The Market Cost of Business Cycle Fluctuations

By Anisha Ghosh, Christian Julliard, and Michael J. Stutzer*

Draft: June 14, 2019

We propose a novel measure of the costs of aggregate economic fluctuations, that does not require the specification of consumers' preferences or the dynamics of the data generating process. Using data on consumption and asset prices, we rely on an information-theoretic approach to recover the *information kernel* (I-SDF). The I-SDF accurately prices broad cross-sections of assets, in- and out-of-sample, and has a strong business cycle component. Using the I-SDF, we find that the welfare benefits of eliminating *all* consumption fluctuations are large on average, and are strongly time-varying and countercyclical. Moreover, the cost of business cycle fluctuations is substantial, accounting for about a quarter to a third of the cost of all consumption fluctuations.

JEL: E3, E2, G12, C5. Keywords: Aggregate Uncertainty, Business Cycle Risk, Pricing Kernel, Empirical Likelihood, Smoothed Empirical Likelihood.

I. Introduction

In his seminal 1987 monograph, Robert E. Lucas Jr. concludes that the welfare benefit of eliminating *all* consumption fluctuations in the U.S. economy is trivially small, hence challenging the desirability of policies aimed at insulating the economy from cyclical fluctuations. As Lucas emphasizes,[1] this result is obtained without taking a stand on the origins of aggregate fluctuations, and it relies solely on the specifications of preferences (a representative agent with time and state separable power utility preferences with a constant coefficient of relative risk aversion) and the data generating process (log-normal aggregate consumption growth rate).

Nevertheless, it is exactly these two assumptions that make Lucas' calculations questionable. This is because evaluating the welfare cost of business cycles is tantamount to *pricing* the risk that households face due to aggregate fluctuations, and an extensive literature has documented how Lucas' specification of preferences and the dynamics of the consumption process grossly underestimate the

* Anisha Ghosh: Desautel Faculty of Management, McGill University, anisha.ghosh@mcgill.ca. Christian Julliard: Department of Finance, London School of Economics, c.julliard@lse.ac.uk. Michael J. Stutzer: Leeds School of Business, The University of Colorado, Boulder, michael.stutzer@colorado.edu.

[1] *"these calculations rest on assumptions about preferences only, and not about any particular mechanism equilibrium or disequilibrium – assumed to generate business cycles"*, Lucas (1987).

market price of risk in the U.S. economy: e.g., the average premium on a broad U.S. stock market index over and above short-term Treasury Bills has been about 7% per year over the last century, while Lucas' specification would imply a premium of less than 1%.[2] Not only does Lucas' specification grossly underestimate the historically observed average return on the aggregate stock market index, it also fails to explain the significant cross-sectional differences in average returns between broad diversified portfolios formed by sorting individual stocks on the basis of observable characteristics (e.g., market value of equity, book-to-market equity) that have been identified to be proxies for underlying sources of systematic risk (see e.g., Lars Peter Hansen and Kenneth J. Singleton (1983), Martin Lettau and Sydney Ludvigson (2001), Jonathan A. Parker and Christian Julliard (2005), Christian Julliard and Anisha Ghosh (2012)).

Indeed, exactly due to the inability of the power utility, log-normal setting to match households' preferences toward risk revealed by the prices of financial assets, a burgeoning literature, based on modifying the preferences of investors and/or the dynamic structure of the economy, has developed. In these models, the resulting pricing kernel (hereafter referred to as the stochastic discount factor or SDF) can be factored into an observable component consisting of a parametric function of consumption growth as with power utility, and a (potentially unobservable) model-specific component. That is, the pricing kernel, M, in these models is of the form

$$M_{t+1} = (C_{t+1}/C_t)^{-\gamma}\,\psi_{t+1}. \tag{1}$$

The Robert E. Lucas (1987) original setting is nested within this family in that it corresponds to the case in which ψ_t is a positive constant and the parameter γ is the Arrow-Pratt relative risk aversion coefficient. Prominent examples of models in this class are: habit formation models (see, e.g., John Y. Campbell and John H. Cochrane (1999), Lior Menzly, Tano Santos and Pietro Veronesi (2004)); long run risks models based on recursive preferences (e.g., Ravi Bansal and Amir Yaron (2004)); models with complementarities in consumption (e.g., Monika Piazzesi, Martin Schneider and Selale Tuzel (2007), Motohiro Yogo (2006)); models in which ψ_t captures departures from rational expectations (e.g. Suleyman Basak and Hongjun Yan (2010)), robust control behavior (e.g. Lars Peter Hansen and Thomas J. Sargent (2010)), aggregation over heterogeneous agents who face uninsurable idiosyncratic shocks to their labor income (e.g. George M. Constantinides and Darrell Duffie (1996), George M. Constantinides and Anisha Ghosh (2017)), as well as solvency constraints (e.g. Hanno N. Lustig and Stijn G. Van Nieuwerburgh (2005)).

Estimates of the cost of business cycles vary widely across these model specifications (see, e.g., Gadi Barlevy (2005) for a survey). More importantly, as with Lucas' original specification, in order for any of the more recent models to constitute a good choice for welfare cost calculations, it should accurately price broad categories of assets. Anisha Ghosh, Christian Julliard and Alex Taylor

[2]This discrepancy is the so-called Equity Premium Puzzle, first identified by Rajnish Mehra and Edward C. Prescott (1985).

(2016*b*) evaluate the pricing performance of several of these consumption-based models and show that they perform quite poorly, producing large pricing errors and low (and often negative) cross-sectional R^2. Therefore, the shortcomings of using Lucas' specification for welfare cost calculations also apply to most of the more recent advances.

In this paper, we do not take a stand on either the preferences of investors, or on the dynamics of the underlying state variables. Rather, we rely on the insight that asset prices contain information about the stochastic discounting of the different possible future states and, therefore, use observed asset prices to recover the SDF. Specifically, we assume that the underlying SDF has the multiplicative form in Equation (1). We use financial asset returns and consumption data to extract, non-parametrically, the *minimum entropy* estimate of the ψ-component of the pricing kernel M such that the resultant M satisfies the unconditional Euler equations for the assets, i.e. successfully prices broad cross sections of assets. This information-theoretic approach, that has its origins in the physical sciences, adds to the standard power utility kernel the *minimum* amount of additional information needed to price assets perfectly, i.e. satisfy the Euler equations. We refer to the estimated M as the *information SDF* (I-SDF) because of the information-theoretic methodology used to recover it.

With this model-free SDF at hand, we obtain the cost of aggregate consumption fluctuations as the ratio of the (shadow) prices of two hypothetical securities – a claim to a *stabilized* version of the aggregate consumption stream from which certain types of fluctuations (e.g., all fluctuations or fluctuations corresponding to business cycle frequencies only) have been removed and a claim to the actual aggregate consumption stream. Fernando Alvarez and Urban J. Jermann (2004) show that, in the context of a representative agent economy, the above ratio measures the *marginal cost* of consumption fluctuations, defined as the per unit benefit of a marginal reduction in consumption fluctuations, expressed as a percentage of lifetime consumption. Our approach allows us to estimate the term structure of the cost of fluctuations, i.e. how the cost (or, the welfare benefit of removing fluctuations) rises with the elimination of aggregate fluctuations over each additional future period.

Our information-theoretic approach to the recovery the SDF corresponds to the empirical likelihood (EL) estimator of Art B. Owen (2001). Using this methodology to recover the (multiplicative) missing component of SDF in a model-free way was originally proposed in Ghosh, Julliard and Taylor (2016*b*). We show that the I-SDF, unlike Lucas' original specification, accurately prices broad cross sections of assets.[3] It, therefore, offers a more reliable choice for assessing investors' attitude toward risk. Also, the I-SDF, unlike Lucas' specification, has a significant business cycle component, suggesting that business cycle risk constitutes an important source of priced risk. Therefore, not surprisingly, we show that the I-SDF implies a larger cost of business cycle fluctuations than those obtained with

[3]See also Anisha Ghosh, Christian Julliard and Alex Taylor (2016*a*) who show that the I-SDF, estimated in a purely out-of-sample fashion, accurately prices the aggregate stock market, broad cross-sections of equity portfolios constructed by sorting stocks on the basis of different observable characteristics (e.g., size, book-to-market-equity, prior returns, industry), as well as currency portfolios and portfolios of commodity futures.

Lucas' specification.

We first apply our methodology to assess the welfare benefits of eliminating *all* consumption fluctuations. This is obtained as the ratio of the price of a claim to the aggregate consumption stream from which all uncertainty has been removed (i.e., where the aggregate consumption growth in each period is replaced with its unconditional mean) and the price of a claim to the risky actual aggregate consumption stream. The I-SDF implies a substantially higher cost of all consumption fluctuations compared to Lucas' original specification. For instance, when the I-SDF is extracted using nondurables and services consumption and with the excess return on the market portfolio as the sole asset, the implied costs of all consumption fluctuations over the next one to five years are 1.5%, 5.2%, 11.8%, 14.3%, and 14.4%, respectively. The corresponding costs for Lucas' specification are typically an order of magnitude smaller at 0.8%, 1.1%, 1.4%, 1.7%, and 1.9%, respectively. These conclusions are robust to the measure of aggregate consumption expenditure used (nondurables and services consumption versus total consumption that also includes expenditure on durables) and the set of test assets used to recover the I-SDF. Our results suggest that economic agents perceive the cost of aggregate economic fluctuations to be quite substantial.

We next use our framework to estimate the costs of business cycle fluctuations. This is obtained as the ratio of the price of a claim to the aggregate consumption stream from which fluctuations corresponding to business cycle frequencies have been removed (using the standard Hodrick-Prescott filter) and the price of a claim to the actual aggregate consumption stream. We find that the cost of business cycle fluctuations is large and constitutes between a quarter to a third of the cost of all consumption fluctuations. For instance, when the I-SDF is extracted using nondurables and services consumption and with the excess return on the market portfolio as the sole asset, the cost of all fluctuations over a five-year period is estimated at 14.4%, while the corresponding cost of business cycle fluctuations is 3.6%. When total (instead of nondurables and services) consumption expenditure is used to recover the SDF, the costs of all fluctuations and business cycle fluctuations over a 5-year period are both estimated to be even higher at 19.7% and 5.1%, respectively. Our results are in stark contrast to those in Alvarez and Jermann (2004) who argue that, while the cost of all consumption fluctuations is very high (they report a baseline value of 28.6% in the infinite-horizon case), the cost of business cycle fluctuations in consumption is miniscule, varying from 0.1% to 0.5%. The difference can be attributed, at least in part, to the strong business cycle component of the I-SDF.

Finally, note that, the above results pertain to the welfare benefits of economic stabilization *on average*. We rely on an extension of the information-theoretic methodology – specifically, the smoothed empirical likelihood (SEL) estimator of Yuichi Kitamura, Gautam Tripathi and Hyungtaik Ahn (2004) – to recover the missing component of the SDF, ψ, in a state-contingent fashion and use it to obtain the cost of all consumption fluctuations in each time period (i.e., in each possible state of the economy). This amounts to calculating the ratio of the time-t prices of the claims to the stabilized consumption stream and the actual risky consumption stream, for each time period t. As with the average cost obtained using the EL estimator, the time series of the cost estimated using the SEL approach also does not require assumptions about the investors' preferences or the dynamics of the data generating process. We find that the cost of consumption fluctuations is strongly time-varying and countercyclical. The cost of one-year fluctuations varies from 0.15% to 8.0%. Also, the cost is strongly countercyclical, rising sharply during recessionary episodes. The correlation between the cost and a dummy variable that takes the value 1 in a year if there was an NBER-designated recession in any of its quarters and 0 otherwise is 36.1%. This finding also helps explain the high cost of business cycle fluctuations that we estimate on average.

Our paper lies at the interface of two, albeit mostly distinct, strands of literature. It contributes to a growing literature that uses an information-theoretic (or, relative-entropy minimizing) alternative to the standard generalized method of moments approach to address a variety of questions in economics and finance. Information-theoretic approaches were first introduced in financial economics by Michael Stutzer (1995, 1996) and Y. Kitamura and M. Stutzer (1997) (see Yuichi Kitamura (2006) for a survey of these methods). Subsequently, these approaches have been used to assess the empirical plausibility of the rare disasters hypothesis in explaining asset pricing puzzles (see, e.g., Julliard and Ghosh (2012)), construct diagnostics for asset pricing models (see, e.g., Caio Almeida and Ren Garcia (2012), David Backus, Mikhail Chernov and Stanley E. Zin (2013)), construct bounds on the SDF and its components and recover the missing component from a candidate kernel (see, e.g., Jaroslav Borovicka, Lars P. Hansen and Jose A. Scheinkman (2016), Ghosh, Julliard and Taylor (2016*b*), Mirela Sandulescu, Fabio Trojani and Andrea Vedolin (2018)), price broad cross sections of assets out of sample (see, e.g., Ghosh, Julliard and Taylor (2016*a*)), and recover investors' beliefs from observed asset prices (see, e.g., Lars Peter Hansen (2014), Anisha Ghosh and Guillaume Roussellet (2019)).

Our paper also contributes to the literature that tries to assess the welfare costs of aggregate economic fluctuations (see, e.g., Lucas (1987), Ayse Imrohoroglu (1989), Andrew Atkeson and Christopher Phelan (1994), Maurice Obstfeld (1994), James Pemberton (1996), Jim Dolmas (1998), Thomas Tallarini (2000), Paul Beaudry and Carmen Pages (2001), Christopher Otrok (2001), Kjetil Storesletten, Chris I. Telmer and Amir Yaron (2001), Alvarez and Jermann (2004), Tom Krebs (2007), Per Krusell and Anthony A. Smith (2009)). Most of this literature assumes particular parametric forms for preferences as well as the dynamics of the underlying data generating process. Our paper, on the other hand, is model-free, not requiring us to take a stance on either of the above. Our approach is similar in spirit to Alvarez and Jermann (2004) that, to the best of our knowledge, are the first to have used asset prices to infer bounds on the welfare cost of business cycle fluctuations. However, we do not need to impose parametric restrictions on either the data generating process for consumption, or on the level and time series variation of interest rates, and do not rely on approximation results.

The reminder of the paper is organized as follows. Section II defines the cost of aggregate consumption fluctuations and describes an information-theoretic methodology to estimate this cost. Section III provides simulation evidence on the power of the information-theoretic methodology to recover the underlying pricing kernel accurately. Section IV contains a description of the data used. The welfare gains from eliminating all consumption fluctuations and fluctuations corresponding to business cycle frequencies are presented in Sections V and VI, respectively. Section VII presents a host of robustness checks. Section VIII relies on an extension of our information-theoretic methodology to provide evidence that the welfare gains from eliminating all consumption uncertainty vary substantially over the business cycle, a finding that is subsequently used in Section IX to explain the main factors driving our results. Finally, Section X concludes with suggestions for future research.

II. Pricing Aggregate Economic Fluctuations

This section defines the cost of fluctuations in aggregate consumption and proposes a novel procedure to measure the cost. Specifically, in Subsection II.A, we define the cost of consumption fluctuations, for two alternative definitions of aggregate consumption fluctuations. These definitions follow Alvarez and Jermann (2004). In Subsection II.B, we propose a novel information-theoretic procedure to measure the cost of fluctuations, for the two different definitions of the fluctuations. Our methodology does not require taking a stance on either investors' preferences or the dynamics of consumption, thereby delivering robust estimates of the cost of consumption fluctuations.

A. *The Cost of Aggregate Fluctuations*

The cost (or, the market price) of consumption fluctuations, ω_0, is defined as the ratio of the prices of two securities: a claim to a *stable* version of the aggregate consumption stream from which certain fluctuations have been removed, and a claim to the actual aggregate consumption stream,

$$\omega_0 = \frac{V_0\left[\left\{C_t^{stab}\right\}_{t\geq 1}\right]}{V_0\left[\left\{C_t\right\}_{t\geq 1}\right]} - 1. \tag{2}$$

In the above equation, $V_0\left[\left\{C_t\right\}_{t\geq 1}\right]$ and $V_0\left[\left\{C_t^{stab}\right\}_{t\geq 1}\right]$ denote the time-0 prices of claims to the future consumption stream and the future stabilized consumption stream, respectively. Therefore, the cost of consumption fluctuations measures how much extra investors would be willing to pay in order to replace the aggregate consumption stream with its stabilized counterpart.

