Identifying Beliefs from Asset Prices*

Anisha Ghosh† Guillaume Roussellet‡

June 10, 2020

Abstract

We propose a novel procedure to identify the marginal stock market investor's beliefs from observed asset prices. Our approach recovers *price-consistent beliefs*, i.e. the 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 assumptions about the data generating process or investor rationality. The recovered beliefs about consumption growth exhibit non-Gaussianity, strong procyclicality in the mean, flat volatility, and bimodality in recessions. Beliefs about the stock market are countercyclical and correlate with survey data on institutional investors' expectations.

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

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

---

*We benefited from helpful comments from Francesco Bianchi, Marine Carrasco, Xiaohong Chen, George Constantinides, Richard Crump, Andras Fulop, René Garcia, Marco Del Negro, Domenico Giannone, Lars Hansen, Scott Joslin, Hanno Lustig, Stefan Nagel, Jean-Paul Renne, Taisuke Otsu, Romeo Tedongap, Fabio Trojani and seminar participants at the University of Cambridge, McGill University, London Business School, the Forecasting and Empirical Methods in Macro and Finance Workshop at the NBER's 2019 Summer Institute, CEU-Essec Conference in Budapest, and the York CAMF Asset Pricing Conference. We thank Qing Xu for excellent research assistance. All errors and omissions are our own.

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

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

"Asset markets are forward-looking. They encode information about investor beliefs about the future and about investors' concerns for risk. How can we use financial market data to extract information about these two components in a reliable way?"

— Lars P.Hansen[1]

I Introduction

Asset prices 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 investors 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 (DGP henceforth, see Barberis and Thaler (2001) for a survey of behavioral finance).

Whichever assumption is used, these models 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 DGP, or the way investors' beliefs depart from it, or both. These frameworks are rather driven by analytic tractability than being empirically grounded, and are prone to potentially large misspecification errors. Misspecification is hard to detect in theory since the true beliefs of investors and the true DGP are unknown. It however translates into the models' inability to fit the time series of asset returns, producing large conditional pricing errors.

[1] https://larspeterhansen.org/research/

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 true DGP or regarding investor rationality or lack thereof. Given a pricing kernel summarizing investors' risk preferences, 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* as they enforce that each test asset at each point in time satisfies the conditional Euler equation restrictions.

Our empirical methodology relies on the non-parametric smoothed empirical likelihood (SEL hereafter) estimator developed by Kitamura et al. (2004). This method approximates the joint 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 simultaneously estimates the multinomial probabilities, i.e. investors' beliefs, and the pricing kernel parameters to maximize the non-parametric log-likelihood, enforcing the constraint that the test assets are perfectly priced. By construction, our methodology thus produces zero pricing errors in both the time series and the cross-sectional dimensions and inherits the desirable properties of a parametric likelihood-based approach such as consistency and asymptotic efficiency (Chamberlain (1987)), while not requiring any parametric assumptions about the underlying DGP. Existing models, on the other hand, often produce large and volatile conditional pricing errors (see, e.g., Nagel and Singleton (2011)) and are prone to potential DGP misspecification.

Since the application of SEL estimation to asset pricing is relatively new, we confirm its empirical performance in benchmark simulations of realistic sample size. We simulate economies corresponding to the standard consumption-CAPM, the external habit model of Campbell and Cochrane (1999) and the long-run risk model of Bansal and Yaron (2004), and show that our estimator successfully and accurately recovers both the preference parameters and the conditional moments of consumption, i.e. investors' beliefs. After recognizing that either beliefs distortions or misspecification of preferences modify the asymptotic behavior of the SEL estimator (see e.g. Kitamura (2003)), we assess the magnitude of these alterations through both a theoretical example where the SEL estimator is obtained in closed-form, and simulations. Our results suggest that, if anything, the SEL tends to underestimate beliefs distortions when present, and the economic plausibility of the preference parameter estimates can serve as a reference for detecting misspecification. We conclude that our empirical results on historical data are likely to be conservative with respect to the amplitude of beliefs distortions.

We apply the methodology on U.S. postwar quarterly data for three alternative preference specifications: time- and state-separable power utility preferences with a constant coefficient of relative risk aversion (CRRA), the external habit formation preferences (Habits), and the Epstein and Zin (1989) recursive preferences (EZ). Our benchmark results use the aggregate stock market returns, the risk-free rate, and returns on portfolios of small, big, growth and value stocks as test assets, and the investors' conditioning set is proxied by past consumption growth and a volatility proxy.

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 latter being negative in almost all time periods. During recessionary episodes, heightened macroeconomic uncertainty manifests itself neither in increased consumption volatility nor in tail risk, but rather in the the appearance of a second mode between the mean and the left tail. These results call into question the widely used assumption of a conditionally Gaussian heteroskedastic DGP in structural asset pricing models.

Second, we show that the price-consistent beliefs are fairly robust to alternative pricing kernels, cross-sections of test assets, and conditioning sets. Pricing kernels implied by recursive preferences and the external habit preferences give similar results to those obtained with the standard power utility preferences, albeit with economically plausible parameter values unlike the latter. Specifically, for the power utility preferences, the SEL point estimates of the CRRA and time discount factor are 145.8 and 1.51, respectively – both arguably outside the corresponding economically plausible ranges. For the habit and recursive preferences, on the other hand, the estimates of the utility curvature parameter are more plausible at 0.51 and 0.64, respectively and at 1.94, and 0.99, respectively, for the time discount factor. For the latter preferences, the estimate of the intertemporal elasticity of substitution is significantly above 1, suggesting preference for the early resolution of uncertainty. To mitigate the potential concern that what we recover is simply a component, unrelated to beliefs, that is missing from the pricing kernel, we show that our price-consistent beliefs about the stock market are positively and significantly correlated with the U.S. Institutional One-Year Confidence Index from Robert Shillers investor survey. All these features are robust to the set of test assets, and 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 qualitative difference to the results.[2]

Third, we provide evidence that commonly assumed consumption growth DGPs cannot reproduce the empirical features captured by the price-consistent beliefs. We consider two widely used specifications: an ARMA(1,1) model and a regime-switching model where the mean of consumption growth varies across latent regimes. For both models, the implied expected consumption growth series miss the high persistence implicit in the price-consistent beliefs, and the conditional skewness is too close to zero and fails to exhibit the cyclical variations of the price-consistent distributions. Our work thus provides an empirical assessment of what these assumed DGPs are missing to explain the behavior of asset prices.

Fourth, we present evidence suggesting the recovered beliefs reflect departures from rationality for the representative investor. We compare the price-consistent beliefs with an *objective* DGP candidate, defined as a standard

[2]Due to space constraints, our complete set of robustness checks is presented in our companion website at https://guillaumeroussellet.shinyapps.io/MeasuringBeliefs/

non-parametric estimator of the conditional density of consumption growth obtained without any asset pricing restrictions. The price-consistent beliefs can be interpreted as the ones that are minimally distorted relative to the objective distribution while satisfying the pricing restrictions. Our estimates show that even relative to this agnostic choice of the objective DGP, price-consistent consumption dynamics are distorted. Our results are indicative of investor exuberance – stemming from overestimating the expected consumption growth rate and underweighting the left tail of the distribution of consumption growth relative to the objective measure.

Our work contributes to two distinct strands of literature. First, the paper contributes to a growing literature departing from the Muth (1961) rational expectations paradigm to explain various aspects of asset market data. In behavioral models, for instance, economic agents are assumed to have certain behavioral biases relative to the rational benchmark (see, e.g., Kahneman and Tversky (1979)) Robust control preferences represent another departure from rational expectations, where investors make consumption-investment decisions from the perspective of the worst-case parametric data-generating process (see, e.g., Hansen and Sargent (2001, 2016)). The literature has shown that robust control can help explain some asset pricing puzzles with economically reasonable risk aversion parameter values (Barillas et al. 2009), volatility of bond risk premia (Piazzesi et al. 2015), or price a broad cross-section of stock protfolios (Wang 2017). Several papers have employed surveys of professional investors to assess beliefs distortions (see, e.g., (Greenwood and Shleifer 2014)). While this undoubtedly represents an important initial step, this approach suffers from noise inherent in this type of data and the fact that investors are rarely asked to report their entire forecasted distribution. Our paper differs from the aforementioned works by using a non-parametric estimation technique to extract possible beliefs distortions, if any, from observed asset prices, thereby avoiding the need for restrictive and potentially misspecified parametric dynamics, or heavy reliance on survey data.

Second, our work builds on a burgeoning literature trying to recover risk premia components with as few assumptions as possible, starting with Ross (2015). Borovicka et al. (2016) show that, under fairly restrictive assumptions, one can recover simultaneously the investors risk preferences and beliefs using a set of Arrow-Debreu securities.[3] The literature has progressed towards relaxing restrictive assumptions, to arrive at an (almost) model-free recovery (Schneider and Trojani 2019), which is akin to identifying the pricing kernel under rational expectations.[4] Unlike this literature, our approach relies on the class of empirical likelihood methods (see Owen (2001)) and does not require complete information on conditional risk neutral probabilities.

While conserving the appealing properties of parametric maximum likelihood estimation (Kitamura 2007), this type of non-parametric estimation method has gained popularity in financial economics. Notable examples of applications include assessing the degree of misspecification of asset pricing models (Almeida and Garcia 2012), generalizing the pricing kernel bounds

[3]see also Carr and Yu (2012), Jensen et al. (2018), Walden (2017) and Qin and Linetsky (2016) and Qin and Linetsky (2017) for related approaches.

[4]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 et al. (2013), Schneider (2015, 2018), Schneider and Trojani (2018) and Orlowski et al. (2018).

(see Backus et al. (2014), Almeida and Garcia (2016)), recovery of the pricing kernel in the presence of pricing errors (Korsaye et al. 2019), or performance evaluation of funds (Almeida et al. 2019). Our approach differs from these works in at least two ways. First, all the above papers interpret the wedge between the objective and the pricing probability measures as the degree of model misspecification, while our method further extracts potential beliefs distortions from the wedge. Second, most works rely on unconditional pricing equations and can only recover unconditional distributions (Ghosh et al. 2016). In turn, our approach is conditional and enforce that the whole time series and cross-section of assets are priced without errors, yielding conditional distributions as a result. A notable exception is the contemporary paper of Chen et al. (2020), who present the theoretical underpinnings of the method used in this paper. They show that information about investors' (potentially distorted) beliefs can be inferred from observed asset prices, given a candidate pricing kernel and a statistical measure of divergence from a family that includes the empirical likelihood minimum distortion measure used in this paper.

II Estimation of Beliefs: Methodology

In this section, we describe the details of our procedure to recover beliefs from asset prices. Throughout, uppercase letters denote random variables, and lowercase letters denote particular realizations of these variables.

II.1 General framework

We assume the absence of arbitrage opportunities, such that a strictly positive stochastic discount factor (SDF), denoted by M_{t+1}, stemming from the representative investor's preferences, exists. The equilibrium gross returns $\mathbf{R}_{t+1} \in \mathbb{R}^k$ of any set of k traded assets, possibly including riskfree rates, satisfy the conditional Euler equations:

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

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 $\underline{\mathcal{F}_t}$, with respect to the probability measure $\mathcal{P}$.[5] Under our assumptions, there also exists a risk neutral measure $\mathcal{Q}$ such that:

$$\frac{\mathrm{d}\mathcal{Q}}{\mathrm{d}\mathcal{P}} = \frac{M_{t+1}}{\mathbb{E}^{\mathcal{P}} \left[M_{t+1} | \underline{\mathcal{F}_t} \right]} \, . \tag{2}$$

Suppose that $\mathcal{P}_0$ denotes the physical (or, objective) probability measure, describing the true data generating process of all variables in the economy. In rational expectations models, the subjective probability measure $\mathcal{P}$ in Equation (1) is equivalent to the objective probability measure $\mathcal{P}_0$, and the SDF is the key variable explaining the behavior of asset prices. However, if investors have any behavioral biases that make their beliefs deviate from rationality, we have:

$$\frac{\mathrm{d}\mathcal{Q}}{\mathrm{d}\mathcal{P}_0} = \frac{\mathrm{d}\mathcal{Q}}{\mathrm{d}\mathcal{P}} \times \frac{\mathrm{d}\mathcal{P}}{\mathrm{d}\mathcal{P}_0} \, . \tag{3}$$

[5] Probability measures are denoted without a $\cdot_t$ subscript by a slight abuse of notation even though they describe joint conditional probability distributions.

The first term of Equation (3) can be identified by the SDF. The change of measure from objective to subjective, $\mathrm{d}\mathcal{P}/\mathrm{d}\mathcal{P}_0$, becomes the missing piece of Equation (3) to price assets. Our objective is to recover this change of measure from observed asset prices at each point in time to uncover the conditional distributions formed by investors given $\underline{\mathcal{F}_t}$, with the least possible assumptions besides the SDF specification summarizing investors' risk preferences.

II.2 Estimating beliefs using the smoothed empirical likelihood approach

Our identification approach relies on the *smoothed empirical likelihood* (SEL) estimation approach, developed by Kitamura et al. (2004). We detail below the general procedure and how it fits into our framework.

We assume that the SDF of economic agents can be identified as a parametric function of consumption growth, denoted by $G_{t+1} \equiv C_{t+1}/C_t$, and a set of other risk factors that we denote by Y_{t+1}:

$$M_{t+1} = M\left(G_{t+1},\, Y_{t+1};\, \theta_0\right)\,, \tag{4}$$

where θ_0 is the true value of the vector of parameters driving the SDF. We assume that the information set of the investors at time t can be summarized by a finite set of risk factors $X_t \in \mathbb{R}^n$. At each point in time, the representative investor observes the realizations of consumption growth g_t, other variables in the SDF y_t, asset returns $\mathbf{r}_t$, and the conditioning variables x_t (that may also contain g_t, y_t and $\mathbf{r}_t$). The econometrician observes a sample of these realizations, from $t = \{1, \ldots, T\}$, and we assume that she knows the true parameters θ_0 driving the SDF (we relax this assumption in the next section).

We assume that the probability measures $\mathcal{P}$ and $\mathcal{P}_0$ are absolutely continuous with respect to Lebesgue. In the spirit of non-parametric density estimators, the SEL approach approximates these conditional probability measures by $(T \times T)$ transition matrices, with as many possible states as observations: the entry (i, j) provides the probability to switch from the state observed at $t = i$ to that observed at $t' = j$. We denote by $p_{i,j}$ and $p_{i,j}^{(0)}$ the subjective and objective conditional probabilities, respectively, of observing the joint outcome $(g_j, y_j, \mathbf{r}_j, x_j)$ at time $t+1$ given that $(g_i, y_i, \mathbf{r}_i, x_i)$ was realized at date t.

As stated above, our goal is to identify the change of measure $\mathrm{d}\mathcal{P}/\mathrm{d}\mathcal{P}_0$. Since $\mathcal{P}_0$ represents the true data-generating process of the variables in our system, we can identify it by a standard non-parametric kernel density estimator. In particular, we have:

$$\forall\, i, j \in \{1, \ldots, T\}\,, \quad p_{i,j}^{(0)} = \frac{\mathcal{K}\left(\frac{x_i - x_j}{b_T}\right)}{\sum_{t=1}^{T} \mathcal{K}\left(\frac{x_i - x_t}{b_T}\right)}, \tag{5}$$

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.[6] The intuition of this estimator is that realizations of conditioning variables that are close to each other tend to produce similar outcomes next period, where the notion of closeness is controlled by the bandwidth parameter. In practice, we will use the Epanechnikov kernel, but our results are relatively insensitive to that

[6] $\mathcal{K}$ should satisfy Assumption 3.3 in Kitamura et al. (2004), that is restated here for convenience. For $X = (X^{(1)}, X^{(2)}, ..., X^{(n)})$, let $\mathcal{K} = \prod_{i=1}^{n} \kappa(X^{(i)})$. Here $\kappa : \mathbb{R} \to \mathbb{R}_+$ is a continuously differentiable p.d.f. with support $[-1, 1]$. κ is symmetric about the origin, and for some $\alpha \in (0, 1)$ is bounded away from zero on $[-a, a]$. b_T is a null sequence of positive numbers such that $Tb_T \to \infty$. See Assumption 3.7 in Kitamura et al. (2004) for additional restrictions on the choice of b_T.

choice.

To identify the subjective conditional probability measure $\mathcal{P}$, we rely on the information provided by asset returns. In particular, using the transition matrix formulation and an assumed specification of investors' preferences $M\left(G_{t+1}\,, Y_{t+1}\,;\,\theta_0\right)$, the empirical Euler equation (1) can be rewritten:

$$\mathbb{E}^{\mathcal{P}}\left[M\left(G_{t+1}\,, Y_{t+1}\,;\,\theta_0\right)\mathbf{R}_{t+1}\middle|\underline{\mathcal{F}_t}\right]=\mathbf{1}_k\Rightarrow\sum_{j=1}^{T}p_{t,j}\cdot M\left(g_j\,, y_j\,;\,\theta_0\right)\mathbf{r}_j=\mathbf{1}_k\,. \quad (6)$$

When the number of assets k is lower than the number of observed dates T, there are many sets of probabilities $\{p_{t,j}(\theta_0)\}_{t,j=1}^{T}$ that satisfy Equation (6). The idea of the SEL estimator is to select the probabilities that minimize the Kullback-Leibler Information Criterion divergence (KLIC, or relative entropy) between the two measures $\mathcal{P}_0$ and $\mathcal{P}$, such that the constraints (6) are satisfied (see White (1982)). More formally, the conditional probabilities $\{p_{t,j}\}_{t,j=1}^{T}$ are such that they belong to the T-simplex Δ and that: $\forall t\in\{1,\ldots,T\}$,

$$\widehat{p}_{t,\cdot}(\theta_0)=\underset{\{p_{t,\cdot}\}\in\Delta}{\arg\min}\;\sum_{j=1}^{T}p_{t,j}^{(0)}\cdot\log\left(\frac{p_{t,j}^{(0)}}{p_{t,j}}\right)\;\text{ s.t. }\sum_{j=1}^{T}p_{t,j}\cdot M\left(g_j,\,y_j;\,\theta_0\right)\mathbf{r}_j=\mathbf{1}_k, \quad (7)$$

where $p_{t,\cdot}$ denotes the $T-$dimensional vector of probabilities $(p_{t,1},...,p_{t,T})$. Note that this KLIC divergence is non-negative and will be exactly equal to zero if and only if the objective and subjective probabilities coincide. Thus, the SEL approach searches for an estimate of subjective beliefs that is as close as possible – in the information-theoretic sense – to the objective one, while also requiring that the estimated beliefs satisfy the pricing restrictions given by the conditional Euler equations. Note that the SEL estimator is the one that maximizes the non-parametric (multinomial) likelihood function subject to the asset pricing constraints.[7] The recovered conditional probabilities $\{\widehat{p}_{t,j}(\theta_0)\}_{t,j=1}^{T}$ represent the beliefs of the representative investor that are consistent with observed asset prices. Hereafter, we refer to this transition matrix as the *price-consistent beliefs*.

The problem formulated in Equation (7) has a closed-form solution (see Appendix A.1), given by: $\forall t, j \in \{1, \ldots, T\}$,

$$\widehat{p}_{t,j}(\theta_0) = \frac{1}{1 + \widehat{\lambda}_t(\theta_0)' \left[M(g_j,\, y_j;\, \theta_0)\, \mathbf{r}_j - \mathbf{1}_k\right]} \times p_{t,j}^{(0)}\,, \tag{8}$$

where $\widehat{\lambda}_t(\theta_0) \in \mathbb{R}^k$ are the Lagrange multipliers associated with the conditional Euler equation constraints, and solve the following unconstrained maximization problem:

$$\widehat{\lambda}_t(\theta_0) = \underset{\lambda_t \in \mathbb{R}^k}{\arg\max} \sum_{j=1}^{T} p_{t,j}^{(0)} \log\left[1 + \lambda_t'\big(M(g_j, y_j; \theta_0)\, \mathbf{r}_j - \mathbf{1}_k\big)\right]. \tag{9}$$

Looking at Equation (8) we see that the SEL estimator distorts the objective conditional probabilities $p_{t,j}^{(0)}$ such that the asset pricing constraints (6) are met. If these constraints held with $p_{t,j}^{(0)}$, the Lagrange multipliers would be null and Equation (8) would provide equality between the objective and subjective probabilities.

