The Market Cost of Business Cycle Fluctuations *

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

Draft: August 8, 2019

We propose a novel approach to measure the costs of aggregate economic fluctuations, that does not require specification of preferences or the data generating process. Using data on consumption and asset prices, we use an information-theoretic approach to recover an *information kernel* (I-SDF). The I-SDF accurately prices broad cross-sections of assets 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 strongly countercyclical. Moreover, the cost of business cycle fluctuations is substantial, accounting for a quarter to a third of the cost of all 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.

* We benefited from helpful comments from Jaroslav Borovicka, George Constantinides, Janice Eberly, Deborah Lucas, Sydney Ludvigson, and seminar participants at the NBER's 2019 Summer Institute Asset Pricing Workshop and the Capital Markets and the Economy (EFEL) Workshop. All errors and omissions are our own.

† 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).

And, an extensive literature has documented how Lucas' specification grossly underestimates the 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:

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

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

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

broad categories of assets. Anisha Ghosh, Christian Julliard and Alex Taylor (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 asset returns and consumption data to extract, non-parametrically, the *minimum relative 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 the 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

[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.

I-SDF implies a larger cost of business cycle fluctuations than those obtained with 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 a sure 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, in our baseline 1929–2015 sample, 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 increase from 1.5% at the one-year horizon to 14.4% for a five-year time period. The corresponding costs obtained with Lucas' specification are typically an order of magnitude smaller at 0.8% 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 expenditures on durables) and the set of assets used to recover the I-SDF. Our results suggest that economic agents perceive the cost of aggregate fluctuations to be substantial.

We next use our framework to estimate the cost 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, 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, in our baseline case, the cost of all fluctuations over a five-year horizon is estimated at 14.4% using the I-SDF, while the corresponding cost of business cycle fluctuations is 3.6%. When total (instead of nondurables and services) consumption expenditures is used to recover the I-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.

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 EL 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 investors' preferences or the dynamics of the data generating process. We find that the cost of consumption fluctuations is strongly time-varying and countercyclical. For instance, in our baseline case, the cost of all one-year fluctuations varies from 0.15% to 8.0%. Also, the cost is strongly countercyclical, rising sharply during recessionary episodes.

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), Ian Martin (2008), 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, unlike these authors, 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. Moreover, Alvarez and Jermann (2004) focus on an infinite time horizon, which makes their estimates very sensitive to calibrations of the real growth rate as well as the discount rates for the infinite-horizon sure and risky consumption claims (their estimates of the cost of all consumption fluctuations vary from 28.0%–1535.7%). Our approach, on the other hand, offers a term structure of the cost of fluctuations, i.e. how the welfare benefits rise with the elimination of fluctuations over each additional future period. This makes our results less sensitive to the choice of discount rates.

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 ability of the information-theoretic methodology to recover the underlying pricing kernel accurately. Section IV contains a description of the data used. Section V reports the empirical results. In particular, the welfare gains from eliminating all consumption fluctuations and fluctuations corresponding to business cycle frequencies are presented in Sections V.A and V.B, respectively. Section V.C presents a host of robustness checks. Section VI 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. Section VII discusses the main factors driving our results. Finally, Section VIII concludes with suggestions for future research.

II. Pricing Aggregate Economic Fluctuations

This section defines the welfare 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 aggregate consumption fluctuations, for two alternative definitions of 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 eliminating 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)).

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. Under fairly general conditions, 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 it from which certain types of fluctuations (e.g., all fluctuations or business cycle fluctuations) have been removed.

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.

Alvarez and Jermann (2004) show that the marginal cost of *all* consumption fluctuations, i.e. the scenario where $C_t^{stab} = E_0(C_t) = (1+\mu_c)^t C_0$ for $t = 1, 2, ..., \infty$, where μ_c denotes the unconditional mean of consumption growth, is given by:

$$\omega_0 = \frac{r_0 - \mu_c}{y_0 - \mu_c} - 1. \tag{3}$$

In the above equation, y_0 and r_0 denote the yields to maturity on claims to the stabilized sure consumption stream and the risky consumption stream, respectively. Calibrating $\mu_c = 2.3\%$, $y_0 = 3.0\%$ and $r_0 - y_0 \geq 0.2\%$, they obtain a very high estimate of the cost of at least 28.6%. However, the above equation highlights that the estimate of the cost is very sensitive to the values of y_0, r_0, and g. Specifically, as $y_0 \to \mu_c$, we have $\omega_0 \to \infty$, and the approach breaks down. Moreover, Olivier J. Blanchard (2019) points out that, at the current time, the nominal rate on a 10-year government bond is 2.7%, while the expected nominal growth rate is 4.0%, causing $y_0 - \mu_c$ to be negative, thereby negating the use of Equation (3). And this is not just a feature of the US, but also other developed economies such as the UK and the Euro Zone. Also, Blanchard (2019) highlights that the current situation is more the norm rather than the exception in the US – the average nominal growth rate and the rate on 1-year government bonds have been 6.3% and 4.7%, respectively, since 1950, and 5.3% and 4.6%, respectively, since 1870, and, in fact, $y_0 - \mu_c$ has been negative in all decades except the 1980s. This reveals the fragility of the results obtained using Equation (3).

Therefore, in this paper, instead of attempting to measure the welfare cost of eliminating consumption fluctuations over an infinite time horizon, we focus on the term structure of finite horizon consumption risk. In other words, we characterize the welfare gains from stabilizing the next $j = 1, ..., J$ periods of consumption uncertainty. This makes our results more robust to the choice of discount rates.

The absence of arbitrage opportunities implies that

$$V_0\left[\{C_t\}_{t=1}^{j}\right] = \sum_{t=1}^{j} V_0\left(C_t\right), \tag{4}$$

for $j \geq 1$, 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, $V_0\left[\left\{C_t^{stab}\right\}_{t=1}^{j}\right]$ can be written as the sum, over time periods 1-j, of the prices of claims to single payoffs equal to the stabilized consumption in each of these future periods. Therefore, the cost of one-period fluctuations is given by:

$$\frac{V_0\left(C_1^{stab}\right)}{V_0\left(C_1\right)} - 1. \tag{5}$$

The (cumulative) cost of two-period fluctuations is given by

$$\frac{V_0\left(C_1^{stab}\right) + V_0\left(C_2^{stab}\right)}{V_0\left(C_1\right) + V_0\left(C_2\right)} - 1. \tag{6}$$

And so on, for any number j of future periods..

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 Equations (5)-(6)) is directly traded in financial markets, their prices are not directly observed. Therefore, the values of these claims need to be 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 proven to be quite sensitive to these two assumptions. 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 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

$$V_t\left(C_{t+1}\right) = \mathbb{E}_t\left[M_{t+1}C_{t+1}\right], \tag{7}$$

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 the 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 (7) by C_t to make both sides stationary, taking unconditional expectations, and using the definition of $\mathbf{z}$, we have

$$\tilde{pc}_1 := \mathbb{E}\left[\frac{V_t\left(C_{t+1}\right)}{C_t}\right] = \int_{\mathbf{z}} M(\mathbf{z})\Delta C(\mathbf{z})d\mathbb{P}(\mathbf{z}), \tag{8}$$

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

$$V_t\left(C_{t+1}^{stab}\right) = \mathbb{E}_t\left[M_{t+1}C_{t+1}^{stab}\right], \tag{9}$$

implying that

$$\tilde{pc}_1^{stab} := \mathbb{E}\left[\frac{V_t\left(C_{t+1}^{stab}\right)}{C_t}\right] = \int_{\mathbf{z}} M(\mathbf{z})\Delta C^{stab}(\mathbf{z})d\mathbb{P}(\mathbf{z}), \tag{10}$$

where $\Delta C_{t+1}^{stab} = \frac{C_{t+1}^{stab}}{C_t}$. In the scenario where we want to obtain the welfare benefit of eliminating *all* consumption uncertainty in the next period, we set $C_{t+1}^{stab} = (1+\mu_c)\,C_t$. Therefore, in this case, $\Delta C_{t+1}^{stab} = (1+\mu_c)$. On the other hand, to assess the cost of 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 determined, the cost of one-period consumption fluctuations is then given by

$$\frac{\tilde{pc}_1^{stab}}{\tilde{pc}_1} - 1. \tag{11}$$

If a history of $M(\mathbf{z}_t) \equiv M_t$, $t = 1, ..., T$, were observable, we could estimate the prices in Equations (8) and (10) and, therefore, the cost of one-period consumption fluctuations in Equation (11): 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.

