP are successful representation of the true probability measure $\mathcal{P}_0$. 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. One could be tempted to interpret deviations of the price-consistent beliefs $\mathcal{P}$ from our assumed $\mathcal{P}_0$ as beliefs distortions. However, we view the results in this section as indicative of either beliefs distortions or misspecification of the true DGP. Also, we acknowledge that the time series models that we compare our price-consistent 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:

$$\log\left(G_{t+1}\right) = (1 - \psi)g + \psi \log\left(G_t\right) + \nu_{t+1} + \theta \nu_t, \quad \text{where} \quad \nu_t \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma^2\right). \tag{23}$$

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 when short- and long-run shocks to consumption growth are perfectly correlated (see, e.g., Bansal and Yaron (2004)). 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 11 plots the historical time series of consumption growth (black line) along with the ARMA(1,1) model-implied conditional mean (red line). 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.

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

$$\log\left(G_{t+1}\right) = \mu_c(s_{t+1}) + \sigma\epsilon_{t+1}, \quad \text{where} \quad \epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}\left(0, 1\right), \tag{24}$$

Figure 11 – Expected and Realized Consumption Growth (%): ARMA(1,1) and Regime Switching Models
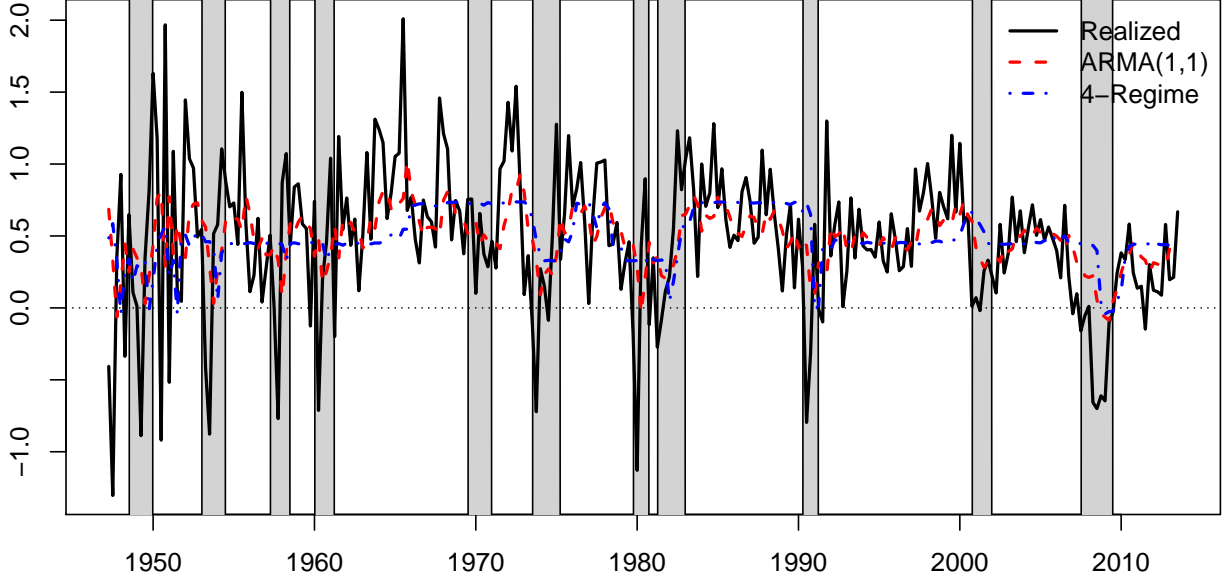


*Notes:* The figure plots the historical time series (black line) along with its model-implied conditional mean (red line), of consumption growth over the period 1947:Q1-2013:Q4. The model parameters are estimated via quasi maximum likelihood.

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, the regime-switching model generates time varying conditional moments and 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 reasonable for this choice of the number of regimes. We estimate the model through standard filtering-based maximum-likelihood. Figure 11 plots its model-implied conditional mean (blue line). The figure shows that the specification offers a good fit for the observed dynamics of consumption growth.

## VI.2 Recovered Beliefs versus Commonly Assumed Dynamics

Consider first the ARMA(1,1) model for consumption growth. Panels A-C of Figure 12 plot the time series of the conditional mean, volatility, and skewness, respectively, of consumption growth implied by the price-consistent beliefs (black line), the ARMA(1,1) specification (red

line), and the regime switching model (blue 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 SEL-based price-consistent beliefs. 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.

Panel B suggests that the conditional homoscedasticity assumption of the ARMA(1,1) model is close to that implied by the price-consistent beliefs. Finally, Panel C suggests that the conditional Gaussianity assumption of the ARMA(1,1) model starkly contrasts with 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. Including GARCH effects and using the historical skewness of the standardized residuals show very similar results since the parameters driving the volatility process end up non-significant. Results for the ARMA-GARCH are provided in Appendix A.3.

We next consider the regime-switching model. As with the ARMA(1,1), 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 countercyclicality 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). 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.