Several properties of the estimator are worth noting. First, the SEL approach only requires the estimation of $(T \times k)$ Lagrange multipliers, and not of $T \times T$ probabilities. 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

[7] This can be easily checked noticing that minimizing $\sum_{j=1}^{T} p_{t,j}^{(0)} \times \log\left(\frac{p_{t,j}^{(0)}}{p_{t,j}}\right)$ is equivalent to maximizing $\sum_{j=1}^{T} p_{t,j}^{(0)} \log(p_{t,j})$.

price (see Equations (8) and (9)).[8] Second, the SEL estimator does not require any parametric assumption about the conditional distribution $\mathcal{P}$. Instead, it remains as agnostic as possible.

II.3 Parameter estimation and inference

In the previous section, we have assumed that the econometrician knows the true parameters θ_0 driving the SDF. This section presents the estimation and inference procedure for θ_0, and the statistical properties of the estimated conditional probabilities. We assume hereby that there is no misspecification, and discuss this assumption in the next section.

There exists a set Θ of admissible parameters, such that $\theta_0 \in \Theta$. The SEL estimator of θ_0 is defined through the concentrated likelihood as:

$$\widehat{\theta} = \arg\max_{\theta \in \Theta} \quad \sum_{t=1}^{T} \sum_{j=1}^{T} \mathbb{T}_{t,j} \times p_{t,j}^{(0)} \log(\widehat{p}_{t,j}(\theta)). \tag{10}$$

where $\mathbb{T}_{t,j}$ is a sequence of trimming functions, incorporated in the log-likelihood to deal with the well-known denominator problem associated with kernel estimators.[9] Kitamura et al. (2004) show that the SEL approach delivers a consistent and asymptotically efficient estimator of θ, i.e. the estimator achieves the semi-parametric efficiency bound in Chamberlain (1987) for conditional moment restriction models, and we have:

$$\sqrt{T}\left(\widehat{\theta} - \theta_0\right) \xrightarrow{d} \mathcal{N}\left(0, I^{-1}\left(\theta_0\right)\right) , \tag{11}$$

[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)).

[9]In practice, to avoid numerical instability, Owen (2001) recommends a smooth transformation of the log-likelihood presented in Appendix A.2.

where $I(\theta_0)$ is the Fisher information matrix, whose estimator is $\widehat{I(\theta)}$ (see Appendix A.3). In this case, the authors also show that the conditional moments derived from the conditional probabilities $\widehat{p}_{t,j}(\widehat{\theta})$ are consistent and provide an efficiency gain from a standard non-parametric kernel density estimator (see also Brown and Newey (1998), Brown and Newey (2002)). The superior numerical properties of an SEL-type estimator of the conditional distribution function are also demonstrated via simulations in Hall et al. (1999).

II.4 Sources and consequences of misspecification

Misspecification modifies the behavior of the estimators presented above. We list three sources of misspecification in our setup: *(i)* the information set spanned by X_t is only a subset of the investors' information set $\mathcal{F}_t$, *(ii)* the investors' preferences are misspecified, and *(iii)* the subjective beliefs $\mathcal{P}$ are different from the true DGP $\mathcal{P}_0$. We explain below the potential consequences of these misspecifications.

Accessing the complete information set of economic agents is a daunting task, since investors may rely on multiple private and public sources to perform their investment decisions. Fortunately, misspecifying the information set of the agents is the most innocuous of the three possible sources listed above. In this case, the SEL estimator only suffers from an asymptotic efficiency loss resulting from a suboptimal instruments choice, in the spirit of GMM estimators. Consistency is unaffected, and conditional moments are defined with respect to X_t instead of the full set $\mathcal{F}_t$.

In comparison, the misspecification of the preferences and/or beliefs distortion modify more profoundly the behavior of the SEL estimators. These two sources of misspecification essentially act in the same fashion, in that they drive a wedge between $\mathcal{P}$ and $\mathcal{P}_0$. Under correct specification, the Euler equation (1) holds under the physical measure $\mathcal{P}_0$. When there are beliefs distortions and potential SDF misspecification, on the other hand, the change of measure (3) can be decomposed:

$$\frac{\mathrm{d}\mathcal{Q}}{\mathrm{d}\mathcal{P}_0} = \frac{\mathrm{d}\mathcal{Q}}{\mathrm{d}\mathcal{P}}(\theta) \times \frac{\mathrm{d}\mathcal{P}}{\mathrm{d}\mathcal{P}^*}(\theta) \times \frac{\mathrm{d}\mathcal{P}^*}{\mathrm{d}\mathcal{P}_0}\,, \tag{12}$$

where $\mathrm{d}\mathcal{Q}/\mathrm{d}\mathcal{P}^*$ represents the change of measure implied by the true SDF of the representative agent. The three terms in Equation (12) represent, respectively, the (misspecified) SDF, the misspecification error, and the actual beliefs distortion $\mathrm{d}\mathcal{P}^*/\mathrm{d}\mathcal{P}_0$. If there is only SDF misspecification but no beliefs distortion, then $\mathcal{P}^* = \mathcal{P}_0$. If there is only beliefs distortions but no SDF misspecification, then $\mathcal{P} = \mathcal{P}^*$. In both cases, the Euler equation holds under a measure different from $\mathcal{P}_0$. However, since the SEL is a minimum contrast estimator, it remains asymptotically consistent and converges to a pseudo-true value θ^*. As shown by Kitamura (2003), under certain regularity conditions:

$$\sqrt{T}\left(\widehat{\theta} - \theta^*\right) \xrightarrow{d} \mathcal{N}\left(0, \Omega^{-1}\left(\theta^*\right)\right)\,, \tag{13}$$

where $\Omega\left(\theta^*\right)$ has a sandwich form provided in Appendix A.3 (see also Almeida and Garcia (2012)). Note that, in general, the pseudo-true value θ^* will be different from the true value θ_0. In this case, the SEL-estimated distribution $\mathcal{P}$ adjusts the standard kernel density estimator so that it converges to the pseudo-true measure. This property parallels that of parametric maximum likelihood estimators for misspecified models (see, e.g., White (1982), Vuong

(1989)).

II.5 An analytical example of misspecification

To further illustrate the consequences of these potential misspecifications, Appendix A.4 presents an example of the SEL estimator in a case where all quantities can be derived analytically. We consider a model where agents have CRRA preferences such that the true SDF is given by $\delta G_{t+1}^{-\theta_0}$. Consumption growth has two states, g_H and g_L, which have respective objective probabilities $(\pi_0,\, 1-\pi_0)$, to be drawn independently over time. Subjective beliefs are represented by $(\pi,\, 1-\pi)$, potentially different from the objective probabilities, and preference misspecification is obtained when $\theta \neq \theta_0$. Last, we only consider a real riskfree asset whose equilibrium gross rate is given by $\frac{1}{\delta\left[\pi g_H^{-\theta_0}+(1-\pi)g_L^{-\theta_0}\right]}$. We show in Appendix A.4 that the SEL produces a closed-form estimate for the subjective probabilities that does not depend on sample size and, for a given θ, is equal to:

$$\widehat{\pi}(\theta) = \frac{g_L^{-\theta_0} - g_H^{-\theta_0}}{g_L^{-\theta} - g_H^{-\theta}} \times \pi + \frac{g_L^{-\theta} - g_L^{-\theta_0}}{g_L^{-\theta} - g_H^{-\theta}}\,. \tag{14}$$

When there is no preference misspecification ($\theta = \theta_0$), the SEL probability estimate will equal the true subjective beliefs π. On the other hand, any other value of θ will, in general, produce a bias. In this economy, the log-likelihood associated with the estimation of θ_0 (Equation (7)) is easily expressed but its maximization problem has no closed-form solution. We can, however, easily check whether θ_0 asymptotically verifies the first-order conditions. The log-likelihood score, evaluated at θ_0, is of the same sign as $\ell_T(\theta_0)$, which converges to (see Appendix A.4):

$$\ell_T(\theta_0) \underset{T\to+\infty}{\overset{\mathcal{P}_0}{\longrightarrow}} (\pi - \pi_0)\, \mathbb{E}^{\mathcal{P}_0}\left[M_t(G_t;\, \theta_0) \log\left(G_t\right)\right] . \tag{15}$$

Equation (15) shows that, asymptotically, θ_0 maximizes the log-likelihood if and only if there are no beliefs distortions, i.e. $\pi = \pi_0$. Conversely, if beliefs are distorted with e.g. $\pi > \pi_0$, the right-hand side of Equation (15) is positive and the maximum likelihood procedure will increase $\widehat{\theta}$ to reproduce the increase of the riskfree rate.[10] We also show in Appendix A.4 that the estimated probability $\widehat{\pi}(\widehat{\theta})$ decreases to get closer to the objective probability π_0, as the formulation of the estimator suggests. Thus, if anything, provided our SDF has the correct functional form, the SEL will tend to underestimate the size of beliefs distortions.

III SEL performance in simulated economies

This section presents an empirical exploration of the performance of the SEL estimator in simulated economies, for both well-specified and misspecified models. We consider all types of misspecification described above, i.e. conditioning set, preferences, and distorted beliefs.

III.1 Simulation details

We consider three baseline representative agent economies. Model (I) is an economy where the representative agent has power utility preferences with a

[10] It is easy to check that the term $\mathbb{E}^{\mathcal{P}_0}\left[M_t(G_t;\, \theta_0) \log\left(G_t\right)\right]$ is positive and equal to the differential of the risk free rate with respect to θ, evaluated at θ_0.

constant coefficient of relative risk aversion (CRRA) and the aggregate consumption growth is i.i.d. Gaussian. Model (II) is the external habit model of Campbell and Cochrane (1999). Model (III) is the Bansal and Yaron (2004) long-run risks model. Details of these models and the corresponding SDFs are provided in Appendix A.5.[11]

We calibrate Model (I) using the moments of consumption growth in the historical quarterly sample and set the CRRA parameter to 10, and use the Campbell and Cochrane (1999) and Bansal and Yaron (2004) original calibrations for Models (II) and (III), respectively. All simulated trajectories are of length 267, the number of quarters we observe in our post-war data. For each case, we simulate 500 trajectories, perform the SEL estimation of both the SDF parameters and subjective distributions using the equilibrium market return and the risk free rate as the test assets. In Model (I), conditioning is irrelevant since the economy is i.i.d. For models (II) and (III), our conditioning set consists of exponentially-weighted moving averages of past consumption growth on the one hand, and of squared residuals obtained from an AR(1) regression on consumption growth. This accommodates for the unobservable state variables in the investors' information set.[12]

We report the mean, median, and 90% confidence intervals of the time series of the conditional mean, volatility, and skewness of consumption growth

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

[12] Our exponentially weighted variables are computed as $X_t^{(EW)} = \alpha X_t + (1-\alpha) X_{t-1}^{(EW)}$. We set $\alpha = 0.28$, whereby the past 13 quarters receive 99% of the weight. Our results are fairly insensitive to the value of α.

in each sample. We use the estimated conditional probabilities to compute the moments as:

$$\widehat{\mathbb{E}}^{\mathcal{P}}\left[\{\log(G_{t+1})\}^k \left| X_t\right.\right] = \sum_{j=1}^{T} \widehat{p}_{t,j}(\widehat{\theta}) \cdot \{\log(g_j)\}^k \ . \tag{16}$$

When the true conditional moments are time-varying (Model III), we report the statistics on the time series of errors, that is the difference between the estimated and true moments: $\widehat{\mathbb{E}}^{\mathcal{P}}\left[\{\log(G_{t+1})\}^k \left| X_t\right.\right] - \mathbb{E}^{\mathcal{P}_0}\left[\{\log(G_{t+1})\}^k \left| \underline{\mathcal{F}_t}\right.\right]$.

We explore the results with respect to several dimensions. First, we ask whether the SEL is able to recover the conditional moments of consumption growth in the three models when beliefs are not distorted. Second, we investigate the performance of the method when beliefs are distorted but the preferences are well-specified. Third, we explore the effects of misspecification of preferences in the absence of any beliefs distortions. The cases considered are summarized in Table 6 in Appendix A.5.4.

III.2 Well-specified models

We first consider the case of well-specified models. We initially focus on the cases where consumption growth is i.i.d. Gaussian, i.e. Models (I.1) and (II.1) of Table 6. Because there is neither SDF misspecification nor belief distortions, these cases present an upper bound for the finite sample performance of the SEL estimators.

The results are presented in Table 1. Consider first Model (I.1). Panel A, Row 1 shows that the SEL approach accurately estimates the CRRA parameter in empirically realistic sample sizes – the median value of this parameter

Table 1 – Estimation results for well-specified models

Model	(A) $\widehat{\theta}$ estimate					(B) Consumption mean				
	true	mean	5%	50%	95%	true	mean	5%	50%	95%
(I.1)	10.0	10.0	9.97	10.0	10.03	0.48	0.48	0.48	0.48	0.48
(II.1)	2.0	2.15	1.40	2.10	3.20	0.48	0.48	0.44	0.48	0.52
(III)	10.0	8.49	0.00	8.10	17.61	0.00	0.00	-0.21	0.00	0.22

Model	(C) Consumption volatility					(D) Consumption skewness				
	true	mean	5%	50%	95%	true	mean	5%	50%	95%
(I.1)	0.51	0.51	0.48	0.51	0.55	0.0	0.0	-0.25	0.01	0.25
(II.1)	0.51	0.50	0.47	0.51	0.54	0.0	-0.00	-0.16	0.00	0.17
(III)	0.00	0.00	-0.11	0.00	0.12	0.00	0.017	-0.19	0.004	0.23

The table presents the statistics of the SDF parameter (Panel A), and the mean (Panel B), volatility (Panel C), and skewness (Panel D) of consumption growth, obtained using the SEL estimator, computed across 500 simulated samples. Each row presents the results for a specific hypothetical representative agent economy. Specifically, Model (I.1) corresponds to an economy with i.i.d. lognormal consumption growth, in which a representative agent has power utility preferences with a constant CRRA. Model (II.1) corresponds to the external habit model of Campbell and Cochrane (1999) and Model (III) corresponds to the long run risks model of Bansal and Yaron (2004).

across the simulated samples coincides with the true value of 10, and its 90% confidence interval, covering the range $[9.97, 10.03]$, is very tight. Panels B–D show that the SEL approach is also very successful at recovering the underlying distribution of consumption growth, including its mean, volatility, and skewness.

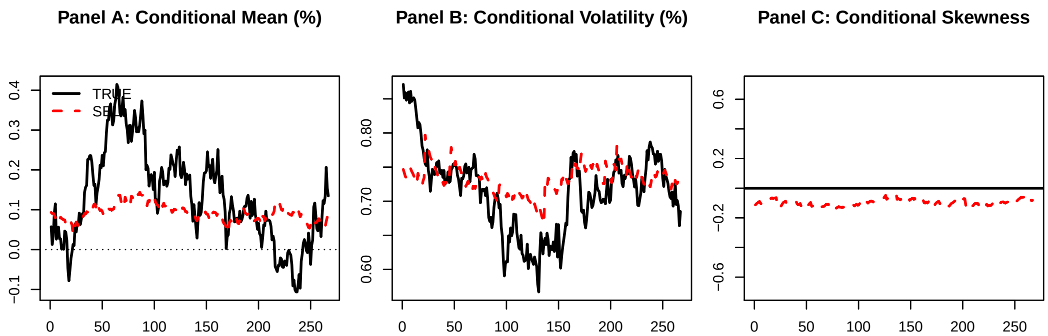
Next we turn to model Model (II.1). Table 1, Panel A, Row 2 shows that the SEL accurately estimates the utility curvature parameter – the mean and median values of the parameter are 2.15 and 2.10, respectively, very close to the true value of 2.0. Panels B–D show that the SEL also accurately recovers the true underlying distribution of consumption growth. Specifically, the

SEL-implied time series of the expected consumption growth and volatility lie in the 90% confidence intervals $[0.44\%, 0.52\%]$, and $[0.47\%, 0.54\%]$, respectively (non-annualized quarterly figures). The SEL also accurately identifies the Gaussianity of the true dynamics – the mean and median SEL-implied skewness of consumption growth are both 0.00 across the simulations and the 90% confidence interval is tight. Thus, the SEL correctly identifies conditional moments as virtually constant over time.

We next turn to the performance of the SEL estimator when the model is well-specified but the moments of consumption growth are time-varying. This corresponds to the long-run risks economy with Epstein-Zin preferences (Model III). Table 1, Row 3 presents the finite-sample statistics of the estimation errors for the moments of consumption growth, aggregated over the simulated samples at each point in time, as well as the estimated risk aversion parameter. As with the external habit economy, the table shows that the SEL approach is quite successful at recovering the risk aversion parameter as well as the marginal moments of consumption growth, although confidence bounds are wider due to the high persistence of the consumption growth process.

Figure 1 presents, for a randomly chosen sample, the time series of the SEL-recovered moments (red-dashed line) along with the true time series of these moments (black solid line). The correlations between the recovered mean and volatility estimates and the true series are high at 56.0% and 54.1%, respectively. For the conditional skewness, the true value is identically equal to zero in all time periods, compared to our slightly negative but flat estimate. Overall, the results suggest that the SEL approach can recover the consumption growth dynamics from asset prices with sufficient precision.

Figure 1 – Estimation results on simulated long-run risk economy with well-specified preferences (Model III)



Notes: This figure plot the time series of true conditional moments of consumption growth (black solid line) and the corresponding SEL-estimated time series (red-dashed line), on one simulated trajectory of the long-run risk model. Our calibration uses parameters from Bansal and Yaron (2004).

III.3 Bias implied by belief distortions

We now investigate what happens when beliefs are distorted. We assume that Model (I) holds (i.i.d. consumption growth and CRRA preferences), but the representative investor has distorted beliefs about the mean and volatility of consumption growth. Specifically, she underestimates the mean consumption growth and overestimates the volatility of consumption growth by χ_μ and χ_σ percent, respectively, where $\chi_\mu = \{15\%, 50\%, 75\%\}$ and $\chi_\sigma = \{0\%, 15\%, 75\%\}$. The results are presented in Table 2.