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 the 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

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

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, in practice, 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 (8) can be rewritten as

$$p\tilde{c}_1 = \mathbb{E}\left[M(\mathbf{z})\right] \int \Delta C(\mathbf{z}) q(\mathbf{z}) d\mathbf{z}. \tag{12}$$

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:

$$\min_{\mathbb{Q}} \int \mathbf{log}\left(\frac{d\mathbb{P}}{d\mathbb{Q}}\right) d\mathbb{P} = \int \mathbf{log}\left(\frac{p(\mathbf{z})}{q(\mathbf{z})}\right) p(\mathbf{z}) d\mathbf{z} \text{ s.t. } \mathbf{0} = \int \mathbf{R}^e(\mathbf{z}) q(\mathbf{z}) d\mathbf{z}. \tag{13}$$

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

$$M_{t+1} = \left(\Delta C_{t+1}\right)^{-\gamma} \psi_{t+1}, \tag{14}$$

leads to the reformulation of Equation (13) as:

$$\min_{\mathbb{F}} \int \mathbf{log}\left(\frac{d\mathbb{P}}{d\mathbb{F}}\right) d\mathbb{P} = \int \mathbf{log}\left(\frac{p(\mathbf{z})}{f(\mathbf{z})}\right) p(\mathbf{z}) d\mathbf{z} \text{ s.t. } \mathbf{0} = \int \mathbf{R}^e(\mathbf{z}) \left(\Delta C(\mathbf{z})\right)^{-\gamma} f(\mathbf{z}) d\mathbf{z}, \tag{15}$$

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 (15), the pricing kernel, M, can be obtained using Equation (14). 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 (15) 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 minimization problem in Equation (15), after dropping redundant terms, can be rewritten as

$$\max_{\psi} \mathbb{E}^{\mathbb{P}}\left[\ln \psi(\mathbf{z})\right] \text{ s.t. } \mathbf{0} = \int \mathbf{R}^{e}(\mathbf{z})\left(\Delta C(\mathbf{z})\right)^{-\gamma} \psi(\mathbf{z}) p(\mathbf{z}) d\mathbf{z}. \tag{16}$$

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 (15), 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 relative entropy is measured in units of information. To provide some intuition, suppose that the consumption growth component of the pricing kernel, $(\Delta C_t)^{-\gamma}$, were sufficient to price assets perfectly. Then $\psi_t \equiv 1, \forall t$, and we have that $\mathbb{F} \equiv \mathbb{P}$, delivering a KLIC divergence $\int \mathbf{log}\left(\frac{d\mathbb{P}}{d\mathbb{F}}\right) d\mathbb{P} = 0$ in Equation (15). 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 $\mathbb{P}$, i.e. the KLIC divergence is positive: $\int \mathbf{log}\left(\frac{d\mathbb{P}}{d\mathbb{F}}\right) d\mathbb{P} > 0$. 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, 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.[4]

Fifth, this approach is numerically simple to implement. Given a history of excess returns and consumption growth $\{\mathbf{r}_t^e, \Delta c_t\}_{t=1}^T$, Equation (16) can be made operational by replacing the expectation with a sample analogue, as is customary for moment based estimators:[5]

$$\arg\max_{\{\psi_t\}_{t=1}^T} \frac{1}{T}\sum_{t=1}^{T} \ln \psi_t \text{ s.t. } \frac{1}{T}\sum_{t=1}^{T} \left(\Delta c_t\right)^{-\gamma} \psi_t \mathbf{r}_t^e = \mathbf{0}. \tag{17}$$

An application of Fenchel's duality theorem to the above problem (see, e.g., Imre

[4] 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.

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

Csiszár (1975), Owen (2001)), delivers the estimates (up to a positive constant scale factor):

$$\hat{\psi}_t = \frac{1}{T(1 + \hat{\theta}(\gamma)'\mathbf{r}_t^e (\Delta c_t)^{-\gamma})} \quad \forall t, \tag{18}$$

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

$$\hat{\theta}(\gamma) = \arg\min_{\theta} - \sum_{t=1}^{T} \log(1 + \theta'\mathbf{r}_t^e (\Delta c_t)^{-\gamma}). \tag{19}$$

Sixth, 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 (15)). 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 EL methodology is quite successful in recovering the ψ-component of the pricing kernel for empirically realistic sample sizes. Details of the simulation design and the performance of the estimator are presented in Section III.

With the recovered ψ-component, the I-SDF is obtained (up to a positive scale factor) as

$$\widehat{M_t} = \kappa\,(\Delta c_t)^{-\gamma}\,\widehat{\psi}_t. \tag{20}$$

The proportionality constant, κ, can be recovered from the Euler equation for the risk free rate. Equation (20) makes clear that our estimator of the ψ-component, as any Generalized Empirical Likelihood approach (see e.g. Kitamura (2006) for a survey), approximates the true unknown ψ distribution with a multinomial with support points given by the sample realizations of the observable variables (in this case, consumption growth and asset returns).

Armed with the I-SDF, we can now estimate the welfare benefits of eliminating consumption fluctuations. Specifically, the value of eliminating *all* consumption fluctuations in the next period is obtained as:

$$\widehat{pc_1^{stab}/p\tilde{c}_1} - 1 = \frac{\sum_{t=1}^{T} \widehat{M_t}\,(1+\mu_c)}{\sum_{t=1}^{T} \widehat{M_t}\Delta c_t} - 1. \tag{21}$$

Similarly, the value of eliminating business cycle fluctuations in the next period is:

$$\widehat{pc_1^{stab}/\tilde{pc}_1} - 1 = \frac{\sum_{t=1}^{T} \widehat{M}_t \Delta c_t^{stab}}{\sum_{t=1}^{T} \widehat{M}_t \Delta c_t} - 1, \tag{22}$$

where Δc_t^{stab} denotes a time-varying stabilized consumption growth from which the business cycle variations have 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.

We will soon see that estimates obtained by using the I-SDF differ markedly from estimates obtained by Lucas' method (summarized on p.10 herein). To help explain this, recall that Lucas' method presumes a complete markets exchange economy in which the (unique) SDF M_t is the normalized marginal rate of substitution (MRS) from the discounted power utility functional. The MRS depends only on consumption growth and model parameters. Under the complete markets assumption, all assets must satisfy the pricing condition $E[M_t\mathbf{R_t^e}] = \mathbf{0}$, including the risk free asset. Yet the Equity Premium Puzzle and the variance and entropy bounds literatures cited herein all establish that the excess returns of popular equity indices will *not* satisfy these constraints when the Lucas SDF is specified with economically plausible parameters. In light of this, subsequent work proposed other consumption-based asset pricing models, but Ghosh, Julliard and Taylor (2016*b*) show that these are similarly problematic when the excess returns of Fama-French equity factor portfolios are included in $\mathbf{R^e}$.

In contrast, the I-SDF satisfies these pricing constraints *by construction* while still including consumption growth in its makeup. This provides a method of pricing consumption fluctuations in a way that is consistent with the pricing of equity portfolios, albeit without the theoretical desideratum of first specifying an exchange or other economic model from which it was derived. Theorists who maintain the complete markets assumption can view our approach as a data-driven procedure to estimate the unknown unique SDF, with the aforementioned desirable properties.

Finally, note that Equations (21) and (22) 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 j periods into the future can be expressed as

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

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

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

The one-period I-SDF, recovered in Equation (20), can be compounded to recover the j-period discount factor:

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

Using $M_{t:t+j}$, we can estimate the price-consumption ratio $\tilde{pc}_j$ for a single consumption claim j periods in the future. And this can be done for any $j = 2, 3, 4, ...$ 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 (21)-(22).

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 exchange 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 successfully recovers 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:

$$\log\left(\Delta C_t\right) \overset{\mathbb{P}}{\sim} \mathcal{N}\left(\mu_c, \sigma_c^2\right). \tag{23}$$

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

$$\log\left(\Delta C_t\right) \overset{\widetilde{\mathbb{P}}}{\sim} \mathcal{N}\left(\widetilde{\mu}_c, \sigma_c^2\right), \tag{24}$$

where $\widetilde{\mu}_c = (1-\lambda)\mu_c$ and $\widetilde{\mathbb{P}}$ denotes the investor's subjective measure. We assume that there are no distortions in the beliefs about the volatility or the higher moments of consumption growth.

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

$$0 = \mathbb{E}^{\widetilde{\mathbb{P}}}\left[\left(\Delta c_{t+1}\right)^{-\gamma}\left(R_{m,t+1} - R_{f,t+1}\right)\right], \tag{25}$$

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 (25), the expectation is evaluated under the investor's subjective measure $\widetilde{\mathbb{P}}$ (instead of the physical measure $\mathbb{P}$). Under weak regularity conditions, Equation (25) may be rewritten as

$$0 = \mathbb{E}^{\mathbb{P}} \left[(\Delta c_{t+1})^{-\gamma} \psi_{t+1} \left(R_{m,t+1} - R_{f,t} \right) \right], \tag{26}$$

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:

$$\left\{ \widehat{\psi}_t \right\}_{t=1}^{T} = \underset{\{\psi_t\}_{t=1}^{T}}{\arg\max} \sum_{t=1}^{T} \log(\psi_t) \quad \text{s.t.} \quad \frac{1}{T} \sum_{t=1}^{T} (\Delta c_{t+1})^{-\gamma} \psi_{t+1} (R_{m,t+1} - R_{f,t+1}) = 0. \tag{27}$$

Using the recovered ψ and approximating the physical measure 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 (27), 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 $\frac{P_t}{D_t} = \nu$, a constant, where

$$\nu = \frac{\exp\left[\log(\delta) + (1-\gamma)\widetilde{\mu}_c + \dfrac{(1-\gamma)^2 \sigma_c^2}{2} \right]}{1 - \exp\left[\log(\delta) + (1-\gamma)\widetilde{\mu}_c + \dfrac{(1-\gamma)^2 \sigma_c^2}{2} \right]}, \tag{28}$$

and the equilibrium risk free rate is also constant at:

$$R_f = \frac{1}{\exp\left(\log(\delta) - \gamma \tilde{\mu}_c + \dfrac{\gamma^2 \sigma_c^2}{2} \right)}. \tag{29}$$

To perform our simulation exercise, we calibrate μ_c and σ_c^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 (23). Using the simulated consumption growth, we obtain the market return as:

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

where ν is defined in Equation (28). The time series of the risk free rate is simply a constant, given by Equation (29).

Using the above time series, we recover the subjective beliefs using the EL approach in Equation (27). 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 $1,000$ simulated samples. We report the averages 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 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. Row 1 shows that the EL method is successful at capturing these subjective beliefs. Specifically, the EL-implied mean consumption growth is on average 2.61% across the $1,000$ simulations, close to the true value of the mean under the subjective beliefs. The EL implied volatility of consumption growth has an average of 3.47% across the 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 under the physical measure. 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 consumption growth – the EL method accurately 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 the 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 the historical sample size in terms of the average mean, volatility, and skewness across the simulations, although the confidence bands are tighter for bigger sample sizes.