If stabilized consumption, C_t^{stab}, is defined as the expected value of future consumption, i.e. $C_t^{stab} = E_0\left(C_t\right)$, then Equation (2) measures the cost of *all* consumption fluctuations. In other words, it measures the benefit of elimination of all consumption uncertainty.

If, on the other hand, stabilized consumption, C_t^{stab}, is defined as the long-term trend consumption, from which fluctuations corresponding to business cycle frequencies have been removed, then Equation (2) measures the cost of business cycle fluctuations in consumption. Business cycles are typically defined as fluctuations that last for no longer than 8 years. A stabilized consumption series from which fluctuations corresponding to business cycle frequencies have been removed can be constructed using smoothing filters like the Hodrick-Prescott filter (see also Morten O. Ravn and Harald Uhlig (2002)).

The absence of arbitrage opportunities implies that

$$V_0\left[\{C_t\}_{t\geq 1}\right] = \sum_{t=1}^{\infty} V_0\left(C_t\right), \tag{3}$$

where $V_0\left(C_t\right)$ denotes the time-0 price of a claim to a single payoff equal to the aggregate consumption at time t. Similarly, $V_0\left[\left\{C_t^{stab}\right\}_{t\geq 1}\right]$ can be written as the sum, across all future periods, of the prices of claims to single payoffs equal to the stabilized consumption in each future period. Therefore, the cost of one-period fluctuations is given by $\frac{V_0\left(C_1^{stab}\right)}{V_0(C_1)} - 1$, the cost of two-period fluctuations is given by $\frac{V_0\left(C_1^{stab}\right)+V_0\left(C_2^{stab}\right)}{V_0(C_1)+V_0(C_2)} - 1$, and so on.

In the context of a representative agent economy, Alvarez and Jermann (2004) show that ω_0 in Equation (2) measures the *marginal cost of consumption fluctuations*, defined as the per unit benefit of a marginal reduction in consumption fluctuations expressed as a percentage of lifetime consumption. The marginal cost provides an upper bound on the *total cost of consumption fluctuations*, where the latter is defined as the additional lifetime consumption, expressed as a percentage of consumption, that the representative agent would demand in order to be indifferent between the risky aggregate consumption stream and a stabilized version of the aggregate consumption stream from which certain types of fluctuations (e.g., all fluctuations or business cycle fluctuations) have been removed.[4]

The benefits of focusing on the marginal cost are two-fold. First, it can be estimated using observed asset prices and the assumption of the absence of arbitrage opportunities, unlike the total cost that requires a fully-specified utility function. Second, it enables the assessment of the welfare benefits of a unit reduction in consumption fluctuations when consumers are bearing all the fluctuations, thereby shedding light on the desirability or lack thereof of policies aimed at only moving partially in the direction of eliminating certain types of aggregate fluctuations.

Note that since neither of the two assets – namely, the claims to aggregate consumption or its stabilized counterpart – that characterize the marginal cost of consumption fluctuations (see Equation (2)) is directly traded in financial markets, their prices are not directly observed. Therefore, the values of these two securities need to estimated in order to obtain the cost of consumption fluctuations. Historically, this has involved taking a stance on investors' preferences, i.e. their stochastic discounting of the various possible future states of the world, and the dynamics of the data generating process, i.e. the likelihood of the states being realized. The resultant estimates of the cost of economic fluctuations have

[4]Alvarez and Jermann (2004) also provide an alternative interpretation of ω_0 in the case of incomplete markets economies with agents possibly subject to uninsurable idiosyncratic risks.

proven to be quite sensitive to these two assumptions (see, e.g. Barlevy (2005) for a survey of the literature). The following subsection outlines a novel econometric methodology for estimating the cost of consumption fluctuations, that does not require any specific functional-form assumptions either about investors' preferences or the dynamics of the data generating process.

B. *Measuring the Cost of Aggregate Fluctuations*

Consider an economy characterized by an augmented state vector $\mathbf{z}_t \in \mathbf{Z}$, augmented by, adding to the beginning of period state variables, the time t realization of the shocks that influence equilibrium quantities. Then, all equilibrium quantities can be viewed as functions of $\mathbf{z}$. For instance, the equilibrium realization of the aggregate consumption growth rate is simply $C_{t+1}/C_t \equiv \Delta c_{t+1} = \Delta c(\mathbf{z_{t+1}})$. That is, consumption growth can be viewed as just a mapping from $\mathbf{z}$ to the (positive) real line i.e. $\Delta c : \mathbf{z} \to \mathbb{R}_+$.

Note that the (shadow) value of a claim to the aggregate consumption next period can be generally expressed as

$$V_t\left(C_{t+1}\right) = \mathbb{E}_t\left[M_{t+1}C_{t+1}\right], \tag{4}$$

where M_t is the pricing kernel. The existence of a (strictly positive) pricing kernel is guaranteed by the assumption of the absence of arbitrage opportunities. For the particular case of a representative agent economy, M can be thought of as the intertemporal marginal rate of substitution of a (fictitious) representative agent who derives utility from the consumption flow C. Note, however, that such a representation is not restricted to representative agent economies but can also obtain in incomplete-markets economies inhabited by heterogeneous agents (as, e.g., in Constantinides and Ghosh (2017)).

By the definition of $\mathbf{z}$, we have $M_t \equiv M(\mathbf{z}_t)$, i.e. in equilibrium $M : \mathbf{z} \to \mathbb{R}_+$. Therefore, dividing Equation (4) by C_t to make both sides stationary, taking unconditional expectations, and using the definition of $\mathbf{z}$, we have

$$\tilde{pc}_1 := \mathbb{E}\left[\frac{V_t\left(C_{t+1}\right)}{C_t}\right] = \int_{\mathbf{z}} M(\mathbf{z})\Delta c(\mathbf{z})d\mathbb{P}(\mathbf{z}), \tag{5}$$

where $\mathbb{P}$ is the (true) underlying physical probability measure and we have used the assumption that $\mathbf{z}$ has a time invariant unconditional distribution. $\tilde{pc}_1$ can be interpreted as the average price (expressed as a fraction of current consumption) of an asset with a single payoff equal to the aggregate consumption next period.

Similarly, the (shadow) value of a claim to a *stabilized* version of the aggregate consumption next period can be expressed as

$$V_t\left(C_{t+1}^{stab}\right) = \mathbb{E}_t\left[M_{t+1}C_{t+1}^{stab}\right], \tag{6}$$

implying that

$$\tilde{pc}_1^{stab} := \mathbb{E}\left[\frac{V_t\left(C_{t+1}^{stab}\right)}{C_t}\right] = \int_{\mathbf{z}} M(\mathbf{z})\Delta c^{stab}(\mathbf{z})d\mathbb{P}(\mathbf{z}), \tag{7}$$

where $\Delta c_{t+1}^{stab} = \frac{C_{t+1}^{stab}}{C_t}$. In the scenario where we want to obtain the cost of *all* consumption uncertainty in the next period, we set $C_{t+1}^{stab} = (1+\mu_c)\,C_t$, where μ_c denotes the unconditional mean of the aggregate consumption growth rate. Therefore, in this case, $\Delta c_{t+1}^{stab} = (1+\mu_c)$. On the other hand, to assess the cost of removing business cycle fluctuations in consumption, we set $C_{t+1}^{stab} = C_{t+1}^{bc}$, where C_{t+1}^{bc} refers to a smoothed version of the aggregate consumption at time $t+1$ from which fluctuations corresponding to business cycle frequencies have been removed.

Once the prices of the claims to the aggregate consumption and the stabilized aggregate consumption next period have been obtained, the cost of one-period consumption fluctuations is then given by

$$\frac{\tilde{pc}_1^{stab}}{\tilde{pc}_1} - 1. \tag{8}$$

If a history of $M(\mathbf{z}_t) \equiv M_t,\ t = 1, ..., T$, were observable, we could estimate the prices in Equations (5) and (7) and, therefore, the cost of one-period consumption fluctuations in Equation (8): in this case the integrals (unconditional expectations) with respect to the physical measure would be replaced by the sums of observations weighted by $1/T$, invoking ergodicity of the processes involved. If the pricing kernel M were a known function of a vector of unknown parameters, these parameters could first be estimated using method of moments approaches, prior to evaluating the cost as above.

For instance, assuming a representative agent endowed with power utility preferences with a constant CRRA, $\tilde{pc}_1$ can be estimated as $\frac{1}{T}\sum_{t=1}^{T}\delta\left(\Delta c_t\right)^{1-\gamma}$, where γ denotes the relative risk aversion coefficient and δ the subjective discount factor. Moreover, assuming log-normality of aggregate consumption growth as in Lucas (1987), we would have

$$\tilde{pc}_1 = \mathbb{E}\left[\delta(\Delta c_t)^{1-\gamma}\right] = e^{\ln(\delta)+(1-\gamma)\mathbb{E}[\ln(\Delta c_t)]+.5(1-\gamma)^2 Var[\ln(\Delta c_t)]}.$$

Similarly, the price of a claim to sure consumption next period, $C_{t+1}^{stab} = (1+\mu_c)\,C_t$, is given by

$$\tilde{pc}_1^{stab} = \mathbb{E}\left[\delta(\Delta c_t)^{-\gamma}\left(1+\mu_c\right)\right] = \left(1+\mu_c\right)e^{\ln(\delta)-\gamma\mathbb{E}[\ln(\Delta c_t)]+.5\gamma^2 Var[\ln(\Delta c_t)]}.$$

The first two moments of log consumption growth, $\mathbb{E}[\ln(\Delta c_t)]$ and $Var[\ln(\Delta c_t)]$, required to obtain $\tilde{pc}_1$ and $\tilde{pc}_1^{stab}$, can be estimated as the respective sample analogs of the underlying unconditional expectations and, therefore, the price of one-period consumption fluctuations can be obtained.

However, the pricing kernel M is not directly observable. Using the above specification of the pricing kernel and lognormal assumption for the dynamics of consumption growth, Lucas estimates a very small cost of consumption fluctuations. Subsequently, researchers have proposed alternative specifications of preferences as well as the dynamics of the consumption growth rate and other variables entering the pricing kernel. The resulting estimates of the cost of aggregate fluctuations have proven to be quite sensitive to these assumptions, varying wildly across these studies.

In this paper, we do not make any assumptions either about the preferences of consumers, or the dynamics of the data generating process. Rather, our methodology is based on the observation that, albeit not directly observable, information about $M(\mathbf{z})$ is available in financial markets. This is because, for any vector of excess returns $\mathbf{R}_t^e \in \mathbb{R}^N$ on N traded assets, the following set of Euler equations must hold in the absence of arbitrage opportunities:

$$\mathbf{0} = \mathbb{E}\left[M_t \mathbf{R}_t^e\right] = \int M(\mathbf{z})\mathbf{R}^e(\mathbf{z})d\mathbb{P}(\mathbf{z}) \equiv \int \mathbf{R}^e(\mathbf{z})d\mathbb{Q}(\mathbf{z}) \equiv \mathbb{E}^{\mathbb{Q}}\left[\mathbf{R}_t^e\right]$$

where $\mathbf{0}$ is an N-dimensional vector of zeros and, by definition, $\mathbf{R}^e : \mathbf{z} \to \mathbb{R}^N$. The so-called risk neutral measure $\mathbb{Q}$ (absolutely continuous with respect to the physical measure $\mathbb{P}$) satisfies the Radon-Nikodym derivative $\frac{d\mathbb{Q}(\mathbf{z})}{d\mathbb{P}(\mathbf{z})} = \frac{M(\mathbf{z})}{\mathbb{E}[M(\mathbf{z})]}$. Note also that, in absence of arbitrage opportunities, if a risk free asset exist, it must satisfy $\mathbb{E}\left[1/R_t^f\right] = \mathbb{E}\left[M_t\right]$.

Let $p(\mathbf{z})$ and $q(\mathbf{z})$ denote, respectively, the pdf's associated with the measures $\mathbb{P}$ and $\mathbb{Q}$. We then have that, by the definition of the measure $\mathbb{Q}$: $q(\mathbf{z})\mathbb{E}\left[M(\mathbf{z})\right] = M(\mathbf{z})p(\mathbf{z})$. Therefore, Equation (5) can be rewritten as

$$\tilde{pc}_1 = \mathbb{E}\left[M(\mathbf{z})\right] \int \Delta c(\mathbf{z})q(\mathbf{z})d\mathbf{z}. \tag{9}$$

The above formulation can be made operational, thanks to the fact that, using asset returns data, we can actually estimate the q distribution. In particular, the q distribution can be estimated to minimize the Kullback-Leibler Information Criterion (KLIC) divergence (or the relative entropy) between the physical and risk neutral measures:

$$\min_{\mathbb{Q}} \int \mathbf{log}\left(\frac{d\mathbb{P}}{d\mathbb{Q}}\right) d\mathbb{P} = \int \mathbf{log}\left(\frac{p(\mathbf{z})}{q(\mathbf{z})}\right) p(\mathbf{z})d\mathbf{z} \text{ s.t. } \mathbf{0} = \int \mathbf{R}^e(\mathbf{z})q(\mathbf{z})d\mathbf{z}. \tag{10}$$

Adding to the above problem the theoretical restriction that the pricing kernel, M, is of the form:

$$M_{t+1} = \left(\Delta c_{t+1}\right)^{-\gamma} \psi_{t+1}, \tag{11}$$

leads to the reformulation of Equation (10) as:

$$\min_{\mathbb{F}} \int \mathbf{log}\left(\frac{d\mathbb{P}}{d\mathbb{F}}\right) d\mathbb{P} = \int \mathbf{log}\left(\frac{p(\mathbf{z})}{f(\mathbf{z})}\right) p(\mathbf{z})d\mathbf{z} \text{ s.t. } \mathbf{0} = \int \mathbf{R}^e(\mathbf{z})\left(\Delta c(\mathbf{z})\right)^{-\gamma} f(\mathbf{z})d\mathbf{z}, \tag{12}$$

where $\frac{d\mathbb{F}(\mathbf{z})}{d\mathbb{P}(\mathbf{z})} = \frac{\psi(\mathbf{z})}{E(\psi(\mathbf{z}))}$ is the Radon-Nikodym derivative of $\mathbb{F}$ with respect to $\mathbb{P}$, and $f(\mathbf{z})$ denotes the pdf associated with the measure $\mathbb{F}$. This is the Empirical Likelihood (EL) estimator of Owen (2001), originally proposed in Ghosh, Julliard and Taylor (2016*b*) to recover the multiplicative missing component of the pricing kernel. Once the $\mathbb{F}$-measure, or, from the expression for the Radon-Nikodym derivative, the missing component, ψ, of the pricing kernel, is estimated as the solution to Equation (12), the pricing kernel, M, can be obtained using Equation (11). We refer to this kernel as the *Information*-SDF, or I-SDF, because of the information-theoretic approach used to recover it.

Ghosh, Julliard and Taylor (2016*b*) point out several reasons why relative entropy minimization is an attractive criterion for recovering the pricing kernel. These are restated here for convenience.

First, the KLIC minimization in Equation (12) is equivalent to maximizing the (expected) ψ nonparametric likelihood function in an unbiased procedure for finding the ψ_t component of the pricing kernel. To see this, note that the maximization problem in Equation (12), after dropping redundant terms, can be rewritten as

$$\max_{\psi} \mathbb{E}^{\mathbb{P}}\left[\ln \psi(\mathbf{z})\right] \text{ s.t. } \mathbf{0} = \int \mathbf{R}^e(\mathbf{z})\left(\Delta c(\mathbf{z})\right)^{-\gamma} \psi(\mathbf{z}) p(\mathbf{z}) d\mathbf{z}. \tag{13}$$

Note also that this is the rationale behind the *principle of maximum entropy* (see e.g. E. T. Jaynes (1957*a*, 1957b)) in physical sciences and Bayesian probability that states that, subject to known testable constraints – the asset pricing Euler restrictions in our case – the probability distribution that best represent our knowledge is the one with maximum entropy, or minimum relative entropy in our notation.