Consider first $\chi_\mu = 15\%$, i.e. the subjective mean is 0.41% compared with the objective value of 0.48%. Panel (A), Row 2 shows that, when there are no

Table 2 – Estimation results on simulated economy with distorted beliefs (Model I.2)

		(A) $\widehat{\theta}$ estimate					(B) Consumption mean				
χ_μ	χ_σ	θ_0	mean	5%	50%	95%	μ^*	mean	5%	50%	95%
0%	0%	10	10	9.97	10	10.03	0.48	0.48	0.48	0.48	0.48
15%	0%	10	9.99	9.95	9.99	10.03	0.41	0.41	0.41	0.41	0.41
50%	0%	10	9.84	9.69	9.85	9.97	0.24	0.25	0.24	0.25	0.26
75%	0%	10	9.49	9.14	9.51	9.77	0.12	0.13	0.13	0.13	0.14
15%	15%	10	10.08	10.03	10.08	10.13	0.41	0.4	0.4	0.4	0.41
50%	15%	10	10	9.84	10	10.13	0.24	0.24	0.23	0.24	0.25
75%	15%	10	9.73	9.25	9.76	10.09	0.12	0.13	0.12	0.13	0.14
15%	75%	10	10.64	10.57	10.64	10.7	0.41	0.36	0.36	0.36	0.37
50%	75%	10	10.94	10.66	10.96	11.17	0.24	0.2	0.2	0.2	0.21
75%	75%	10	11.77	10.86	11.81	12.61	0.12	0.09	0.08	0.09	0.11

		(C) Consumption volatility					(D) Consumption skewness				
χ_μ	χ_σ	σ^*	mean	5%	50%	95%	s^*	mean	5%	50%	95%
0%	0%	0.51	0.51	0.48	0.51	0.55	0	0	-0.25	0.01	0.25
15%	0%	0.51	0.52	0.48	0.52	0.55	0	-0.01	-0.28	-0.01	0.24
50%	0%	0.51	0.59	0.53	0.58	0.66	0	-0.12	-0.58	-0.1	0.28
75%	0%	0.51	0.64	0.57	0.64	0.73	0	-0.08	-0.52	-0.06	0.32
15%	15%	0.59	0.52	0.48	0.52	0.56	0	0	-0.3	0.01	0.26
50%	15%	0.59	0.59	0.53	0.58	0.65	0	-0.1	-0.59	-0.07	0.29
75%	15%	0.59	0.65	0.57	0.64	0.75	0	-0.09	-0.6	-0.07	0.31
15%	75%	0.89	0.53	0.49	0.53	0.58	0	-0.01	-0.34	0.01	0.25
50%	75%	0.89	0.61	0.54	0.6	0.69	0	-0.11	-0.6	-0.08	0.28
75%	75%	0.89	0.67	0.58	0.66	0.76	0	-0.06	-0.55	-0.04	0.37

The table presents the statistics of the SDF parameter (Panel A), and the mean (Panel B), volatility (Panel C), and skewness (Panel D) of consumption growth, obtained using the SEL estimator, computed across 500 simulated samples. The samples are simulated from a hypothetical endowment economy with i.i.d. lognormal consumption growth, in which a representative agent with power utility preferences is pessimistic and underestimates the mean consumption growth by χ_μ percent and overestimates the volatility by χ_σ percent. Variables with a * represents the subjective moments of consumption growth.

belief distortions about the volatility, the median CRRA parameter estimate across the simulations is 9.99, virtually indistinguishable from the true value of 10. Panels (B)–(D) show that the SEL method is successful at capturing the investors' subjective beliefs: the SEL-implied mean of consumption growth has a median value of 0.41% (Panel B, Row 2), identical to the subjective mean and smaller than the objective value of 0.48%; the estimated volatility of consumption growth has a median of 0.52% across the simulated samples, very close to the historical value of 0.51 (Panel C, Row 2); and the median coefficient of skewness across the simulations is -0.01, very close to the true value of 0.00 (Panel D, Row 2).

Rows 5 and 8 show that very similar results are obtained when there are also 15% and 75% beliefs distortions in the volatility, respectively. Specifically, the median estimates of the CRRA parameter remain very close to the true value of 10 with very tight confidence bands, the median SEL-implied means of consumption growth are very close to the subjective mean and statistically and economically smaller than the true objective mean, and the median skewness is close to zero. Note that, when there are simultaneous beliefs distortions in the mean and the volatility of consumption growth, the SEL successfully identifies the former but not the latter because the volatility of consumption growth has a smaller impact on asset prices in this economy. The remaining rows of the table show that the above performance of the SEL estimator is maintained in the presence of more severe belief distortions.

III.4 Bias implied by misspecified preferences

Finally, we turn to the performance of the SEL estimator in the presence of misspecification in preferences. Specifically, we consider Model (II), i.e. the external habit model, but the econometrician erroneously applies the SEL methodology to a power utility SDF. Results are qualitatively similar when the simulated model is model (III).

Table 3 – Estimation results on simulated economy with misspecified SDF (Model II.2)

(a) $\widehat{\theta}$ estimate					(b) Consumption mean				
true	mean	5%	50%	95%	true	mean	5%	50%	95%
2.0	75.10	64.85	74.50	87.15	0.48	0.48	0.43	0.48	0.54
(c) Consumption volatility					(d) Consumption skewness				
true	mean	5%	50%	95%	true	mean	5%	50%	95%
0.51	0.50	0.47	0.51	0.54	0.00	-0.002	-0.25	0.003	0.23

The table presents the statistics of the SDF parameter (Panel A), and the mean (Panel B), volatility (Panel C), and skewness (Panel D) of consumption growth, obtained using the SEL estimator, computed across 500 simulated samples. The samples are simulated from a hypothetical endowment economy with i.i.d. lognormal consumption growth, in which a representative agent has external habit preferences. The econometrician wrongly assumes power utility preferences and applies the SEL method.

The simulation results, presented in Table 3, show that the method retains its ability to accurately recover the dynamics in the misspecified setting by estimating a much higher value of the utility curvature parameter (that corresponds to the CRRA in the misspecified power utility SDF): the median value of the parameter estimate across simulations is 74.5 and its 90% confidence interval covers the range $[64.85, 87.15]$, while its true value is only 2.0. Indeed, the external habit SDF is more volatile than the one implied by power utility for reasonable CRRA parameter values. Thus the SEL inflates the curvature parameter to reproduce the increased SDF volatility. Despite these implausible estimates, consumption mean, volatility and skewness possess very tight 90% confidence bands, always centered around the true value. The estimated conditional moments are still virtually constant over time, such that the i.i.d. feature of the data is well identified despite the misspecification. This provides anecdotal evidence that the SEL is able to recover consumption dynamics despite misspecification in preferences, and that implausible parameter estimates are merely a sign of such misspecification.

IV Price-consistent beliefs in the U.S.

IV.1 Estimation details

We select quarterly post-war U.S. data from 1947:Q1 to 2013:Q4. Real aggregate consumption entering the SDF specifications is proxied as the per capita real personal consumption expenditures on non-durables and services. Our baseline specification follows the inputs used for the simulated economies. For the basis assets, we use the real market returns computed as the value-weighted index of all stocks on the NYSE, AMEX, and NASDAQ from CRSP, the real risk free rate obtained as a deflated measure of the 3-month nominal sovereign Tbill yield, and portfolios of small, big, growth, and value stocks obtained from Ken French website. More details about the data can be found in Appendix A.6.

We consider the three types of preferences used in the simulated economies (power utility with a constant CRRA, external habit, and Epstein-Zin, as detailed in Appendix A.7), and the same two exponentially-smoothed conditioning variables for all three models (past consumption growth and past squared consumption shocks). 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.

IV.2 Pricing kernel parameter estimates

Table 4 presents the SEL point estimates of the preference parameters, along with standard errors in parentheses, for the three different SDF specifications considered. Column 2 shows that, for the standard power utility preferences, the estimate of the risk aversion coefficient is 145.8. Although the standard error of the estimate is also high at 53.67, the 90% confidence interval does not include values lower than 57.5. This result echoes the well documented fact that this model requires implausibly high levels of risk aversion to explain the equity premium (e.g. Mehra and Prescott (1985), Weil (1989)). What our finding adds in this context is the observation that the model cannot be resurrected even by allowing investors' subjective beliefs to differ from the physical data generating process. The estimate of the time discount factor is greater than 1 at 1.51, that may also be argued as being outside the economically meaningful range.

Column 3 presents the results for the external habit SDF. Note that this SDF depends not only on consumption growth, but also on the growth in the surplus consumption ratio. To avoid making any additional assumptions other than the functional form of the SDF, we proxy the unobserved habit level as an exponentially-weighted moving average of past consumption levels, where the moving average (or, mean-reversion) coefficient is estimated, alongside the utility curvature parameter and the time discount factor. The parameter estimates are more reasonable at 0.513 for the utility curvature parameter and 0.635 for the time discount factor, with standard errors of 0.13 and 0.14, respectively. The moving average coefficient for the habit level is estimated at

Table 4 – SEL parameter estimates by SDF

	CRRA	HABITS	EPSTEIN-ZIN
δ	1.51	0.635	0.989
	[0.118]	[0.139]	[0.004]
γ	145.81	0.513	1.942
	[53.7]	[0.125]	[0.849]
Mean-reversion		0.5	
		[–]	
EIS			3
			[–]
KLIC	5.558	4.639	4.668

Notes: This table presents point estimates from the SEL maximum likelihood estimation along with the standard deviation of estimation in brackets. δ is the subjective discount factor parameter, γ is the utility curvature parameter, "Mean-reversion" is the mean-reversion of habits, and "EIS" is the intertemporal elasticity of substitution. Standard deviations are estimated by computation of the hessian of the log-likelihood. The '–' indicates that the estimate reaches the bound that we imposed for estimation.

0.5, suggesting that habits are formed by placing 99% weight on the immediate past 7 quarters, i.e. close to 2 years, of consumption realizations.

Column 4 presents the results for the SDF implied by recursive preferences. In this case the SDF depends on the return on total wealth that is unobservable and hard to proxy because a large fraction of total wealth is accounted for by human capital that is not directly traded. Since we want to avoid making any extra assumptions, and in the absence of an obvious proxy, we simply proxy the return on total wealth with the return on the stock market. Our results should, therefore, be interpreted with caution. The estimates of the risk aversion and the elasticity of intertemporal substitution (EIS) parameters are 1.94 and 3.0, respectively. The risk aversion is higher than the inverse of the EIS, suggesting a preference for the early resolution of uncertainty. However, the standard error for this latter parameter is large.

As discussed in the previous section, we can use these parameter estimates to gauge the misspecification of the SDF by assessing the economic plausibility of the estimates. In light of the implausible parameter estimates obtained with the CRRA preferences, we focus on the external habit and Epstein-Zin preferences hereafter.

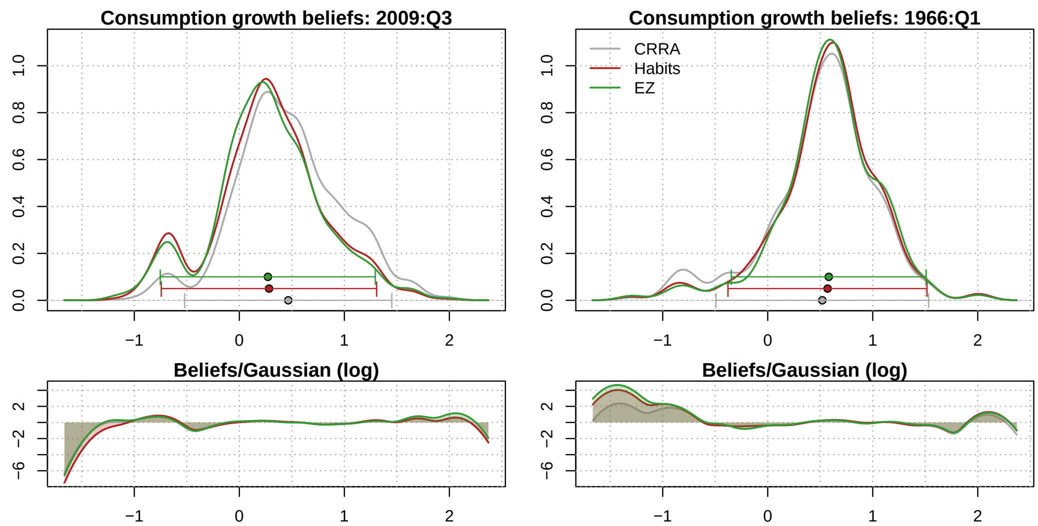
IV.3 Beliefs about consumption growth

To provide an initial snapshot of the results, we present the one-quarter ahead conditional distributions of consumption growth, recovered for all three choices of preferences, in two extreme states – the peak of the recent financial crisis (2009:Q3 as of Q2) and high economic growth (1966:Q1 as of 1965:Q4).

The (smoothed) distributions are presented in Figure 2 along with their deviations relative to Gaussian distributions with the same mean and variance (bottom panel). Blatant from Figure 2 is the strong non-Gaussianity of the price-consistent densities of consumption growth, whatever the choice of SDF. For both good and bad states, there are significant deviations of the estimated beliefs with respect to a Gaussian distribution with the same mean and volatility. Thus, any DGP featuring conditional Gaussianity is unlikely to reproduce investors' beliefs.

The nature of these deviations is, nonetheless, time-dependent. Given the implausible risk-aversion estimated with CRRA preferences, we focus on the results obtained with the habit preferences (similar results are obtained with the recursive preferences). Figure 2 shows that for the habit preferences, in

Figure 2 – Price-consistent conditional densities of consumption growth



Notes: The figure plots the conditional densities of consumption growth in 2009:Q3 and 1966:Q1, and their log-divergence with a Gaussian density with the same mean and variance. Points represent the mean of these distributions and horizontal bars are the 95% confidence bands implied by a Gaussian distribution. The conditional densities are obtained using the SEL probabilities estimated for these quarters. The pricing kernels correspond to CRRA (grey), habits (red) and Epstein-Zin (green). Test assets include market returns, Tbill rate and small, big, growth and value portfolios. The conditioning set includes past consumption growth and a consumption growth volatility proxy. Sample is 1947:Q2-2013:Q4.

the wake of the 2008 financial crisis, the SEL-implied density is bi-modal, with the lower mode occurring at about 2 standard deviations below the mean (left panel). The probability of being between 1.5 and 2 standard deviations below the mean is 8.4% for the SEL-implied density – about twice the corresponding Gaussian-implied probability. There is barely any bimodality observed in the good consumption state (right panel).

Overall, the bi-modality is indicative of heightened macroeconomic uncertainty during bad times, in line with an extensive literature that empirically documents countercyclical variation in uncertainty (see, e.g., Jurado et al. (2015)). However, our results suggest that the increased uncertainty during bad times reflects the fattening of the close left tail rather than an increase in the volatility or in the probability of catastrophes (see also the results of Backus et al. (2014) using index-options).

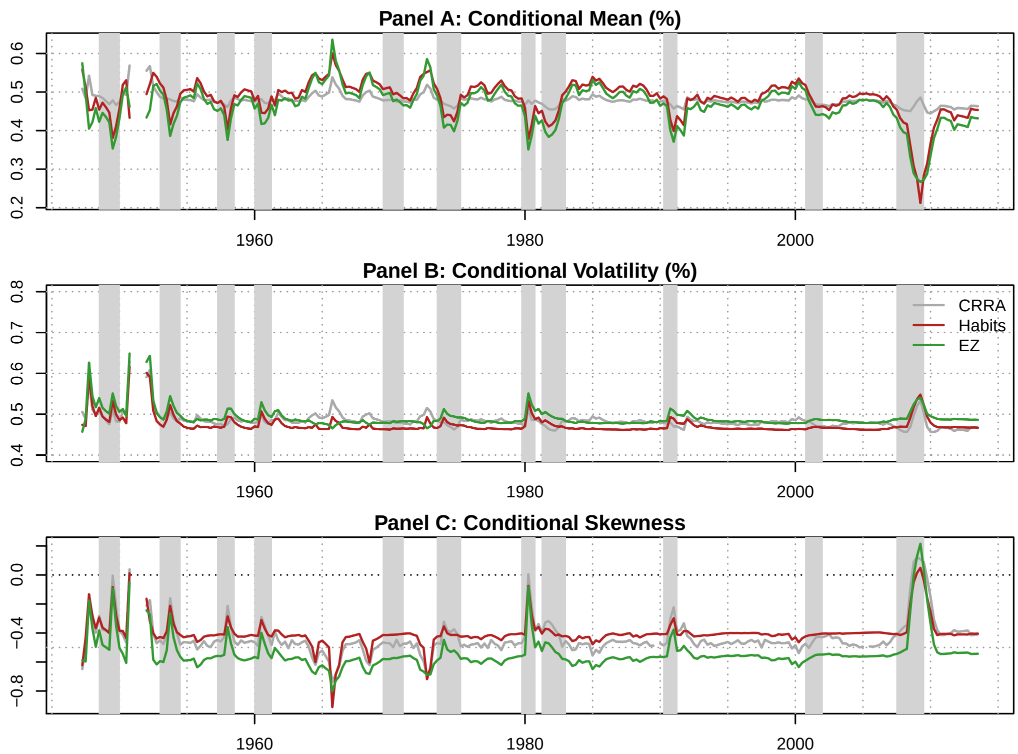
IV.4 Price-consistent moments of consumption growth

The SEL-estimated conditional probabilities allow us to compute the time series of beliefs, i.e. the conditional moments of the variables included in the estimation (see Equation 16). We present the time series of the mean, volatility and normalized skewness of consumption growth in Figure 3.

Panel A first shows that the conditional mean is strongly procyclical, for all three preference specifications, in line with the asset-pricing literature. It is at its peak at – or shortly before – the onset of a recessionary episode (denoted by grey-shaded areas), declines steadily through the recession, reaches its trough around the end of the recession before rising back up. The correlations between the conditional means based on the CRRA, habits and recursive preferences and a recession indicator are -0.27, -0.48, and -0.48, respectively.

Second, beliefs about expected consumption growth are more persistent and less volatile than realized consumption growth. For both the habit and recursive preferences, the conditional mean has an annualized volatility of 0.1%, ten times smaller than the 1.0% volatility of realized consumption growth, and their first-order autocorrelation coefficient is 0.88 compared to only 0.31 for realized consumption growth. Regressions of consumption growth onto the

Figure 3 – Price-consistent conditional moments of consumption growth



Notes: The figure plots the time series of price-consistent conditional mean, volatility and skewness of consumption growth (Panels A, B and C, resp.). Shaded areas correspond to NBER recessions. The conditional moments are obtained using the estimated SEL distributions. The pricing kernels correspond to CRRA (grey), habits (red) and Epstein-Zin (green). Test assets include market returns, Tbill rate and small, big, growth and value portfolios. The conditioning set includes past consumption growth and a consumption growth volatility proxy. Sample is 1947:Q2-2013:Q4.

price-consistent forecasts produce statistically significant slope coefficients for both preference specifications, and an R^2 of about 5%.

Third, Panel B of Figure 3 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, for all three specifications of preferences, the time-variation in the volatility is much smaller than the variation in the mean. For instance, for the habit preferences, the conditional mean varies over [0.21%, 0.60%] whereas the volatility varies over [0.46%, 0.62%], and a similar range can be observed for recursive preferences.

Fourth, Panel C of Figure 3 reveals that investors perceive the skewness of consumption growth to be negative in almost all states with cyclical fluctuations. 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. These last two features are in line with the findings of Colacito et al. (2016), who find that conditional moments of growth obtained through cross-sectional aggregation of surveys of professional forecasters feature highly time-varying skewness whereas the volatility is roughly time invariant.

Table 5 – Correlation of beliefs: alternative pricing kernels

	(A): Mean		*(B): Volatility*		*(C): Skewness*	
	Habits	EZ	Habits	EZ	Habits	EZ
Crra	0.525	0.548	0.790	0.708	0.927	0.965
Habits	1.000	0.979	1.000	0.958	1.000	0.925

Notes: The table reports the correlations between the price-consistent 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.

Finally, we perform a comparison between the time series implied by the three preferences. All preferences tend to agree on a virtually time-invariant consumption growth volatility. For habits and recursive preferences, all time series correlations are above 90%, showing the strong similarities for consumption beliefs implied by the two preferences. CRRA preferences produce expected consumption growth with correlations of about 50% with the other two, which grows above 70% and 90% for volatility and skewness, respectively. We thus interpret our results on consumption beliefs as largely robust to the choice of preferences.

Appendix A.8 shows that the recovered beliefs are robust to several alternative choices for the bandwidth parameter, the set of assets, and the conditioning set. This notably alleviates the concern that the estimation is largely driven by the choice of the kernel density weights (i.e., the choice of the bandwidth parameter) rather than the Euler constraints.