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.

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	
$\tilde{\mu} = 0.90\mu$	2.59 [2.34,2.87]	3.44 [3.05,3.82]	$-.016$ [$-.44$,.41]
$\tilde{\mu} = 0.85\mu$	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 $1,000$ 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.

IV. Data Description

The extraction of the 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 to mitigate concerns that certain possible states may not have been realized in the sample. 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 $S\&P$ 500 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 EL procedure outlined in Section II, to obtain the cost of aggregate consumption fluctuations, i.e. the welfare benefits of eliminating *all* consumption uncertainty as well as removing only business cycle fluctuations in consumption.

Before presenting the empirical results, we turn to a discussion of the SDF parameter γ that enters the welfare cost calculations (see, e.g., Equations (21)–(22)). As highlighted in the introduction, the multiplicative decomposition of the SDF, $M_t = \left(\frac{C_t}{C_{t-1}}\right)^{-\gamma}\psi_t$, is motivated by the observation that virtually all structural asset pricing models proposed in the literature imply this form for the SDF. Different models offer different economic interpretations of the ψ-component and the utility curvature parameter γ. For example, in the time and state separable power utility model, $\psi = \delta$, the subjective discount factor, and γ is the CRRA of the representative agent. An upper bound of 10 is generally considered plausible for the CRRA parameter. However, much higher levels of risk aversion are needed for the model to explain several observed features of financial market data. In models with Larry G. Epstein and Stanley E. Zin (1989) recursive preferences, $M_t = \delta^\eta \left(\frac{C_t}{C_{t-1}}\right)^{-\frac{\eta}{\rho}} R_{c,t}^{\eta-1} = \delta^\eta \left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} \left(\frac{\frac{P_{c,t}}{C_t}+1}{\frac{P_{c,t-1}}{C_{t-1}}}\right)^{\eta-1}$, where γ denotes the CRRA, ρ the elasticity of intertemporal substitution, $\eta = \frac{1-\gamma}{1-\frac{1}{\rho}}$, and $R_{c,t} = \frac{P_{c,t}+C_t}{P_{c,t-1}}$ denotes the unobservable return on total wealth. These models typically calibrate $\gamma = 10$ (see, e.g., Bansal and Yaron (2004)). Some models with recursive preferences calibrate γ to much larger values (e.g., Monika Piazzesi and Martin Schneider (2007)). In models with external habit formation (see, e.g., Campbell and Cochrane (1999)), $M_t = \delta \left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} \left(\frac{S_t}{S_{t-1}}\right)^{-\gamma}$, where S_t is the surplus consumption ratio and γ the utility curvature parameter that is a determinant of the time-varying risk aversion $\frac{\gamma}{S_t}$. Campbell and Cochrane (1999) calibrate $\gamma = 2$. However, Ghosh, Julliard and Taylor (2016*b*) show that the model needs a higher γ (typically in excess of 7) to satisfy entropy bounds for admissible SDFs, that are tighter than the seminal variance bounds of Lars Peter Hansen and Ravi Jagannathan (1991). In models with complementarities in consumption, see e.g., Piazzesi, Schneider and Tuzel (2007), $M_t = \delta \left(\frac{C_t}{C_{t-1}}\right)^{-\gamma} \left(\frac{A_t}{A_{t-1}}\right)^{\frac{\gamma\zeta-1}{\zeta-1}}$, where A_t is the expenditure share on non-housing consumption, γ^{-1} is the intertemporal elasticity of substitution, and ζ is the intratemporal elasticity of substitution between housing services and non-housing consumption. The authors' consider two alternative calibrations of $\gamma = 5$ and $\gamma = 16$. However, Ghosh, Julliard and Taylor (2016*b*) show that the model needs a higher γ (typically in excess of 20) to satisfy entropy bounds for admissible SDFs. To summarize, most models in the literature either calibrate the SDF parameter γ to 10 or higher values and/or require such values of the parameter to explain asset prices.

Also, in addition to recovering the ψ-component of the SDF, our information-theoretic EL procedure offers a way to estimate γ. Specifically, the EL estimator of γ is defined as (see Kitamura and Stutzer (1997)):

$$\widehat{\gamma}^{EL} = \max_{\gamma} \max_{\psi} \mathbb{E}^{\mathbb{P}} [\ln \psi(\mathbf{z})] \text{ s.t. } \mathbf{0} = \int \mathbf{R}^e(\mathbf{z}) \left(\Delta C(\mathbf{z})\right)^{-\gamma} \psi(\mathbf{z}) p(\mathbf{z}) d\mathbf{z}. \tag{30}$$

Kitamura and Stutzer (1997) show that the EL estimator is consistent and asymptotically normal, with its asymptotic distribution given by:

$$\sqrt{T}\left(\widehat{\gamma}^{EL} - \gamma_0\right) \xrightarrow{d} N(0, (D' S^{-1} D)^{-1}),$$

where $S = E^{\mathbb{F}}\left[(C_t/C_{t-1})^{-\gamma_0} \mathbf{R}_t^e \mathbf{R}_t^{e'} (C_t/C_{t-1})^{-\gamma_0}\right]$ is the covariance matrix of the sample moment restrictions and $D = E^{\mathbb{F}}\left[\left.\frac{\partial\left\{(C_t/C_{t-1})^{-\gamma}\mathbf{R}_t^e\right\}}{\partial\gamma}\right|_{\gamma=\gamma_0}\right]$ is the derivative of the moments with respect to the γ parameter.

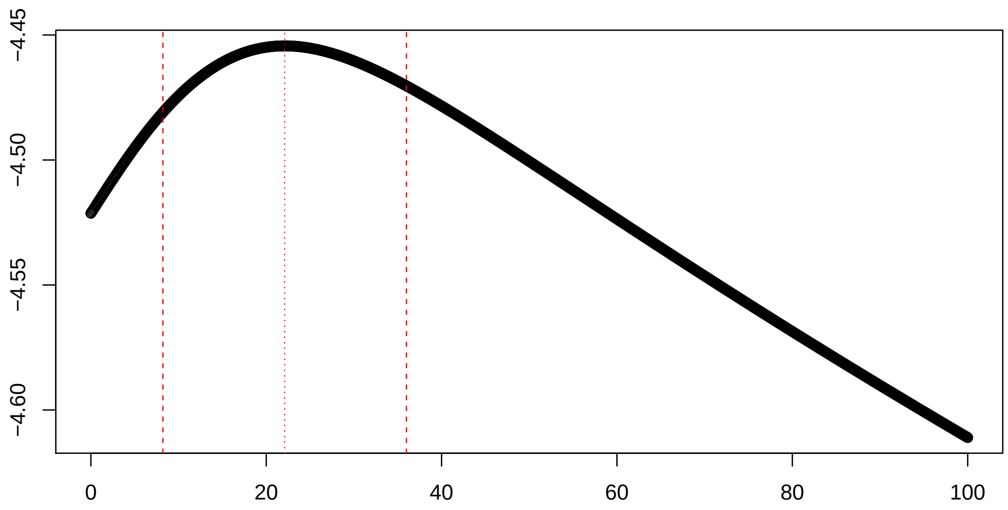
We estimate γ in our baseline 1929–2015 sample, using total consumption expenditures as the measure of aggregate consumption and the excess return on the market as the sole test asset. Figure 1 plots the objective function in Equation

(30) as a function of γ. The point estimate of γ is 22.1 (red dotted line). The estimated asymptotic standard error is 7.09. Thus, the 95% confidence interval covers the range $[8.21, 36.0]$ (red dashed lines). A similar point estimate of 26.2 with a standard error of 8.20 is obtained when nondurables and services consumption is used as the measure of aggregate consumption expenditures. In this case, the 95% confidence interval for γ is $[10.1, 42.3]$.

Motivated by the observations that most theoretical models calibrate γ to 10 or higher values and that the estimated 95% confidence interval for this parameter typically has around 10 as the lower bound, we set $\gamma = 10$ in our baseline results. Note that higher values of γ serve to further increase the marginal utility of the representative agent in bad states of the world with low consumption growth rate and, therefore, would further increase the estimates of the cost of consumption fluctuations. In Section VI.C, we assess the sensitivity of our results to alternative choices of γ.

We next proceed to estimate the cost of fluctuations. Section V.A presents the cost of all consumption fluctuations. Section V.B presents the cost of business cycle fluctuations in consumption. Finally, in Section V.C, we present a host of robustness checks.

Figure 1. Profile Likelihood



Note: The figure plots the EL objective function as a function of the SDF parameter γ. Consumption denotes the real personal total consumption expenditure (includes durables, nondurables, and services). The excess return on the market portfolio is the sole test asset. The sample is annual, covering the period 1929-2015.

A. The Cost of All Consumption Uncertainty

Recall that, rather than estimating the cost of aggregate consumption fluctuations over an infinite time horizon, we focus on the term structure of the cost for finite time periods. Specifically, we estimate the (cumulative) cost for one- to five-year time horizons.

Equation (11) defines the cost of one-period fluctuations, i.e. the welfare benefit of removing fluctuations in the next period alone. The cost is the ratio of the prices of two hypothetical securities: 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, $\tilde{pc}_1$. Equations (21)–(22) reveal that the prices of these two securities and, therefore, the cost of all one-period consumption fluctuations, depend on the SDF. We use the I-SDF, recovered using the EL approach, to measure this cost. The costs of multi-year fluctuations are obtained by compounding the I-SDF, as explained in Section II.B. Note that the recovered I-SDF depends on the particular measure of the aggregate consumption expenditures as well as on the set of assets used (see Equations (18)-(19)). To ensure robustness, we estimate the I-SDF using two different measures of consumption expenditures and two alternative sets of assets.