Second, the use of relative entropy, due to the presence of the logarithm in the objective function in Equation (12), naturally imposes the non-negativity of the pricing kernel.

Third, our approach to recover the ψ_t component of the pricing kernel satisfies the Occam's razor, or law of parsimony, since it adds the *minimum amount of information* needed for the pricing kernel to price assets. This is due to the fact that the relative entropy is measured in units of information. To provide some intuition, suppose that the consumption growth component of the pricing kernel, $(\Delta c_t)^{-\gamma}$, is sufficient to price assets perfectly. Then $\psi_t \equiv 1, \forall t$, and we have that $\mathbb{F} \equiv \mathbb{P}$. However, if the consumption growth component is not sufficient to price assets (as is the case in reality), then the estimated measure $\mathbb{F}$ is distorted relative to the physical measure $\mathbb{P}$. And, our estimator searches for a measure $\mathbb{F}$ that is as close as possible, in an information-theoretic sense, to the physical measure $\mathbb{P}$. In other words, the approach distorts the physical probabilities as little as possible in order to satisfy the Euler equation restrictions. And the estimator is non-parametric in the sense that it does not require any parametric functional-form assumptions about the ψ-component of the kernel or the distribution of the data.

Fourth, there is no ex-ante restriction on the number of assets that can be used in constructing ψ_t, and the approach can naturally handle assets with negative expected rates of return (cf. Fernando Alvarez and Urban J. Jermann (2005)).

Fifth, as implied by the work of Donald E. Brown and Robert L. Smith (1990), the use of entropy is desirable if we think that tail events are an important component of the risk measure.[5]

Sixth, this approach is numerically simple to implement. Given a sample of size T and a history of excess returns and consumption growth $\{\mathbf{R}_t^e, \Delta c_t\}_{t=1}^T$, Equation (13) can be made operational by replacing the expectation with a sample analogue, as is customary for moment based estimators:[6]

$$\arg\max_{\{\psi_t\}_{t=1}^T} \frac{1}{T}\sum_{t=1}^{T} \ln \psi_t \ \text{ s.t. } \frac{1}{T}\sum_{t=1}^{T} (\Delta c_t)^{-\gamma}\, \psi_t \mathbf{R}_t^e = \mathbf{0}. \tag{14}$$

A standard application of Fenchel's duality theorem to the above problem (see, e.g., Imre Csiszár (1975), Owen (2001)), delivers the estimates (up to a positive constant scale factor):

$$\hat{\psi}_t = \frac{1}{T(1+\hat{\theta}'\mathbf{R}_t^e\,(\Delta c_t)^{-\gamma})} \quad \forall t, \tag{15}$$

where $\hat{\theta} \in \mathbb{R}^N$ is the vector of Lagrange multipliers that solves the unconstrained dual optimization problem:

$$\hat{\theta} = \arg\min_{\theta} -\sum_{t=1}^{T} \log(1+\theta'\mathbf{R}_t^e\,(\Delta c_t)^{-\gamma}). \tag{16}$$

Seventh, and perhaps most importantly, the I-SDF successfully prices assets. Note that this result is not surprising *in sample*, because the I-SDF is constructed to price the test assets in-sample (see Equation (12)). However, Ghosh, Julliard and Taylor (2016*a*) show that the good pricing performance of the I-SDF also obtains out-of-sample for broad cross-sections of assets, including domestic and international equities, currencies, and commodities. The out-of-sample performance of the I-SDF is superior to not only the single factor CAPM and the Consumption-CAPM, but also the more recent Fama-French 3 and 5 factor models. This suggests that the I-SDF is more successful at capturing the relevant sources of priced risk and, therefore, offers a more reliable candidate kernel with which to measure the cost of aggregate economic fluctuations.

Finally, we show, via simulation exercises, that the information-theoretic methodology is remarkably successful in recovering the ψ-component of the pricing kernel. Details of the simulation design and the performance of the estimator are presented in Section III.

Under the assumption that the physical measure can be approximated with an empirical distribution that assigns probability weight $1/T$ to every sample

[5] Brown and Smith (1990) develop what they call "a Weak Law of Large Numbers for rare events;" that is, they show that the empirical distribution observed in a very large sample converges to the distribution that minimizes the relative entropy.

[6] This amounts to assuming ergodicity for both the pricing kernel and asset returns.

realization t, $\frac{d\mathbb{Q}(\mathbf{z})}{d\mathbb{P}(\mathbf{z})} = \frac{M(\mathbf{z})}{\mathbb{E}[M(\mathbf{z})]} = \frac{(\Delta c(\mathbf{z}))^{-\gamma}\psi(\mathbf{z})}{\mathbb{E}[M(\mathbf{z})]}$ implies that the risk neutral measure is proportional to the pricing kernel M and that the proportionality constant can be recovered from the Euler equation for the risk free rate. Thus, the risk neutral measure is given by

$$\hat{q}_t \equiv \widehat{q(\mathbf{z}_t)} = \kappa \left(\Delta c(\mathbf{z}_t)\right)^{-\gamma} \psi(\mathbf{z_t}) = \frac{\kappa \left(\Delta c_t\right)^{-\gamma}}{T(1 + \hat{\theta}' \mathbf{R}^e_t \left(\Delta c_t\right)^{-\gamma})} \quad \forall t, \tag{17}$$

where κ is a strictly positive normalization constant chosen such that $\sum_{t=1}^{T} \hat{q}_t = 1$.

Equation (17) makes clear that our estimator of the risk neutral measure, as any Generalized Empirical Likelihood approach (see e.g. Kitamura (2006) for a survey), approximates the true unknown q distribution with a multinomial with support points given by the realizations of the observable variables (in this case, consumption growth and asset returns) at the various dates in the sample.

The result in Equation (17) implies that the value of switching from the stochastic consumption growth $\Delta c(\mathbf{z})$ to a constant growth $(1 + \mu_c)$ can be estimated, given a history of consumption growth and asset returns of length T, via the estimated percentage increase in the price-consumption ratio

$$\widehat{pc_1^{stab}/p\tilde{c}_1} = \frac{\sum_{t=1}^{T} \left(1 + \mu_c\right) \hat{q}_t}{\sum_{t=1}^{T} \Delta c_t \hat{q}_t} = \frac{1 + \mu_c}{\sum_{t=1}^{T} \Delta c_t \hat{q}_t}. \tag{18}$$

Note that Equation (18) represents the value of eliminating *all* consumption fluctuations in the next period. If aggregate consumption uncertainty is caused by both business cycle and lower frequency fluctuations, the value of eliminating the former can be estimated via a simple modification to Equation (18). Specifically, we replace the constant consumption growth rate μ_c with a time-varying stabilized consumption growth from which the business cycle variation has been removed. This stabilized version of consumption, C^{stab}, can be obtained by an application of the Hodrick-Prescott filter to the original consumption series. The value of eliminating business cycle fluctuations can, therefore, be expressed as

$$\widehat{pc_1^{stab}/p\tilde{c}_1} = \frac{\sum_{t=1}^{T} \Delta c_t^{stab} \hat{q}_t}{\sum_{t=1}^{T} \Delta c_t \hat{q}_t}. \tag{19}$$

Note that Equations (18) and (19) represent the costs of all consumption fluctuations and business cycle fluctuations, respectively, for one period alone. It is straightforward to extend the analysis to obtain the cost of fluctuations for multiple periods. For instance, the (shadow) value of a claim to the aggregate consumption two periods into the future can be expressed as

$$V_t \left(C_{t+2}\right) = \mathbb{E}_t \left[M_{t:t+2} C_{t+2}\right],$$

where $M_{t:t+2}$ denotes the two-period SDF. Thus, the expected price-consumption ratio of a security that delivers a single payoff equal to the two-period consumption is given by

$$\tilde{pc}_2 := \mathbb{E}\left[\frac{V_t\left(C_{t+2}\right)}{C_t}\right] = \mathbb{E}\left[M_{t:t+2}\frac{C_{t+2}}{C_t}\right].$$

The one-period I-SDF recovered in Equation (17) can be compounded to estimate $\tilde{pc}_2$:[7]

$$M_{t:t+2} = \prod_{j=1}^{2} M_{t+j}.$$

Similarly, we can estimate the price-consumption ratio $\tilde{pc}_j$ for a single consumption claim j periods in the future, for any $j = 2, 3, 4, \ldots$. Using the estimated price-consumption ratios of the claims to single future payoffs, we can estimate the price-consumption ratio of an asset that delivers the stochastic consumption in each of the next J periods i.e. $\tilde{pc}_{1:J} := \sum_{j=1}^{J} \tilde{pc}_j$. Hence, it is straightforward to compute the value of removing all or business cycle fluctuations in consumption over J periods with expressions analogous to the ones in Equations (18)-(19).

III. Performance of the EL Estimator: An Example Economy

In this section, we provide simulation evidence on the performance of the EL estimator in recovering the ψ-component of the pricing kernel. Specifically, we consider a hypothetical economy in which the representative investor's subjective beliefs diverge from the true underlying (or, physical) distribution of the data. As we show below, in this economy, the ψ-component of the kernel captures the divergence between the subjective and physical measures. We then show that the EL estimator is remarkably successful in recovering ψ and, therefore, the subjective beliefs of the investor. The details of the simulation design are presented below.

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

$$\log\left(\Delta c_t\right) \overset{\mathbb{P}}{\sim} \mathcal{N}\left(\mu, \sigma^2\right). \tag{20}$$

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

$$\log\left(\Delta c_t\right) \overset{\widetilde{\mathbb{P}}}{\sim} \mathcal{N}\left(\widetilde{\mu}, \sigma^2\right), \tag{21}$$

where $\widetilde{\mu} = (1-\lambda)\mu$ and $\widetilde{\mathbb{P}}$ denotes the investor's subjective measure. We assume

[7] Alternatively, using the approach in Equation (12), one could estimate directly the two period risk neutral probability measure using a change of measure with respect to the variable $M_{t,t+1} := M_t M_{t+1}$ and two period returns. In this case, the notation in Equation (12) would be reinterpreted as $\mathbf{z}$ being the state vector for two contiguous periods and $\mathbf{R}^e$ denoting two period excess returns.

that there are no distortions in the beliefs about the volatility and the higher moments of consumption growth.

In this economy, the following Euler equation holds in equilibrium:

$$0 = \mathbb{E}^{\widetilde{\mathbb{P}}}\left[(\Delta c_{t+1})^{-\gamma}\left(R_{m,t+1} - R_{f,t+1}\right)\right], \tag{22}$$

where $R_{m,t}$ and $R_{f,t}$ denote the market return and the risk free rate, respectively, at time t. Note that, in Equation (22), the expectation is evaluated under the investor's subjective measure (instead of the physical measure $\mathbb{P}$). Under weak regularity conditions, Equation (22) may be rewritten as

$$0 = \mathbb{E}^{\mathbb{P}}\left[(\Delta c_{t+1})^{-\gamma}\,\psi_{t+1}\left(R_{m,t+1} - R_{f,t}\right)\right], \tag{23}$$

where $\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} = \frac{\psi}{E(\psi)}$ is the Radon-Nikodym derivative of $\widetilde{\mathbb{P}}$ with respect to $\mathbb{P}$. Thus, in this economy, the ψ-component of the kernel captures the divergence between the subjective and physical measures.

Note that this example economy fits into the framework described in Section II. Therefore, given time series data on consumption growth, the market return, and risk free rate, the EL approach can be used to estimate (up to a strictly positive constant scale factor) the ψ-component of the kernel:

$$\left\{\widehat{\psi}_t\right\}_{t=1}^{T} = \arg\max_{\{\psi_t\}_{t=1}^{T}} \sum_{t=1}^{T}\log(\psi_t) \quad \text{s.t.} \quad \frac{1}{T}\sum_{t=1}^{T}(\Delta c_{t+1})^{-\gamma}\,\psi_{t+1}(R_{m,t+1}-R_{f,t+1}) = 0. \tag{24}$$

Using the recovered ψ and under the assumption that the physical measure can be approximated with an empirical distribution that assigns probability weight $1/T$ to every sample realization, i.e., $\widehat{\mathbb{P}} = \{\widehat{p}_t\}_{t=1}^{T} = \frac{1}{T}$, the subjective measure $\widehat{\widetilde{\mathbb{P}}} = \left\{\widehat{\widetilde{p}}_t\right\}_{t=1}^{T}$ can be obtained from the definition of the Radon-Nikodym derivative.

We show, via simulations, that the EL approach successfully recovers ψ and, therefore, $\widetilde{\mathbb{P}}$. In order to perform the EL estimation in Equation (24), we need the time series of consumption growth, the market return, and the risk free rate. Note that, in this economy, equilibrium asset prices reflect the subjective beliefs of the investor. In particular, the equilibrium price-dividend ratio is $\frac{P_t}{D_t} = z$, a constant, where

$$z = \frac{\exp\left[\log(\delta) + (1-\gamma)\widetilde{\mu} + \dfrac{(1-\gamma)^2\sigma^2}{2}\right]}{1 - \exp\left[\log(\delta) + (1-\gamma)\widetilde{\mu} + \dfrac{(1-\gamma)^2\sigma^2}{2}\right]}, \tag{25}$$

and the equilibrium risk free rate is also constant at:

$$R_f = \frac{1}{\exp\left(\log(\delta) - \gamma\tilde{\mu} + \frac{\gamma^2\sigma^2}{2}\right)}. \tag{26}$$

To perform our simulation exercise, we calibrate μ and σ^2 to the sample mean and variance, respectively, of (log) consumption growth in our data (real per capita total consumption over 1929-2015). The preference parameters are calibrated at $\delta = 0.99$ and $\gamma = 10$. We simulate a time series of consumption growth using Equation (20). Using the simulated consumption growth, we obtain the market return as:

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

where z is defined in Equation (25). The time series of the risk free rate is simply a constant given by Equation (26).

Using the above time series, we recover the subjective beliefs using the EL approach. Armed with the subjective probabilities, we compute the mean, volatility, and skewness of consumption growth. Note that these are the moments of consumption growth that are consistent with the asset prices, i.e. the moments as perceived by the representative investor. We repeat the above estimation for 500 simulated samples. We report the means and 90% confidence intervals of the moments of consumption growth across these simulations. To demonstrate the power of the estimation approach, we present results for different magnitudes of the beliefs distortion, i.e. for $\lambda = \{0.10, 0.15, 0.20\}$, and for different simulated sample sizes, i.e. $T_{sim} = \{85, 200, 500\}$. The first choice of sample size, $T_{sim} = 85$, corresponds to the size of the historical sample that we use in our empirical analysis.

The results are reported in Table 1. Panel A presents results for $T_{sim} = 85$. Consider first Row 1, where investors are assumed to underestimate the mean of consumption growth by 10%, i.e. the mean of 2.55% under subjective beliefs is 10% below the historical mean of 2.83%. The equilibrium market return and risk free rate reflect these subjective beliefs of investors. Row 1 shows that the EL method is successful at capturing these subjective beliefs of investors. Specifically, the EL-implied mean of consumption growth has a mean of 2.61% across the 500 simulations, close to the true value of the mean under subjective beliefs. The EL implied volatility of consumption growth has a mean of 3.47% across the 500 simulations – once again quite close to the historical value. Note that, in our experiment, there are no beliefs distortions in the volatility and the EL method successfully identifies the volatility observed in the historical data. Finally, the average of the coefficient of skewness across the simulations is -0.003, very close to the true value of 0.