Overall, our results suggest that cyclical variations in the first and third moments of consumption growth are important components of the beliefs process. Time-variation in the second moment (volatility), on the other hand, seems to be economically small. Incorporation of time-varying skewness, either about the true DGP or the subjective beliefs process, is typically missing from most models.

IV.5 Price-consistent moments of market returns

We turn now to price-consistent beliefs about the stock market. Figure 4 presents the time series of the conditional mean, volatility, and Sharpe ratio of the equity premium, for the three different choices of preferences (Panels A, B, and C).

Panel A shows that the expected equity premium is strongly countercyclical for all three preference specifications. The correlations between the expected premium and a recession dummy are 0.46, 0.43, and 0.38 for the CRRA, habit, and recursive preferences, respectively. The average (annualized) levels of the premia implied by these three SDFs are 4.81%, 5.37%, and 7.13%, respectively. In comparison, the historical average equity premium is 7.5%.

Similar features are observed in Panel B for the time series of conditional volatilities of the market return. Specifically, the volatility is countercyclical with correlations with a recession dummy between 0.41 and 0.47 for all preferences; and has an average annualized values of 17% for each preference specification compared with the historical volatility of 16.6%.

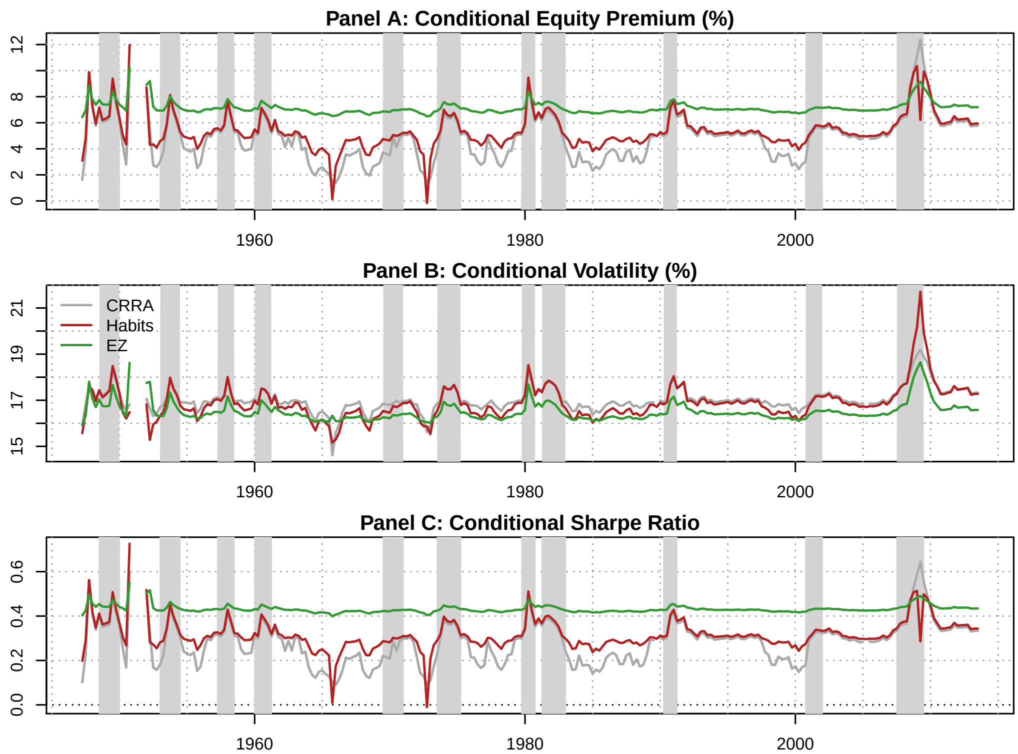
The countercyclicality of the volatility of the market returns in less pronounced than that of the conditional mean, such that the Sharpe ratio is countercyclical as well (see Panel C of Figure 4). The Sharpe ratio roughly follows the same pattern as the equity premium, ranging from 0 to 0.7 for the habit preferences, and averaging between 0.3 and 0.5 depending on the preferences.

To summarize, our results suggest that the representative investor's beliefs about the market return are strongly countercyclical. This conclusion is robust to the specification of the risk preferences of the investor.

IV.6 Are price-consistent beliefs really beliefs?

In this section, we address the concern that all our chosen preferences might be misspecified, such that what we recover is actually the multiplicative misspecification component of the true (unknown) pricing kernel instead of actual beliefs.

Figure 4 – Time series of Conditional Moments of Market Excess Return



Notes: The figure plots the time series of price-consistent conditional mean, volatility and Sharpe ratio of market excess returns (Panels A, B and C, resp.). Shaded areas correspond to NBER recessions. The conditional moments are obtained using the estimated SEL distributions. The pricing kernels correspond to CRRA (grey), habits (red) and Epstein-Zin (green). Test assets include market returns, Tbill rate and small, big, growth and value portfolios. The conditioning set includes past consumption growth and a consumption growth volatility proxy. Sample is 1947:Q2-2013:Q4.

We compare the time series of the price-consistent expected stock market returns (Figure 4, Panel A) with survey data. Specifically, we consider Robert Shiller's institutional investor survey, released by the Investor Behavior Project.[13] The survey, conducted at six-month intervals over the period

[13]We choose institutional investors surveys as opposed to surveys of individual investors. Greenwood and Shleifer (2014) show that expectations of stock market returns in surveys of individual investors are procyclical. However, as we show here, survey data on institutional

July 1989 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 One-Year Confidence Index is the percentage of institutional investors expecting an increase in the Dow in the coming year.[14]

The scatterplots of the Confidence Index against the price-consistent beliefs implied by the three SDFs are presented in Figure 5. The correlations of the beliefs with the Confidence Index are positive at 0.27, 0.23, and 0.18, respectively, for the power utility, habit and EZ price-consistent beliefs. Regressions of the Confidence Index on the price-consistent beliefs produce statistically significant slope coefficients with t-statistics of 2.37 and 1.99, respectively, for the CRRA and habit preference specifications. The recursive preferences produce insignificant results, with a t-statistics of 1.55, because of the smoothness of the time series noted above.

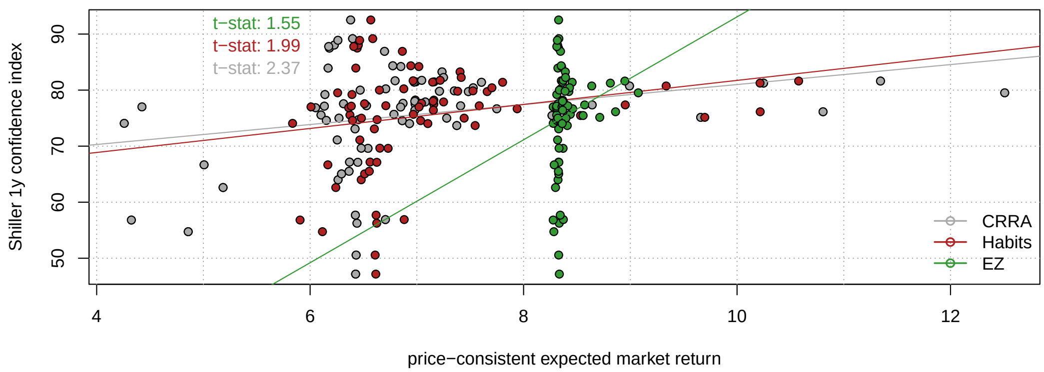
The main conclusion we draw from this exercise is that the recovered beliefs co-move positively with institutional investors' beliefs about the stock market captured in survey data. This lends additional support to the conjecture that the recovered distribution represents the beliefs of the representative investor about the aggregate equity market.

As a second piece of evidence, we reproduce the analysis performed in Ghosh et al. (2019) in Appendix A.9, Figure 12 – namely, we separately recover the beliefs of investors who allocate heavily to large market capitalization stocks (akin to the market portfolio) and of investors who allocate heavily to

investors' expectations suggest that their beliefs about the stock market are countercyclical.

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

Figure 5 – Comparison of Expected Excess Market Returns with Survey Forecasts



Notes: The figure plots the scatterplot of the U.S. Institutional One-Year Confidence Index on the price-consistent expected market return. The pricing kernel correspond respectively to CRRA (grey), external habit (red), and Epstein-Zin (green) recursive preferences. Test assets include market returns, Tbill rate and small, big, growth and value portfolios. The conditioning set includes past consumption growth and a consumption growth volatility proxy. Sample is 1989:Q2-2013:Q4.

small market cap and growth stocks (often regarded as the public market analog of private venture capital backed companies, and whose behavior has proven anomalous with respect to most asset pricing models). We find that the beliefs of the former about the stock market are strongly countercyclical, whereas those of the latter are procyclical. These findings offer a potential reconciliation of the evidence of procyclical expected market returns in individual investor survey data with the countercyclical expected returns implied by rational expectations representative agent models. Importantly, for the purposes of the current paper, this finding further bolsters our claim that the SEL approach successfully recovers investors' beliefs from observed asset prices.

V Are price-consistent beliefs rational?

In this section, we compare our price-consistent beliefs to potential candidates for the data generating process of consumption growth $\mathcal{P}_0$.

V.1 Are price-consistent beliefs comparable to commonly assumed data generating processes?

We first compare our price-consistent beliefs to some commonly assumed data generating processes in the macro and finance literature. We restrain from interpreting deviations of the price-consistent beliefs from these benchmarks as beliefs distortions, since they might just be the result of the misspecification of the true DGP.

Our first choice of DGP is a standard ARMA(1,1) model for consumption growth, perhaps the most extensively used specification in the literatures.[15] We include GARCH volatility in the specification for the sake of completeness. Our second choice is a regime-switching model, where the mean of consumption growth differs across latent regimes (see, e.g., David and Veronesi (2013), Ghosh and Constantinides (2017)). 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 four,

[15] Wachter (2006) assumes an ARMA(1,1) 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 (see, e.g., Bidder and Dew-Becker (2016) and Szoke (2017)).

based on information criterion.[16]

We estimate both models using historical data on consumption growth alone. Figure 13 in Appendix A.10 presents the time series of consumption growth over our sample period, along with the estimated conditional means implied by both models. Both models perform reasonably well in capturing the bulk of business cycle fluctuations in consumption growth.

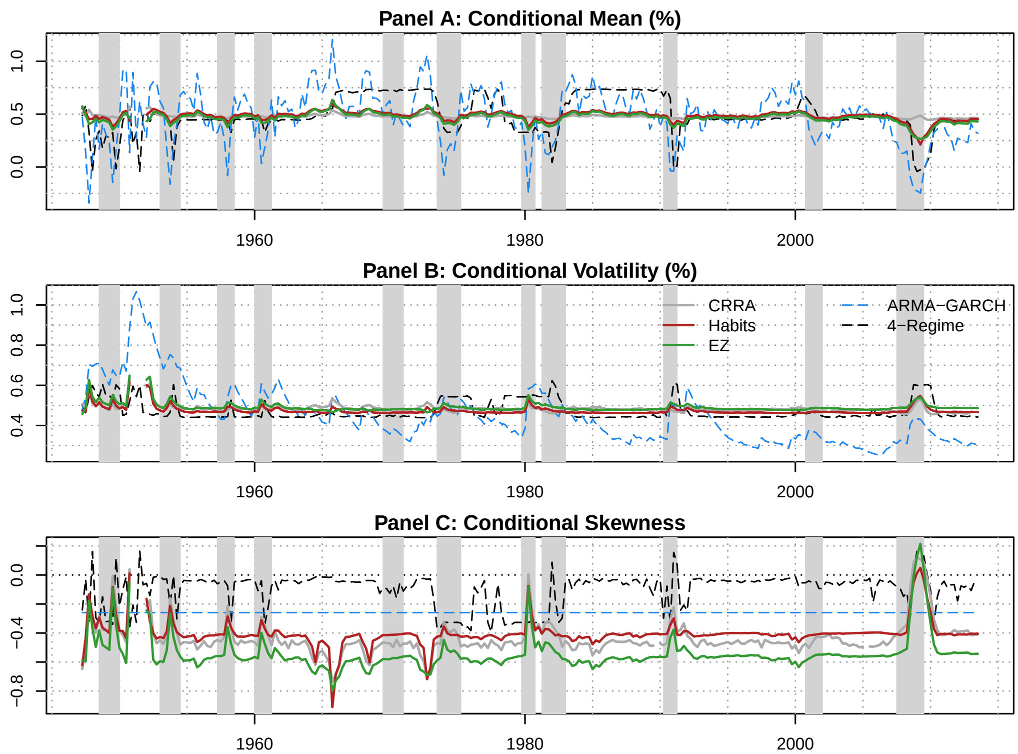
We then compare the conditional moments of consumption growth implied by these two DGPs with the price-consistent beliefs. Panels A-C of Figure 6 plot the time series of the conditional mean, volatility, and skewness, respectively, of consumption growth implied by the price-consistent beliefs, the ARMA-GARCH specification, and the regime switching model.

Panel A shows that both the ARMA and the regime switching models imply much more volatile and less persistent forecasts about future consumption growth compared to the price-consistent beliefs. Specifically, the series of expected consumption growth implied by the models have an annualized volatility of 0.38% and 0.35%, respectively, almost four times the volatility of 0.1% of the price-consistent beliefs. The ARMA-implied conditional mean has a lower persistence, measured by a first-order autocorrelation coefficient of 0.75, while the regime switching model produces consumption forecasts with 0.83 autocorrelation, compared to 0.86 for the SEL. 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

[16] We also considered an alternative with Markov switching volatility and obtained similar results.

explaining asset prices compared to the persistence that would be estimated using historical consumption data alone.

Figure 6 – Comparison of ARMA(1,1), Regime Switching, and SEL



Notes: The figure plots the time series of price-consistent conditional mean, volatility and skewness of consumption growth (Panels A, B and C, resp.) along with ARMA-GARCH (dashed-blue) and regime switching (dashed-black) model estimates. Shaded areas correspond to NBER recessions. The conditional moments are obtained using the estimated SEL distributions. The pricing kernels correspond to CRRA (grey), habits (red) and Epstein-Zin (green). Test assets include market returns, Tbill rate and small, big, growth and value portfolios. The conditioning set includes past consumption growth and a consumption growth volatility proxy. Sample is 1947:Q2-2013:Q4.

Turning to volatility estimates, Panel B suggests that low time-variation of the volatility of the regime switching model is in line with that implied by the price-consistent beliefs. In turn, the ARMA-GARCH specification implies too large variations, but the parameters are very imprecisely estimated (see Appendix A.11).

Finally, Panel C suggests that the conditional Gaussianity assumption of the ARMA-GARCH model contrasts with that implied by the recovered beliefs. The historical skewness of the standardized residuals is consistently higher than the price-consistent skewness, for all choices of preferences. The regime switching model implied skewness differs even more: the skewness implied by the model is close to zero on average and shows countercyclical fluctuations, contrary to our price-consistent estimates.

In sum, our results suggest that popular DGPs used in macro-finance models are likely missing key empirical features of investors' beliefs. Our results can offer guidance on how to formulate and calibrate these types of models.

V.2 Non-parametric benchmark for objective beliefs

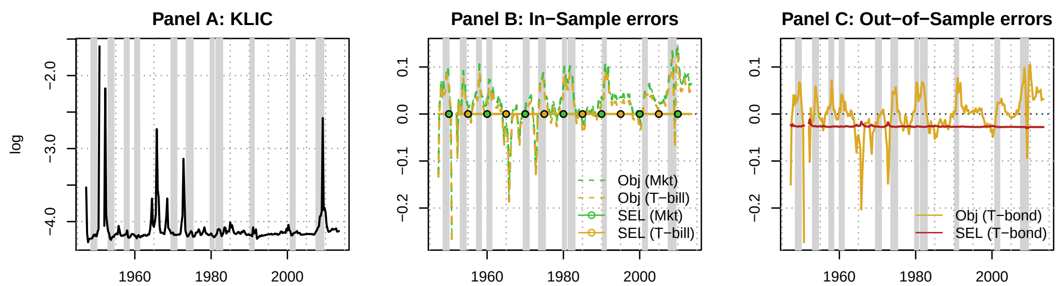
As mentioned above, differences between the price-consistent beliefs and commonly assumed DGPs could be the result of misspecification of consumption growth dynamics and not necessarily indicative of beliefs distortions. The question, however, naturally arises regarding whether investors' beliefs are rational or whether they show significant distortions. The rational expectations hypothesis and behavioral alternatives 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 potential deviations from rationality represents an important advance in this debate.

To mitigate concerns of misspecification in the choice of objective DGP, we choose the non-parametric kernel density estimator $p_{i,j}^{(0)}$ of Equation (5) as the objective DGP for consumption growth. Recall that our SEL estimator produces conditional densities that have the least possible distortions with respect to this objective benchmark, while enforcing the Euler equation constraints.

We first compare the pricing performance of this non-parametric objective measure to that of the price-consistent measure, with habit preferences. We chose these preferences because they do not suffer from implausibility of the parameter estimates (CRRA), or imperfect proxies for components of the SDF (EZ), and report results for the other preferences in Appendix A.12. Figure 15 plots the time series of the conditional pricing errors for the stock market return and the T-bill rate, under each of the two measures (Panel B). The objective measure produces large and highly volatile conditional pricing errors, peaking at an absolute value of about 30% in the war aftermath. The pricing errors are larger during bad states of the world, and the 2008 financial crisis sees both pricing errors jump to 15%. In contrast, the price-consistent probabilities produce conditional pricing errors that are identically equal to zero in all periods, by construction.

To further quantify the time-variation in the beliefs distortions, Figure 15 (Panel A) plots the time series of the KLIC divergence between the objective and subjective probability measures – $\widehat{\mathcal{P}}_0$ and $\widehat{\mathcal{P}}(\theta)$, respectively – as given by the objective function in Equation (7). The KLIC follows the pattern of the time series of pricing errors, and are particularly large in the aftermath of the war, in the late sixties and seventies recession, and during the 2008 financial

Figure 7 – Conditional Pricing Errors and Kullback-Leibler Divergence: Habit preferences



Notes: The figure plots the (log) KLIC divergence (A), the in-sample pricing errors (B) for market returns (green) and T-bill (yellow) with and without pricing constraints (solid dotted and dashed, resp.), and out-of-sample errors on the 10y T-bond real quarterly returns (C) with and without pricing constraints (yellow and red, resp.). Shaded areas correspond to NBER recessions. The conditional moments are obtained using the estimated SEL distributions. The pricing kernel corresponds to habits. Test assets include market returns, Tbill rate and small, big, growth and value portfolios. The conditioning set includes past consumption growth and a consumption growth volatility proxy. Sample is 1947:Q2-2013:Q4.

crisis.

We also compare the objective and subjective probability estimates by comparing their pricing errors on an asset that was not used in the estimation of price-consistent beliefs, namely the real quarterly returns on a 10y Treasury bond.[17] Panel C of Figure 15 shows that the pricing errors under the objective measure are large and volatile (varying from -30% to 10%), while being small and virtually constant under the subjective measure. As a result, the RMSE of pricing errors under the subjective measure is 2.6%, about half of that obtained with the objective measure (4.3%). This out-of-sample exercise further confirms the ability of our methodology to capture a price-consistent

[17]We calculate returns on a 10y T-bond by combining smoothed Nelson-Siegel-Svensson estimates of Gurkaynak et al. (2007), available from 1961, and the yearly time series from Aswath Damadoran (http://pages.stern.nyu.edu/~adamodar/).