Figure 12 – Comparison of ARMA(1,1), Regime Switching, and SEL
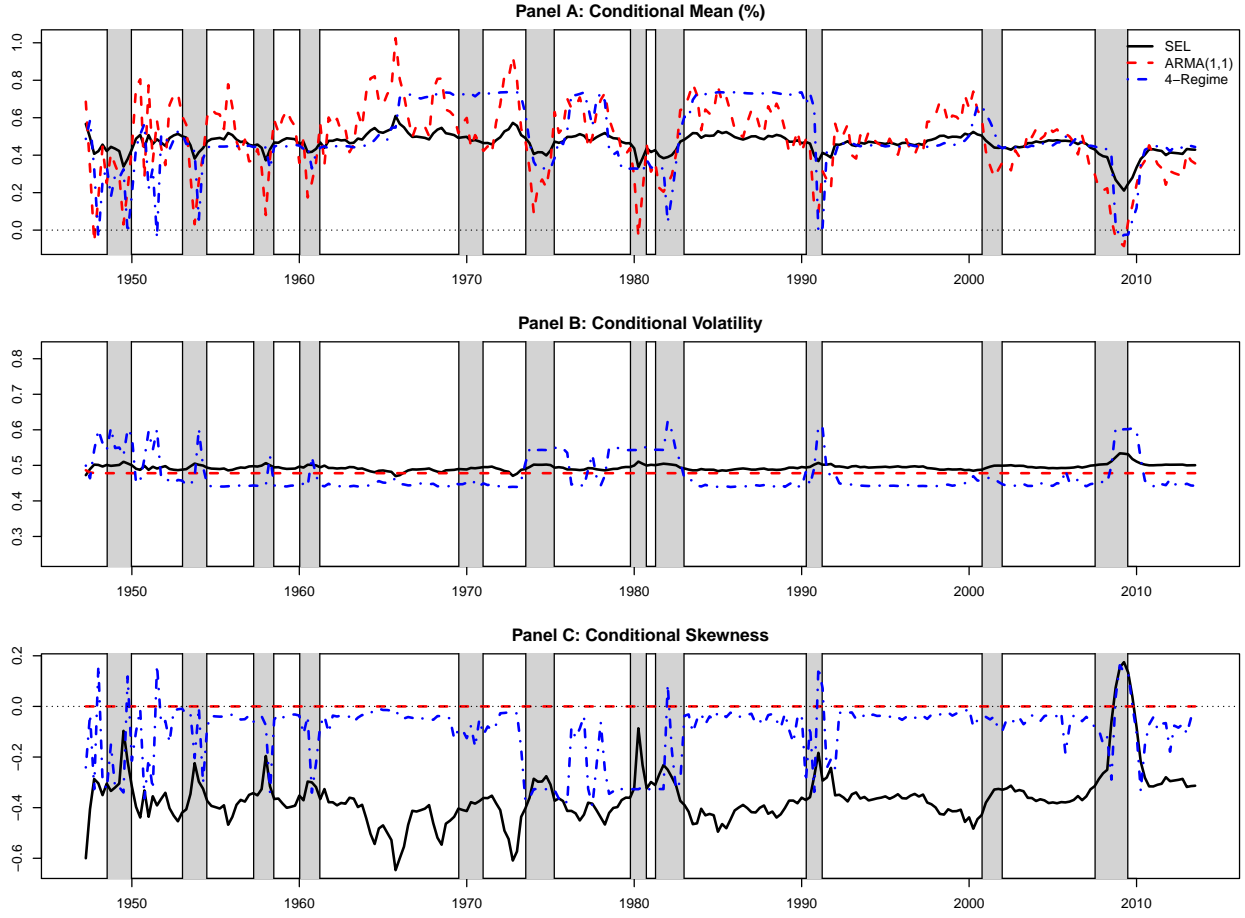
*Notes:* The figure plots the time series of the conditional mean (Panel A), conditional volatility (Panel B), and conditional coefficient of skewness (Panel C) of consumption growth obtained using the estimated SEL distributions (black line), from the ARMA(1,1) specification (red line) and from the regime switching model (blue 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.

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 modeling guidelines for macro finance models in order to improve their ability to match the observed dynamics of asset prices.

# VII    Are Price-Consistent Beliefs Rational?

In Section V, we recovered investors' beliefs $\mathcal{P}$ about future consumption growth 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 mentioned earlier, these DGPs may be misspecified and may not represent the true probability measure $\mathcal{P}_0$ accurately. As a next step, the question naturally arises regarding whether investors' beliefs are rational ($\mathcal{P} = \mathcal{P}_0$) or whether they are distorted relative to rationality ($\mathcal{P} = \mathcal{P}^* \neq \mathcal{P}_0$). 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. The SEL approach used in this paper offers a way to address possible DGP misspecifications. Section VII.1 presents an

alternative, information-theoretic interpretation of the SEL estimator. This property makes the SEL objective function measure deviations from a benchmark. This benchmark is non-parametric, not requiring any functional-form assumptions on the conditional distribution of the variables of interest. It may thus be argued that this benchmark constitutes an attractive candidate for a *objective DGP* 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 price-consistent 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 provide an information-theoretic interpretation to the SEL estimator (see, e.g., Kitamura and Stutzer (1997)).