Rows 2 and 3 show that similar, albeit stronger, results are obtained for more severe beliefs distortions in the mean of consumption growth – the EL method correctly identifies the subjective mean and the 90% confidence intervals do not contain the corresponding values of the mean under the physical measure, and the estimated volatility and skewness are very close to their historical values with tight confidence bands. Finally, Panels B and C show the effect of increasing sample size on the performance of the EL estimator – the performance at samples sizes of 200 and 500 are quite similar to those observed for available sample sizes in the historical data in terms of the average mean, volatility, and skewness across the simulations, although the confidence bands are tighter for longer sample sizes.

Table 1: Estimating Subjective Beliefs

	Mean (%)	Volatility (%)	Skewness
	true values		
$\tilde{\mu} = \mu$	2.83	3.39	0
$\tilde{\mu} = 0.90\mu$	2.55	3.39	0
$\tilde{\mu} = 0.85\mu$	2.41	3.39	0
$\tilde{\mu} = 0.80\mu$	2.27	3.39	0
	Panel A: T=85		
$\tilde{\mu} = 0.90\mu$	2.61 [2.35,2.89]	3.47 [3.09,3.84]	−.003 [−.45,.38]
$\tilde{\mu} = 0.85\mu$	2.47 [2.23,2.76]	3.48 [3.15,3.86]	.008 [−.40,.41]
$\tilde{\mu} = 0.80\mu$	2.35 [2.10,2.66]	3.51 [3.14,3.88]	.001 [−.43,.43]
	Panel B: T=200		
$\tilde{\mu} = 0.90\mu$	2.60 [2.42,2.80]	3.46 [3.21,3.73]	−.011 [−.34,.29]
$\tilde{\mu} = 0.85\mu$	2.50 [2.33,2.71]	3.52 [3.27,3.77]	−.049 [−.37,.26]
$\tilde{\mu} = 0.80\mu$	2.40 [2.19,2.61]	3.56 [3.29,3.81]	−.063 [−.40,.24]
	Panel C: T=500		
$\tilde{\mu} = 0.90\mu$	2.61 [2.49,2.73]	3.47 [3.31,3.63]	−.036 [−.24,.16]
$\tilde{\mu} = 0.85\mu$	2.51 [2.39,2.64]	3.52 [3.36,3.68]	−.054 [−.26,.15]
$\tilde{\mu} = 0.80\mu$	2.41 [2.26,2.57]	3.57 [3.39,3.75]	−.068 [−.33,.13]

The table presents the average of the mean (Column 2), volatility (Column 3), and skewness (Column 4) of consumption growth, along with the 90% confidence intervals (in square brackets below), computed from 500 simulated samples. The samples are simulated from a hypothetical endowment economy in which a representative agent with power utility preferences is pessimistic and underestimates the mean consumption growth. Panels A, B, and C present results for different sample sizes, whereas Rows 1-3 in each panel present results for different degrees of pessimism. The expectations underlying the calculation of the moments of consumption growth are evaluated under the subjective measure recovered using the EL approach.

Overall, the results suggest that the EL estimator performs quite well at identifying the ψ-component of the pricing kernel for empirically realistic sample sizes.

This lends further support for its use in the recovery of the pricing kernel for welfare cost calculations.

IV. Data Description

The extraction of the information kernel (I-SDF) for use in welfare cost calculations requires data on the aggregate consumption expenditures and returns on a set of traded assets. Ideally, we would like to use the longest available time series of these variables in the estimation. At the same time, to assess the robustness of our key results, we would like to repeat our analysis for different measures of consumption expenditures as well as different sets of assets. While data on total consumption is available from 1890 onwards, disaggregated expenditures on different consumption categories (e.g., durables, nondurables, and services) are only available from 1929 onwards. Moreover, data on broad cross sections of asset returns are also not available prior to the late 1920s. Therefore, we focus on a baseline data sample starting at the onset of the Great Depression (1929-2015).

For the 1929-2015 data sample, we consider two alternative measures of consumption: (i) the personal consumption expenditure on nondurables and services, and (ii) the personal consumption expenditure on durables, nondurables and services. The consumption data are obtained from the Bureau of Economic Analysis. Nominal consumption is converted to real using the Consumer Price Index (CPI).

We use different sets of assets to extract the I-SDF: (i) the market portfolio, proxied by the Center for Research in Security Prices (CRSP) value-weighted index of all stocks on the NYSE, AMEX, and NASDAQ, and (ii) the 6 equity portfolios formed from the intersection of two size and three book-to-market-equity groups. The proxy for the risk-free rate is the one-month Treasury Bill rate. The returns on all the above assets are obtained from Kenneth French's data library. Annual returns for the assets are computed by compounding monthly returns within each year and converted to real using the CPI. Excess returns on the portfolios are then computed by subtracting the risk free rate.

To further assess the robustness of our results, we also repeat our analysis using two alternative data sets: (i) total personal consumption expenditure over the 1890-2015 sample and the excess return on the $S\&P500$ as the sole asset, and (ii) the personal consumption expenditure on nondurables and services along with the excess return on the CRSP value-weighted market portfolio, over the entire available quarterly sample 1947:Q1-2015:Q4.

V. The Market Value of Aggregate Uncertainty

In this section, we use the I-SDF, extracted using the information-theoretic procedure outlined in Section II, to obtain the cost of *all* consumption fluctuations, i.e. the welfare benefits of removing all fluctuations (or uncertainty) in consumption.

Equation (18) defines the cost for one-period fluctuations, i.e. the benefit of removing fluctuations in the next period alone. The cost is the ratio of the prices of two hypothetical securities (expressed as a fraction of current consumption):

a claim to a deterministic (or, sure) consumption in the next period, pc_1^{stab}, and a claim to the actual aggregate consumption next period, $\tilde{pc}_1$. Equation (18) reveals that the prices of the two securities and, therefore, the cost of one-period consumption fluctuations, depend on the underlying risk neutral measure. The risk neutral measure estimated using the EL procedure depends on the particular measure of the aggregate consumption expenditures as well as on the set of assets used (see Equations (16)-(17)). To ensure that the results are robust, we estimate the risk neutral measure using two different measures of consumption expenditures and two alternative sets of assets.

Table 2: Cumulative Cost of Nondurables & Services Consumption Fluctuations

	All Fluctuations					B. C. Fluctuations				
	1 Yr	2 Yr	3 Yr	4 Yr	5 Yr	1 Yr	2 Yr	3 Yr	4 Yr	5 Yr
Panel A: Market Portfolio										
I-SDF	1.53	5.15	11.75	14.28	14.44	.556	1.48	3.39	3.90	3.57
CRRA Kernel	.933	2.08	3.73	4.87	5.03	.457	.854	1.32	1.52	1.40
Lucas	.751	1.09	1.40	1.68	1.94	-	-	-	-	-
Panel B: FF 6 Portfolios										
I-SDF	1.29	3.52	6.65	10.63	11.20	.462	1.03	2.07	3.03	2.90
CRRA Kernel	.933	2.08	3.73	4.87	5.03	.457	.854	1.32	1.52	1.40
Lucas	.751	1.09	1.40	1.70	1.94	-	-	-	-	-

The table reports the (cumulative) cost of *all* aggregate consumption fluctuations (Columns 2-6) and the cost of business cycle fluctuations in consumption (Columns 7-11), for 1-5 years. Consumption denotes the real per capita personal consumption expenditure of nondurables and services. The costs are calculated using the I-SDF (Row 1), the kernel implied by power utility preferences with a constant CRRA (Row 2), and Lucas' original specification that involves power utility preferences and i.i.d. lognormal aggregate consumption growth dynamics (Row 3). Panel A presents results when the excess return on the market portfolio is the sole asset used to recover the I-SDF. In Panel B, on the other hand, the I-SDF is estimated using the 6 Fama-French size and book-to-market-equity sorted portfolios. The sample is annual covering the period 1929-2015.

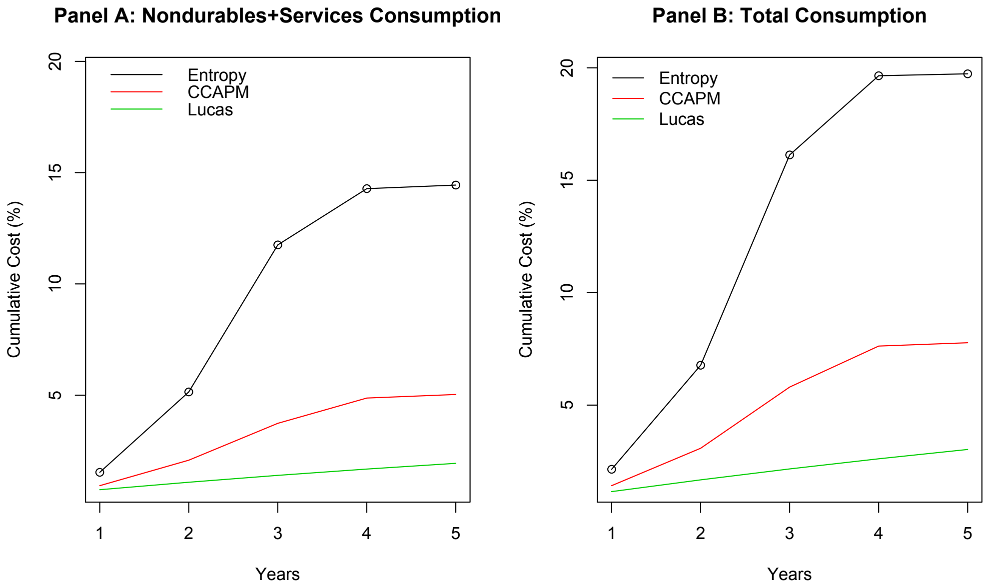
Table 2 presents the cost of consumption fluctuations when consumption refers to the expenditure on nondurables and services. Panel A presents the results when the market portfolio alone is used in the extraction of the risk neutral measure with the EL approach while Panel B does the same when the six size and book-to-market-equity sorted portfolios of Fama-French are used to extract $\mathbb{Q}$. Consider first Column 2 of Panel A. Row 1 shows that when the market portfolio alone is used in the extraction of the I-SDF, the cost of one-period consumption fluctuations is estimated to be 1.5%. Row 2 shows that the corresponding cost, estimated using the pricing kernel implied by power utility preferences with a constant CRRA (hereafter referred to as the CRRA kernel) equal to 10, is an order of magnitude smaller at .93%.[8] Row 3 shows that, if the assumption of lognormal consumption growth is imposed on the CRRA kernel – this corresponds to Lucas' original specification – the cost of one-period consumption fluctuations further reduces to .75%.

Note that the above results pertain to the cost of fluctuations in one-period consumption alone. Columns 3, 4, 5, and 6 of Panel A present the cost of consumption fluctuations over two, three, four, and five years, respectively. Row

[8]This result is robust to values of the CRRA between 1 and 10.

1 shows that, using the I-SDF, the cost of consumption fluctuations over two, three, four, and five years increases to 5.2%, 11.8%, 14.3%, and 14.4%, respectively. Note that the cost of consumption fluctuations over two years is more than three times higher than the cost of fluctuations over one year alone (5.2% versus 1.5%). Similarly, the cost of consumption fluctuations over a three-year period is more than seven times higher than the cost over one year alone (11.8% versus 1.5%); and the costs over four- and five-year periods are each almost ten times higher than the cost over one year (14.3% and 14.4%, respectively, versus 1.5%). This suggests that consumption responds slowly to news and that agents' marginal utility and, therefore, the (true) underlying pricing kernel is a function not only of current consumption but also expected future consumption, consistent with the evidence in Parker and Julliard (2005).

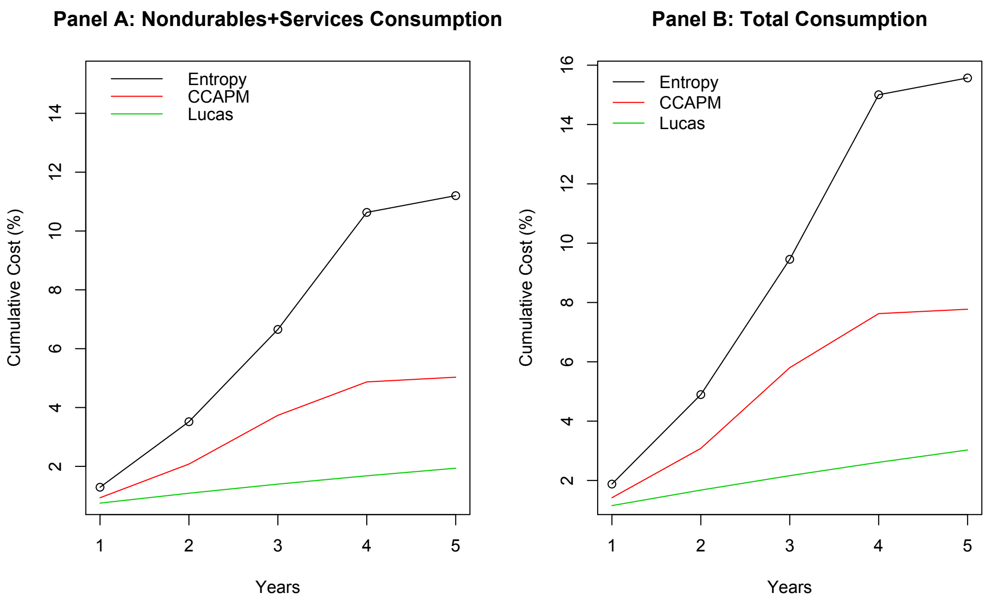
Figure 1. Marginal Cost of All Consumption Fluctuations, 1929-2015



Notes: The figure plots the cumulative cost of *all* aggregate consumption fluctuations over 1-5 years, for different choices of the pricing kernel and measures of consumption. Panel A presents results when consumption refers to the real per capita personal consumption expenditure of nondurables and services, while Panel B does the same when consumption denotes real per capita total personal consumption expenditure. The costs are presented for the I-SDF extracted using the excess return on the market portfolio as the sole test asset (black line), the pricing kernel implied by power utility preferences with a constant CRRA (red line), and Lucas' original specification that involves power utility preferences and i.i.d. lognormal aggregate consumption growth dynamics (green line).

Row 2 shows that the CRRA kernel implies much smaller costs of two, three, four, and five year consumption fluctuations of 2.1%, 3.7%, 4.9%, and 5.0%, respectively. In fact, the costs are an order of magnitude smaller than the costs implied by the I-SDF (with the exception of the two-year fluctuations that is also less than half of that implied by the I-SDF). Lucas' kernel in Row 3 implies even smaller costs of two, three, four, and five years fluctuations at 1.1%, 1.4%, 1.7%, and 1.9%, respectively. Figure 1, Panel A plots the cost of 1-5 year consumption fluctuations implied by the I-SDF recovered using the market portfolio alone (black line), CRRA kernel (red line), and Lucas' specification of power utility preferences and lognormal consumption growth (green line).

FIGURE 2. MARGINAL COST OF ALL CONSUMPTION FLUCTUATIONS, 1929-2015



Notes: The figure plots the cumulative cost of *all* aggregate consumption fluctuations over 1-5 years, for different choices of the pricing kernel and measures of consumption. Panel A presents results when consumption refers to the real per capita personal consumption expenditure of nondurables and services, while Panel B does the same when consumption denotes real per capita total personal consumption expenditure. The costs are presented for the I-SDF extracted using the cross-section of 6 size and book-to-market-equity sorted portfolios of Fama and French as test asset (black line), the pricing kernel implied by power utility preferences with a constant CRRA (red line), and Lucas' original specification that involves power utility preferences and i.i.d. lognormal aggregate consumption growth dynamics (green line).

Similar results are obtained in Table 2, Panel B when the six size and book-to-market-equity sorted portfolios are used in the extraction of the I-SDF. The information kernel implies a cost of 1.3% for one-period consumption fluctuations, an order of magnitude bigger than the .9% implied by the CRRA kernel. The cost of fluctuations for two-, three-, four-, and five-year periods are also substantially higher for the I-SDF compared to the CRRA kernel – 3.5% versus 2.1% for two periods, 6.7% versus 3.7% for three periods, 10.6% versus 4.9% for four periods, and 11.2% versus 5.0% for five periods. Lucas' specification implies even smaller costs. Figure 2, Panel A plots the cost of 1-5 year consumption fluctuations implied by the I-SDF recovered using the 6 FF portfolios (black line), CRRA kernel (red line), and Lucas' specification (green line).