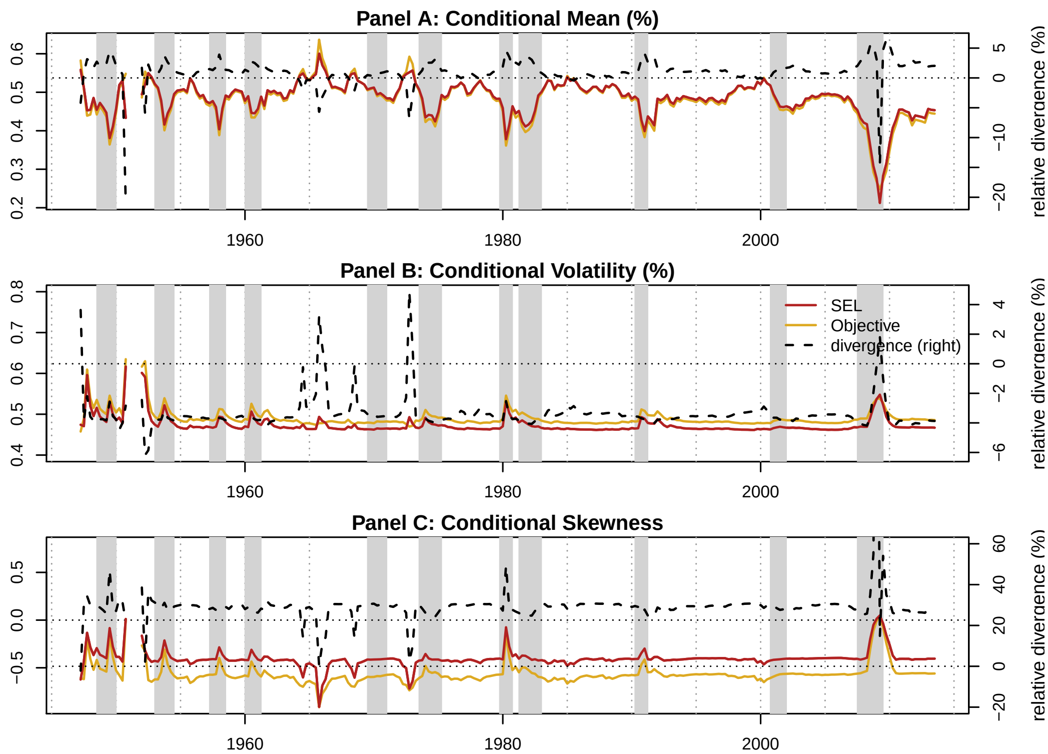
measure.

To identify the precise nature of the beliefs distortions, Figure 8 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 higher expected consumption growth compared to that implied by the objective measure for most states of the world. Thus, investors seem to consistently overestimate by between 1% and 5% consumption growth forecasts during good and bad times. If anything, the distortions are countercyclical showing slight increases during recessions.

In contrast, Panel B shows that the price-consistent beliefs typically imply smaller volatility of consumption growth relative to the objective measure, by about 4%. Moreover, the distortions do not show meaningful cyclical variations during our sample. Finally, in Panel C, we plot the time series of the conditional skewness of consumption growth under the two measures and the percentage change between them. The discrepancies are the largest of all moments and mostly flat over the business cycle. The average divergence between these series is above 20%, showing significant positive peaks during the 1980s and the 2008 crises.

To summarize, investors' beliefs about consumption growth seem distorted relative to the objective benchmark in two important respects. First, the conditional mean and volatility are consistently over and underestimated, respectively, under the price-consistent beliefs compared to the objective measure.

Figure 8 – Time Series of Moments Under SEL and Objective Measures: Habit preferences



Notes: The figure plots the time series of price-consistent conditional mean, volatility and skewness of consumption growth (Panels A, B and C, resp., red), along with a non-parametric estimate of the same quantities (yellow). The relative divergence between the two measures is represented by black dashed lines and measured on the right axis. Shaded areas correspond to NBER recessions. The conditional moments are obtained using the estimated SEL distributions. The pricing kernel corresponds to habits (red). Test assets include market returns, Tbill rate and small, big, growth and value portfolios. The conditioning set includes past consumption growth and a consumption growth volatility proxy. Sample is 1947:Q2-2013:Q4.

Second, the conditional skewness is less negative under the price-consistent beliefs compared to the objective measure. Both these dimensions of distortions are indicative of investor exuberance, with investors believing that the mean consumption growth were higher and the left-tail of the distribution were truncated relative to what is implied by the objective measure. And, the magnitudes of the distortions in the skewness far exceed those in the mean. 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.

VI Conclusion and extensions

Asset prices reflect investors' beliefs about future economic and financial outcomes. Relying on this insight, we propose an information-theoretic methodology to recover investors' beliefs from observed asset prices. Our approach relies on the smoothed empirical likelihood (SEL) estimator of Kitamura et al. (2004), and does not require any functional-form assumptions about the true DGP or assumptions regarding investor rationality or lack thereof. 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 that investors use to form their beliefs.

The recovered beliefs suggest that the expected 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, including cyclical variation in the skewness and bimodality during bad times. We show that these findings are robust to alternative inputs.

Our results suggest that investors' beliefs are distorted relative to commonly assumed objective benchmarks. When the latter is taken as a non-parametric kernel density estimate, investors seem to overestimate both the conditional mean and skewness and underestimate the volatility of consumption growth. These distortions are indicative of investor exuberance.

References

Almeida, C., K. Ardison, and R. Garcia. 2019. "Nonparametric Assessment of Hedge Fund Performance". *Journal of Econometrics (forthcoming).*

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.

Backus, D., M. Chernov, and S. Zin. 2014. "Sources of Entropy in Representative Agent Models". *The Journal of Finance* 69 (1): 51–99.

Bansal, R., and A. Yaron. 2004. "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles". *Journal of Finance* 59, no. 4 (): 1481–1509.

Barberis, N., and R. Thaler. 2001. "A Survey of Behavioral Finance". In *Handbook of the Economics of Finance.* Elsevier Science B.V.

Barillas, F., L.-P. Hansen, and T. 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, no. 9 (): 2494–2527.

Borovicka, J., L.-P. Hansen, and J. Scheinkman. 2016. "Misspecified Recovery". *The Journal of Finance.*

Borwein, J., and A. Lewis. 1991. "Duality Relationships for Entropy-type Minimization Problems". *SIAM Journal of Control and Optimization* 29:325–338.

Brown, B., and W. Newey. 1998. "Efficient Semiparametric Estimation of Expectations". *Econometrica* 66, no. 2 (): 453–464.

Brown, B., and W. Newey. 2002. "Generalized Method of Moments, Efficient Bootstrapping, and Improved Inference". *Journal of Business and Economic Statistics* 20:507–517.

Campbell, J., and J. Cochrane. 1999. "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior". *Journal of Political Economy* 107, no. 2 (): 205–51.

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

Chamberlain, G. 1987. "Asymptotic Efficiency in Estimation with Conditional Moment Restrictions". *Journal of Econometrics* 34 (3): 305–334.

Chen, X., L. P. Hansen, and P. G. Hansen. 2020. *Robust Identification of Investor Beliefs.* Tech. rep.

Colacito, R., E. Ghysels, J. Meng, and W. Siwasarit. 2016. "Skewness in Expected Macro Fundamentals and the PRedictability of Equity Returns: Evidence and Theory". *Review of Financial Studies* 29 (8).

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*, no. 121: 682–746.

Epstein, L., and S. Zin. 1989. "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework". *Econometrica* 57:937–968.

Ghosh, A., and G. Constantinides. 2017. "What Information Drives Asset Prices?"

Ghosh, A., C. Julliard, and A. 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.*

Ghosh, A., A. Korteweg, and Q. Xu. 2019. "Recovering Heterogeneous Beliefs and Preferences from Asset Prices". Working paper.

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

Gurkaynak, R. S., B. Sack, and J. H. Wright. 2007. "The U.S. Treasury yield curve: 1961 to the Present". *Journal of Monetary Economics* 54, no. 8 (): 2291–2304.

Hall, P., R. Wolff, and Q. Yao. 1999. “Methods for Estimating a Conditional Distribution Function”. *Journal of the American Statistical Association* 94 (): 154–163.

Hansen, L. P., and T. J. Sargent. 2016. “Set of Models and Prices of Uncertainty”. NBER Working Papers.

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

Jensen, C., D. Lando, and L. Pedersen. 2018. “Generalized Recovery”. *Journal of Financial Economics* forthcoming.

Jurado, K., S. Ludvigson, and S. Ng. 2015. “Measuring Uncertainty”. *American Economic Review* 105:1177–1216.

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

Kitamura, Y. 2003. “A Likelihood-Based Approach to the Analysis of a Class of Nested and Non-Nested Models”. Working Paper.

— . 2007. “Nonparametric Likelihood: Efficiency and Robustness”. *The Japanese Economic Review* 58 (1).

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

Korsaye, S., A. Quaini, and F. Trojani. 2019. “Smart SDFs”.

Kozhan, R., A. Neuberger, and P. Schneider. 2013. “The Skew Risk Premium in the Equity Index Market”. *Review of Financial Studies* 26:2174–2203.

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. Prescott. 1985. “The Equity Premium: A Puzzle”. *Journal of Monetary Economics* 15, no. 2 (): 145–61.

Muth, J. 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*, no. 3: 873–909.

Orlowski, P., A. Sali, and F. Trojani. 2018. *Arbitrage Free Dispersion*. Tech. rep.

Owen, A. B. 2001. *Empirical Likelihood*. Chapman / 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).

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

Schneider, P. 2018. “An Anatomy of the Market Return”. *Journal of Financial Economics* Forthcoming.

— . 2015. “Generalized risk premia”. *Journal of Financial Economics* 116 (3): 487–504.

Schneider, P., and F. Trojani. 2019. “(Almost) Model-Free Recovery”. *Journal of Finance* 74 (1).

— . 2018. “Divergence and the Price of Uncertainty”. *The Journal of Financial Econometrics* 1 (56).

Szoke, B. 2017. “Estimating Robustness”. *Working paper*, NYU.

Vuong, Q. H. 1989. “Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses”. *Econometrica*: 307–333.

Wachter, J. 2006. “A Consumption-Based Model of the Term Structure of Interest Rates”. *Journal of Financial Economics*: 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 (for online publication)

A.1 Minimum distance estimation problem

The general estimation problem of the probabilities can be formulated as follows. We are looking for a change of measure $\mathrm{d}\mathcal{P}^*/\mathrm{d}\mathcal{P}_0$, such that:

$$\frac{\mathrm{d}\mathcal{P}^*}{\mathrm{d}\mathcal{P}_0} = \text{arg min} \quad \mathbb{E}^{\mathcal{P}_0}\left[\phi\left(\frac{\mathrm{d}\mathcal{P}}{\mathrm{d}\mathcal{P}_0}\right)|\underline{\mathcal{F}_t}\right] \quad \text{s.t.} \quad \mathbb{E}^{\mathcal{P}_0}\left[\frac{\mathrm{d}\mathcal{P}}{\mathrm{d}\mathcal{P}_0}M_{t+1}\mathbf{R}_{t+1}|\underline{\mathcal{F}_t}\right] = \mathbf{1}_k\,, \tag{17}$$

and that $\mathcal{P}$ is indeed a probability measure, where $\phi(\cdot)$ is a discrepancy function (see Almeida and Garcia (2016)). Cressie and Read (1984) propose the discrepancy function as:

$$\phi_\gamma\left(\frac{\mathrm{d}\mathcal{P}}{\mathrm{d}\mathcal{P}_0}\right) = \frac{1}{\gamma(\gamma+1)}\left[\left(\frac{\mathrm{d}\mathcal{P}}{\mathrm{d}\mathcal{P}_0}\right)^\gamma - 1\right] \tag{18}$$

such that the empirical counterpart of Equation (17) is given by:

$$\mathbb{E}^{\mathcal{P}_0}\left[\phi_\gamma\left(\frac{\mathrm{d}\mathcal{P}}{\mathrm{d}\mathcal{P}_0}\right)\right] = \frac{1}{\gamma(\gamma+1)}\sum_{j=1}^{T} p_{t,j}\left[\left(\frac{p_{t,j}}{p_{t,j}^{(0)}}\right)^\gamma - 1\right]. \tag{19}$$

For values of $\gamma = 0$, this discrepancy is given by:

$$\mathbb{E}^{\mathcal{P}_0}\left[\phi_0\left(\frac{\mathrm{d}\mathcal{P}}{\mathrm{d}\mathcal{P}_0}\right)\right] = \sum_{j=1}^{T} p_{t,j}\log\left(\frac{p_{t,j}}{p_{t,j}^{(0)}}\right), \tag{20}$$

and for $\gamma = -1$ this discrepancy is called *empirical likelihood* and given by,

$$\mathbb{E}^{\mathcal{P}_0}\left[\phi_{-1}\left(\frac{\mathrm{d}\mathcal{P}}{\mathrm{d}\mathcal{P}_0}\right)\right] = -\sum_{j=1}^{T} p_{t,j}^{(0)}\log\left(\frac{p_{t,j}}{p_{t,j}^{(0)}}\right). \tag{21}$$

With the Cressie-Read family of discrepancies, the problem can be rewritten:

$$\widehat{p}_{t,\cdot} = \arg\min \quad \frac{1}{\gamma(\gamma+1)} \sum_{j=1}^{T} p_{t,j} \left[\left(\frac{p_{t,j}}{p_{t,j}^{(0)}} \right)^{\gamma} - 1 \right] \quad \text{s.t.} \quad \sum_{j=1}^{T} p_{t,j} M_j \mathbf{r}_j = \mathbf{1}_k \quad \text{and} \quad \sum_{j=1}^{T} p_{t,j} = 1\,. \tag{22}$$

The Lagrangian associated to this problem is given by:

$$\mathcal{L}_t = \frac{1}{\gamma(\gamma+1)} \sum_{j=1}^{T} p_{t,j} \left[\left(\frac{p_{t,j}}{p_{t,j}^{(0)}} \right)^{\gamma} - 1 \right] + \lambda_t' \left[\sum_{j=1}^{T} p_{t,j} \left(M_j \mathbf{r}_j - \mathbf{1}_k \right) \right] + \delta_t \left[\sum_{j=1}^{T} p_{t,j} - 1 \right]. \tag{23}$$

The first-order condition is obtained as:

$$\frac{\partial \mathcal{L}_t}{\partial p_{t,j}} = \frac{1}{\gamma} \left(\frac{p_{t,j}}{p_{t,j}^{(0)}} \right)^{\gamma} - \frac{1}{\gamma(\gamma+1)} + \lambda_t' \left(M_j \mathbf{r}_j - \mathbf{1}_k \right) + \delta_t = 0\,. \tag{24}$$

Thus, we obtain:

$$p_{t,j;\gamma} = \left[\frac{1}{\gamma+1} - \gamma \left[\lambda_t' \left(M_j \mathbf{r}_j - \mathbf{1}_k \right) + \delta_t \right] \right]^{1/\gamma} p_{t,j}^{(0)}\,, \tag{25}$$

For $\gamma = 0$ and $\gamma = -1$, the solution converges to:

$$\begin{aligned} p_{t,j;0} &= p_{t,j}^{(0)} \times e^{-1-\lambda_t'(M_j \mathbf{r}_j - \mathbf{1}_k) - \delta_t} \\ p_{t,j;-1} &= p_{t,j}^{(0)} \times \frac{1}{\lambda_t' \left(M_j \mathbf{r}_j - \mathbf{1}_k \right) + \delta_t} \end{aligned} \tag{26}$$

In general, δ_t and λ_t can be obtained numerically, such that:

$$\sum_{j=1}^{T} p_{t,j}^{(0)} \left[\frac{1}{\gamma+1} - \gamma\left[\lambda_t'\left(M_j\mathbf{r}_j - \mathbf{1}_k\right) + \delta_t\right]\right]^{1/\gamma} \left(M_j\mathbf{r}_j - \mathbf{1}_k\right) = \mathbf{0}_k$$
$$\sum_{j=1}^{T} p_{t,j}^{(0)} \left[\frac{1}{\gamma+1} - \gamma\left[\lambda_t'\left(M_j\mathbf{r}_j - \mathbf{1}_k\right) + \delta_t\right]\right]^{1/\gamma} = 1 \quad (27)$$

Note that for $\gamma = -1$, it is easy to show that $\delta_t = 1$ and λ_t is given by:

$$\sum_{j=1}^{T} p_{t,j}^{(0)} \frac{1}{\lambda_t'\left(M_j\mathbf{r}_j - \mathbf{1}_k\right) + 1} \left(M_j\mathbf{r}_j - \mathbf{1}_k\right) = \mathbf{0}_k$$
$$\iff \quad \lambda_t = \arg\max \sum_{j=1}^{T} p_{t,j}^{(0)} \log\left[1 + \lambda_t'\left(M_j\mathbf{r}_j - \mathbf{1}_k\right)\right]$$

KLIC is one member of Cressie-Read (CR) divergence measures between probability distributions with $\gamma = -1$ (see Cressie and Read (1984)). Members of the CR-family also include Euclidean likelihood (see e.g. Antoine et al. (2007), entropy (Backus et al. (2014) and Bakshi and Chabi-Yo (2018) for instance), and a more general approach is developed by Almeida and Garcia (2012, 2016) and Sandulescu et al. (2018) for deriving admissible conditions for SDFs.

A.2 Owen normalization

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 λ_i a corner solution to the optimization problem (9). In order to avoid this case, we use the normalization introduced by Owen (2001). In order to avoid numerical issues associated with the estimation of the Lagrange multipliers, he proposes the following transformation.

$$\widehat{\lambda}_i^{(owen)}(\theta) = \underset{\lambda_i \in \mathbb{R}^k}{\operatorname{argmax}} \sum_{j=1}^{T} p_{i,j}^{(0)} \cdot \Psi_\nu \left[1 + \lambda_i' \left(M(g_j, y_j; \theta)\, \mathbf{r}_j - \mathbf{1}_k\right)\right] \quad (28)$$

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

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:

$$\widehat{p}_{i,j}^{(owen)}(\theta) = \frac{p_{i,j}^{(0)}}{1 + \lambda_i^{(owen)\prime} \left(M(g_j, y_j; \theta)\, \mathbf{r}_j - \mathbf{1}_k\right)} \times \left(\sum_{j=1}^{T} \frac{p_{i,j}^{(0)}}{1 + \lambda_i^{(owen)\prime} \left(M(g_j, y_j; \theta)\, \mathbf{r}_j - \mathbf{1}_k\right)} \right)^{-1}, \quad (30)$$

A.3 Variance-covariance of estimation

In the well-specified case, Kitamura et al. (2004) show that the variance-covariance matrix of the estimator of θ is given by:

$$
\begin{aligned}
I(\theta_0) \quad = \quad & \mathbb{E}^{\mathcal{P}_0}\left\{\mathbb{E}^{\mathcal{P}_0}\left(\left.\frac{\partial M(G_{t+1},Y_{t+1};\theta)\,\mathbf{R}_{t+1}}{\partial\theta}\right|_{\theta=\theta_0}\middle| x_t\right)\right.\\
& \times \quad \mathbb{E}^{\mathcal{P}_0}\left([M(G_{t+1},Y_{t+1};\theta_0)\mathbf{R}_{t+1}-\mathbf{1}_k]\,[M(G_{t+1},Y_{t+1};\theta_0)\mathbf{R}_{t+1}-\mathbf{1}_k]'\,\middle|x_t\right)\\
& \quad \times \quad \left.\mathbb{E}^{\mathcal{P}_0}\left(\left.\frac{\partial M(G_{t+1},Y_{t+1};\theta)\,\mathbf{R}_{t+1}}{\partial\theta'}\right|_{\theta=\theta_0}\middle| x_t\right)\right\}.
\end{aligned}
$$

If the model is misspecified, either resulting from SDF misspecification of beliefs distortion, then the SEL estimator converges to its pseudo-true value θ^* and the asymptotic variance covariance of estimation is given by:

$$
\Omega(\theta^*) = E\left[D(x_t)V^{-1}(x_t)D'(x_t)\right]\,. \tag{31}
$$

where:

$$
\begin{aligned}
D(x_t) \quad & \equiv \quad \mathbb{E}\left[\left.\frac{\partial}{\partial\theta}\frac{M(G_{t+1},Y_{t+1};\theta)\,\mathbf{R}_{t+1}}{1+\lambda_t(\theta)'\,(M(G_{t+1},Y_{t+1};\theta)\,\mathbf{R}_{t+1}-\mathbf{1}_k)}\right|_{\theta=\theta^*}\middle| x_t\right],\\
V(x_t) \quad & \equiv \quad \mathbb{E}\left[S(X_t)S(X_t)'\middle|x_t\right]\\
& \qquad \text{where} \quad S(X_t) \equiv \frac{M(G_{t+1},Y_{t+1};\theta^*)\,\mathbf{R}_{t+1}-\mathbf{1}_k}{1+\lambda_t(\theta^*)'\,(M(G_{t+1},Y_{t+1};\theta^*)\,\mathbf{R}_{t+1}-\mathbf{1}_k)}\,.
\end{aligned}
$$

Although unclear from our notations, $S(X_t)$ depends on X_t only through the dependence of λ_t.