To see this, let $\mathbb{P}$ be the set of all conditional probability measures defined on $\mathbb{R}^{q+k}$, where $q$ denotes the dimension of the variables $(G_{t+1}, Y_{t+1}, X_{t+1})$ 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 $\mathcal{P}_0$, that satisfy the conditional Euler equations:

$$\mathbb{P}(\theta) := \left\{ \mathcal{P} \in \mathbb{P} \, : \, \mathbb{E}^{\mathcal{P}} \left[ M\left( G_{t+1}, Y_{t+1}; \theta \right) \mathbf{R}_{t+1}^e \big| X_t \right] = \mathbf{0} \right\}, \forall t \in \{1, \dots, T\} \,. \tag{25}$$

Therefore, $\mathbb{P}(\theta)$ is the set of all the conditional probability measures that are consistent with the asset pricing model characterized by the conditional Euler equation restrictions, for a given value of the SDF parameters.

The SEL estimation can then be shown to select a probability measure $\widehat{\mathcal{P}}(\theta)$ such that:

$$\widehat{\mathcal{P}}(\theta) = \inf_{\mathcal{P} \in \mathbb{P}(\theta)} \text{KLIC}(\mathcal{P}_0, \mathcal{P}) \quad \equiv \quad \inf_{\mathcal{P} \in \mathbb{P}(\theta)} \int \log \left( \frac{\mathrm{d}\,\mathcal{P}_0^{(\cdot|X_t)}}{\mathrm{d}\,\mathcal{P}^{(\cdot|X_t)}} \right) \mathrm{d}\,\mathcal{P}_0^{(\cdot|X_t)}$$

$$\text{s.t.} \quad \mathbb{E}^{\mathcal{P}} \left[ M\left( G_{t+1}, Y_{t+1}; \theta \right) \mathbf{R}_{t+1}^e \big| X_t \right] = \mathbf{0}, \tag{26}$$

where the superscript $\mathcal{P}_0^{(\cdot|X_t)}$ is added to emphasize the conditional nature of the probability measures, and $\text{KLIC}(\mathcal{P}_0, \mathcal{P})$ is the Kullback-Leibler Information Criterion (KLIC) divergence (or, relative entropy) between the two measures $\mathcal{P}_0$ and $\mathcal{P}$ (see White (1982)).[15]

---

[15]KLIC actually belongs to the broader class of Cressie-Read divergence measures between probability distributions (see Cressie and Read (1984)). This family of divergence criteria has gained popularity in financial econometrics both for robustness purposes and escaping parametric specifications for asset pricing. Members of the CR-family include Euclidean likelihood (see e.g. Antoine, Bonnal, and Renault (2007),

Note that the KLIC divergence is non-negative and is exactly equal to zero if and only if $\widehat{\mathcal{P}}(\theta) = \mathcal{P}_0$ 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 $\mathcal{P}$ that makes the estimated beliefs as close as possible – in the information-theoretic sense – to the objective one $\mathcal{P}_0$, while also requiring that the estimated beliefs satisfy the pricing restrictions given by the conditional Euler equations.

When the objective measure is assumed to be described by a $T \times T$ transition matrix, it can be shown that the SEL estimator is the one solving both Equations (6) and (26) The same minimization problem can also be considered without the asset pricing constraints. In this case, it is easy to show that our estimated $\widehat{\mathcal{P}}_0(\theta)$ is defined by the kernel weights $\widehat{p}_{i,j}^{(0)}(\theta) = \omega_{i,j}$ of Equation (7) in lieu of Equation (8).

This is a natural choice for the objective measure because it maximizes the log-likelihood of the data in Equation (6), 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, $\mathcal{P}(\theta)$, that satisfies the Euler restrictions, while deviating as little as possible from the objective measure.

The SEL procedure thus enables the characterization of the deviations (if any) from $\widehat{\mathcal{P}}_0$ to $\widehat{\mathcal{P}}(\theta)$. Moreover, it does so without the need for any parametric distributional assumptions on either the objective measure, or the nature of the beliefs distortions relative to the objective measure. 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.

Since $\widehat{\mathcal{P}}_0$ can be interpreted as an estimator of objective *rational* beliefs, any deviations from $\widehat{\mathcal{P}}(\theta)$ to $\widehat{\mathcal{P}}_0$ can be interpreted as investors' distortions from rationality. In this case, the optimal KLIC distance of Equation (26) is positive, and the precise nature of the deviations

---

entropy (Backus, Chernov, and Zin (2014) and Bakshi and Chabi-Yo (2018) for instance), and a more systematic approach is developed by Almeida and Garcia (2012, 2016) and Sandulescu, Trojani, and Vedolin (2018) for deriving admissible conditions for SDFs.

can be characterized by examining the behavior of the Radon-Nikodym derivative $\frac{\mathrm{d}\widehat{\mathcal{P}}(\theta)}{\mathrm{d}\widehat{\mathcal{P}}_0}$.