The results in Table 2 were obtained using personal consumption expenditures on nondurables and services as the measure of consumption. Table 3, that uses the total consumption expenditures (including durables) as the measure of consumption, produces results qualitatively similar to those in Table 2. Note that, not surprisingly, the costs of fluctuations are bigger with total consumption compared to nondurables and services consumption. Figure 1, Panel B compares the cost of 1-5 year consumption fluctuations implied by the I-SDF, the CRRA kernel, and Lucas' original specification, when the the market portfolio alone is used in the extraction of the I-SDF. Figure 2, Panel B compares the cost for the three kernel specifications, when the 6 FF portfolios are used to recover the I-SDF.

Table 3: Cumulative Cost of Total Consumption Fluctuations, 1929-2015

	All Fluctuations					B. C. Fluctuations				
	1 Yr	2 Yr	3 Yr	4 Yr	5 Yr	1 Yr	2 Yr	3 Yr	4 Yr	5 Yr
Panel A: Market Portfolio										
I-SDF	2.15	6.77	16.13	19.65	19.73	.896	2.09	4.85	5.55	5.12
CRRA Kernel	1.42	3.08	5.80	7.63	7.77	.761	1.32	2.08	2.40	2.21
Lucas	1.15	1.68	2.16	2.61	3.03	-	-	-	-	-
Panel B: FF 6 Portfolios										
I-SDF	1.88	4.89	9.46	15.00	15.57	.770	1.60	3.05	4.35	4.14
CRRA Kernel	1.42	3.08	5.80	7.63	7.77	.761	1.32	2.08	2.40	2.21
Lucas	1.15	1.68	2.16	2.61	3.03	-	-	-	-	-

The table reports the (cumulative) cost of *all* aggregate consumption fluctuations (Columns 2-6) and the cost of business cycle fluctuations in consumption (Columns 7-11), for 1-5 years. Consumption denotes the real per capita *total* personal consumption expenditure (includes durables, nondurables, and services). The costs are calculated using the I-SDF (Row 1), the kernel implied by power utility preferences with a constant CRRA (Row 2), and Lucas' original specification that involves power utility preferences and i.i.d. lognormal aggregate consumption growth dynamics (Row 3). Panel A presents results when the excess return on the market portfolio is the sole asset used to recover the I-SDF. In Panel B, on the other hand, the I-SDF is estimated using the 6 Fama-French size and book-to-market-equity sorted portfolios. The sample is annual covering the period 1929-2015.

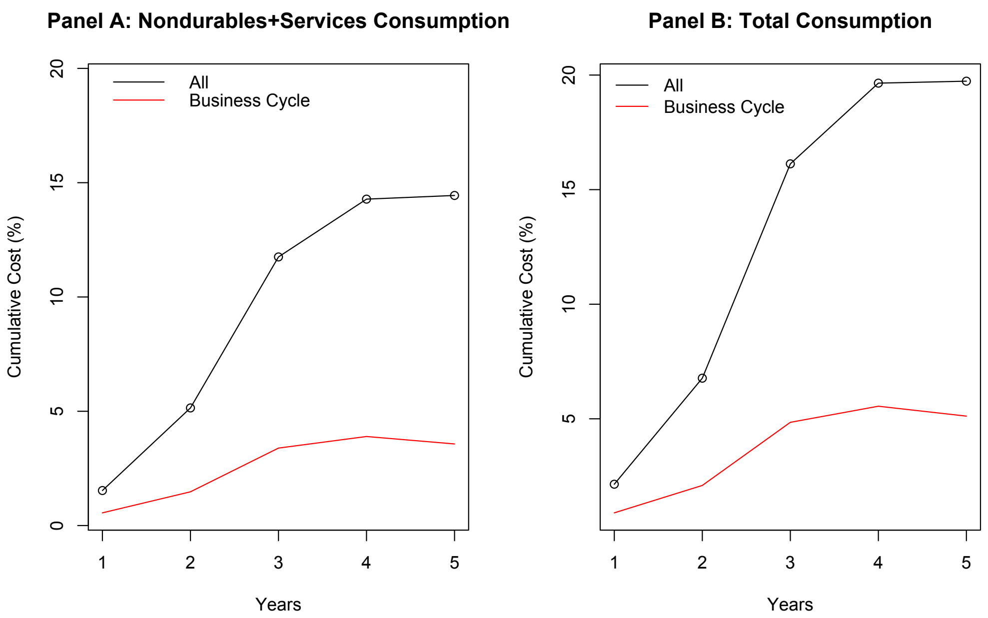
The results of this section suggest that economic agents perceive the cost of aggregate economic fluctuations to be quite substantial. The cost is substantially higher than that originally obtained by Lucas. Costs higher than Lucas' estimates have been more recently reported in the literature – Alvarez and Jermann (2004) report a baseline average cost of 28.6% for all consumption fluctuations in all future periods for an infinitely-lived agent. Our estimates of the cost are in line with Alvarez and Jermann (2004). Specifically, our estimates of the cost of all consumption fluctuations for 5 periods alone vary from 11.2%-19.7%, depending on the measure of aggregate consumption expenditure or the set of assets used to recover the I-SDF. Comparing our numbers with that in Alvarez and Jermann (2004) suggests that the cost of 5-year fluctuations alone accounts for 39.2%-68.9% of the cost of lifetime consumption fluctuations for an infinitely lived agent.

Moreover, our approach also offers a term structure of the costs of fluctuations, i.e. how the welfare benefits rise with the elimination of aggregate fluctuations over each additional future period.

VI. Business Cycle vs. Long Run Uncertainty

While Section V focused on the cost of *all* consumption fluctuations, in this section we obtain the cost of *business cycle* fluctuations in consumption. Just like the cost of all consumption uncertainty, the cost of business cycle fluctuations in consumption can be obtained as the ratio of the prices of two hypothetical securities: a claim to a stabilized consumption path, pc^{stab}, and a claim to the actual aggregate consumption, $\tilde{pc}$. Stabilized consumption in this case refers to the residual after the business cycle component has been removed from the aggregate consumption series. We compute the stabilized consumption series using the widely used Hodrick-Prescott filter. Since our empirical analysis uses annual data, we use a smoothing parameter of 6.25 in the application of the Hodrick-Prescott filter, following the suggestions in Ravn and Uhlig (2002).

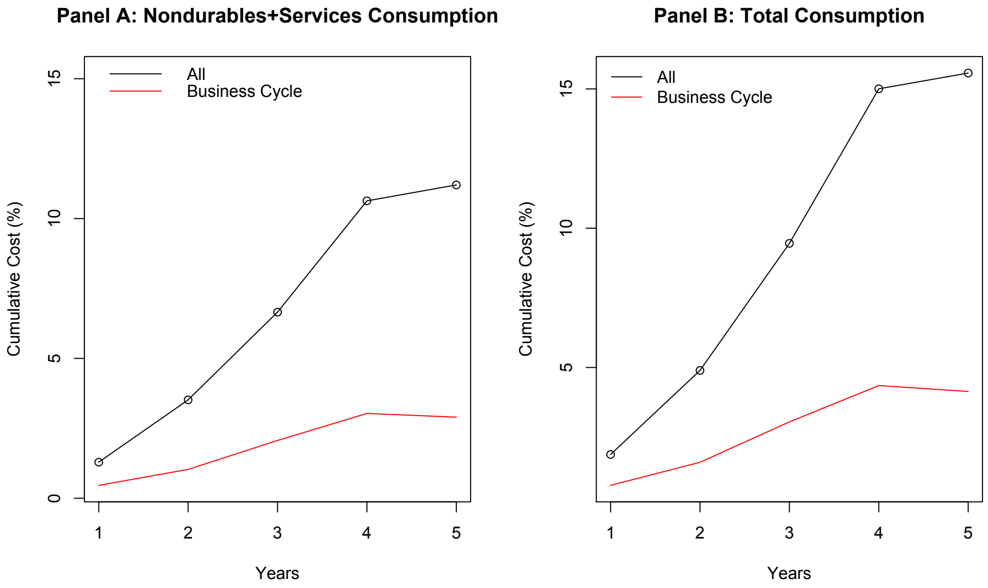
Figure 3. Marginal Cost of All versus Business Cycle Consumption Fluctuations, 1929-2015



Notes: The figure plots the cumulative cost of all aggregate consumption fluctuations (black line) and business cycle fluctuations in consumption (red line), over 1-5 *years, obtained using the I-SDF for different choices of the measures of consumption. Panel A presents results when consumption refers to the real per capita personal consumption expenditure of nondurables and services, while Panel B does the same when consumption denotes real per capita total personal consumption expenditure. The I-SDF* *is extracted using the excess return on the market portfolio as the sole test asset. The sample is annual covering the period* 1929-2015.

The results are presented in the last five columns of Table 2, for nondurables and services consumption. Panel A, Row 1 shows that, using the I-SDF extracted from the market portfolio alone, the cost of business cycle fluctuations in consumption over a one-year time horizon is estimated to be 0.6%. The costs of business cycle fluctuations over two, three, four, and five year horizons increase to 1.5%, 3.4%, 3.9%, and 3.7%, respectively. Row 2 shows that, for the CRRA kernel, while the cost of business cycle fluctuations over a one-year period is only slightly smaller than that obtained with the I-SDF (0.5% versus 0.6%), the cost increases little for multi-year horizons in case of the former. For instance, the cost of five-year fluctuations is only 1.4% – less than half of the cost of 3.7% implied by the I-SDF. Panel B shows that similar results are obtained when the six size and book-to-market-equity sorted portfolios are used in the extraction of the I-SDF.

Figure 4. Marginal Cost of All versus Business Cycle Consumption Fluctuations, 1929-2015



Notes: The figure plots the cumulative cost of all aggregate consumption fluctuations (black line) and business cycle fluctuations in consumption (red line), over 1-5 *years, obtained using the I-SDF for different choices of the measures of consumption. Panel A presents results when consumption refers to the real per capita personal consumption expenditure of nondurables and services, while Panel B does the same when consumption denotes real per capita total personal consumption expenditure. The I-SDF is extracted using the excess returns on the 6 Fama-French size and book-to-market-equity sorted portfolios. The sample is annual covering the period* 1929-2015.

An important point to note is that while the estimates of the cost of business cycle fluctuations are smaller than the cost of all consumption uncertainty, the former, nonetheless, represents a substantial fraction of the latter. For instance, Panel A, Row 1 shows that, when the market portfolio is used in the extraction of the I-SDF, the cost of business cycle fluctuations constitutes 36.3% of the cost of all consumption fluctuations over a one-year horizon. The cost of business cycle fluctuations over two, three, four, and five years account for 28.7%, 28.9%, 27.3%, and 24.7%, respectively, of the cost of all consumption fluctuations over these time horizons. Figure 3, Panel A plots the cost of all consumption fluctuations (black line) and the cost of business cycle fluctuations (red line) over 1-5 years.

Similar results are obtained in Table 2, Panel B when the six size and book-to-market-equity sorted portfolios are used in the extraction of the I-SDF. Specifically, the cost of business cycle fluctuations over one to five years accounts for 35.8%, 29.3%, 31.1%, 31.0%, and 25.9%, respectively, of the cost of all consumption fluctuations over these time horizons. This is further demonstrated in Figure 4, Panel A.

Finally, the results remain largely unchanged when total consumption expenditure (instead of nondurables and services expenditure) is used as the measure of consumption in recovering the I-SDF. These are presented in Table 3, Rows 7-11. Panel B of Figures 3-4 plot these costs of all and business cycle fluctuations when the set of assets used to recover the I-SDF consists of the market alone and the 6 FF portfolios, respectively.

Overall, we find that the costs of business cycle fluctuations are large and constitute between a quarter to a third of the cost of all consumption fluctuations. Our results are in contrast to those in Alvarez and Jermann (2004) who argue that while the cost of all consumption fluctuations is very high with a baseline value of 28.6% for an infinitely lived agent, the cost of business cycle fluctuations in consumption is miniscule, varying from 0.1% to 0.5%. Our estimates of the costs of business cycle fluctuations over a cumulative five-year period alone are as high as 5.1% – between ten and fifty times higher than the estimates in Alvarez and Jermann (2004). Also note that the estimates in the latter, unlike our estimates, correspond to eliminating business cycle fluctuations for all (infinite) future periods, not just for a five-year time horizon.

VII. Robustness

In this section, we perform a number of checks to establish the robustness of our estimates of the cost of all consumption uncertainty as well as the cost of business cycle fluctuations in consumption reported in Sections V and VI. For all the robustness tests, consumption refers to the per capital total personal consumption expenditure.[9] The results are presented in Table 4.

First, we present the estimates for an alternative definition of relative entropy. Equation (12) reveals that relative entropy is not symmetric. Therefore, we can reverse the roles of the physical measure $\mathbb{P}$ and the tilted measure $\mathbb{F}$ so as to obtain an alternative definition of relative entropy. This alternative criterion function can be minimized to estimate the measure, $\mathbb{F}$, and, therefore, the missing component,

[9]Very similar results are obtained using nondurables and services consumption and are omitted for brevity.

ψ, of the pricing kernel:

$$\min_{\mathbb{F}} \int \mathbf{log}\left(\frac{d\mathbb{F}}{d\mathbb{P}}\right) d\mathbb{F} = \int \mathbf{log}\left(\frac{f(\mathbf{z})}{p(\mathbf{z})}\right) f(\mathbf{z})d\mathbf{z} \text{ s.t. } \mathbf{0} = \int \mathbf{R}^e(\mathbf{z}) \left(\Delta c(\mathbf{z})\right)^{-\gamma} f(\mathbf{z})d\mathbf{z}, \tag{27}$$

This is the Exponentially-Tilted (ET) estimator of Kitamura and Stutzer (1997) (see also Susanne M. Schennach (2005)). As with the EL estimator, the ET estimator is also numerically simple to implement. Specifically, the ψ-component is estimated (up to a positive constant scale factor) as:

$$\hat{\psi}_t = \frac{e^{\hat{\theta}'\mathbf{R}_t^e(\Delta c_t)^{-\gamma}}}{\frac{1}{T}\sum_{t=1}^T e^{\hat{\theta}'\mathbf{R}_t^e(\Delta c_t)^{-\gamma}}} \quad \forall t, \tag{28}$$

where $\hat{\theta} \in \mathbb{R}^N$ is the vector of Lagrange multipliers that solves the unconstrained dual problem:

$$\hat{\theta} = \arg\min_{\theta} \left[\log\left(\frac{1}{T}\sum_{t=1}^T e^{\hat{\theta}'\mathbf{R}_t^e(\Delta c_t)^{-\gamma}}\right)\right]. \tag{29}$$

We recover the I-SDF using the ET approach and use it to calculate the costs of consumption fluctuations. The results are presented in Row 1 of each panel in Table 4. Panel A, Row 1 reports the results when the market portfolio is the sole test asset used to extract the I-SDF. The results are very similar to those obtained using the EL approach in Table 3 – the cumulative costs of all 1 to 5 year fluctuations in consumption are 2.1%, 6.1%, 14.5%, 17.3%, and 17.3%, respectively, remarkably close to the corresponding values (2.2%, 6.8%, 16.1%, 19.7%, and 19.7%, respectively) obtained using the EL approach. The costs of 1 to 5 year business cycle fluctuations in consumption are also very similar for the two approaches – 0.90%, 2.0%, 4.5%, 5.0%, and 4.6%, respectively, for the ET approach versus 0.90%, 2.1%, 4.9%, 5.6%, and 5.1%, respectively, for the EL. Therefore, for both approaches, the cost of business cycle fluctuations constitutes between a quarter to a third of the cost of all consumption fluctuations. Finally, Panel B, Row 1 shows that the results for the two approaches remain quite similar when the six size and book-to-market-equity sorted portfolios of Fama-french are used to recover the I-SDF.