A.4 Misspecification example

We consider an economy where consumption growth can take either a high state g_H with probability π_0 and a low state with probability $1-\pi_0$. States are drawn independently over time. The representative agent has CRRA preferences with subjective discount factor δ and risk aversion parameter θ_0:

$$M_{t+1}(G_{t+1};\,\theta_0) = \delta G_{t+1}^{-\theta_0} \tag{32}$$

The representative agent possesses distorted beliefs, such that she believes that the high and low state probabilities are given by $(\pi, 1-\pi)$ where $\pi \neq \pi_0$. In this model, the gross equilibrium riskfree rate is given by:

$$R_f = \frac{1}{\mathbb{E}^{\mathcal{P}}\left[M_{t+1}(G_{t+1};\,\theta_0)\right]} = \frac{1}{\delta\left[\pi g_H^{-\theta_0} + (1-\pi) g_L^{-\theta_0}\right]}\,. \tag{33}$$

We assume below that the econometrician does not know the true θ_0 that was used to price assets. Instead, he uses θ possibly different from θ_0. In this case we have that:

$$M_{t+1}(G_{t+1};\,\theta) R_f = \frac{G_{t+1}^{-\theta}}{\pi g_H^{-\theta_0} + (1-\pi) g_L^{-\theta_0}}\,. \tag{34}$$

Let us consider first the non-parametric estimator of the states probabilities without the asset pricing constraints. It is trivial to show that:

$$p_t^{(0)} = \frac{n_T}{T}\mathbb{1}\left\{g_t = g_H\right\} + \left(1 - \frac{n_T}{T}\right)\mathbb{1}\left\{g_t = g_L\right\} \quad \text{where} \quad n_T = \sum_{t=1}^{T}\mathbb{1}\left\{g_t = g_H\right\}\,. \tag{35}$$

Using Equation (9), we have:

$$
\begin{aligned}
\widehat{\lambda} &= \operatorname{argmax} \quad \sum_{t=1}^{T} p_t^{(0)} \times \log\left[1+\lambda\left(m_t(\theta) g_t - 1\right)\right] \\
&= \frac{n_T}{T}\log\left[1+\lambda\frac{g_H^{-\theta}-\pi g_H^{-\theta_0}-(1-\pi)g_L^{-\theta_0}}{\pi g_H^{-\theta_0}+(1-\pi)g_L^{-\theta_0}}\right] \\
&+ \quad \left(1-\frac{n_T}{T}\right)\log\left[1+\lambda\frac{g_L^{-\theta}-\pi g_H^{-\theta_0}-(1-\pi)g_L^{-\theta_0}}{\pi g_H^{-\theta_0}+(1-\pi)g_L^{-\theta_0}}\right].
\end{aligned}
$$

Let us compute the first-order conditions:

$$
\begin{aligned}
0 &= \frac{n_T}{T}\times\frac{g_H^{-\theta}-\pi g_H^{-\theta_0}-(1-\pi)g_L^{-\theta_0}}{\pi g_H^{-\theta_0}+(1-\pi)g_L^{-\theta_0}}\times\left[1+\lambda\frac{g_H^{-\theta}-\pi g_H^{-\theta_0}-(1-\pi)g_L^{-\theta_0}}{\pi g_H^{-\theta_0}+(1-\pi)g_L^{-\theta_0}}\right]^{-1} \\
&+ \quad \left(1-\frac{n_T}{T}\right)\times\frac{g_L^{-\theta}-\pi g_H^{-\theta_0}-(1-\pi)g_L^{-\theta_0}}{\pi g_H^{-\theta_0}+(1-\pi)g_L^{-\theta_0}}\times\left[1+\lambda\frac{g_L^{-\theta}-\pi g_H^{-\theta_0}-(1-\pi)g_L^{-\theta_0}}{\pi g_H^{-\theta_0}+(1-\pi)g_L^{-\theta_0}}\right]^{-1}.
\end{aligned}
$$

Normalizing to the same denominators, this equation is equivalent to:

$$
\begin{aligned}
0 &= \frac{n_T}{T}\times\left(g_H^{-\theta}-\pi g_H^{-\theta_0}-(1-\pi)g_L^{-\theta_0}\right)\times\left[1+\lambda\frac{g_L^{-\theta}-\pi g_H^{-\theta_0}-(1-\pi)g_L^{-\theta_0}}{\pi g_H^{-\theta_0}+(1-\pi)g_L^{-\theta_0}}\right] \\
&+ \quad \left(1-\frac{n_T}{T}\right)\times\left(g_L^{-\theta}-\pi g_H^{-\theta_0}-(1-\pi)g_L^{-\theta_0}\right)\times\left[1+\lambda\frac{g_H^{-\theta}-\pi g_H^{-\theta_0}-(1-\pi)g_L^{-\theta_0}}{\pi g_H^{-\theta_0}+(1-\pi)g_L^{-\theta_0}}\right].
\end{aligned}
$$

Isolating the terms in λ, we obtain:

$$\begin{aligned}
&\frac{\lambda}{\pi g_H^{-\theta_0} + (1-\pi) g_L^{-\theta_0}} \Bigg\{ \frac{n_T}{T} \left[g_H^{-\theta} - \pi g_H^{-\theta_0} - (1-\pi) g_L^{-\theta_0} \right] \left[g_L^{-\theta} - \pi g_H^{-\theta_0} - (1-\pi) g_L^{-\theta_0} \right] \\
&\quad + \quad \left(1 - \frac{n_T}{T}\right) \left[g_L^{-\theta} - \pi g_H^{-\theta_0} - (1-\pi) g_L^{-\theta_0} \right] \left[g_H^{-\theta} - \pi g_H^{-\theta_0} - (1-\pi) g_L^{-\theta_0} \right] \Bigg\} \\
= \quad & -\frac{n_T}{T} \left[g_H^{-\theta} - \pi g_H^{-\theta_0} - (1-\pi) g_L^{-\theta_0} \right] - \left(1 - \frac{n_T}{T}\right) \left[g_L^{-\theta} - \pi g_H^{-\theta_0} - (1-\pi) g_L^{-\theta_0} \right] \\
\Longleftrightarrow \quad & \lambda \times \frac{\left[g_H^{-\theta} - \pi g_H^{-\theta_0} - (1-\pi) g_L^{-\theta_0} \right] \left[g_L^{-\theta} - \pi g_H^{-\theta_0} - (1-\pi) g_L^{-\theta_0} \right]}{\pi g_H^{-\theta_0} + (1-\pi) g_L^{-\theta_0}} \\
= \quad & -\frac{n_T}{T} \left[g_H^{-\theta} - \pi g_H^{-\theta_0} - (1-\pi) g_L^{-\theta_0} \right] - \left(1 - \frac{n_T}{T}\right) \left[g_L^{-\theta} - \pi g_H^{-\theta_0} - (1-\pi) g_L^{-\theta_0} \right] .
\end{aligned}$$

In the end we obtain:

$$\begin{aligned}
\widehat{\lambda} \quad = \quad & -\left[\pi g_H^{-\theta_0} + (1-\pi) g_L^{-\theta_0} \right] \times \Bigg\{ \frac{n_T}{T} \times \frac{1}{g_L^{-\theta} - \pi g_H^{-\theta_0} - (1-\pi) g_L^{-\theta_0}} \\
& + \quad \left(1 - \frac{n_T}{T}\right) \times \frac{1}{g_H^{-\theta} - \pi g_H^{-\theta_0} - (1-\pi) g_L^{-\theta_0}} \Bigg\}
\end{aligned} \tag{36}$$

Using Equation (8), the probability estimate $\widehat{\pi}(\theta)$ is obtained as:

$$\widehat{\pi}(\theta) = \frac{n_T}{T} \times \frac{1}{1 + \widehat{\lambda}(\theta) \left[m_H(\theta) R_f - 1 \right]} . \tag{37}$$

The denominator of this expression can be expressed:

$$
\begin{aligned}
1+\widehat{\lambda}(\theta)\left[m_H(\theta)R_f-1\right] &= 1+\widehat{\lambda}(\theta)\times\frac{g_H^{-\theta}-\pi g_H^{-\theta_0}+(1-\pi)g_L^{-\theta_0}}{\pi g_H^{-\theta_0}+(1-\pi)g_L^{-\theta_0}} \\
&= 1-\left\{\frac{n_T}{T}\times\frac{1}{g_L^{-\theta}-\pi g_H^{-\theta_0}-(1-\pi)g_L^{-\theta_0}}\right. \\
&\quad+\left.\left(1-\frac{n_T}{T}\right)\times\frac{1}{g_H^{-\theta}-\pi g_H^{-\theta_0}-(1-\pi)g_L^{-\theta_0}}\right\}\times\left[g_H^{-\theta}-\pi g_H^{-\theta_0}+(1-\pi)g_L^{-\theta_0}\right] \\
&= 1-\left(1-\frac{n_T}{T}\right)-\frac{n_T}{T}\times\frac{g_H^{-\theta}-\pi g_H^{-\theta_0}+(1-\pi)g_L^{-\theta_0}}{g_L^{-\theta}-\pi g_H^{-\theta_0}-(1-\pi)g_L^{-\theta_0}} \\
&= \frac{n_T}{T}\left[1-\frac{g_H^{-\theta}-\pi g_H^{-\theta_0}+(1-\pi)g_L^{-\theta_0}}{g_L^{-\theta}-\pi g_H^{-\theta_0}-(1-\pi)g_L^{-\theta_0}}\right] \\
&= \frac{n_T}{T}\times\frac{g_L^{-\theta}-g_H^{-\theta}}{g_L^{-\theta}-\pi g_H^{-\theta_0}-(1-\pi)g_L^{-\theta_0}}\,.
\end{aligned}
$$

Therefore,

$$
\widehat{\pi}(\theta)=\frac{g_L^{-\theta}-\pi g_H^{-\theta_0}-(1-\pi)g_L^{-\theta_0}}{g_L^{-\theta}-g_H^{-\theta}}\,. \tag{38}
$$

We can factor out π to obtain:

$$
\widehat{\pi}(\theta)=\pi\times\frac{g_L^{-\theta_0}-g_H^{-\theta_0}}{g_L^{-\theta}-g_H^{-\theta}}+\frac{g_L^{-\theta}-g_L^{-\theta_0}}{g_L^{-\theta}-g_H^{-\theta}}\,. \tag{39}
$$

The first-order differential of the log-probability is given by:

$$
\frac{\partial\log\left[\widehat{\pi}(\theta)\right]}{\partial\theta} = \frac{-\log(g_L)g_L^{-\theta}}{g_L^{-\theta}-\pi g_H^{-\theta_0}-(1-\pi)g_L^{-\theta_0}}-\frac{\log(g_H)g_H^{-\theta}-\log(g_L)g_L^{-\theta}}{g_L^{-\theta}-g_H^{-\theta}}
$$

When $\theta = \theta_0$, we have:

$$\begin{aligned}
\left.\frac{\partial \log\left[\widehat{\pi}(\theta)\right]}{\partial \theta}\right|_{\theta=\theta_0} &= \frac{-\log(g_L)g_L^{-\theta_0}}{g_L^{-\theta_0} - \pi g_H^{-\theta_0} - (1-\pi)g_L^{-\theta_0}} - \frac{\log(g_H)g_H^{-\theta_0} - \log(g_L)g_L^{-\theta_0}}{g_L^{-\theta_0} - g_H^{-\theta_0}} \\
&= \frac{-\log(g_L)g_L^{-\theta_0}}{\pi\left(g_L^{-\theta_0} - g_H^{-\theta_0}\right)} - \frac{\log(g_H)g_H^{-\theta_0} - \log(g_L)g_L^{-\theta_0}}{g_L^{-\theta_0} - g_H^{-\theta_0}} \\
&= \frac{(\pi-1)\log(g_L)g_L^{-\theta_0} - \pi\log(g_H)g_H^{-\theta_0}}{\pi\left(g_L^{-\theta_0} - g_H^{-\theta_0}\right)} \\
&\propto -\mathbb{E}^{\mathcal{P}}\left[\log(G_t) \times G_t^{-\theta_0}\right] \\
&\propto -\mathbb{E}^{\mathcal{P}}\left[M_t(G_t;\, \theta_0)\log(G_t)\right] .
\end{aligned}$$

Then we compute the first order differential of the riskfree rate:

$$\begin{aligned}
\frac{\partial R_f}{\partial \theta} &\propto -\left[\pi g_H^{-\theta}(-\log g_h) + (1-\pi)g_L^{-\theta}(-\log g_L)\right] \\
&= \mathbb{E}^{\mathcal{P}}\left[M_t(G_t;\, \theta)\log(G_t)\right]
\end{aligned}$$

Let us turn now to the maximum likelihood problem. The log-likelihood is given by:

$$\mathcal{L}_T(\theta) = \frac{n_T}{T}\log\left[\widehat{\pi}(\theta)\right] + \left(1 - \frac{n_T}{T}\right)\log\left[1 - \widehat{\pi}(\theta)\right] . \tag{40}$$

The first-order differential is given by:

$$\begin{aligned}
\frac{\partial \mathcal{L}_T(\theta)}{\partial \theta} &= \frac{n_T}{T} \times \frac{-\log(g_L)g_L^{-\theta}}{g_L^{-\theta} - \pi g_H^{-\theta_0} - (1-\pi)g_L^{-\theta_0}} + \left(1 - \frac{n_T}{T}\right)\frac{\log(g_H)g_H^{-\theta}}{\pi g_H^{-\theta_0} + (1-\pi)g_L^{-\theta_0} - g_H^{-\theta}} \\
&\quad - \frac{\log(g_H)g_H^{-\theta} - \log(g_L)g_L^{-\theta}}{g_L^{-\theta} - g_H^{-\theta}}
\end{aligned}$$

Applying this function in $\theta = \theta_0$, we obtain:

$$\begin{aligned}
\frac{\partial \mathcal{L}_T(\theta_0)}{\partial \theta} &= \frac{n_T}{T} \times \frac{-\log(g_L) g_L^{-\theta_0}}{g_L^{-\theta_0} - \pi g_H^{-\theta_0} - (1-\pi) g_L^{-\theta_0}} + \left(1 - \frac{n_T}{T}\right) \frac{\log(g_H) g_H^{-\theta_0}}{\pi g_H^{-\theta_0} + (1-\pi) g_L^{-\theta_0} - g_H^{-\theta_0}} \\
&\quad - \frac{\log(g_H) g_H^{-\theta_0} - \log(g_L) g_L^{-\theta_0}}{g_L^{-\theta_0} - g_H^{-\theta_0}} \\
&= -\frac{n_T}{T} \times \frac{\log(g_L) g_L^{-\theta_0}}{\pi \left(g_L^{-\theta_0} - g_H^{-\theta_0}\right)} + \left(1 - \frac{n_T}{T}\right) \frac{\log(g_H) g_H^{-\theta_0}}{(1-\pi)\left(g_L^{-\theta_0} - g_H^{-\theta_0}\right)} \\
&\quad - \frac{\log(g_H) g_H^{-\theta_0} - \log(g_L) g_L^{-\theta_0}}{g_L^{-\theta_0} - g_H^{-\theta_0}} \\
&\propto \left(1 - \frac{n_T/T}{\pi}\right) \log(g_L) g_L^{-\theta_0} + \left(\frac{1 - n_T/T}{1-\pi} - 1\right) \log(g_H) g_H^{-\theta_0}
\end{aligned}$$

Letting the number of observations tend to infinity, we obtain:

$$\begin{aligned}
\frac{\partial \mathcal{L}_T(\theta_0)}{\partial \theta} &\xrightarrow[T \to +\infty]{\mathcal{P}_0} \left(1 - \frac{\pi_0}{\pi}\right) \log(g_L) g_L^{-\theta_0} + \left(\frac{1-\pi_0}{1-\pi} - 1\right) \log(g_H) g_H^{-\theta_0} \\
&\propto (1-\pi)(\pi - \pi_0) \log(g_L) g_L^{-\theta_0} + \pi (\pi - \pi_0) \log(g_H) g_H^{-\theta_0} \\
&= (\pi - \pi_0) \left[(1-\pi) \log(g_L) g_L^{-\theta_0} + \pi \log(g_H) g_H^{-\theta_0}\right] \\
&= (\pi - \pi_0) \mathbb{E}^{\mathcal{P}} \left[M_t(G_t; \theta_0) \log(G_t)\right] \qquad (41)
\end{aligned}$$

A.5 Simulated Economies

A.5.1 Model (I)

We consider an endowment economy where a representative agent has power utility preferences with a constant coefficient of relative risk aversion. Suppose that consumption growth is *i.i.d.* log-normal:

$$\log(G_{t+1}) \overset{\mathcal{P}_0}{\sim} \mathcal{N}\left(\mu, \sigma^2\right). \qquad (42)$$

We assume that the representative investor is possibly pessimistic and acts as if the average consumption growth were lower than μ. Specifically, she acts as if consumption growth has a mean of $\mu^* = (1 - \chi_\mu)\mu$ and volatility $\sigma^* = (1 + \chi_\sigma)\sigma$, where χ_μ, $\chi_\sigma \in (0, 1)$ is the severity of pessimism, and there are no distortions in the beliefs about the volatility:

$$\log(G_{t+1}) \overset{\mathcal{P}}{\sim} \mathcal{N}\left((1 - \chi_\mu)\mu, (1 + \chi_\sigma)^2\sigma^2\right). \tag{43}$$

In the above scenario, equilibrium asset prices reflect the subjective beliefs of investors. Solving the equilibrium, we have that the price-dividend ratio $P_t/D_t = Z$ is constant and equal to:

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

where $Z(\chi)$ is defined by:

$$Z(\chi) = \frac{1}{\frac{1}{\delta} \cdot \exp\left\{(1 - \theta_0)(1 - \chi_\mu)\mu + \frac{(1-\theta_0)^2(1+\chi_\sigma)^2\sigma^2}{2}\right\} - 1}\,. \tag{44}$$

In this economy, the Euler equation holds under the subjective probability measure $\mathcal{P}$. We have:

$$\delta \mathbb{E}^{\mathcal{P}}\left[G_t^{-\theta_0} R_{m,t}(\chi)\right] = 1\,.$$

We can use the EL approach (without conditioning variables) to estimate the distorted unconditional distribution of consumption growth, $\mathcal{P}$. Our probabilities solve the following problem:

$$\{\widehat{p}(\theta)\} = \arg\max_{\{p_t\}\in\Delta} \sum_{t=1}^{T} \log(p_t) \quad \text{s.t.} \quad \sum_{t=1}^{T} p_t \cdot g_t^{-\theta_0} \cdot r_{m,t} = 1\,. \tag{45}$$

A.5.2 Model (II)

We consider the external habit model of Campbell and Cochrane (1999). In this model, the representative investor has power utility preferences over consumption in excess of a time-varying habit (or, subsistence) level. The aggregate consumption growth is assumed to follow an *i.i.d.* process:

$$\log(G_{t+1}) \overset{\mathcal{P}_0}{\sim} \mathcal{N}\left(\mu, \sigma^2\right). \tag{46}$$

The stochastic discount factor is given by Equation (51). The log surplus consumption ratio evolves as a heteroskedastic AR(1) process:

$$\log(S_t) = (1-\phi)\,\overline{\log(S)} + \phi \log(S_{t-1}) + \lambda\left(\log(S_{t-1})\right) v_t, \tag{47}$$

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

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

$$s_{max} = \overline{\log(S)} + \frac{1}{2}\left(1-\overline{S}^2\right), \qquad \overline{S} = \sigma\sqrt{\frac{\theta_0}{1-\phi}}.$$

In this model, Campbell and Cochrane (1999) show that the riskfree rate is constant over time and that the price-consumption and price-dividend ratios vary over time and depend on S_t.