Table 2: Cumulative Cost of 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: Nondurables & Services Consumption										
I-SDF (Mkt)	1.53	5.15	11.75	14.28	14.44	.556	1.48	3.39	3.90	3.57
I-SDF (FF6)	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	1.68	1.94	-	-	-	-	-
Panel B: Total Consumption										
I-SDF (Mkt)	2.15	6.77	16.13	19.65	19.73	.896	2.09	4.85	5.55	5.12
I-SDF (FF6)	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) costs of *all* aggregate consumption fluctuations (Columns 2-6) and the costs of business cycle fluctuations in consumption (Columns 7-11), over one-to five-year horizons. Panel A presents results when consumption denotes the real personal consumption expenditure of nondurables and services, while Panel B does the same for total personal consumption expenditure (that includes durables). In each panel, the costs are calculated using the I-SDF recovered from the market portfolio alone (Row 1), the I-SDF recovered from the six size and book-to-market-equity sorted portfolios of Fama and French (Row 2), the kernel implied by power utility preferences with a constant CRRA (Row 3), and Lucas' original specification that involves power utility preferences and i.i.d. lognormal aggregate consumption growth dynamics (Row 4). The sample is annual covering the period 1929-2015.

The results are presented in Table 2. Panel A presents results when consumption refers to the expenditure on nondurables and services, while Panel B does the same for total consumption expenditures (including durables). Consider first Panel A. In Row 1, the market portfolio alone is used in the extraction of the I-SDF. Row 1, Column 2 shows that the cost of all one-period consumption fluctuations is estimated to be 1.5%. Row 2, Column 2 shows that, when the six size and book-to-market-equity sorted portfolios of Fama-French are used to recover the I-SDF, the estimated cost of all one-period consumption fluctuations is quite similar at 1.3%. Row 3 shows that the one-year cost, estimated using the pricing kernel implied by power utility preferences with a constant CRRA (hereafter referred to as the CRRA kernel), is an order of magnitude smaller at .93%. And, Row 4 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%.

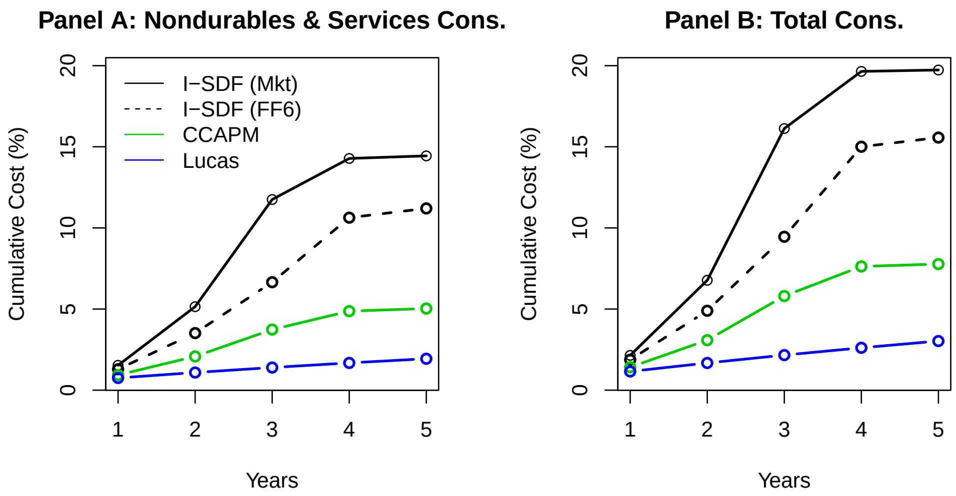
Note that the above results pertain to the cost of fluctuations in one-period consumption alone. Columns 3, 4, 5, and 6 of Panel A present the costs of consumption fluctuations over two, three, four, and five year horizons, respectively. Row 1 shows that, when the market portfolio alone is used to recover the I-SDF, the costs of consumption fluctuations over two, three, four, and five years increase 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).

Row 3 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 4 implies even smaller costs of 1.1%, 1.4%, 1.7%, and 1.9% at two-, three-, four-, and five-year horizons, respectively. Finally, very similar results are obtained when the 6 FF portfolios are used to recover the I-SDF. Row 2 shows that the costs of fluctuations for two-, three-, four-, and five-year periods are substantially higher for the I-SDF compared to the CRRA kernel – 3.5% versus 2.1% for two years, 6.7% versus 3.7% for three years, 10.6% versus 4.9% for four years, and 11.2% versus 5.0% for five years. And the costs are even higher compared to Lucas' specification.

The results in Table 2, Panel A were obtained using personal consumption expenditures on nondurables and services as the measure of consumption. Panel B, that uses the total consumption expenditures (including durables) as the measure of consumption, produces results similar to those in Panel A. Note that, not surprisingly, the costs of fluctuations are bigger with total consumption compared to nondurables and services consumption. Specifically, the I-SDF implies that the costs of fluctuations increase from 2.2% to 19.7% from one-year to five-year horizons, when recovered from the market portfolio alone, and from 1.9% to 15.6% when the 6 FF portfolios are used as test assets. By contrast, the CRRA kernel and Lucas' specification imply much smaller costs that increase from 1.4% to 7.8% and from 1.2% to 3.0%, respectively, from one- to five-year time periods.

Figure 2 plots the term structure of the costs of fluctuations over one- to five-year horizons. Panel A presents the results for nondurables and services consumption, while Panel B focuses on total consumption expenditures. In each panel, the black solid line corresponds to the costs obtained with the I-SDF recovered with the market portfolio as the sole test asset. The black-dashed line, on the other hand, denotes the costs implied by the I-SDF extracted from the 6 FF portfolios. The green and blue lines denote the costs estimated with the CRRA kernel and Lucas' specification, respectively. The figure highlights the higher costs implied by the I-SDF relative to those obtained with Lucas' specification of preferences and the dynamics of the consumption growth rate.

Figure 2. Marginal Cost of All Consumption Fluctuations, 1929-2015



Notes: The figure plots the cumulative costs of *all* aggregate consumption fluctuations over one- to five-year horizons, for different choices of the pricing kernel and measures of consumption. Panel A presents results when consumption refers to the real personal consumption expenditure of nondurables and services, while Panel B does the same when consumption denotes the 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 I-SDF extracted using the excess returns on the 6 FF portfolios as test assets (black dashed line), the pricing kernel implied by power utility preferences with a constant CRRA (green line), and Lucas' original specification that involves power utility preferences and i.i.d. lognormal aggregate consumption growth dynamics (blue line).

Overall, the results of this section suggest that economic agents perceive the cost of aggregate economic uncertainty to be quite substantial. For instance, our estimates of the cost of all consumption fluctuations over a five-year horizon 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. The cost is substantially higher than that originally obtained by Lucas. Note that costs higher than Lucas' estimates have been reported in the literature – Alvarez and Jermann (2004) report very high costs of all consumption fluctuations over an infinite time horizon. However, as argued in Section II.A, the focus on an infinite time horizon makes their estimates very sensitive to calibrations of the real growth rate as well as the discount rates for the infinite-horizon sure and risky consumption claims. As an illustration of this sensitivity, their estimate of the cost varies from 28.0%–1535.7% based on different calibrations. Our approach, on the other hand, 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. This makes our results less sensitive to the choice of discount rates. Also, an attractive feature of our method is that it seems to have well-defined asymptotics – the cumulative welfare costs seem to stabilize with the increase in the number of time periods (see Figure 2). These aspects of the methodology make our quantitative estimates more reliable.

B. Business Cycle vs. Long Run Uncertainty

While Section V.A 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 stream and a claim to the actual aggregate consumption. 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).

The results are presented in the last five columns of Table 2, Panel A for nondurables and services consumption. 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. Similar results are obtained in Row 2 when the six size and book-to-market-equity sorted portfolios are used in the recovery of the I-SDF – the costs of business cycle fluctuations increase from 0.5% at the one-year horizon to 2.9% for a five-year time period.

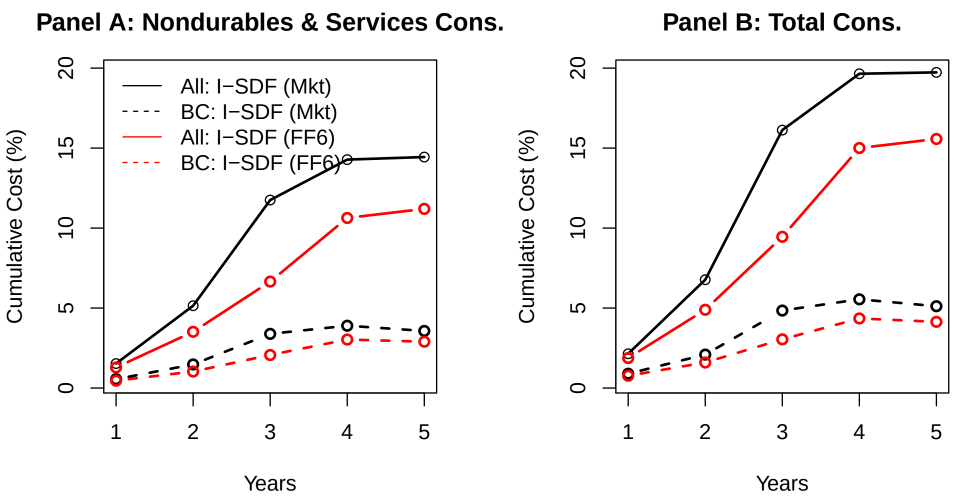
Row 3 shows that, for the CRRA kernel, while the cost of business cycle fluctuations over a one-year period is similar to that obtained with the I-SDF (0.5% versus 0.5%–0.6%), the cost increases little for multi-year horizons in the case of the former. For instance, the cost of five-year fluctuations is only 1.4% – less than half of the costs of 3.7% and 2.9% implied by the I-SDF in Rows 1 and 2, respectively.