## VII.2　Empirical Results

We estimate the non-parametric *objective* measure for our benchmark conditioning set. We start by comparing the pricing performance of the objective measure $\widehat{\mathcal{P}}_0$ with the price-consistent measure $\widehat{\mathcal{P}}(\theta)$. Figure 13 (left panel) 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 – the correlation between the pricing errors and a NBER-recession dummy variable is 47.5%. The price-consistent probabilities on the other hand, produce conditional pricing errors that are identically equal to zero in all periods.

Figure 13 – Conditional Pricing Errors and Kullback-Leibler Divergence
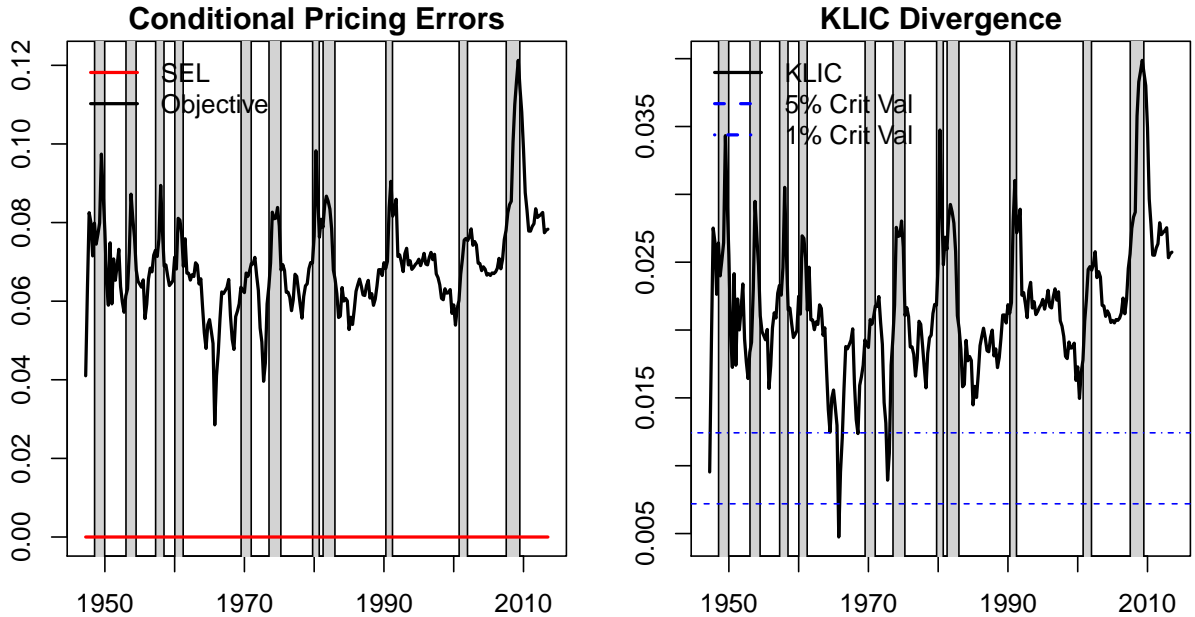


*Notes:* The figure plots the time series of the conditional pricing errors for the excess return on the market portfolio (left panel) and the KLIC divergence between the SEL-implied probability measure and the objective measure given by the kernel weights (right panel). The conditional expectation underlying the calculation of the pricing error is evaluated using the SEL probabilities (red line) or using the kernel weights (black line). For the KLIC divergence, we plot the 5% and 1% critical value of the G-test, i.e. for a given level of confidence $\alpha$: $F^{-1}_{\chi^2}(\alpha, df = 1)/(2 \times T)$. 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.

Figure 13 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 13 (right panel) plots the time series of the KLIC divergence between the probability measures $\widehat{\mathcal{P}}_0$ and $\widehat{\mathcal{P}}(\theta)$, as given by the empirical counterpart of the objective function in Equation (26):

$$\mathrm{KLIC}_t\left(\widehat{\mathcal{P}}_0,\,\widehat{\mathcal{P}}(\theta)\right) = \sum_{j=1}^{T} \omega_{t,j} \times \log\left(\frac{\omega_{t,j}}{\widehat{p}_{t,j}^{SEL}}\right). \tag{27}$$

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 14 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. Panel A shows that the price-consistent probability weights imply a lower expected consumption growth compared to that implied by the objective measure for most states of the world. Thus, investors seem to consistently underestimate consumption growth forecasts during good and bad times alike. However, 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 significantly greater during recessionary episodes. During the recent financial crisis, for instance, the underestimation of the conditional mean was as high as 16.7%.

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 and countercyclical discrepancies in the skewness between the two measures.

Interestingly, we do not find much evidence of increased persistence under the distorted 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.