A.5.3 Model (III)

Model (III) is the long run risks model of Bansal and Yaron (2004). In this model, under the true probability measure $\mathcal{P}_0$, aggregate consumption and dividend growth rates have a small persistent predictable component, F_t and stochastic volatility, Σ_t, that captures time-varying economic uncertainty:

$$\begin{aligned}
\log\left(G_{t+1}\right) &= \mu_c + F_t + \Sigma_t \epsilon_{c,t+1}, \\
\log\left(G_{d,t+1}\right) &= \mu_d + \phi F_t + \phi_d \Sigma_t \epsilon_{d,t+1}, \\
F_{t+1} &= \rho F_t + \phi_e \Sigma_t \epsilon_{f,t+1}, \\
\Sigma_{t+1}^2 &= (1-\nu)\sigma^2 + \nu \Sigma_t^2 + \sigma_w \epsilon_{w,t+1},
\end{aligned} \tag{48}$$

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:

$$\mathbb{E}^{\mathcal{P}_0}\left[\delta^{\psi_0}\left(G_{t+1}\right)^{-\frac{\psi_0}{\rho_0}} R_{c,t+1}^{\psi_0-1} \cdot \mathbf{R}_{t+1} \,\Big|\, X_t, \Sigma_t^2\right] = \mathbf{1}, \tag{49}$$

where $R_{c,t+1}$ is the unobservable return on total wealth. We solve for the equilibrium as in the original article and set the model parameters equal to the authors' calibrated values.

Bansal and Yaron (2004) show that, in equilibrium, the (log) price-consumption ratio is an affine function of the two state variables (F_t, Σ_t^2), with coefficients that are known functions of the underlying model parameters. This renders the return on total wealth, $R_{c,t+1}$ a known function of the two state variables. Therefore, the SDF is a known function of the two state variables in addition to consumption growth. The equilibrium risk free rate is also a linear function of the two state variables. Finally, the price-dividend ratio is a known function of the two state variables and, therefore, the market return is a function of the two state variables in addition to the dividend growth rate.

A.5.4 Summary of simulated models

Table 6 – Simulated models and potential misspecifications

	Model (I)		Model (II)		Model (III)
	(I.1)	*(I.2)*	*(II.1)*	*(II.2)*	*(III)*
G_{t+1} DGP $\mathcal{P}_0$	$\mathcal{N}(\mu, \sigma^2)$		$\mathcal{N}(\mu, \sigma^2)$		long-run risk
True SDF M_{t+1}	CRRA		External habit		Epstein-Zin
Misspecification	X	beliefs: $\mathcal{N}\left(\mu^*, {\sigma^*}^2\right)$	X	SDF: CRRA	X

The table presents the various simulated economies we consider to assess the performance of the SEL estimator in empirically realistic sample sizes ($T = 267$). For case (I.2), we assume $\mu^* = (1 - \chi_\mu)\mu$ and $\sigma^* = (1 + \chi_\sigma)\sigma$, where $\chi_\mu = \{15\%, 50\%, 75\%\}$ and $\chi_\sigma = \{0\%, 15\%, 75\%\}$.

A.6 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 return 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.

A.7 Different Stochastic Discount Factors

A.7.1 CRRA preferences

Our first choice of SDF corresponds to power utility preferences 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).

$$M_{t+1} = \delta \cdot G_{t+1}^{-\theta_0} \tag{50}$$

where δ denotes the subjective discount factor and θ_0 the relative risk aversion coefficient.

A.7.2 External habit formation

Second, we consider 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:

$$M_{t+1}(G_{t+1}, Y_{t+1}; \theta_0) = \delta \left(\frac{S_{t+1}}{S_t} \cdot G_{t+1} \right)^{-\theta_0}, \tag{51}$$

where δ is the subjective time discount factor, θ_0 is a utility curvature parameter that provides a lower bound on the time varying CRRA, $S_t = \frac{C_t - H_t}{C_t}$ denotes the surplus consumption ratio, and H_t is the habit level.

Note that the SDF depends on the surplus consumption ratio, S_t, which is not directly observed. In the simulation experiment where, we use the simulated series of S_t and we consider the habits are directly observable to the econometrician (see appendix A.5). For the application on historical data, we extract the time series of the surplus consumption ratio from observed consumption data as follows. We measure H_t by a weighted average of past real consumption in non-durables and in services.

$$H_t = (1 - \varphi) \cdot H_{t-1} + \varphi \cdot C_{t-1}\,. \tag{52}$$

The above dynamics of log habit would emerge as a log-linear approximation around the non-stochastic steady-state for a variety of parametric assumptions about the dynamics of the consumption growth rate and the surplus consumption ratio. This renders the SDF fully observable and, therefore, our SEL approach can be applied to recover the investors' beliefs.

A.7.3 Epstein-Zin preferences

Our last specification of the pricing kernel is that implied by the recursive preferences of Epstein and Zin (1989) and Weil (1989), for which the SDF is given by

$$M_{t+1}(G_{t+1}, Y_{t+1}; \theta_0) = \delta^{\psi} \left(G_{t+1}\right)^{-\frac{\psi}{\rho}} R_{c,t+1}^{\psi-1},$$

where $R_{c,t+1}$ is the unobservable gross return on an asset that delivers aggregate consumption as its dividend each period, δ is the subjective time discount factor, ρ_0 is the elasticity of intertemporal substitution, $\psi := \frac{1-\gamma}{1-1/\rho}$, and γ_0 is the relative risk aversion coefficient. In this case, following our formulation, $\theta = (\rho, \psi)$.

We approximate the unobservable $R_{c,t+1}$ with the observed return on the aggregate stock market. As with the external habit model, the SDF, therefore, becomes fully observable rendering it amenable to SEL estimation of the investors' beliefs.

A.8 Robustness of price-consistent beliefs

All of our robustness cases can be explored and downloaded on our companion website at `https://guillaumeroussellet.shinyapps.io/MeasuringBeliefs/`.

A.8.1 Robustness to Choice of Bandwidth Parameter

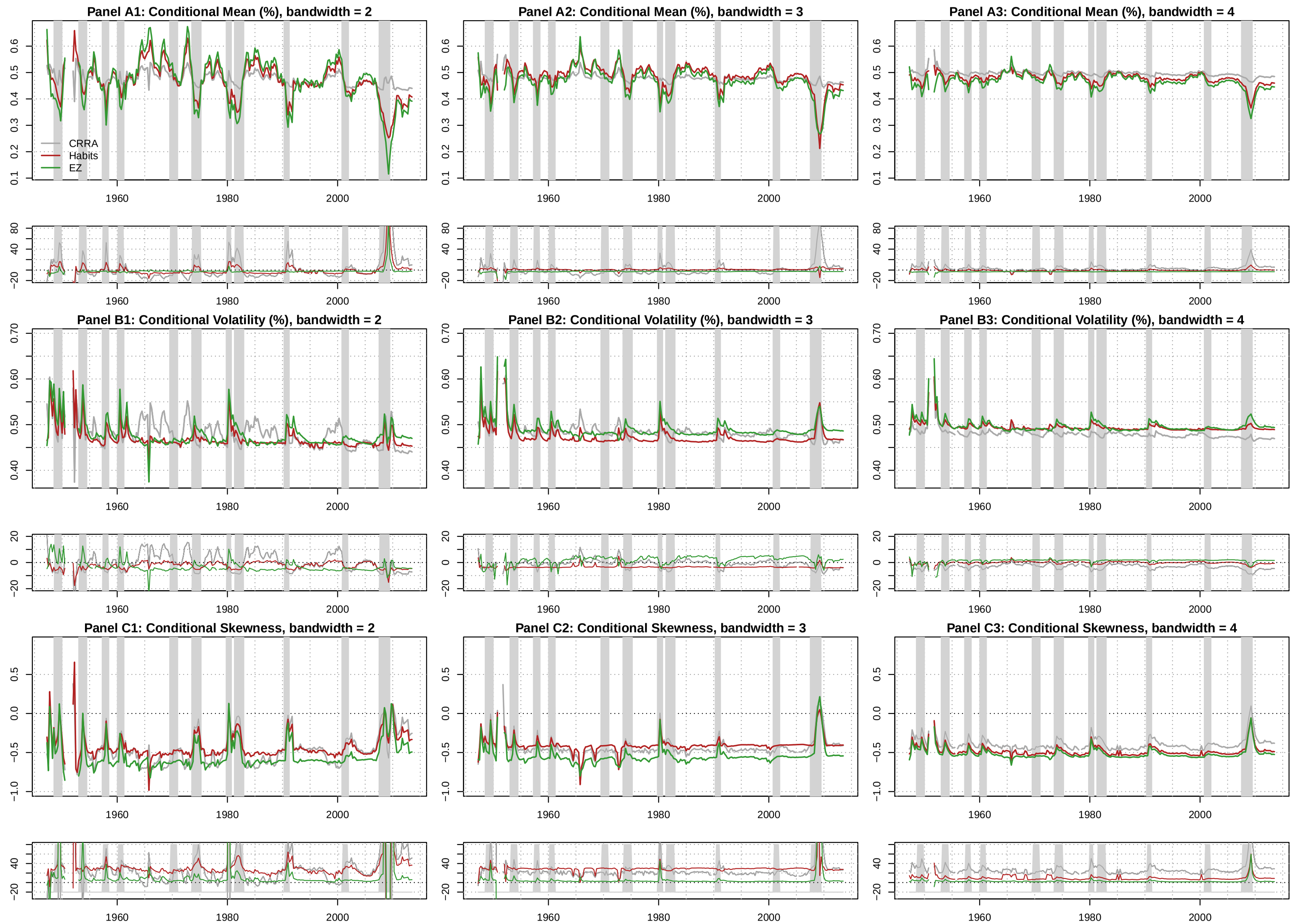
Figure 9 presents the time series of the conditional mean (Panel A), conditional volatility (Panel B), and conditional skewness (Panel C) of consumption growth obtained using the estimated SEL distributions, for different choices of the bandwidth parameter, b_T. The middle panel (2) presents the results for our baseline case, $b_T = 3$, while left and right panels present cases of $b_T = 2$ and $b_T = 4$, respectively.

The figure shows that the recovered beliefs about consumption growth are fairly robust to plausible choices of the bandwidth parameter. In particular, the strongly cyclical beliefs about expected consumption growth and the skewness of consumption growth obtain for all these choices of the bandwidth parameter, with smaller values of this parameter leading to bigger cyclical variation in these moments. Meanwhile, the flat conditional volatility of consumption growth is also a feature of the recovered beliefs for each value of the bandwidth, although lower bandwidth naturally produce larger volatilities of all time series.

Note that the above choices of the bandwidth parameter may be argued to be fairly broad. To provide further intuition regarding this, note that $b_T = 4$ may be interpreted as a four-standard deviation neighborhood around the current state being regarded as encompassing the possible states of the world that may be realized in the subsequent period. Thus, the set of possible states allowed for is quite large and accommodates the possibility of rare disaster states being realized. Consider, for instance, the quarter with the highest realized consumption growth in the post war period – 1965:Q4 with a quarterly consumption growth of 2.0%. The quarterly volatility of consumption growth over the post war period is 0.5%. Therefore, a four standard error interval around the realized state includes $[-.03\%, 4.0\%]$, i.e. even though the realized consumption growth is very high, the SEL method still incorporates the possibility of a severe economic downturn in the subsequent period. Conversely, consider the quarter with the lowest realized consumption growth over this period – the aftermath of World War II in 1947:Q4 with a consumption growth of -1.3%. A four standard error interval around the realized state now includes $[-3.3\%, .7\%]$, i.e. even during very bad times, the method allows for the possibility of substantial recoveries in the subsequent period.

To summarize, the above findings suggest that our results are unlikely to be driven by specific choices of the bandwidth parameter.

Figure 9 – Robustness to Choice of Bandwidth Parameter



Notes: The figure plots the time series of price-consistent conditional mean, volatility and skewness of consumption growth (Panels A, B and C, resp.) for bandwidth values of 2 (panel 1), 3 (benchmark case, panel 2) and 4 (panel 3). Smaller interpanels present the relative divergence with respect to the objective measure. Shaded areas correspond to NBER recessions. The conditional moments are obtained using the estimated SEL distributions. The pricing kernels correspond to CRRA (grey), habits (red) and Epstein-Zin (green). Test assets include market returns, Tbill rate and small, big, growth and value portfolios. The conditioning set includes past consumption growth and a consumption growth volatility proxy. Sample is 1947:Q2-2013:Q4.

A.8.2 Robustness to Cross Sections of Assets

We assess the robustness of the results by considering an alternative set of test assets, namely combinations of the excess returns on the market portfolio, the 'Small' and 'Big' portfolios (the bottom and top deciles of portfolios formed by sorting stocks on the basis of market capitalization) and the '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), along with the T-bill rate.

For all sets of preferences, price-consistent means of consumption growth are all highly correlated, and nearly indistinguishable from each other for the habits and EZ preferences. The effect of adding the riskless asset to the set of test assets can be clearly seen for the volatility estimates using CRRA preferences. Considering the T-bill increases the volatility and shows less trivial cyclical fluctuations. The same effect can be observed for Habits, but not for recursive preferences where all time series are confounded. Last, our conclusions of a negative average skewness spiking up during some recessions are largely unchanged by the choice of alternative test assets, although some magnitude differences are observed for CRRA and habit preferences.

A.8.3 Robustness to Instruments

Next, we show that the recovered beliefs are also robust to the specification of the conditioning set. Ideally, we would like to include as many variables as possible in the conditioning set. However, as with all non-parametric estimators, our approach suffers from the curse of dimensionality. We, thus, consider several alternative specifications for the conditioning set.

Recall that, in our baseline specification, the conditioning set consists of variables related to the history of consumption and a proxy for the consumption volatility, namely the squared past consumption shocks as measure by an AR(1) residual.

We add additional macro variables to the conditioning set. Our choices for these additional variables draw 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.

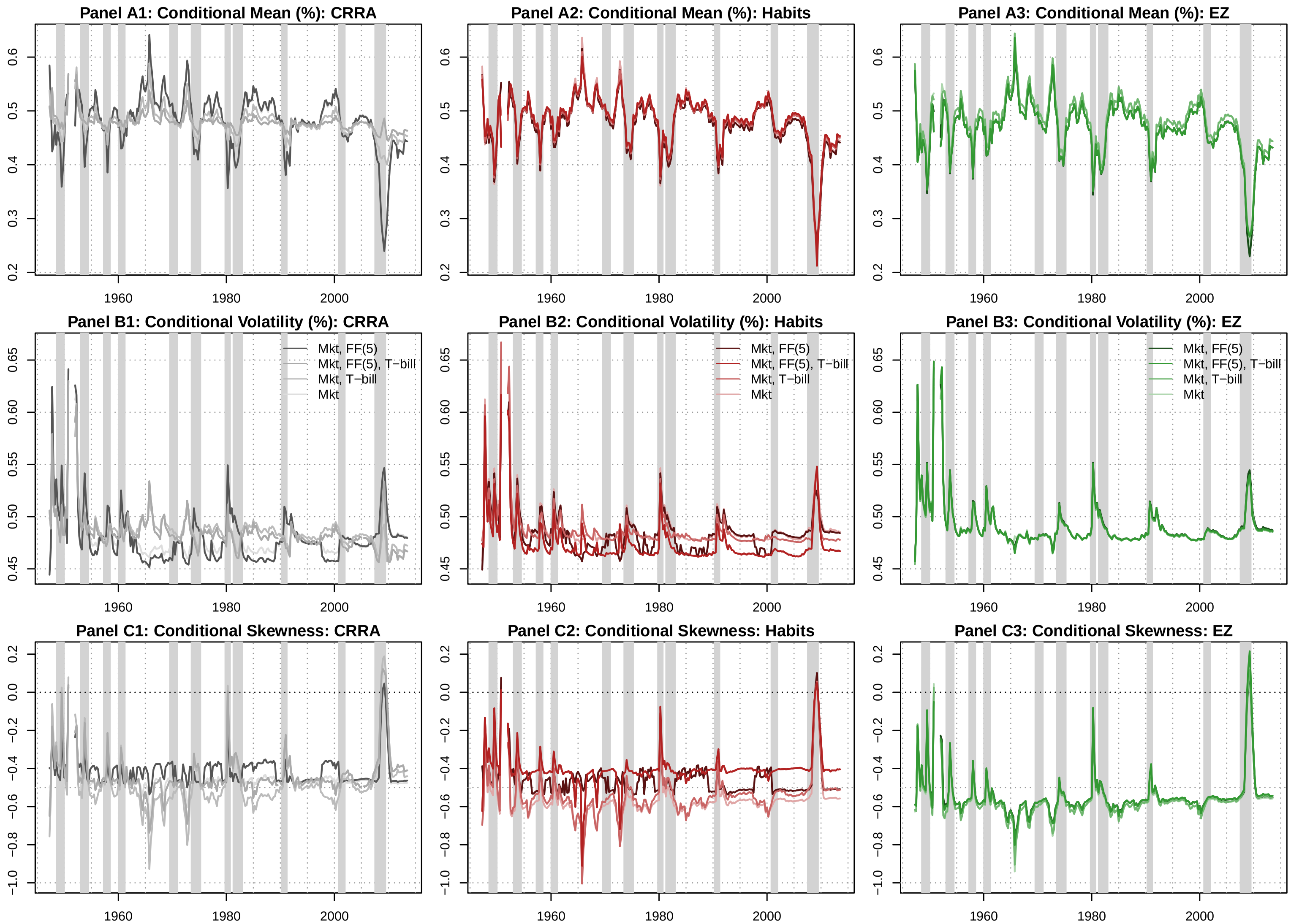
Finally, to address the concern that additional variables, not included in the above choices, may be in the information set of investors, we extract principal components (PCs) from a broad cross-section of over a hundred macro variables. The variables cover six broad categories of macroeconomic data: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

Figure 11 presents the time series of the first three moments of consumption growth recovered from eight different choices of the conditioning set. The choices include an exponentially-weighted moving average of past consumption growth (Cg), variance proxy (var), inflation (inf), growth in the average hourly earnings of production in private non farm payrolls (ear), returns on test assets used in the estimation (ret), first and second principal components from a broad cross section of over a hundred macro variables (PC1 and PC2, resp.)

The figure shows that our results are quite robust to the choice of the conditioning set. Most of the business cycle fluctuations are similar across conditioning sets, with some differences in the average level of moments. For consumption growth mean, we see that our baseline conditioning set tends to convey less volatile fluctuations than the other cases, while the latter stay in a constrained range compared to the realized consumption growth. Volatility series all show contained movements, and the use of the principal components in the conditioning set creates the biggest difference by lowering the average volatility by about 10bps (non annualized). Last, all skewness estimates are consistently negative, except during recessions where they spike up, showing procyclicality.