An important point to note is that while the estimates of the costs of business cycle fluctuations are smaller than the costs 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. Similarly, when the 6 FF portfolios are used for the recovery of the I-SDF, the costs of business cycle fluctuations over one to five years account for 35.8%, 29.3%, 31.1%, 31.0%, and 25.9%, respectively, of the cost of all consumption fluctuations over these time horizons.

Figure 3, Panel A plots the term structure of the cost of all consumption fluctuations (solid line) and business cycle fluctuations in consumption (dashed line) over 1-5 years. The black lines present the estimates obtained when the market portfolio alone is used as the test asset to recover the I-SDF. The red lines, on the other hand, are based on the estimates obtained when the I-SDF is recovered from the 6 FF portfolios. The fairly large ratio of the cost of business cycle fluctuations to the cost of all consumption fluctuations, at all time horizons, is evident from the figure. Moreover, as with the cost of all fluctuations, the cost of business cycles seems to stabilize with increase in the time horizon, thereby suggesting well-defined asymptotics.

Finally, the results remain largely unchanged when total consumption expenditure (instead of nondurables and services expenditure) is used as the measure of consumption. These are presented in Table 2, Panel B and Figure 3, Panel B.

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

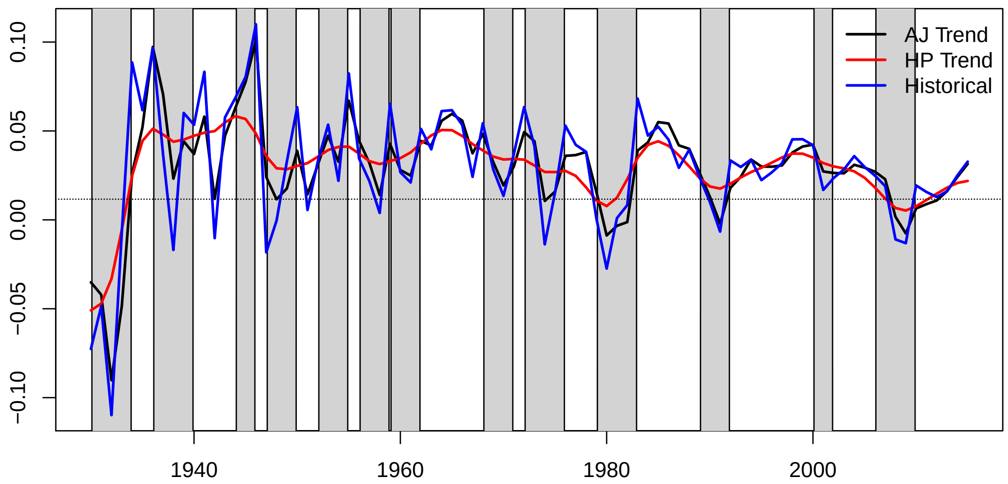


Notes: The figure plots the term structure of the (cumulative) cost of all aggregate consumption fluctuations (solid line) and business cycle fluctuations in consumption (dashed line), over 1-5 *years, obtained using the I-SDF. Panel A presents results when consumption refers to the real personal consumption expenditure of nondurables and services, while Panel B does the same when consumption denotes total personal consumption expenditure. The I-SDF is extracted using the excess return on the market portfolio as the sole test asset (black lines) and the 6 FF portfolios (red lines). The sample is annual covering the period* 1929-2015.

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, the cost of business cycle fluctuations in consumption is miniscule, varying from 0.1% to 0.5%. Our estimates of the cost of business cycle fluctuations over a cumulative five-year period 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. Therefore, the question naturally arises as to what drives this difference. We show that the discrepancy is driven, at least in part, by the choice of the smoothing filter used to remove business cycle variation from the historical consumption series.[6] We use the widely used Hodrick-Prescott (HP) two-sided filter to obtain a long run trend consumption series from which fluctuations corresponding to business cycle frequencies (fluctuations lasting less than eight years) have been removed. Alvarez and Jermann (2004) (AJ), on the other hand, use a one-sided filter, whereby trend consumption at time-t is expressed as a weighted average of K lags, with the coefficients chosen so as to represent a low-pass filter that lets pass frequencies that correspond to cycles of eight years and more. Figure 4 presents a comparison of the HP and AJ filters. The figure plots the historical consumption growth (blue line), the trend consumption growth obtained using the HP filter (red line), and the trend consumption growth obtained using the AJ filter (black line). Consumption refers to the total personal consumption expenditures.

Figure 4. Comparison of HP and AJ Filters



Notes: The figure plots the historical consumption growth (blue line), the trend consumption growth obtained using the HP filter (red line), and the trend consumption growth obtained using the AJ filter (black line). Consumption refers to the total personal consumption expenditures. The sample is annual over 1929–2015.

[6]We thank Jaroslav Borovicka for pointing this out.

The figure shows that the HP filter delivers a smoother trend consumption growth relative to the AJ filter. In fact, it is known that a one-sided filter of the AJ type, with coefficients chosen to let pass frequencies that correspond to cycles of at least a given length, cannot fully eliminate higher frequency fluctuations. In other words, it also lets pass some fluctuations corresponding to higher frequencies. Consequently, in the context of the present application, the computed trend contains a non negligible amount of business cycle variability. In fact, the trend consumption growth is markedly different between the HP and AJ filters over our sample period. Specifically, the historical real consumption growth has a volatility of 3.4%, while the trend consumption growth obtained with the HP and AJ filters have volatilities of 1.9% and 2.8%, respectively. Thus, the trend growth obtained with the AJ filter has 46% higher volatility than that obtained with the HP filter. Therefore, not surprisingly, the cost of business cycles obtained with the HP filter are higher than those obtained with the AJ filter. As an illustration, when the I-SDF is recovered from the market portfolio, the cost of business cycle fluctuations over one- to five-year horizons takes values 0.9%, 2.1%, 4.8%, 5.5%, and 5.1%, respectively, with the HP filter. The corresponding costs obtained using the AJ filter are 0.5%, 1.2%, 2.6%, 2.2%, and 1.5%, respectively.

Separately, in Section VI, we present further evidence supporting the high cost of business cycle fluctuations, using an approach that does not involve a smoothing filter. Specifically, we rely on an extension of our information-theoretic methodology to obtain the (potentially) time-varying cost of *all* one-period consumption fluctuations in all possible states of the world. We show that the cost is strongly time-varying and countercyclical, reaching as high as 8.0% during one of the years of the Great Depression, having an average value of 5.8% during the four years 1930–1933 of the Great Depression, and having an average of 1.2% during the recent financial crisis of 2008–2009. Note that these estimates are for one-year fluctuations alone. The focus on all (as opposed to business cycle) fluctuations avoids the use of a smoothing filter. And the countercyclical nature of the cost is indicative of the high perceived costs of business cycles.

C. 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.A and V.B. For all the robustness tests, consumption refers to the total personal consumption expenditure.[7] The results are presented in Table 3.

First, we present the estimates for an alternative definition of relative entropy. Equation (15) 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 relative entropy can then be minimized to recover the measure, $\mathbb{F}$, and, therefore, the missing

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

component, ψ, of the pricing kernel:

(31)

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

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:

$$\hat{\psi}_t = \frac{e^{\hat{\theta}(\gamma)'\mathbf{R}_t^e(\Delta c_t)^{-\gamma}}}{\frac{1}{T}\sum_{t=1}^{T} e^{\hat{\theta}(\gamma)'\mathbf{R}_t^e(\Delta c_t)^{-\gamma}}} \quad \forall t, \tag{32}$$

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

$$\hat{\theta}(\gamma) = \arg\min_{\theta}\left[\log\left(\frac{1}{T}\sum_{t=1}^{T} e^{\theta'\mathbf{R}_t^e(\Delta c_t)^{-\gamma}}\right)\right]. \tag{33}$$

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 3. 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 2, Panel B – the cumulative costs of all one- to five-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 one- to five-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 EL and ET 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 3: 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 one- to five-year time horizons.

Consumption denotes the real *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 while satisfying the pricing restrictions (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, except for Row 3 where it extends over 1890-2015.

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

(34)

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

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 (34) 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 (34) and the EL and ET estimators defined in Equations (15) and (31), 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 risk neutral measured implied by a candidate model SDF.

The solution to Equation (34) is obtained as:

$$\hat{q}_t = \frac{e^{\hat{\theta}(\gamma)'\mathbf{R}_t^e} (\Delta c_t)^{-\gamma}}{\frac{1}{T}\sum_{t=1}^{T} e^{\hat{\theta}(\gamma)'\mathbf{R}_t^e} (\Delta c_t)^{-\gamma}} \quad \forall t, \tag{35}$$

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

$$\hat{\theta}(\gamma) = \arg\min_{\theta} \left[\log\left(\frac{1}{T}\sum_{t=1}^{T} e^{\theta'\mathbf{R}_t^e} (\Delta c_t)^{-\gamma}\right)\right]. \tag{36}$$

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 3, Row 1 of each panel) and EL (Table 2, Panel B, Rows 1-2) 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 Table 2 and Rows 1-2 of Table 3). 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 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.[8]

Overall, our results suggest that the estimates of the cost of aggregate economic fluctuations are fairly 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.