Figure 14 – Time Series of Moments Under SEL and Objective Measures and Divergence of Measures
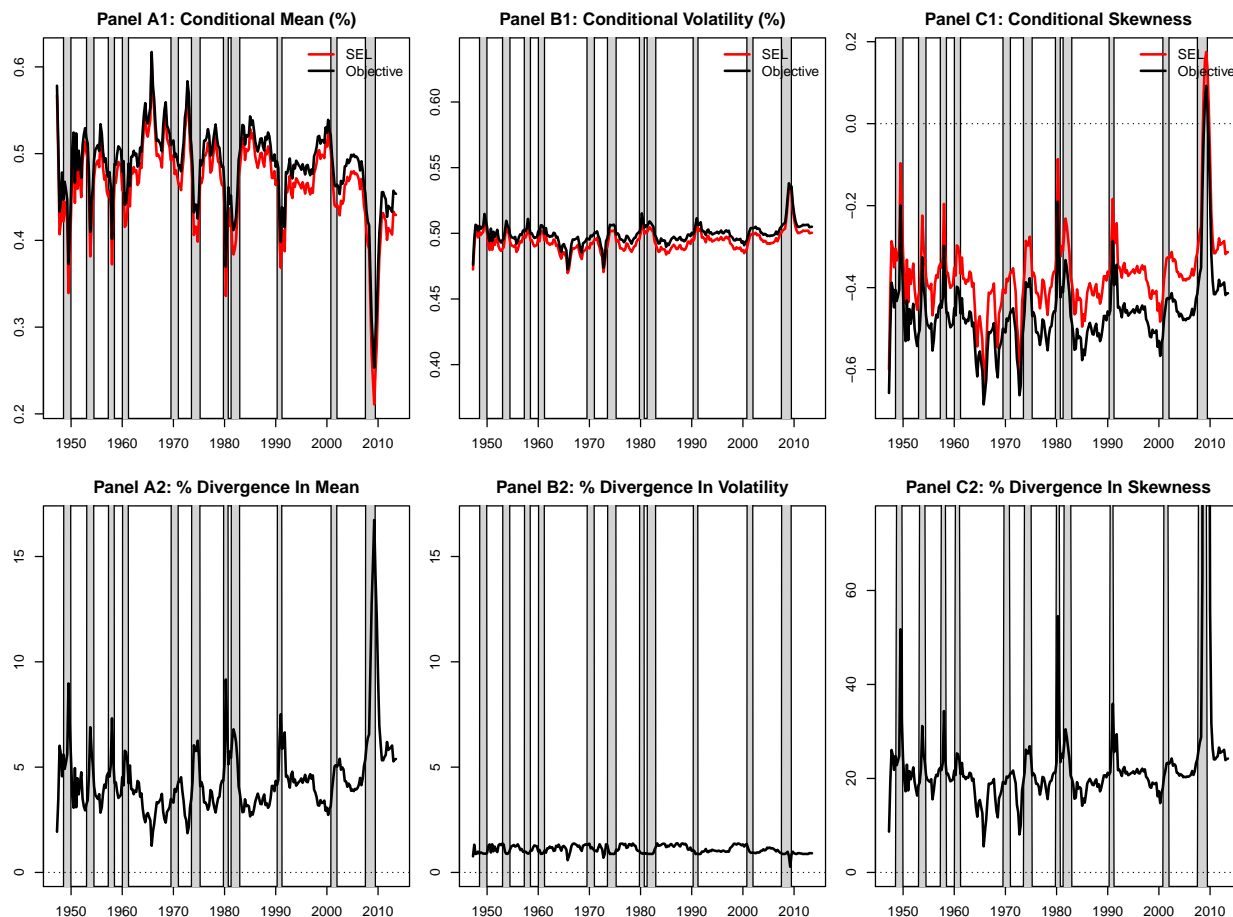
*Notes:* The figure plots the time series of the KLIC divergence between the SEL-implied probability measure and the objective measure given by the kernel weights. The former is obtained using as inputs the SDF implied by power utility preferences with a CRRA, the excess return on the market is used 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.

To summarize, investors' beliefs about consumption growth seem distorted relative to the objective benchmark in two important respects. First, the conditional mean is lower under the price-consistent beliefs compared to the objective measure, 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 price-consistent beliefs compared to the objective measure. This is a consequence of the mean of the distribution shifting to the left under the price-consistent measure but the extreme left and right tails not expanding relative to the objective measure. This can be

interpreted as investors overweighting events slightly below the mean while attaching similar weights to events far out in the left and right tails relative to the objective measure. 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.

# 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' conditional 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 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.

# References

AIT-SAHALIA, Y., AND A. LO (1998): "Nonparametric Estimation of State-Price Densities Implicit In Financial Asset Prices," *Journal of finance*, 53(2), 499–547.

ALMEIDA, C., K. ARDISON, AND R. GARCIA (2018): "Nonparametric Assessment of Hedge Fund Performance," Discussion paper.

ALMEIDA, C., AND R. GARCIA (2012): "Assessing Misspecified Asset Pricing Models with Empirical Likelihood Estimators," *Journal of Econometrics*, 170(2), 519–537.

——— (2016): "Economic Implications of Nonlinear Pricing Kernels," *Management Science*, 63(10), 3147–3529.

ALTISSIMO, F., AND A. MELE (2009): "Simulated Non-Parametric Estimation of Dynamic Models," *Review of Economic Studies*, 76, 413–450.

ANTOINE, B., H. BONNAL, AND E. RENAULT (2007): "On the Efficient Use of the Informational Content of Estimating Equations: Implied Probabilities and Euclidean Empirical Likelihood," *Journal of Econometrics*, 138(2), 461–487.