Table 4: Cumulative Cost of Total Consumption Fluctuations, Robustness Checks

	All Fluctuations					B. C. Fluctuations				
	1 Yr	2 Yr	3 Yr	4 Yr	5 Yr	1 Yr	2 Yr	3 Yr	4 Yr	5 Yr
Panel A: Market Portfolio										
$I\text{-}SDF^{ET}$	2.07	6.13	14.68	17.32	17.31	.904	1.98	4.47	4.98	4.59
$I\text{-}SDF^{Alt}$	1.83	4.92	10.87	12.75	12.70	.851	1.75	3.48	3.82	3.52
1890-2015	1.38	2.69	4.85	6.84	8.24	.931	1.39	2.08	2.54	2.69
Panel B: FF 6 Portfolios										
$I\text{-}SDF^{ET}$	1.83	4.81	8.72	14.09	14.75	.691	1.43	2.73	4.01	3.83
$I\text{-}SDF^{Alt}$	1.76	4.46	8.96	14.12	14.67	.764	1.61	3.02	4.20	4.00
1890-2015	-	-	-	-	-	-	-	-	-	-

The table reports the (cumulative) cost of *all* aggregate consumption fluctuations (Columns 2-6) and the cost of business cycle fluctuations in consumption (Columns 7-11), for 1-5 years. Consumption denotes the real per capita *total* personal consumption expenditure (includes durables, nondurables, and services). The costs are calculated using the I-SDF extracted with the ET approach (Row 1), the risk-neutral measure recovered by minimizing the distance from the CRRA model-implied risk-neutral measure (Row 2), and the I-SDF extracted with the EL approach over the longer 1890-2015 sample (Row 3). Panel A presents results when the excess return on the market portfolio is the sole asset used to recover the I-SDF. In Panel B, on the other hand, the I-SDF is estimated using the 6 Fama-French size and book-to-market-equity sorted portfolios. The sample is annual covering the period 1929-2015.

Our second robustness check uses yet another definition of relative entropy. Specifically, we recover the risk-neutral measure $\mathbb{Q}$ such that:

(30)
$$\widehat{\mathbb{Q}} = \min_{\mathbb{Q}} \int \mathbf{log}\left(\frac{d\mathbb{Q}}{d\mathbb{Q}^m}\right) d\mathbb{Q} = \int \mathbf{log}\left(\frac{q(\mathbf{z})}{q^m(\mathbf{z})}\right) q(\mathbf{z})d\mathbf{z} \text{ s.t. } \mathbf{0} = \int \mathbf{R}^e(\mathbf{z})q(\mathbf{z})d\mathbf{z},$$

where $\frac{d\mathbb{Q}^m}{d\mathbb{P}} = \frac{(\Delta c)^{-\gamma}}{E\left[(\Delta c)^{-\gamma}\right]}$. In other words, $\mathbb{Q}^m$ is the risk neutral measure implied by the power utility model with a constant CRRA. Thus, Equation (30) recovers the risk neutral measure $\mathbb{Q}$ that is minimally distorted relative to the CRRA model implied risk neutral measure $\mathbb{Q}^m$, while also successfully pricing the set of test assets used in the estimation. Note that the main difference between Equation (30) and the EL and ET estimators defined in Equations (10) and (27), respectively, is that while the latter two minimize the relative entropy (or distance) between the recovered measure and the physical measure, the former minimizes the distance between the recovered risk neutral measure and the measured implied by a candidate model SDF.

The solution to Equation (30) is obtained as:

$$\hat{q}_t = \frac{e^{\hat{\theta}'\mathbf{R}^e_t}(\Delta c_t)^{-\gamma}}{\frac{1}{T}\sum_{t=1}^{T} e^{\hat{\theta}'\mathbf{R}^e_t}(\Delta c_t)^{-\gamma}} \quad \forall t, \tag{31}$$

where $\hat{\theta} \in \mathbb{R}^N$ is the vector of Lagrange multipliers that solves the dual problem:

$$\hat{\theta} = \arg\min_{\theta} \left[\log\left(\frac{1}{T}\sum_{t=1}^{T} e^{\hat{\theta}'\mathbf{R}^e_t}(\Delta c_t)^{-\gamma}\right)\right]. \tag{32}$$

We use the recovered risk neutral measure $\hat{q}_t$ to calculate the cost of consumption fluctuations. The results, reported in Row 2 of Panels A and B, for the scenarios when the test assets consist of the market portfolio alone and the six Fama-french portfolios, respectively, are very similar to those obtained with the ET (Table 4, Row 1) and EL (Table 3, Row 1) approaches.

Third, we present the costs of fluctuations using the EL approach with data going back as far as 1890. The excess return on the market is the sole test asset, with the return on the S&P composite index used as a proxy for the market return and the prime commercial paper rate as a proxy for the risk free rate. The data are obtained from Robert Shiller's website. The costs of all and business cycle fluctuations in consumption, presented in Row 3 of Panel A, are smaller than those obtained using the baseline 1929-2015 sample (see Tables 2, 3 and Rows 1-2 of Table 4). The smaller estimates of the cost obtained in this longer data sample can be accounted for, at least partly, by the usage of the commercial paper rate as a proxy for the risk free rate, thereby leading to an underestimation of the magnitude of the equity premium puzzle in this sample. Specifically, the average level of the equity premium is 7.9% in the baseline sample, more than double the value of 3.1% in the longer 1890 onwards sample. Moreover, just as with the baseline sample, the cost of business cycle fluctuations still accounts for a substantial fraction (more than a third) of the cost of all consumption fluctuations for all the horizons considered.[10]

Overall, our results suggest that the estimates of the cost of aggregate economic fluctuations are robust to the measure of consumption expenditures, the set of test assets used to recover the I-SDF, the choice of sample period, as well as the precise definition of relative entropy. This lends further support to the quantitative estimates in the paper.

VIII. Time-Varying Cost of Aggregate Fluctuations

Our analysis, so far, has focused on the *expected* cost of consumption fluctuations, i.e. the average cost over all possible states. This is why the cost was defined as the ratio of the expected (or, average) prices of a claim to a stabilized consumption stream and a claim to the actual aggregate consumption stream (see Equations (5) and (7)). In this section, we provide evidence that the cost of fluctuations varies substantially over time. And, perhaps more importantly, the precise nature of the time-variation helps shed some light on the reasons for the substantial welfare benefits of eliminating not only all consumption uncertainty, but also business cycle fluctuations in consumption that we estimate in Sections V and VI. To our knowledge, this is the first attempt to recover the time-varying cost of aggregate economic fluctuations, without taking a stance on investors' preferences or the dynamics of the data generating process.

Subsection VIII.A describes an extension of the information-theoretic EL approach, namely the smoothed empirical likelihood (SEL) estimator of Kitamura, Tripathi and Ahn (2004), that we use to recover the time-varying cost of fluctuations. Subsection VIII.B presents simulation evidence on the performance of the SEL estimator. Finally, Subsection VIII.C presents the estimated time series of the cost of removing all consumption uncertainty over a one-period time horizon.

A. Smoothed Empirical Likelihood (SEL)

Following the notation in Section II, the time-t cost of *all* one-period consumption fluctuations is defined as

$$\frac{\frac{V_t\left(C_{t+1}^{stab}\right)}{C_t}}{\frac{V_t(C_{t+1})}{C_t}} - 1 = \frac{\mathbb{E}^{\mathbb{P}_t}\left[M_{t+1}\frac{C_{t+1}^{stab}}{C_t}|\underline{\mathcal{F}_t}\right]}{\mathbb{E}^{\mathbb{P}_t}\left[M_{t+1}\frac{C_{t+1}}{C_t}|\underline{\mathcal{F}_t}\right]} - 1 = \frac{\mathbb{E}^{\mathbb{P}_t}\left[M_{t+1}\left(1+\mu_c\right)|\underline{\mathcal{F}_t}\right]}{\mathbb{E}^{\mathbb{P}_t}\left[M_{t+1}\frac{C_{t+1}}{C_t}|\underline{\mathcal{F}_t}\right]} - 1, \tag{33}$$

[10]Since the size and book-to-market-equity sorted portfolios are not available prior to the late 1920s, we cannot recover the I-SDF using these portfolios over the 1890-2015 sample.

where $\underline{\mathcal{F}_t} = \{\mathcal{F}_t, \mathcal{F}_{t-1}, \ldots\}$ denotes the investors' information set at time t, $\mathbb{E}^{\mathbb{P}_t}\left[.|\underline{\mathcal{F}_t}\right]$ refers to the expectation with respect to the physical measure $\mathbb{P}$ conditional on the investors' time-t information set, and the second equality follows from the definition of stabilized consumption as one from which all uncertainty has been removed: $C^{stab}_{t+1} = (1+\mu_c)\, C_t$.

As in Section II, we assume that the pricing kernel M has a multiplicative form, $M_{t+1} = (\Delta c_{t+1})^{-\gamma}\, \psi_{t+1}$. We then rely on an extension of our information-theoretic methodology to estimate the ψ-component of the pricing kernel that now satisfies the *conditional* (not just the unconditional) Euler equation restrictions for a chosen cross section of assets. Recall that our information-theoretic EL approach in Section II recovers a pricing kernel that prices assets unconditionally, i.e. satisfies the unconditional Euler equations producing zero unconditional pricing errors. The extension of the methodology considered in this section recovers an SDF that satisfies the more stringent conditional Euler equation restrictions, thereby producing zero conditional pricing errors. The recovered SDF, therefore, must also price assets unconditionally. Specifically, we use the smoothed empirical likelihood (SEL) estimator of Kitamura, Tripathi and Ahn (2004). As described below, the SEL estimator relies on the same principles as the EL estimator, but incorporates additional constraints through conditional moment restrictions.

The absence of arbitrage opportunities implies the following conditional pricing restrictions:

$$\mathbb{E}^{\mathbb{P}_t}\left[M_{t+1}\mathbf{R}^e_{t+1}|\underline{\mathcal{F}_t}\right] = \mathbb{E}^{\mathbb{P}_t}\left[(\Delta c_{t+1})^{-\gamma}\,\psi_{t+1}\mathbf{R}^e_{t+1}|\underline{\mathcal{F}_t}\right] = \mathbf{0}, \tag{34}$$

where the first equality follows from the assumed multiplicative decomposition of the SDF. Under weak regularity conditions, we have

$$\mathbb{E}^{\mathbb{P}_t}\left[(\Delta c_{t+1})^{-\gamma}\,\frac{\psi_{t+1}}{E^{\mathbb{P}_t}(\psi_{t+1}|\underline{\mathcal{F}_t})}\mathbf{R}^e_{t+1}|\underline{\mathcal{F}_t}\right] = \mathbb{E}^{\mathbb{F}_t}\left[(\Delta c_{t+1})^{-\gamma}\,\mathbf{R}^e_{t+1}|\underline{\mathcal{F}_t}\right] = \mathbf{0}, \tag{35}$$

where $\frac{d\mathbb{F}_t}{d\mathbb{P}_t} = \frac{\psi_{t+1}}{E^{\mathbb{P}_t}(\psi_{t+1}|\underline{\mathcal{F}_t})}$ is the Radon-Nikodym derivative of $\mathbb{F}$ with respect to $\mathbb{P}$.

We assume that the time-t information set of the investors, $\underline{\mathcal{F}_t}$, can be summarized by a finite vector of random variables, that we denote by $X_t \in \mathbb{R}^m$. Suppose that the historical realizations of consumption growth, excess returns, and the conditioning variables are given by $(\Delta c_t, \mathbf{r}^e_t, x_t)^T_{t=1}$,[11] and that these realizations characterize the (finite number of) possible states of the world. Let $f_{i,j}$ denote the conditional probability (under the measure $\mathbb{F}$) of observing the joint outcome $(\Delta c_j, \mathbf{r}^e_j, x_j)$ at time $t+1$, i.e. the probability of state j being realized at time $t+1$, given that state i was realized at time t.

The SEL estimator of the transition matrix $\{f_{i,j}; i,j = 1,\ldots,T\}$ is such that it

[11]Throughout this section, uppercase letters are used to denote random variables and the corresponding lowercase letters to particular realizations of these variables.

belongs to the simplex:

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

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

$$\left(\widehat{f}_{i,\cdot}^{SEL}(\gamma)\right) = \underset{(f_{i,\cdot}) \in \Delta_i}{\arg\max} \sum_{j=1}^{T} \omega_{i,j} \log(f_{i,j}) \quad \text{s.t.} \quad \sum_{j=1}^{T} f_{i,j} \times (\Delta c_j)^{-\gamma}\, \mathbf{r}_j^e = \mathbf{0}, \tag{36}$$

where $f_{i,\cdot}$ denotes the T-dimensional vector $(f_{i,1}, ..., f_{i,T})$, Θ is the set of all admissible parameters γ, and $\omega_{i,j}$ are non-negative weights used to smooth the likelihood objective function. In the spirit of non-parametric estimators:

$$\omega_{i,j} = \frac{\mathcal{K}\left(\dfrac{x_i - x_j}{b_T}\right)}{\displaystyle\sum_{t=1}^{T} \mathcal{K}\left(\dfrac{x_i - x_t}{b_T}\right)}, \tag{37}$$

where $\mathcal{K}$ is a kernel function belonging to the class of second order product kernels,[12] and the bandwidth b_T is a smoothing parameter.[13]

Note that the objective function in Equation (36) is simply a 'smoothed' log-likelihood, with the constraints enforcing the conditional Euler equation restrictions in Equation (35). The weights $\omega_{i,j}$ used to smooth the log-likelihood are standard non-parametric kernel weights. The intuition behind the estimator may be understood as follows. Note that we are interested in recovering $f_{i,j}$, for $i, j = 1, 2, ..., T$. For each value of the current state x_i, the SEL estimator focuses on a fixed neighbourhood around x_i, where the neighbourhood is defined in terms of the distance of other possible values of the state from the current state, i.e. $|x_i - x_j|$, and not in terms of proximity in time. The estimator then assigns positive weights $\omega_{i,j}$ only to those states that lie within the fixed neighbourhood of the current state, with the exact values of the weights determined by the kernel function, the distance $|x_i - x_j|$, and the bandwidth parameter b_T (see Equation 37). The states that lie outside the fixed neighbourhood each receive a weight of zero. Finally, the SEL approach determines the conditional probability of each state with non-zero weight, $\omega_{i,j} > 0$, so as to maximize the smoothed log-likelihood of the data, subject to the constraint that the estimated conditional distribution, $\left\{\widehat{f}_{i,j}^{SEL}; j = 1, 2, ..., T\right\}$, satisfies the conditional Euler equation restrictions (see Equation 36). The states with zero weight, $\omega_{i,j} = 0$, each receive a conditional probability of zero.

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

[13] In theory, b_T is a null sequence of positive numbers such that $Tb_T \to \infty$.

The solution to Equation (36) is analytical and given by:
$\forall i, j \in \{1, \ldots, T\}$,

$$\widehat{f}_{i,j}^{SEL}(\gamma) = \frac{\omega_{i,j}}{1 + (\Delta c_j)^{-\gamma}\,\widehat{\lambda}_i(\gamma)'\,\mathbf{r}_j^e}, \tag{38}$$

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

$$\widehat{\lambda}_i(\gamma) = \arg\max_{\lambda_i \in \mathbb{R}^n} \sum_{j=1}^{T} \omega_{i,j} \log\left[1 + (\Delta c_j)^{-\gamma}\,\lambda_i'\,\mathbf{r}_j^e\right]. \tag{39}$$

Equations (38) and (39) show that the SEL procedure delivers a $(T \times T)$ matrix of probabilities $\left(\widehat{f}_{i,j}^{SEL}(\gamma)\right)$ for each value of the parameter γ. Each row $i : i = \{1, 2, ..., T\}$ contains the probabilities of transitioning to each of the T possible states $j : \{j = 1, 2, ..., T\}$ in the subsequent period, conditional on state i having been realized in the current period. Therefore, the approach recovers the *entire conditional distribution* of the data, under the measure $\mathbb{F}$, that is consistent with observed asset prices, i.e. that satisfies the conditional Euler equations. Moreover, it does so without the need for any parametric functional-form assumptions on the form of the distribution, i.e. on the form of the ψ-component of the SDF. Rather, it approximates the conditional distribution, for each possible value of the current state, as a multinomial on the observed data sample.