VI. 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 of the world. This is why the cost was defined as the ratio of the expected (or, average) prices of claims to a stabilized consumption stream and the actual aggregate consumption stream. For instance, the cost of all one-period consumption fluctuations (or, the welfare benefits of eliminating all consumption uncertainty for one-period) was defined as:

$$\frac{\tilde{pc}_1^{stab}}{\tilde{pc}_1} - 1 = \frac{\mathbb{E}^{\mathbb{P}}\left[\frac{V_t\left(C_{t+1}^{stab}\right)}{C_t}\right]}{\mathbb{E}^{\mathbb{P}}\left[\frac{V_t(C_{t+1})}{C_t}\right]} - 1. \tag{37}$$

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.A and V.B. 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 VI.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 fluctu-

[8]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.

ations. Subsection VI.B presents simulation evidence on the performance of the SEL estimator. Finally, Subsection VI.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

$$\frac{\frac{V_t\left(C_{t+1}^{stab}\right)}{C_t}}{\frac{V_t(C_{t+1})}{C_t}} - 1 = \frac{\mathbb{E}^{\mathbb{P}_t}\left[M_{t+1}\frac{C_{t+1}^{stab}}{C_t}|\underline{\mathcal{F}_t}\right]}{\mathbb{E}^{\mathbb{P}_t}\left[M_{t+1}\frac{C_{t+1}}{C_t}|\underline{\mathcal{F}_t}\right]} - 1 = \frac{\mathbb{E}^{\mathbb{P}_t}\left[M_{t+1}\left(1+\mu_c\right)|\underline{\mathcal{F}_t}\right]}{\mathbb{E}^{\mathbb{P}_t}\left[M_{t+1}\frac{C_{t+1}}{C_t}|\underline{\mathcal{F}_t}\right]} - 1, \tag{38}$$

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: $C_{t+1}^{stab} = (1+\mu_c)\,C_t$. Note that the difference between the average cost in Equation (37) and the time-t cost in Equation (38) is that, while the former involves the evaluation of *unconditional* expectations to obtain the average prices of the consumption claims, the latter requires the computation of the time-t prices of these claims as the *conditional* expectations of their discounted payoffs.

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 (the I-SDF) 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 I-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:

$$\mathbb{E}^{\mathbb{P}_t}\left[M_{t+1}\mathbf{R}_{t+1}^e|\underline{\mathcal{F}_t}\right] = \mathbb{E}^{\mathbb{P}_t}\left[(\Delta C_{t+1})^{-\gamma}\,\psi_{t+1}\mathbf{R}_{t+1}^e|\underline{\mathcal{F}_t}\right] = \mathbf{0}, \tag{39}$$

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

$$\mathbb{E}^{\mathbb{P}_t}\left[(\Delta C_{t+1})^{-\gamma}\frac{\psi_{t+1}}{E^{\mathbb{P}_t}(\psi_{t+1}|\underline{\mathcal{F}_t})}\mathbf{R}^e_{t+1}|\underline{\mathcal{F}_t}\right]=\mathbb{E}^{\mathbb{F}_t}\left[(\Delta C_{t+1})^{-\gamma}\mathbf{R}^e_{t+1}|\underline{\mathcal{F}_t}\right]=\mathbf{0}, \tag{40}$$

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 $\underline{X_t}\in\mathbb{R}^m$. Suppose that the historical realizations of consumption growth, excess returns, and the conditioning variables are given by $(\Delta c_t,\mathbf{r}^e_t,x_t)_{t=1}^T$,[9] 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 $\{f_{i,j};i,j=1,\dots,T\}$ is such that it 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: $\forall i\in\{1,\dots,T\},\quad\forall\gamma\in\Theta,$

$$\left(\widehat{f}^{SEL}_{i,\cdot}(\gamma)\right)=\underset{(f_{i,\cdot})\in\Delta_i}{\arg\max}\sum_{j=1}^T\omega_{i,j}\log(f_{i,j})\quad\text{s.t.}\quad\sum_{j=1}^T f_{i,j}\times(\Delta c_j)^{-\gamma}\,\mathbf{r}^e_j=\mathbf{0}, \tag{41}$$

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:

$$\omega_{i,j}=\frac{\mathcal{K}\left(\frac{x_i-x_j}{b_T}\right)}{\sum_{t=1}^T\mathcal{K}\left(\frac{x_i-x_t}{b_T}\right)}, \tag{42}$$

where $\mathcal{K}$ is a kernel function belonging to the class of second order product kernels,[10] and the bandwidth b_T is a smoothing parameter.[11]

Note that the objective function in Equation (41) is simply a 'smoothed' log-likelihood, with the constraints enforcing the conditional Euler equation restrictions in Equation (40). The weights $\omega_{i,j}$ used to smooth the log-likelihood are

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

[10] $\mathcal{K}$ 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 $[-a,a]$.

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

standard non-parametric kernel weights. The intuition behind the estimator may be understood as follows. Note that we are interested in recovering the conditional probabilities $f_{i,j}$, for $i, j = 1, 2, ..., T$. For each possible 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 states 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 42). 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 41). The states with zero weight, $\omega_{i,j} = 0$, each receive a conditional probability of zero.

The solution to Equation (41) is analytical and given by:

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

$$\widehat{f}_{i,j}^{SEL}(\gamma) = \frac{\omega_{i,j}}{1 + (\Delta c_j)^{-\gamma}\, \widehat{\theta}_i(\gamma)'\, \mathbf{r}_j^e}\,, \tag{43}$$

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

$$\widehat{\theta}_i(\gamma) = \underset{\theta_i \in \mathbb{R}^N}{\arg\max} \sum_{j=1}^{T} \omega_{i,j} \log\left[1 + (\Delta c_j)^{-\gamma}\, \theta_i'\, \mathbf{r}_j^e\right]. \tag{44}$$

Equations (43) and (44) 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 (41) can also be reformulated as:

$$\left(\widehat{f}_{i,\cdot}^{SEL}(\gamma)\right) = \underset{(f_{i,\cdot}) \in \Delta_i}{\arg\min} \sum_{j=1}^{T} \log\left(\frac{\omega_{i,j}}{f_{i,j}}\right) \omega_{i,j} \quad \text{s.t.} \quad \sum_{j=1}^{T} f_{i,j} \times (\Delta c_j)^{-\gamma}\, \mathbf{r}_j^e = \mathbf{0}, \tag{45}$$

The objective function in Equation (45) 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 $\mathbb{F}$ with respect to $\mathbb{P}$. 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, 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:

(46)
$$\frac{\frac{V_t\left(C_{t+1}^{stab}\right)}{C_t}}{\frac{V_t(C_{t+1})}{C_t}} - 1 = \frac{\mathbb{E}^{\mathbb{F}_t}\left[\left(\Delta C_{t+1}\right)^{-\gamma}\left(1+\mu_c\right)|\underline{\mathcal{F}_t}\right]}{\mathbb{E}^{\mathbb{F}_t}\left[\left(\Delta C_{t+1}\right)^{-\gamma}\left(\Delta C_{t+1}\right)|\underline{\mathcal{F}_t}\right]} - 1 = \frac{(1+\mu_c)\sum_{j=1}^{T}\widehat{f}_{t,j}^{SEL}\times(\Delta c_j)^{-\gamma}}{\sum_{j=1}^{T}\widehat{f}_{t,j}^{SEL}\times(\Delta c_j)^{1-\gamma}} - 1.$$

Finally, note that the question naturally arises as to the economic interpretation of the recovered ψ-component of the kernel. For instance, it could capture a misspecification of the power utility SDF. In fact, a large literature has developed based on modifying the preferences of investors, wherein the ψ-component could capture investors' habits (e.g., habit formation preferences), the unobserved return on total wealth (e.g., Epstein and Zin (1989) recursive preferences), or the ratio of durable to non-durable consumption or housing to non-housing consumption (e.g., preferences defined over different consumption bundles) to name a few. Alternatively, the ψ-component could capture investors' subjective beliefs about future macroeconomic and financial outcomes. Ghosh and Roussellet (2019) present evidence in favour of the latter interpretation. Specifically, they show that the recovered component is remarkably similar across a range of preference specifications. Moreover, consistent with the interpretation of ψ as capturing investors' beliefs, they show that the recovered beliefs about consumption growth have strong forecasting power for consumption growth and that the beliefs about the stock market are strongly related to survey data on institutional investors' confidence in the stock market.

B. *Performance of the SEL Estimator*

Before presenting the empirical results, we point out that the SEL approach is quite effective at recovering the ψ-component of the kernel for empirically relevant sample sizes. Ghosh and Roussellet (2019) show, via simulation exercises, that the SEL approach successfully recovers 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:

$$\mathbb{E}^{\mathbb{F}_t}\left[\underbrace{(\Delta c_{t+1})^{-\frac{\eta}{\rho}} R_{c,t+1}^{\eta-1}}_{M_{t+1}} (\mathbf{R}_{m,t+1} - R_{f,t+1}) \,|\underline{\mathcal{F}_t}\right] = \mathbf{0}, \tag{47}$$

where $R_{c,t}$ denotes the return on total wealth, ρ the elasticity of intertemporal substitution (EIS), and $\eta = \frac{1-\gamma}{1-\frac{1}{\rho}}$. 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: $\forall i \in \{1, \ldots, T\}$,

$$\left(\widehat{f}_{i,\cdot}^{SEL}\right) = \underset{(f_{i,\cdot})\in\Delta_i}{\arg\max} \sum_{j=1}^{T} \omega_{i,j} \log(f_{i,j}) \quad \text{s.t.} \quad \sum_{j=1}^{T} f_{i,j} \times (\Delta c_j)^{-\frac{\eta}{\rho}} r_{c,j}^{\eta-1}(r_{m,j} - r_{f,j}) = \mathbf{0}. \tag{48}$$

This procedure is then repeated 500 times to obtain the sampling distribution of the recovered $\mathbb{F}$.