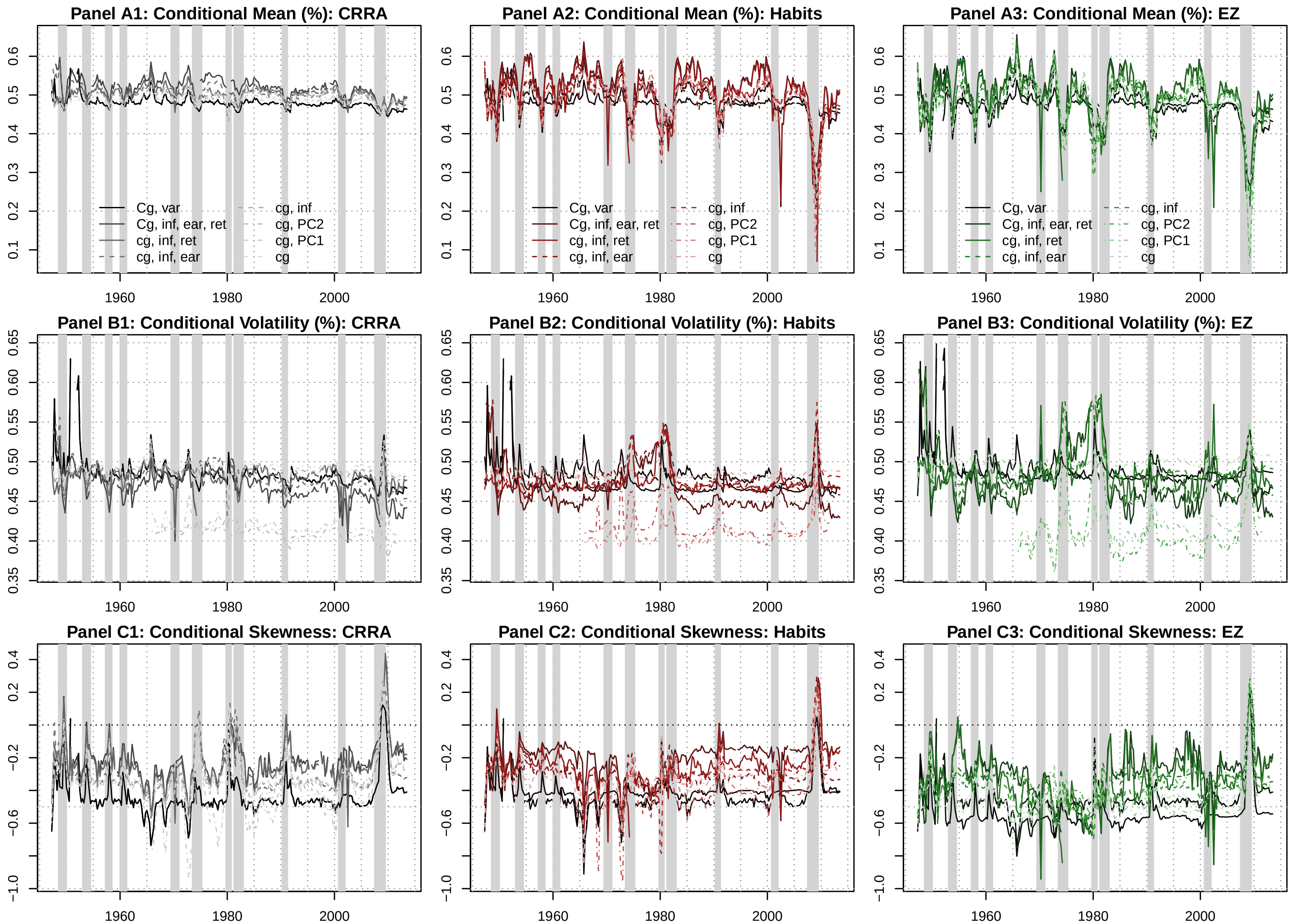
We also present pairwise correlations between the time series of conditional moments in Table 7. Nearly all correlations are above 50%, and most of them lie above 75%.

Figure 10 – Time series of Conditional Moments, Alternative Test Assets



Notes: The figure plots the time series of price-consistent conditional mean, volatility and skewness of consumption growth (Panels A, B and C, resp.). Each plot presents the price-consistent beliefs for 4 sets of test assets: (dark to light) Market + small, big, growth and value portfolios, Market + small, big, growth and value portfolios + T-bill, Market + T-bill, and Market only. Shaded areas correspond to NBER recessions. The conditional moments are obtained using the estimated SEL distributions. The pricing kernels correspond to CRRA (grey), habits (red) and Epstein-Zin (green). The conditioning set includes past consumption growth and a consumption growth volatility proxy. Sample is 1947:Q2-2013:Q4.

Figure 11 – Time series of Conditional Moments, Alternative Instruments



Notes: The figure plots the time series of price-consistent conditional mean, volatility and skewness of consumption growth (Panels A, B and C, resp.). Each plot presents the price-consistent beliefs for 8 sets of conditioning variables combining: past consumption growth (Cg), variance proxy (var), inflation (inf), earnings growth (ear), returns on assets (ret), first and second principal components of macroeconomic variables (PC1 and PC2, resp.). Shaded areas correspond to NBER recessions. The conditional moments are obtained using the estimated SEL distributions. The pricing kernels correspond to CRRA (grey), habits (red) and Epstein-Zin (green). Test assets are Market, small, big, growth and value portfolios, and T-bill. Sample is 1947:Q2-2013:Q4.

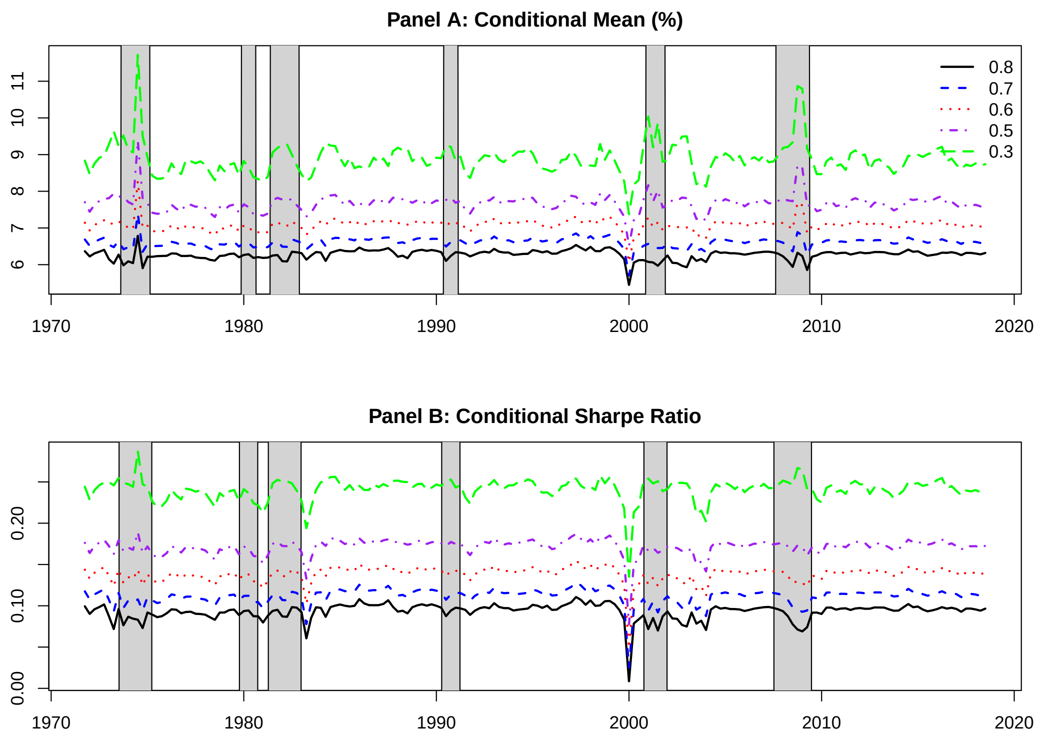
Table 7 – Correlation of beliefs: alternative instruments

Panel A: Conditional Mean																				
A1: CRRA							A2: HABITS							A3: EZ						
(2)	(3)	(4)	(5)	(6)	(7)	(8)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
0.89	0.69	0.61	0.94	0.80	0.16	0.56	0.73	0.41	0.44	0.85	0.34	0.35	0.30	0.82	0.96	0.84	0.94	0.91	0.71	0.87
1.00	0.64	0.76	0.87	0.83	0.34	0.58	1.00	0.26	0.64	0.66	0.36	0.16	0.31	1.00	0.73	0.97	0.87	0.68	0.68	0.80
	1.00	0.92	0.74	0.82	0.22	0.67		1.00	0.87	0.43	0.72	0.20	0.49		1.00	0.77	0.85	0.92	0.66	0.80
		1.00	0.69	0.77	0.39	0.65			1.00	0.49	0.70	0.18	0.48			1.00	0.86	0.75	0.74	0.78
			1.00	0.84	0.20	0.62				1.00	0.44	0.50	0.39				1.00	0.82	0.71	0.93
				1.00	0.28	0.84					1.00	0.27	0.60					1.00	0.72	0.80
					1.00	0.35						1.00	0.51						1.00	0.65
Panel B: Conditional Volatility																				
B1: CRRA							B2: HABITS							B3: EZ						
(2)	(3)	(4)	(5)	(6)	(7)	(8)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
0.84	0.93	0.81	0.95	0.95	0.90	0.95	0.88	0.45	0.49	0.67	0.81	0.57	0.37	0.79	0.86	0.74	0.75	0.67	0.66	0.66
1.00	0.75	0.99	0.79	0.78	0.67	0.78	1.00	0.53	0.65	0.42	0.80	0.27	0.20	1.00	0.59	0.86	0.69	0.56	0.66	0.64
	1.00	0.76	0.90	0.91	0.86	0.90		1.00	0.97	0.11	0.74	0.10	0.35		1.00	0.54	0.83	0.85	0.77	0.79
		1.00	0.77	0.74	0.71	0.77			1.00	0.08	0.75	0.08	0.26			1.00	0.57	0.54	0.59	0.51
			1.00	0.96	0.97	0.99				1.00	0.52	0.81	0.50				1.00	0.90	0.79	0.88
				1.00	0.93	0.95					1.00	0.36	0.64					1.00	0.69	0.86
					1.00	0.97						1.00	0.68						1.00	0.74
Panel C: Conditional Skewness																				
C1: CRRA							C2: HABITS							C3: EZ						
(2)	(3)	(4)	(5)	(6)	(7)	(8)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
0.80	0.99	0.82	0.93	0.95	0.87	0.93	0.77	0.73	0.63	0.55	0.90	0.46	0.27	0.64	0.85	0.56	0.84	0.52	0.74	0.73
1.00	0.78	0.99	0.75	0.77	0.67	0.74	1.00	0.62	0.75	0.28	0.74	0.25	0.10	1.00	0.57	0.90	0.44	0.56	0.42	0.41
	1.00	0.82	0.95	0.96	0.90	0.95		1.00	0.95	0.35	0.84	0.30	0.38		1.00	0.74	0.64	0.68	0.59	0.57
		1.00	0.79	0.79	0.74	0.79			1.00	0.15	0.72	0.13	0.23			1.00	0.25	0.65	0.29	0.20
			1.00	0.95	0.96	1.00				1.00	0.68	0.95	0.46				1.00	0.31	0.84	0.89
				1.00	0.95	0.95					1.00	0.61	0.75					1.00	0.28	0.41
					1.00	0.96						1.00	0.78						1.00	0.82

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 set considered, for the CRRA pricing kernel (Panel 1), Habit (Panel 2), and Epstein-Zin (Panel 3).

A.9 Replication of Ghosh et al. (2019)

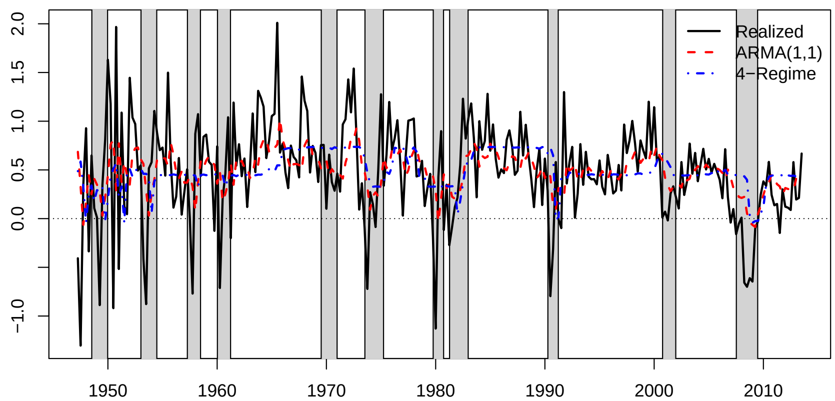
Figure 12 – Beliefs About the Stock Market Across Heterogeneous Investors with Optimal Portfolio $R_W = \omega_{RF} R_f + \omega_{SG} R_{SG} + (1 - \omega_{RF} - \omega_{SG}) R_m$, for $\omega_{RF} = 0.2$ and Different Values of ω_{SG}



Notes: The figure plots the time series of the conditional mean (Panel A) and Sharpe ratio (Panel B) of the market return, as perceived by different investor types. Shaded areas denote NBER designated recession periods. Each investors' conditional moments are obtained using the estimated SEL distributions. The pricing kernel is exponentially affine in the portfolio return for that investor. The test assets consist of the excess returns on the market portfolio and the portfolio held by that investor type. The conditioning set consists of an exponentially-weighted moving average of the returns on the two test assets. The sample is quarterly covering the period 1972:Q1-2018:Q4.

A.10 Conditional means of ARMA and regime switching consumption models

Figure 13 – Expected and Realized Consumption Growth (%): ARMA(1,1) and Regime Switching Models



Notes: The figure plots the historical time series of consumption growth (black line), along with its conditional mean implied by the ARMA (1,1) model (red line) and the regime switching model (blue line), over the period 1947:Q1-2013:Q4. The models' parameters are estimated via quasi maximum likelihood, using consumption data alone.

A.11 Results of the ARMA-GARCH Process for the Consumption Growth

We estimate the following model:

$$\log(G_{t+1}) = (1-\psi)g + \psi \log(G_t) + \sigma_{t+1}\nu_{t+1} + \theta\sigma_t\nu_t\,, \quad \text{where} \quad \nu_t \sim \mathcal{D}(0,\,1)$$

$$\sigma_{t+1}^2 = \omega + (\alpha + \gamma \mathbb{1}\{\nu_t < 0\})\,\sigma_t^2\nu_t^2 + \beta\sigma_t^2\,,$$

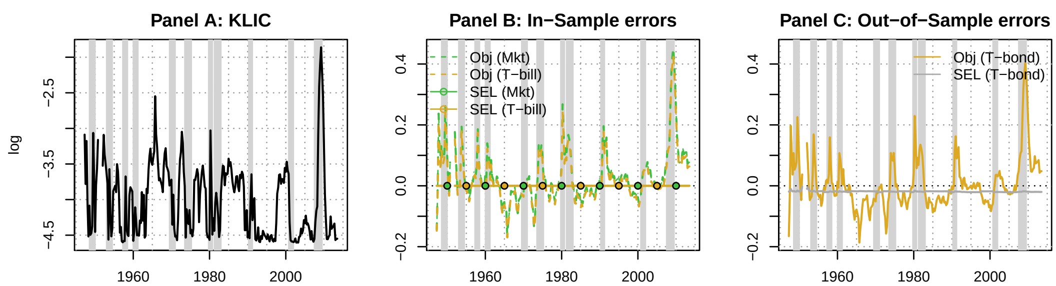
Table 8 – ARMA-GARCH estimated parameters

	g	ψ	θ	ω	α	β	γ
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)

that is an ARMA(1,1) mean model and a GJR-GARCH model. In addition, we do not assume that the standardized shocks ν_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 8.

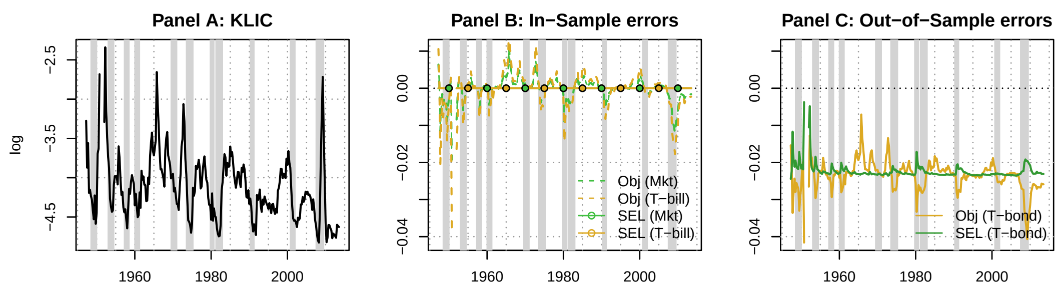
A.12 Pricing errors: CRRA and recursive preferences

Figure 14 – Conditional Pricing Errors and Kullback-Leibler Divergence: CRRA preferences



Notes: The figure plots the (log) KLIC divergence used for estimation (A), the in-sample pricing errors (B) for market returns (green) and T-bill (yellow) with and without pricing constraints (solid dotted and dashed, resp.), and out-of-sample errors on the 10y T-bond real quarterly returns (C) with and without pricing constraints (yellow and red, resp.). Shaded areas correspond to NBER recessions. The conditional moments are obtained using the estimated SEL distributions. The pricing kernel corresponds to CRRA. Test assets include market returns, Tbill rate andsmall, big, growth and value portfolios. The conditioning set includes past consumption growth and a consumption growth volatility proxy. Sample is 1947:Q2-2013:Q4.

Figure 15 – Conditional Pricing Errors and Kullback-Leibler Divergence: EZ preferences



Notes: The figure plots the (log) KLIC divergence used for estimation (A), the in-sample pricing errors (B) for market returns (green) and T-bill (yellow) with and without pricing constraints (solid dotted and dashed, resp.), and out-of-sample errors on the 10y T-bond real quarterly returns (C) with and without pricing constraints (yellow and red, resp.). Shaded areas correspond to NBER recessions. The conditional moments are obtained using the estimated SEL distributions. The pricing kernel corresponds to Epstein-Zin. Test assets include market returns, Tbill rate and small, big, growth and value portfolios. The conditioning set includes past consumption growth and a consumption growth volatility proxy. Sample is 1947:Q2-2013:Q4.

References

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.

Antoine, B., H. Bonnal, and É. 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. 2018. “New Entropy Restrictions and the Quest for Better Specified Asset Pricing Models”. *Journal of Financial and Quantitative Analysis.*

Bansal, R., and A. Yaron. 2004. “Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles”. *Journal of Finance* 59, no. 4 (): 1481–1509.

Breeden, D. T. 1979. “An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities”. *Journal of Financial Economics* 7:265–96.

Campbell, J., and J. Cochrane. 1999. "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior". *Journal of Political Economy* 107, no. 2 (): 205–51.

Cressie, N., and T. Read. 1984. "Multinomial Goodness-of-Fit Tests". *Journal of the Royal Statistics Society Series B* 46:440–464.

Epstein, L., and S. Zin. 1989. "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework". *Econometrica* 57:937–968.

Ghosh, A., and G. Constantinides. 2017. "What Information Drives Asset Prices?"

Ghosh, A., A. Korteweg, and Q. Xu. 2019. "Recovering Heterogeneous Beliefs and Preferences from Asset Prices". Working paper.

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

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

Owen, A. B. 2001. *Empirical Likelihood*. Chapman / Hall.

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*. Tech. rep.

Weil, P. 1989. “The Equity Premium Puzzle and the Risk-Free Rate Puzzle”. *Journal of Monetary Economics* 24:401–421.