Note that the SEL estimator in Equation (36) can also be reformulated as:

$$\left(\widehat{f}_{i,\cdot}^{SEL}(\gamma)\right) = \arg\min_{(f_{i,\cdot}) \in \Delta_i} \sum_{j=1}^{T} \log\left(\frac{\omega_{i,j}}{f_{i,j}}\right)\omega_{i,j} \quad \text{s.t.} \quad \sum_{j=1}^{T} f_{i,j} \times (\Delta c_j)^{-\gamma}\,\mathbf{r}_j^e = \mathbf{0}, \tag{40}$$

The objective function in Equation (40) is the KLIC divergence between the measure $\mathbb{F}_t \equiv (f_{t,j})_{j=1}^T$ that is consistent with asset prices, i.e. satisfies the conditional Euler equations for the test assets, and the physical measure proxied by $\mathbb{P}_t \equiv (\omega_{t,j})_{j=1}^T$. $\frac{f_{t,j}}{\omega_{t,j}} = \frac{\psi_{t,j}}{E^{\mathbb{P}_t}(\psi_{t,j}|\underline{\mathcal{F}_t})}$ is the Radon-Nikodym derivative of $\mathbb{F}_t$ with respect to $\mathbb{P}_t$. Suppose that the consumption growth component of the pricing kernel, $(\Delta c)^{-\gamma}$, is sufficient to price assets perfectly. Then the second component of the pricing kernel $\psi_{t,j} \equiv 1, \forall j = 1, 2, ..., T$, and we have that $f_{t,j} = \omega_{t,j}, \forall j = 1, 2, ..., T$, the latter being the physical measure. However, if the consumption growth component is not sufficient to price assets (as is the case in reality), the estimated measure $\mathbb{F}_t$ is distorted relative to the physical measure $\mathbb{P}_t$. And, the SEL estimator searches for a measure $\mathbb{F}_t$ that is as close as possible, in an information-theoretic sense, to the physical measure $\mathbb{P}_t$. In other words, the approach distorts the physical probabilities as little as possible in order to satisfy the conditional Euler equation restrictions.

Using the SEL-estimated conditional distribution, the cost of one-period consumption fluctuations at each date (or state) t can be calculated as:

(41)
$$\frac{\frac{V_t\left(C_{t+1}^{stab}\right)}{C_t}}{\frac{V_t(C_{t+1})}{C_t}}-1=\frac{\mathbb{E}^{\mathbb{F}_t}\left[(\Delta c_{t+1})^{-\gamma}(1+\mu_c)\,|\underline{\mathcal{F}_t}\right]}{\mathbb{E}^{\mathbb{F}_t}\left[(\Delta c_{t+1})^{-\gamma}(\Delta c_{t+1})\,|\underline{\mathcal{F}_t}\right]}-1=\frac{(1+\mu_c)\sum_{j=1}^{T}\widehat{f}_{t,j}^{SEL}\times(\Delta c_j)^{-\gamma}}{\sum_{j=1}^{T}\widehat{f}_{t,j}^{SEL}\times(\Delta c_j)^{1-\gamma}}-1.$$

B. Performance of the SEL Estimator

Ghosh and Roussellet (2019) show, via simulation exercises, that the SEL approach is quite successful at recovering the conditional distribution of the data that is consistent with asset prices, i.e. $\mathbb{F}$ in our notation. Specifically, they consider a Bansal and Yaron (2004) long run risks economy. Thus, the following conditional Euler equation holds in equilibrium for the excess return on the stock market:

$$\mathbb{E}^{\mathbb{F}_t}\left[\underbrace{(\Delta c_{t+1})^{-\frac{\theta}{\rho}}R_{c,t+1}^{\theta-1}}_{M_{t+1}}(\mathbf{R}_{m,t+1}-R_{f,t+1})\,|\underline{\mathcal{F}_t}\right]=\mathbf{0}, \tag{42}$$

where $R_{c,t}$ denotes the return on total wealth, ρ the elasticity of intertemporal substitution (EIS), and $\theta=\frac{1-\gamma}{1-\frac{1}{\rho}}$. The investors' information set at time-t consists of the two model-implied state variables: $\mathcal{F}_t=\left\{\nu_t,\sigma_t^2\right\}$, where ν_t denotes the expected consumption growth rate and σ_t^2 its stochastic variance. Thus, the SDF in this economy depends not only on consumption growth (as in the standard time and state separable power utility model), but also on the return on total wealth.

Ghosh and Roussellet (2019) set the preference parameters and the parameters governing the dynamics of the consumption and dividend growth processes to the authors' calibrated values. They then simulate a time series, of the same length T as the historical sample, of the two state variables and, therefore, consumption growth and the return on total wealth to recover the time series of the SDF; and they also simulate a time series of the market return and the risk free rate. Using the simulated sample, they then recover the distribution $\mathbb{F}$ using the SEL approach. Note that the implementation of the SEL approach requires specification of two inputs – the test assets and the conditioning set. The authors' use the excess return on the market as the sole test asset and the two model-implied state variables as constituting the conditioning set. Also, the SEL estimation approach, like all other nonparametric procedures, requires specification of the kernel function and the associated bandwidth parameter. All the authors' results are computed with the Epanechnikov kernel function and with the bandwidth parameter $b_{v,T}=3\hat{\sigma_v}$, where $\hat{\sigma_v}$ is the empirical standard deviation of the conditioning variable v.[14]

[14]The results are robust to alternative choices of the kernel function and the smoothing parameter within four standard deviations of the volatility of the conditioning variable.

This leads to the following estimate of the conditional distribution $\mathbb{F}$: $\forall i \in \{1, \ldots, T\}$,
(43)

$$\left(\widehat{f}_{i,\cdot}^{SEL}\right) = \underset{(f_{i,\cdot}) \in \Delta_i}{\arg\max} \sum_{j=1}^{T} \omega_{i,j} \log(f_{i,j}) \quad \text{s.t.} \quad \sum_{j=1}^{T} f_{i,j} \times (\Delta c_j)^{-\frac{\theta}{\rho}} \, r_{c,j}^{\theta-1} \mathbf{r}_{m,j}^{e} = \mathbf{0}.$$

Note that, since the SDF is fully specified and not missing any components and there are no beliefs distortions, the measure $\mathbb{F}$ in Equation (42) coincides with the physical measure $\mathbb{P}$. Ghosh and Roussellet (2019) show that Equation (43) identifies the physical measure very well, i.e. $\left\{\widehat{f}_{i,j}^{SEL}\right\}_{i,j=1}^{T}$ recovers the time series of the conditional moments of the consumption growth rate with a high degree of accuracy.

C. *Empirical Results*

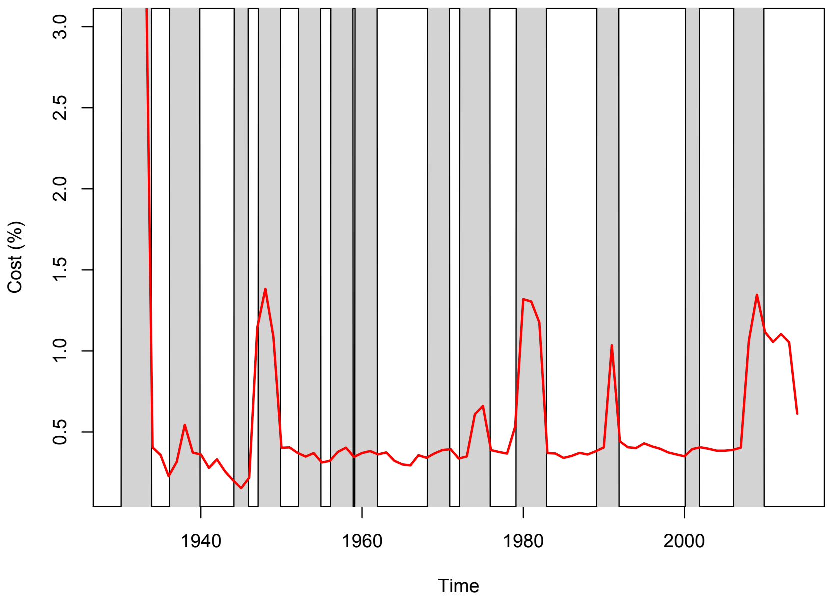
Having shown that the SEL approach is quite successful at recovering the conditional distribution of the data, we now proceed to use the method to estimate the cost of aggregate consumption fluctuations in different states (or, times).

We first estimate the time series of the cost of one-period consumption fluctuations in Equation (41) in our baseline sample covering the period 1930-2015. Each year corresponds to a particular state and the SEL approach estimates the welfare benefits of eliminating consumption uncertainty in the subsequent year. In our implementation, we use nondurables and services consumption as the measure of the aggregate consumption expenditures, the excess return on the market portfolio as the sole test asset, and an exponentially-weighted moving average of lagged consumption growth as the conditioning variable.

Figure 5 presents the time series of the cost. Several features are immediately evident from the figure. First, the cost is strongly time-varying – it varies from 0.15% to 8.0% a year, with an average of 0.75%. Second, the cost is strongly countercyclical, rising sharply during recessionary episodes. The average of the cost over a subsample that corresponds to recession years, where a year is classified as a recession year if there is an NBER-designated recession in any of its quarters, is 1.17%. The estimated costs are particularly high during the period of the Great Depression 1930-1933, with a mean of 5.8% and the maximum as high as 8.0%. As a contrast, the average cost over the subsample comprised of expansionary episodes alone is less than half of that during recessions at 0.53%. The correlation between the cost and a dummy variable that takes the value 1 in a given year if there is an NBER-designated recession in any of its quarters and 0 otherwise is 36.1%. Finally, the estimates of the cost are large, given that they represent the welfare benefits of eliminating consumption uncertainty for one period alone.

Note that the above results are obtained using a weighted average of past consumption growth as the sole conditioning variable. This may potentially raise concerns about the robustness of the findings. In unreported results, we repeat the SEL estimation using additional conditioning variables, such as weighted averages of past inflation, labor market variables, and the stock market return. The results remain largely unaltered, both qualitatively as well as quantitatively.[15]

Figure 5. Time-Varying Cost of One-Period Consumption Fluctuations, 1929-2015



Notes: The figure plots the time series of the cost of one-period consumption uncertainty. The cost is estimated using the SEL approach, using nondurables and services consumption as the measure of the consumption expenditures, the excess return on the market portfolio is the sole test asset, and an exponentially-weighted moving average of lagged consumption growth as the conditioning variable. The sample is annual, covering the period 1930-2015.

Overall, our results suggest that the cost of consumption fluctuations is strongly countercyclical and this offers, at least a partial, explanation of the high costs of business cycle fluctuations that we estimate in Sections V and VI.

IX. What Drives the Results?

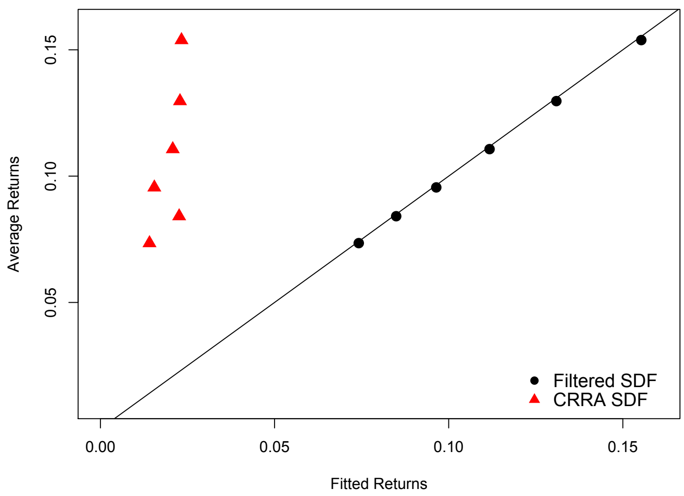
Our results suggest that the welfare benefits of eliminating all consumption uncertainty as well business cycle fluctuations in consumption are substantially bigger than those obtained with the CRRA kernel or with Lucas' original specification that imposes the additional assumption of lognormal consumption growth to the CRRA kernel. Moreover, the cost of consumption uncertainty is strongly time-varying and countercyclical, rising sharply during economic downturns. These results are robust to the precise measure of the aggregate consumption expenditure

[15]These results are available from the authors upon request.

as well as the set of traded assets used to recover the I-SDF. A natural question arises as to which features of the I-SDF drive these results. In this section, we highlight two characteristics of the I-SDF that can help interpret our findings.

First, the I-SDF can successfully explain the historically observed average returns on both the aggregate stock market index as well as returns on broad diversified portfolios formed by sorting stocks on the basis of observable characteristics such as size and the book-to-market-equity ratio, i.e. it accurately prices assets. The CRRA kernel and Lucas' specification, on the other hand, produce large average pricing errors for these assets. Figure 6 plots the historical average excess returns (y-axis) along with the average excess returns implied by a particular pricing kernel (x-axis), for the six size-and book-to-market-equity sorted portfolios of Fama and French. For a candidate pricing kernel M, the average excess return on portfolio i implied by the kernel is obtained as $-\frac{Cov\left(M_t, R^e_{i,t}\right)}{E(M_t)}$. The average excess returns on these portfolios implied by the I-SDF are denoted by black circles, while those implied by the CRRA kernel are denoted by red triangles.

Figure 6. Unconditional Pricing Errors, 1929-2015



Notes: The figure plots the historical average excess returns (y-axis) along with the average excess returns implied by candidate pricing kernels (x-axis), for the six size-and book-to-market-equity sorted portfolios of Fama and French. The average excess returns on these portfolios implied by the I-SDF are denoted by black circles, while those implied by the CRRA kernel are denoted by red triangles. The sample is annual, covering the period 1930-2015.

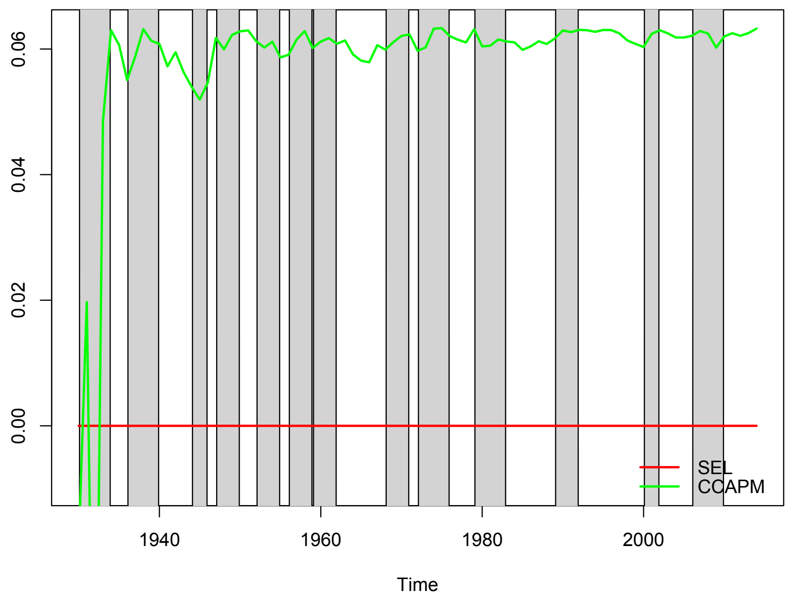
The figure shows that the CRRA kernel grossly underestimates the average excess returns. Specifically, the historical average excess return across the 6 portfolios is 10.8%, whereas the CRRA kernel implies an average of only 2.0%. Also, the kernel fails to explain the substantial cross-sectional differences in average returns across the portfolios. The historical average excess return varies from 7.3% for the portfolio comprised of large market cap and growth stocks to more than double at 15.4% for the portfolio of small market cap and value stocks. To the contrary, the average excess returns implied by the CRRA kernel are 1.4% and 2.3%, respectively, for these two portfolios. The cross-sectional R^2, defined as the ratio of the cross-sectional variance of the average excess returns implied by the CRRA model and the cross-sectional variance of the historical average excess returns is only 1.8%. These shortcomings of the CRRA model have been widely documented in the literature and our results confirm these findings.