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.[12]

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 (47) coincides with the physical measure $\mathbb{P}$, i.e. $\psi \equiv 1$ and $\mathbb{F} \equiv \mathbb{P}$. Ghosh and Roussellet (2019) show that Equation (48) 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

[12]The results are robust to alternative choices of the kernel function and the smoothing parameter within four standard deviations of the volatilities of the conditioning variables.

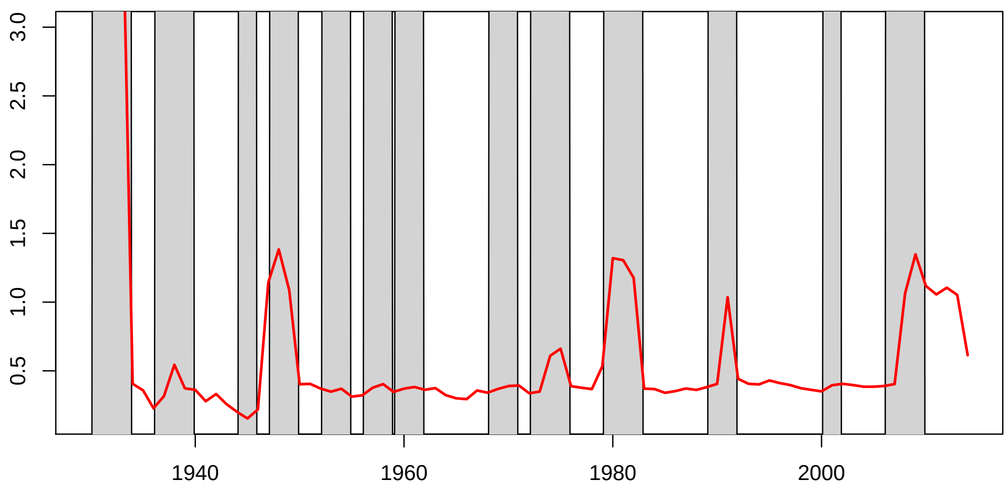
rate with a high degree of accuracy.

C. Empirical Results

We now proceed to use the SEL method to estimate the cost of aggregate consumption fluctuations in different states (or, times). We focus on the cost of all one-period consumption fluctuations, given by Equation (46).

We first estimate the time series of the cost 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 all consumption uncertainty in the subsequent year. In our implementation, we use nondurables and services consumption as the measure of the aggregate consumption expenditures and the excess return on the market portfolio as the sole test asset. Note that the SEL procedure requires the specification of the investors' conditioning set. In our baseline results, we use an exponentially-weighted moving average of lagged consumption growth as the conditioning variable.

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 as 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.

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 a maximum as high as 8.0%. In 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 all consumption uncertainty for one period alone.

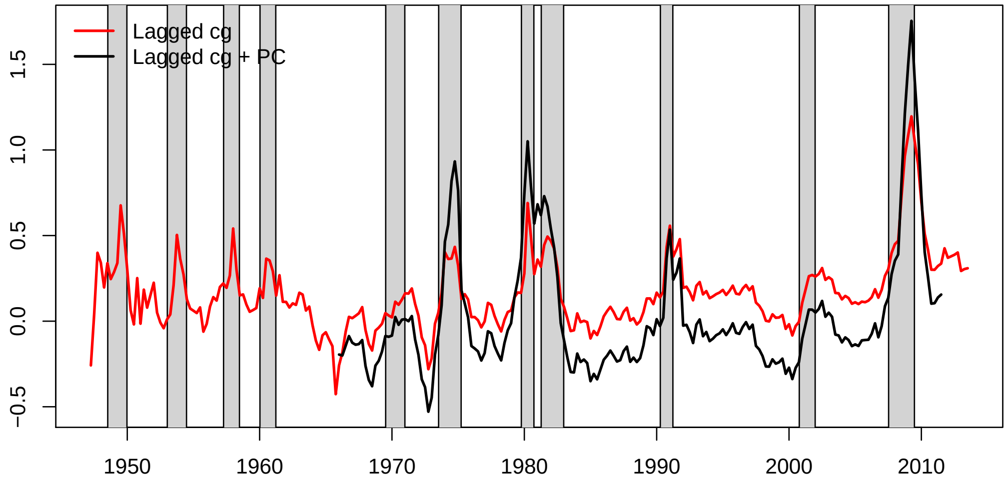
Next, to establish robustness of the results, we present the time series of the cost for alternative choices of the sample period, conditioning set, and the SDF parameter γ. First, note that our baseline results were obtained for the 1929-2015 sample period. This raises the potential concern that our results may be largely driven by the volatile prewar period, that included the episodes of the Great Depression and the aftermath of World War II (the two disaster macroeconomic episodes identified in Robert J. Barro (2006)). To mitigate this concern, Figure 6 presents the time series of the annualized cost (red line) using quarterly data over the postwar period 1947:$Q1$–2015:$Q4$. As in the baseline case, nondurables and services consumption is the measure of the aggregate consumption expenditures, the excess return on the market portfolio is the test asset, and an exponentially-weighted moving average of lagged consumption growth is the conditioning variable. The strong countercyclical variation in the cost is immediately evident from the figure. In fact, the countercyclicality is even more pronounced in the postwar period, compared to the longer 1929–2015 sample – the correlation with the recession dummy is 49.2% over the former period compared with 36.1% in the latter longer sample. Also, the magnitudes of the costs over the postwar subperiod are similar, regardless of whether the full 1929-2015 sample or the postwar period alone is used in the estimation of these costs. For instance, during the period of the Great Depression, 2008–2009, the cost of removing one-year fluctuations is estimated to be 1.20% on average using the longer sample, similar to the average cost of 0.86% obtained using the postwar sample.

As a second robustness check, we present results for an expanded conditioning set. Note that our baseline results were 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. Therefore, we estimate the time series of the cost when the conditioning set includes not only an exponentially-weighted average of past consumption growth, but also an exponentially-weighted average of a principal component extracted from a broad cross section of over a hundred macro variables. Specifically, we obtain panel data on 106 macroeconomic variables from Sydney Ludvigson's web site, based on the Global Insights Basic Economics Database and The Conference Board's Indicators Database. The variables cover six broad categories of macroeconomic data: output, labor market, housing sector, orders and inventories, money and credit, and price levels. We transform each variable to make it stationary and then extract a principal component from the cross section of transformed variables.[13] The time series of the

[13]We refer the reader to Ludvigson's website for a detailed description of these variables and the

cost is presented in Figure 6 (black line). Since data on the broad cross section are only available from the mid-sixties, the cost estimates start from 1966:$Q$1. The figure shows that the recovered time series of the cost seems quite robust to the choice of the conditioning set.

Figure 6. Time-Varying Cost: Robustness to Sample Period and Conditioning Set



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 and the excess return on the market portfolio as the sole test asset. The conditioning set consists of an exponentially-weighted moving average of lagged consumption growth (red line) and lagged consumption growth and a principal component extracted from a broad cross section of 106 *macro variables (black line). The sample is quarterly, covering the period* 1947:$Q$1–2015:$Q$4 *(red line) or* 1966:$Q$1–2015:$Q$4 *(black line).*

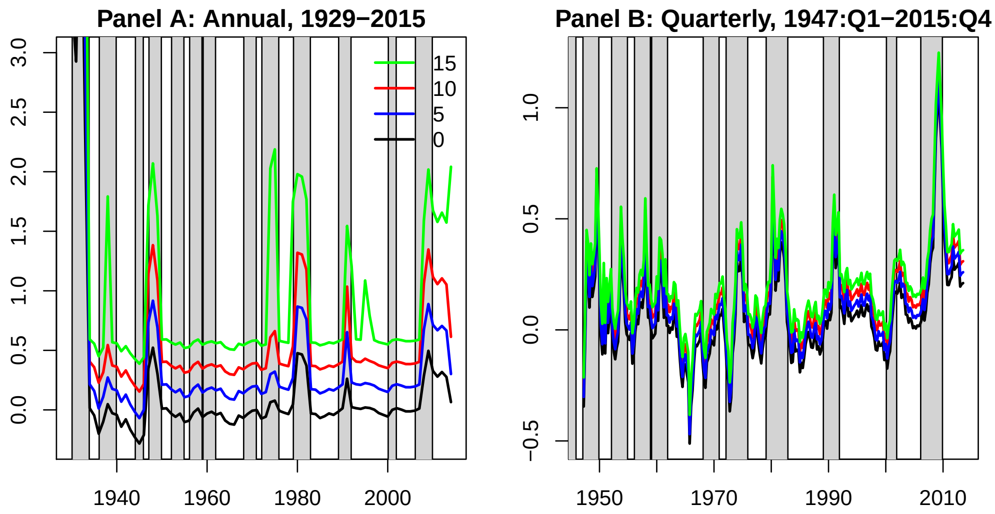
Finally, we present the time series of the cost for alternative choices of the SDF parameter γ. The results so far set $\gamma = 10$. In Figure 7, we plot the time series of the cost for alternative choices of γ. Specifically, we consider $\gamma = 0$ (black line), 5 (blue line), 10 (red line), and 15 (green line). Panels A and B present the results using annual data over 1929–2015, and quarterly data over 1947:$Q$1–2015:$Q$4, respectively.