BACKUS, D., M. CHERNOV, AND S. ZIN (2014): "Sources of Entropy in Representative Agent Models," *The Journal of Finance*, 69(1), 51–99.

BAKSHI, G., AND F. CHABI-YO (2012): "Variance Bounds on the Permanent and Transitory Components of Stochastic Discount Factors," *Journal of Financial Economics*, 105, 191?208.

——— (2018): "New Entropy Restrictions and the Quest for Better Specified Asset Pricing Models," *Journal of Financial and Quantitative Analysis*.

BAKSHI, G., AND D. MADAN (2008): "Investor heterogeneity, aggregation, and the non-monotonicity of the aggregate marginal rate of substitution in the price of market-equity," Discussion paper.

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

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

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

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

BOROVICKA, J., L. P. HANSEN, AND J. A. SCHEINKMAN (2016): "Misspecified Recovery," *The Journal of Finance*, mimeo.

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

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

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

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

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

CARR, P., AND J. YU (2012): "Risk, Return, and Ross Recovery," *Journal of Derivatives*, 20(1), 38–59.

CARRASCO, M., AND J.-P. FLORENS (2000): "Generalization of GMM to a Continuum of Moment Conditions," *Econometric Theory*, 16(6), 797–834.

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

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

CRESSIE, N., AND T. READ (1984): "Multinomial Goodness-of-t Tests," *Journal of the Royal Statistics Society Series B*, 46, 440–464.

CRUDU, F., AND Z. SANDOR (2017): "On the Finite-Sample Properties of Conditional Empirical Likelihood Estimators," *Communication in Statistics?Simulation and Computation*, 46(2), 1520–1545.

CUESDEANU, H., AND J. C. JACKWERTH (2018): "The pricing kernel puzzle: survey and outlook," *Annals of Finance*, 14, 289–329.

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

DOMINGUEZ, M., AND I. LOBATO (2004): "Consistent Estimation of Models Defined by Conditional Moment Restrictions," *Econometrica*, 72(5), 1601–1615.

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

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

GAGLIARDINI, P., C. GOURIÉROUX, AND E. RENAULT (2011): "Efficient Derivative Pricing by the Extended Method of Moments," *Econometrica*, 79(4), 1181–1232.

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

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

GOSPODINOV, N., AND T. OTSU (2012): "Local GMM Estimation of Time Series Models with Conditional Moment Restrictions," *Journal of Econometrics*, 170(2), 476–490.

GREENWOOD, R., AND A. SHLEIFER (2014): "Expectations of Returns and Expected Returns," *Review of Financial Studies*, 27(3), 714–746.

HANSEN, L., AND R. JAGGANATHAN (1997): "Assessing Specification Errors in Stochastic Discount Factor Models," *The Journal of finance*, 52(2), 557–590.

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

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

IMBENS, G., S. DONALD, AND W. NEWEY (2003): "Empirical Likelihood Estimation and Consistent Tests with Conditional Moment Restrictions," *Journal of Econometrics*, 117(1), 55–93.

JACKWERTH, J. (2000): "Recovering Risk Aversion from Option Prices and Realized Returns," *The Review of Financial Studies*, 13(2), 433–451.

JENSEN, C., D. LANDO, AND L. PEDERSEN (2018): "Generalized Recovery," *Journal of Financial Economics*, forthcoming.

KAHNEMAN, D., AND A. TVERSKY (1979): "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47, 263–291.

KISTUL, Y., AND J. WRIGHT (2013): "The economics of options-implied inflation probability density functions," *Journal of Financial Economics*, 110(3), 696–711.

KITAMURA, Y. (2007): "Nonparametric Likelihood: Efficiency and Robustness," *The Japanese Economic Review*, 58(1).

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

KITAMURA, Y., G. TRIPATHI, AND H. AHN (2004): "Empirical Likelihood-Based Inference in Conditional Moment Restriction Models," *Econometrica*, 72(6), 1667–1714.

KOZHAN, R., A. NEUBERGER, AND P. SCHNEIDER (2013): "The Skew Risk Premium in the Equity Index Market," *Review of Financial Studies*, 26, 2174–2203.

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

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

MARTIN, I. (2013): "Consumption-Based Asset Pricing with Higher Cumulants," *Review of Economic Studies*, 80(2), 745–773.

——— (2017): "What is the Expected Return on the Market?," *Quaterly Journal of Economics*, 132(1), 367–433.

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

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

NAGEL, S., AND K. SINGLETON (2011): "Estimation and Evaluation of Conditional AssetPricing Model," *Journal of Finance*, (3), 873–909.

ORLOWSKI, P., A. SALI, AND F. TROJANI (2018): "Arbitrage Free Dispersion," Discussion paper, HEC Montreal.

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

PIAZZESI, M., J. SALOMAO, AND M. SCHNEIDER (2015): "Trend and Cycle in Bond Premia," mimeo.

QIN, L., AND V. LINETSKY (2016): "Positive Eigenfunctions of Markovian Pricing Operators: Hansen-Scheinkman Factorization, Ross Recovery and Long-Term Pricing," *Operations Research*, 64(1), 99–117.