The above observations suggest that the CRRA model misses important components of the underlying sources of systematic risk. Since the welfare costs of consumption fluctuations depend critically on people's attitudes towards different sources of risk, the estimates of this cost obtained using the CRRA kernel should, at best, be interpreted with caution. Moreover, Ghosh, Julliard and Taylor (2016*b*) evaluate the pricing performance of several other prominent consumption-based models (that were intended to overcome the shortcomings of the CRRA model) and show that they too perform quite poorly, producing large pricing errors and low (and often negative) cross-sectional R^2. This may partly account for the wildly different costs of fluctuations obtained using these alternative model specifications. More importantly, it suggests that the concerns with using the CRRA kernel to estimate the cost of aggregate fluctuations may carry over to many of the more recent pricing kernel specifications as well.

Figure 6 shows that the I-SDF, on the other hand, accurately prices assets. This result, per se, is hardly surprising because the I-SDF was constructed to price the assets in-sample (see Equation (12)). This may potentially raise concerns regarding over-fitting and spurious inference. In this regard, Ghosh, Julliard and Taylor (2016*a*) show that the good pricing performance of the I-SDF also obtains out-of-sample for broad cross-sections of assets, including domestic and international equities, currencies, and commodities. The out-of-sample performance of the I-SDF is superior to not only the single factor CAPM and the Consumption-CAPM, but also to the more recent Fama-French 3 and 5 factor models.

Also, not only does the I-SDF price assets unconditionally delivering zero average pricing errors, it also produces zero conditional pricing errors. Figure 7 plots the time series of the conditional pricing errors for the excess stock market return implied by the I-SDF (red line) and the CRRA kernel (green line). The SEL approach, described in Section VIII, can be used to compute the conditional pricing errors implied by the I-SDF. Specifically, the conditional pricing error for the excess market return at each date t is given by $\sum_{j=1}^{T} \widehat{f}_{t,j}^{SEL} \times (\Delta c_j)^{-\gamma}\, \mathbf{r}_{m,j}^{e}$, where $\widehat{f}_{t,j}^{SEL} = \psi_{t,j}\omega_{t,j}$. Figure 7 shows that the pricing errors are identically equal to zero at each time period, demonstrating the strength of the SEL method. The conditional pricing error at date-t implied by the CRRA kernel, on the other hand is given by $\sum_{j=1}^{T} \omega_{t,j}^{SEL} \times (\Delta c_j)^{-\gamma}\, \mathbf{r}_{m,j}^{e}$. The figure shows that the pricing errors are economically large in this case, varying from -7.0% to 6.3%. The CRRA kernel fails to match even the historically observed *average* level of the stock market return, producing a large unconditional pricing error. Not surprisingly, it also generates large conditional pricing errors for the market return.

FIGURE 7. CONDITIONAL PRICING ERRORS, 1929-2015



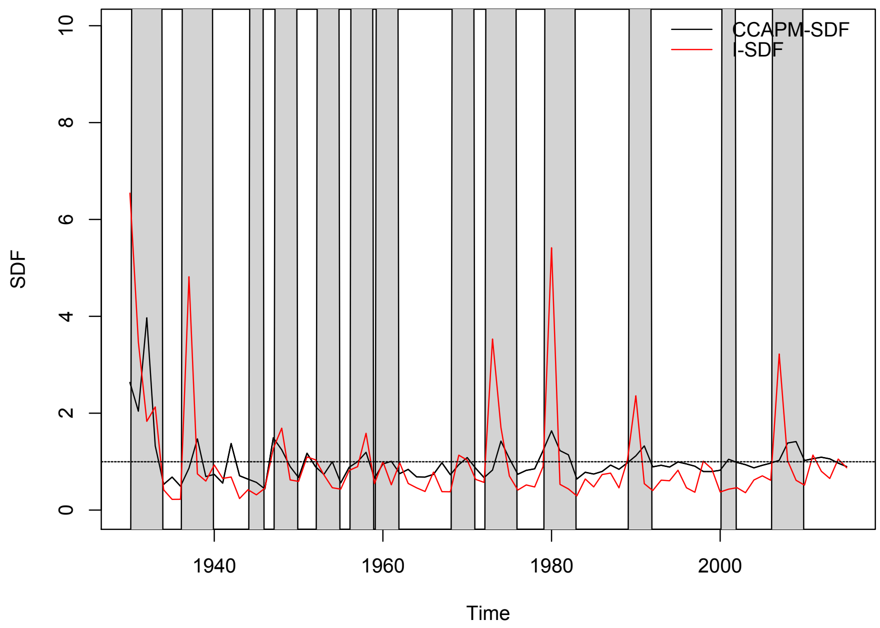
Notes: The figure plots the time series of conditional pricing errors for the excess stock market return, implied by the I-SDF (red line) and the CRRA kernel (green line). The I-SDF is extracted using the SEL method, with nondurables and services consumption as the measure of the consumption expenditures, the excess return on the market portfolio as the sole test asset, and an exponentially-weighted moving average of past consumption growth as the conditioning variable. The sample is annual, covering the period 1930-2015.

Overall, the I-SDF seems to more effectively capture the relevant sources of priced risk and is, therefore, likely to provide more reliable estimates of the welfare costs of aggregate fluctuations.

A second important feature of the I-SDF is that it has a strong business cycle component. Figure 8 plots the time series of the I-SDF (red line) and the CRRA kernel (black line). The more pronounced business cycle component of the I-SDF relative to the CRRA kernel is immediately apparent. The I-SDF is typically substantially higher than the CRRA kernel during recessionary episodes and lower than the former during the expansionary phase of the business cycle. This suggests that business cycle risk is an important source of priced risk, helping interpret our finding that the cost of business cycle fluctuations in consumption constitutes a substantial proportion of the cost of all consumption fluctuations.

Notes: The figure plots the time series of the I-SDF (red line) and the CRRA kernel (black line). The I-SDF is extracted using the EL approach, with nondurables and services consumption as the measure of

FIGURE 8. TIME SERIES OF THE SDF, 1929-2015



the consumption expenditures and the excess return on the market portfolio as the sole test asset. The sample is annual, covering the period 1930-2015.

X. Conclusion

We propose a novel approach to measure the welfare costs of aggregate economic fluctuations. Our methodology does not require specific assumptions regarding either the preferences of consumers or the dynamics of the data generating process. Instead, using data on consumption growth and returns on a chosen set of assets, we rely on an information-theoretic (or relative entropy minimization) approach to estimate the pricing kernel. We refer to the resulting kernel as the *information kernel* or the I-SDF because of the information-theoretic approach used in its recovery. Unlike the CRRA kernel or Lucas' original specification that imposes the additional assumption of lognormality of consumption growth on the CRRA model, the I-SDF accurately prices a broad set of assets, thereby successfully capturing the relevant sources of systematic risk in the economy. Using the I-SDF, we show that the welfare benefits from the elimination of all consumption uncertainty are very large – typically, an order of magnitude bigger than those implied by Lucas' specification. Moreover, contrary to existing literature, the costs of business cycle fluctuations in consumption constitute a substantial fraction –

typically between one-fourth to one-third – of the costs of all consumption uncertainty. These results are robust to the precise measure of consumption used, the cross-section of assets used to extract the I-SDF, and the choice of sample period. Finally, using an extension of our methodology, we present evidence that the welfare benefits of aggregate consumption fluctuations are strongly time-varying and countercyclical.

The difference in the results from earlier literature can be attributed, at least in part, to two factors. First, the I-SDF correctly prices broad cross sections of assets, and thereby identifies the relevant sources of priced risk more effectively than existing approaches. Second, the I-SDF has a strong business cycle component, suggesting that business cycle risk is an important source of priced risk.

REFERENCES

Almeida, Caio, and Ren Garcia. 2012. "Assessing misspecified asset pricing models with empirical likelihood estimators." *Journal of Econometrics*, 170(2): 519–537.

Alvarez, Fernando, and Urban J. Jermann. 2004. "Using Asset Prices to Measure the Cost of Business Cycles." *Journal of Political Economy*, 112(6): pp. 1223–1256.

Alvarez, Fernando, and Urban J. Jermann. 2005. "Using Asset Prices to Measure the Persistence of the Marginal Utility of Wealth." *Econometrica*, 73(6): 1977–2016.

Atkeson, Andrew, and Christopher Phelan. 1994. "Reconsidering the costs of business cycles with incomplete markets." *NBER Macroeconomics Annual*, 187–207.

Backus, David, Mikhail Chernov, and Stanley E. Zin. 2013. "Sources of Entropy in Representative Agent Models." *Journal of Finance.* forthcoming.

Bansal, Ravi, and Amir Yaron. 2004. "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles." *Journal of Finance*, 59(4): 1481–1509.

Barlevy, Gadi. 2005. "The cost of business cycles and the benefits of stabilization." *Economic Perspectives, Federal Reserve Bank of Chicago*, Q(I): pp. 32–49.

Basak, Suleyman, and Hongjun Yan. 2010. "Equilibrium Asset Prices and Investor Behaviour in the Presence of Money Illusion." *The Review of Economic Studies*, 77(3): pp. 914–936.

Beaudry, Paul, and Carmen Pages. 2001. "The cost of business cycles and the stabilization value of unemployment insurance." *European Economic Review*, 45: 1545–1572.

Borovicka, Jaroslav, Lars P. Hansen, and Jose A. Scheinkman. 2016. "Misspecified Recovery." *Journal of Finance*, 71(6): 2493–2544.

Brown, Donald E., and Robert L. Smith. 1990. "A Correspondence Principle for Relative Entropy Minimization." *Naval Research Logistics*, 37(2).

Campbell, John Y., and John H. Cochrane. 1999. "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior." *Journal of Political Economy*, 107(2): 205–51.

Constantinides, George M., and Anisha Ghosh. 2017. "Asset Pricing Tests with Long Run Risks in Consumption Growth." *Journal of Finance*, 72(1): 415–460.

Constantinides, George M., and Darrell Duffie. 1996. "Asset Pricing with Heterogeneous Consumers." *Journal of Political Economy*, 104(2): 219–40.

Csiszár, Imre. 1975. "I-Divergence Geometry of Probability Distributions and Minimization Problems." *Annals of Probability*, 3: 146–158.

Dolmas, Jim. 1998. "Risk preferences and the welfare cost of business cycles." *Review of Economic Dynamics*, 1: 646–676.

Ghosh, Anisha, and Guillaume Roussellet. 2019. "Identifying Beliefs from Asset Prices." Working Paper, McGill University.

Ghosh, Anisha, Christian Julliard, and Alex Taylor. 2016*a*. "An Information-Based One-Factor Asset Pricing Model." mimeo.

Ghosh, Anisha, Christian Julliard, and Alex Taylor. 2016*b*. "What is the Consumption-CAPM missing? An information-Theoretic Framework for the Analysis of Asset Pricing Models." *Review of Financial Studies*, forthcoming.

Hansen, Lars P., and Ravi Jagannathan. 1997. "Asssessing Specification Errors in Stochastic Discount Factor Models." *The Journal of Finance*, 52: 557–590.

Hansen, Lars Peter. 2014. "Nobel Lecture: Uncertainty Outside and Inside Economic Models." *Journal of Political Economy*, 122(5): 945 – 987.

Hansen, Lars Peter, and Kenneth J. Singleton. 1983. "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns." *Journal of Political Economy*, 91: 249–68.

Hansen, Lars Peter, and Thomas J. Sargent. 2010. "Fragile Beliefs and the Price of Uncertainty." *Quantitative Economics*, 1(1): 129–162.

Imrohoroglu, Ayse. 1989. "Cost of business cycles with indivisibilities and liquidity constraints." *Journal of Political Economy*, 97(6): 1364–1383.

Jaynes, E. T. 1957*a*. "Information Theory and Statistical Mechanics." *The Physics Review*, 106(4): 620–630.

Jaynes, E. T. 1957*b*. "Information theory and statistical mechanics II." *The Physical Review*, 108(2): 171–190.

Julliard, Christian, and Anisha Ghosh. 2012. "Can Rare Events Explain the Equity Premium Puzzle?" *Review of Financial Studies*, 25(10): 3037–3076.

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

Kitamura, Yuichi. 2006. "Empirical Likelihood Methods in Econometrics: Theory and Practice." Cowles Foundation, Yale University Cowles Foundation Discussion Papers 1569.

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

Krebs, Tom. 2007. "Job Displacement Risk and the Cost of Business Cycles." *AER*, 97: 664–686.

Krusell, Per, and Anthony A. Smith. 2009. "Revisiting the welfare effects of eliminating business cycles." *Review of Economic Dynamics*, 12(3): 393–404.

Lettau, Martin, and Sydney Ludvigson. 2001. "Consumption, Aggregate Wealth, and Expected Stock Returns." *Journal of Finance*, 56(3): 815–49.

Lucas, Robert E. 1987. *Models of Business Cycles.* Oxford:Blackwell.

Lustig, Hanno N., and Stijn G. Van Nieuwerburgh. 2005. "Housing Collateral, Consumption Insurance, and Risk Premia: An Empirical Perspective." *The Journal of Finance*, 60(3): pp. 1167–1219.

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

Menzly, Lior, Tano Santos, and Pietro Veronesi. 2004. "Understanding Predictability." *Journal of Political Economy*, 112(1): 1–47.

Obstfeld, Maurice. 1994. "Evaluating risky consumption paths: The role of intertemporal substitutability." *European Economic Review*, 38(7): 1471–1486.

Otrok, Christopher. 2001. "On measuring the welfare cost of business cycles." *Journal of Monetary Economics*, 47(1): 61–92.

Owen, Art B. 2001. *Empirical Likelihood.* Chapman and Hall.

Parker, Jonathan A., and Christian Julliard. 2005. "Consumption Risk and the Cross-Section of Expected Returns." *Journal of Political Economy*, 113(1).

Pemberton, James. 1996. "Growth trends, cyclical fluctuations, and welfare with non-expected utility preferences." *Economics Letters*, 50(3): 387–392.

Piazzesi, Monika, Martin Schneider, and Selale Tuzel. 2007. "Housing, consumption and asset pricing." *Journal of Financial Economics*, 83(3): 531–569.

Ravn, Morten O., and Harald Uhlig. 2002. "On adjusting the Hodrick-Prescott filter for the frequency of observations." *The Review of Economics and Statistics*, 84(2): 371–375.

Sandulescu, Mirela, Fabio Trojani, and Andrea Vedolin. 2018. "Model-Free International Stochastic Discount Factors." Working Paper, Boston University.

Schennach, Susanne M. 2005. "Bayesian Exponentially Tilted Empirical Likelihood." *Biometrika*, 92(1): 31–46.

Storesletten, Kjetil, Chris I. Telmer, and Amir Yaron. 2001. "The welfare cost of business cycles revisited: Finite lives and cyclical variation in idiosyncratic risk." *European Economic Review*, 45(7): 1311–1339.

Stutzer, Michael. 1995. "A Bayesian approach to diagnosis of asset pricing models." *Journal of Econometrics*, 68(2): 367 – 397.

Stutzer, Michael. 1996. "A simple Nonparametric Approach to Derivative Security Valuation." *Journal of Finance*, LI(5): 1633–1652.

Tallarini, Thomas. 2000. "Risk sensitive real business cycles." *Journal of Monetary Economics*, 45(3): 507–532.

Yogo, Motohiro. 2006. "A Consumption-Based Explanation of Expected Stock Returns." *Journal of Finance*, 61(2): 539–580.