Consider first the results obtained using quarterly data in Panel B. The panel shows that the estimated time series of the cost is quite robust to the choice of the power utility parameter. The estimates increase marginally with the γ parameter, but the differences are economically small. Somewhat bigger differences are obtained in Panel A that uses annual data over the longer available sample. These differences are largely driven by the very low consumption growth realizations during the Great Depression period that have a very large effect on the marginal

transformations applied to make them stationary.

utility for high values of the utility curvature parameter γ and make investors very concerned about the recurrence of this particularly bad state in future periods as well. However, the estimates of the cost are strongly countercyclical even in the limiting scenario when we set $\gamma = 0$ and thereby rule out the ex ante dependence of the SDF on aggregate consumption growth – the average cost during the four years 1930–1933 of the Great Depression is 3.5% for $\gamma = 0$ versus 5.8% for $\gamma = 10$ with a maximum value of 4.9% for $\gamma = 0$ versus 8.0% for $\gamma = 10$; and the average cost during the two years 2008–2009 of the Great Depression is 0.49% for $\gamma = 0$ versus 1.2% for $\gamma = 10$. Note that the estimates of the cost are large for both values of γ, given that these are the costs of one-year fluctuations alone.

FIGURE 7. TIME-VARYING COST: ROBUSTNESS TO SDF PARAMETER



Notes: The figure plots the time series of the cost of one-period consumption uncertainty, for different values of the SDF parameter γ. The black, blue, red, and green lines plot the cost for $\gamma = 0, 5, 10$, and 15, respectively. 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 as the 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 (Panel A) or quarterly, covering the period 1947:Q1–2015:Q4 (Panel B).

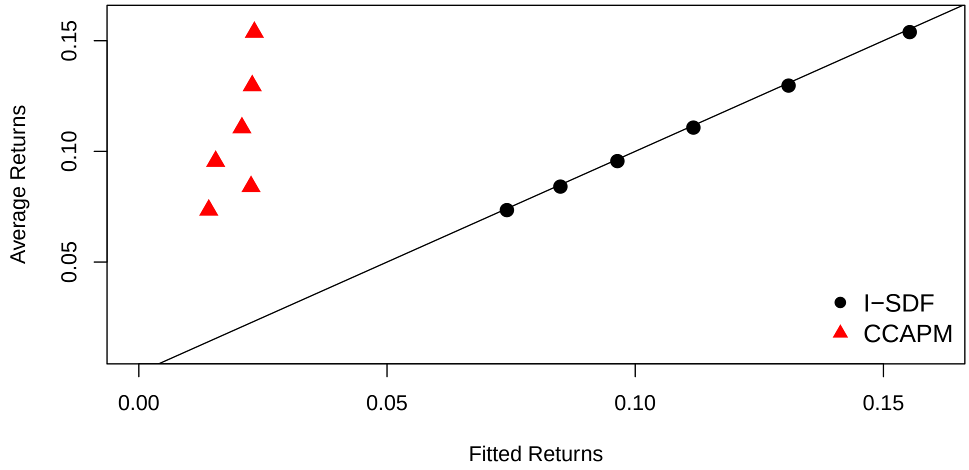
Overall, our results suggest that the cost of consumption fluctuations is strongly countercyclical and this offers, at least a partial, explanation of the high costs of business cycle fluctuations that we estimate in Sections V.A and V.B.

VII. 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. 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 8 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 $-\frac{Cov(M_t, R^e_{i,t})}{E(M_t)}$. 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 8. 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 1929-2015.

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.

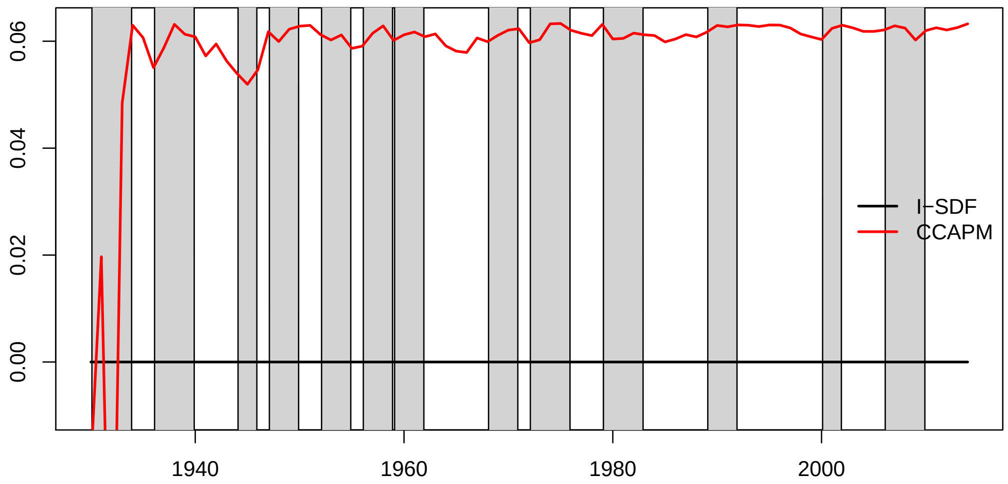
The above observations suggest that the CRRA model misses important components of the underlying sources of systematic risk. Since the welfare costs of consumption fluctuations depend critically on people's attitudes towards different sources of risk, the estimates of this cost obtained using the CRRA 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 8 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 (15)). 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.

We next show that, not only does the I-SDF price assets unconditionally delivering zero average pricing errors, it also produces zero conditional pricing errors. This is an important feature of the I-SDF that lends further support to the claim that it more accurately captures investors' attitude towards risk and, therefore, constitutes an attractive candidate kernel with which to measure the cost of aggregate economic fluctuations. Furthermore, the success of the I-SDF at pricing assets conditionally is not shared by most other candidate kernels. Stefan Nagel and Kenneth Singleton (2011) show that asset pricing models, even the ones that produce small average or unconditional pricing errors, typically produce large and volatile conditional pricing errors. They conclude that models are unable to simultaneously match the cross section and time series of asset returns.

Figure 9 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 VI, is 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 9 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 $\sum_{j=1}^{T} \omega_{t,j} \times (\Delta c_j)^{-\gamma} \mathbf{r}_{m,j}^{e}$. 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 9. Conditional Pricing Errors, 1929-2015

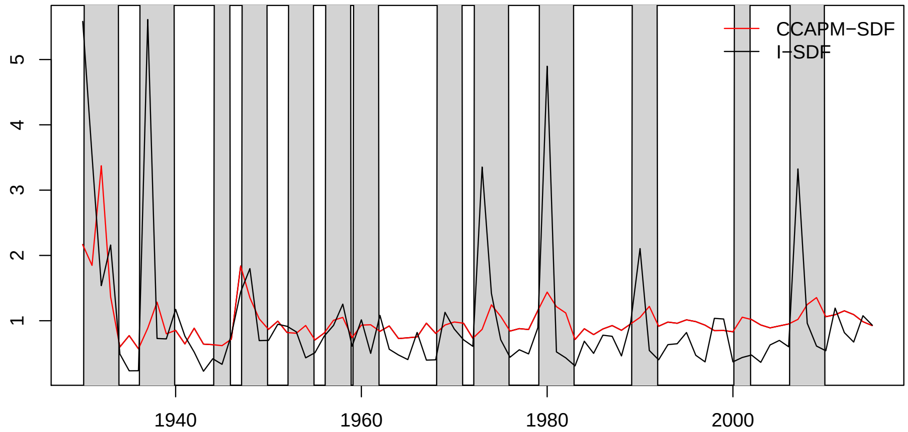


Notes: The figure plots the time series of conditional pricing errors for the excess stock market return, implied by the I-SDF (red line) and the CRRA kernel (green line). The I-SDF is extracted using the SEL method, with nondurables and services consumption as the measure of the consumption expenditures, 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 1929-2015.

Overall, the I-SDF seems to be more effective at capturing 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 10 plots the time series of the I-SDF (red line), recovered using the excess return on the market portfolio as the test asset, 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. The time series of the I-SDF looks very similar when it is recovered using the 6 FF portfolios as test assets. 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.

FIGURE 10. TIME SERIES OF THE SDF, 1929-2015



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 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.

VIII. 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 – unconditionally as well as conditionally, in-sample as well as out-of-sample – 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, the costs of business cycle fluctuations in consumption constitute a substantial proportion – typically between a quarter to a third – of the costs of all consumption uncertainty. Finally, using an extension of our information-theoretic methodology, we present evidence that the welfare benefits of eliminating 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 models. Second, the I-SDF has a strong business cycle component, suggesting that business cycle risk is an important source of priced risk.

Note that, while our results indicate that the cost of business-cycle fluctuations may be higher than previously thought, this does not imply that government policies intended to control these fluctuations are more desirable than previously thought. Even if government policies were effective in curbing fluctuations, one should not assume that the trend growth in consumption will be unaffected by them.

Finally, the present paper focuses on estimating the welfare costs of aggregate consumption uncertainty. However, our methodology is considerably general and may also be applied to obtain the costs of uninsurable idiosyncratic risk, such as labor income risk. This is left for future research.

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.

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.

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

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.

Blanchard, Olivier J. 2019. "Public Debt and Low Interest Rates." AFA Lecture.

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.

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

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 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 Ravi Jagannathan. 1991. "Implications of Security Market Data for Models of Dynamic Economies." *Journal of Political Economy*, 99(2): 225–62.

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.

Martin, Ian. 2008. "Disasters and the Welfare Cost of Uncertainty." *aerpp*, 98(2).

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.

Nagel, Stefan, and Kenneth Singleton. 2011. "Estimation and Evaluation of Conditional AssetPricing Model." *Journal of Finance*, , (3): 873–909.

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, and Martin Schneider. 2007. "Equilibrium Yield Curves." Vol. 21. Cambridge: MIT Press.

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.