QIN, L., AND V. LINETSKY (2017): "Long Term Risk: A Martingale Approach," *Econometrica*, 85(1).

QIU, C., AND T. OTSU (2017a): "Information Theoretic Approach to High Dimensional Multiplicative Models: Stochastic Discount Factor and Treatment Effect," Discussion paper, LSE.

——— (2017b): "Nonparametric Stochastic Discount Factor For High-Dimensional Financial Returns," Discussion paper, LSE (mimeo).

ROSENBERG, J. V., AND R. F. ENGLE (2002): "Empirical pricing kernels," *Journal of Financial Economics*, 64(3), 341–372.

ROSS, S. A. (2015): "The Recovery Theorem," *The Journal of Finance*, 70(2), 615–648.

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

SANDULESCU, M., F. TROJANI, AND A. VEDOLIN (2018): "Model-Free International Stochastic Discount Factors," Discussion paper.

SCHNEIDER, P. (2015): "Generalized risk premia," *Journal of Financial Economics*, 116(3), 487–504.

——— (2018): "An Anatomy of the Market Return," *Journal of Financial Economics*, Forthcoming.

SCHNEIDER, P., AND F. TROJANI (2017): "(Almost) Model-Free Recovery," *Journal of Finance*, Forthcoming.

——— (2018): "Divergence and the Price of Uncertainty," *The Journal of Financial Econometrics*, 1(56).

SONG, Z., AND D. XIU (2016): "A Tale of Two Option Markets: Pricing Kernels and Volatility Risk," *Journal of Econometrics*, 190(1), 176–196.

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

WACHTER, J. (2006): "A Consumption-Based Model of the Term Structure of Interest Rates," *Journal of Financial Economics*, pp. 365–399.

WALDEN, J. (2017): "Recovery with Unbounded Diffusion Processes," *Review of Finance*, 21(4), 1403–1444.

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

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

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

# A Appendix

## A.1 Owen normalization

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

$$\widehat{\lambda}_i^{(o)}(\theta) \quad = \quad \underset{\lambda_i \in \mathbb{R}^k}{\operatorname{argmax}} \sum_{j=1}^{T} \omega_{i,j} \cdot \Psi_\nu \left[ 1 + M(g_j, y_j; \theta) \cdot \lambda_i' \, \mathbf{r}_j^e \right] \tag{28}$$

$$\text{where} \quad \Psi_\nu(x) = \begin{cases} \log(x) & \text{if } x > \nu \\ \log(\nu) - \dfrac{3}{2} + 2\dfrac{x}{\nu} - \dfrac{1}{2}\left(\dfrac{x}{\nu}\right)^2 & \text{if } x \leqslant \nu \end{cases} \tag{29}$$

Equation (29) 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 $\widehat{\lambda}_i^{(o)}$ (see Equation (8)) deviate from unity. Again, Owen (2001) suggests to normalize the probabilities *ex-post* so that they add up to one:

$$\widehat{p}_{i,j}^{SEL^{(o)}}(\theta) = \frac{\omega_{i,j}}{1 + M(g_j,\, y_j;\, \theta) \cdot \widehat{\lambda}_i^{(o)}(\theta)' \, \mathbf{r}_j^e} \times \left( \sum_{j=1}^{T} \frac{\omega_{i,j}}{1 + M(g_j,\, y_j;\, \theta) \cdot \widehat{\lambda}_i^{(o)}(\theta)' \, \mathbf{r}_j^e} \right)^{-1}, \tag{30}$$

## 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

$$M_t = \delta \cdot G_t^{-\gamma} \left( S_t / S_{t-1} \right)^{-\gamma}, \tag{31}$$

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, $S$, 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:

$$\log(G_t) = g + v_t, \quad v_t \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma^2\right).$$

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

$$\log(S_t) = (1 - \phi)\,\overline{\log(S)} + \phi \log(S_{t-1}) + \lambda\left(\log(S_{t-1})\right) v_t, \tag{32}$$

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

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

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

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, $\widehat{v}_t = \log(G_t) - g$, to extract the implied time series of the surplus consumption ratio using Equation (32). 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

$$M_{t+1} = \delta^\theta \left(G_{t+1}\right)^{-\frac{\theta}{\rho}} R_{c,t+1}^{\theta-1},$$

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, $\log G_{t+1}$ and $\Delta d_{t+1}$, respectively, are modeled as containing a small persistent expected growth rate component, $x_t$, that follows a heteroscedastic 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

$$\ln M_{t+1} = c_1 + c_2 \log G_{t+1} + c_3 x_{t+1} + c_4 \sigma_{t+1}^2 + c_5 x_t + c_6 \sigma_t^2 \tag{33}$$

Table 4 – ARMA-GARCH estimated parameters

| | $g$ | $\psi$ | $\theta$ | $\omega$ | $\alpha$ | $\beta$ | $\gamma$ |
|---|---|---|---|---|---|---|---|
| Estimate | 0.451 | 0.809 | $-0.424$ | 0.011 | 0.049 | 0.825 | 0.150 |
| Robust Stdev | (0.076) | (0.070) | (0.097) | (0.007) | (0.051) | (0.068) | (0.088) |

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.

## A.3 Results of the ARMA-GARCH Process for the Consumption Growth

We estimate the following model:

$$
\begin{aligned}
\log\left(G_{t+1}\right) &= (1-\psi)g + \psi\log\left(G_t\right) + \sigma_{t+1}\nu_{t+1} + \theta\sigma_t\nu_t\,, \quad \text{where} \quad \nu_t \sim \mathcal{D}\left(0,1\right) \\
\sigma_{t+1}^2 &= \omega + \left(\alpha + \gamma\mathbb{1}\{\nu_t < 0\}\right)\sigma_t^2\nu_t^2 + \beta\sigma_t^2\,,
\end{aligned}
$$

that is an ARMA(1,1) mean model and a GJR-GARCH model. In addition, we do not assume that the standardized shocks $\nu_t$ are *i.i.d.* Gaussian but rather that they are martingale difference with a particular distribution $\mathcal{D}(\cdot)$.

The estimated parameters are given in Table 4 and the conditional moments are provided on Figure ??

Figure 15 – Comparison of ARMA-GARCH and SEL
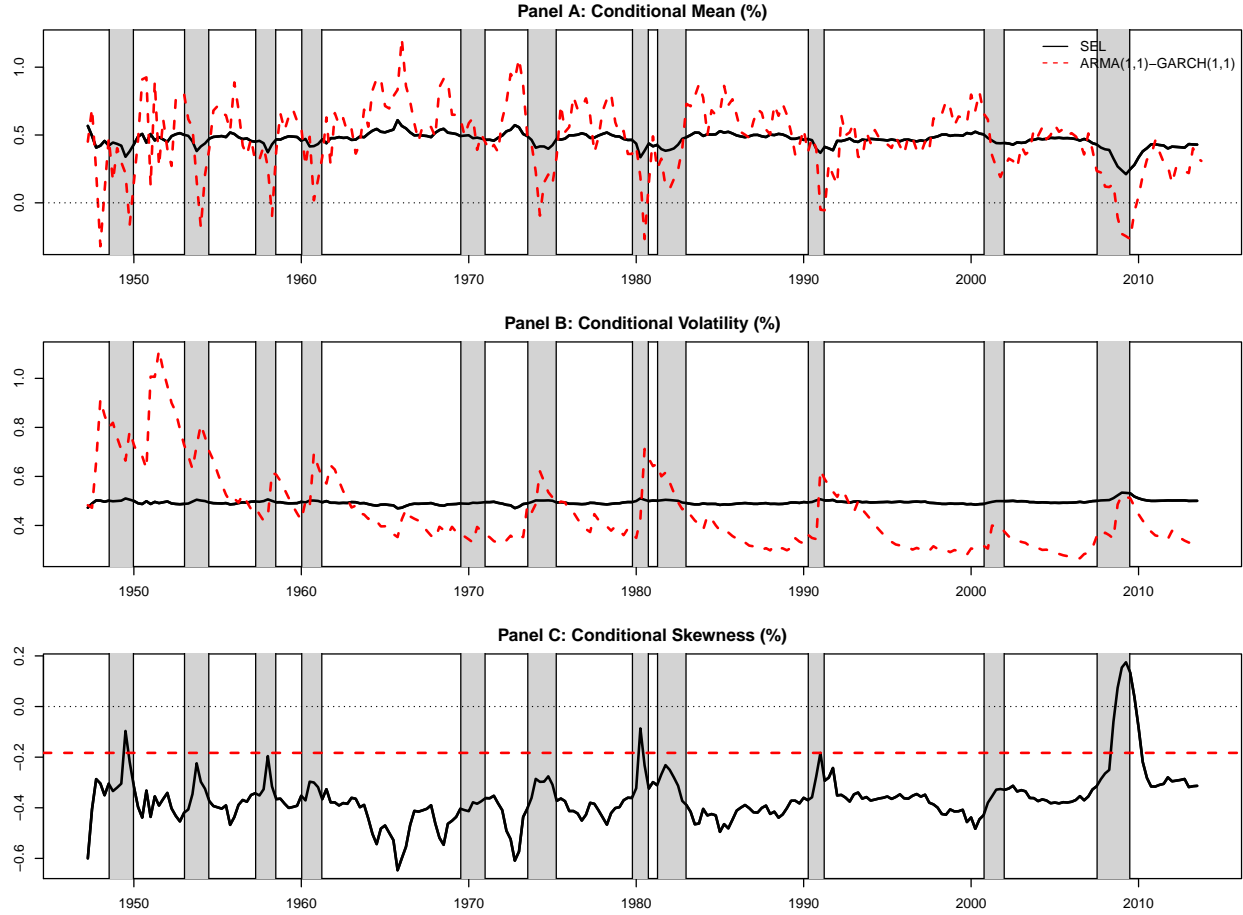
*Notes:* The figure plots the time series of the conditional mean (Panel A), conditional volatility (Panel B), and conditional coefficient of skewness (Panel C) of consumption growth obtained using the estimated SEL distributions (black line) and from the ARMA(1,1) specification (red line). Shaded areas 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.